https://doi.org/10.15388/vu.thesis.830
https://orcid.org/0000-0002-7027-394X

VILNIUS UNIVERSITY

Robertas Jurkus

Maritime Traffic Awareness Evaluation Using Deep Neural Networks

DOCTORAL DISSERTATION
Technological Sciences,
Informatics Engineering (T 007)

VILNIUS 2025

The dissertation was prepared between 2020 and 2024 at Vilnius University.

Academic Supervisor – Prof. Dr. Povilas Treigys (Vilnius University, Technological Sciences, Informatics Engineering – T 007).

This Doctoral Dissertation will be Defended at a Public Meeting of the Dissertation Defence Panel:
Chairman – Prof. Dr. Julius Žilinskas (Vilnius University, Technological Sciences, Informatics Engineering – T 007).
Members:

Prof. Dr. Olga Kurasova (Vilnius University, Technological Sciences, Informatics Engineering – T 007),

Prof. Dr. Pilar Martinez Ortigosa (University of Almeria, Spain, Technological Sciences, Informatics Engineering – T 007),

Prof. Dr. Dalius Mažeika (Vilnius Gediminas Technical University, Technological Sciences, Informatics Engineering – T 007),

Assoc. Prof. Dr. Viktor Medvedev (Vilnius University, Technological Sciences, Informatics Engineering – T 007).

The dissertation shall be defended at a public meeting of the Dissertation Defence Panel at 12:00 p.m. on the 26th of September 2025 in Room 203 of the Institute of Data Science and Digital Technologies, Vilnius University. Address: Akademijos g. 4, LT-04812, Vilnius, Lithuania, Tel. +370 5 210 9300; e-mail: info@mii.vu.lt.

The text of this dissertation can be accessed at the libraries of Vilnius University, as well as on the website of Vilnius University:
`https://www.vu.lt/naujienos/ivykiu-kalendorius`

https://doi.org/10.15388/vu.thesis.830
https://orcid.org/0000-0002-7027-394X

VILNIAUS UNIVERSITETAS

Robertas Jurkus

Jūros eismo informuotumo vertinimas naudojant giliuosius neuroninius tinklus

DAKTARO DISERTACIJA
Technologijos mokslai,
Informatikos inžinerija (T 007)

VILNIUS 2025

Disertacija rengta 2020–2024 metais Vilniaus universitete.

Mokslinis vadovas – prof. dr. Povilas Treigys (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – T 007).

Gynimo taryba:
Pirmininkas – prof. dr. Julius Žilinskas (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – T 007).
Nariai:

prof. dr. Olga Kurasova (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – T 007),

prof. dr. Pilar Martinez Ortigosa (Almerijos universitetas, Ispanija, technologijos mokslai, informatikos inžinerija – T 007),

prof. dr. Dalius Mažeika (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija – T 007),

doc. dr. Viktor Medvedev (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – T 007).

Disertacija ginama viešame Gynimo tarybos posėdyje 2025 m. rugsėjo 26 d. 12:00 val. Vilniaus universiteto Duomenų mokslo ir skaitmeninių technologijų instituto 203 auditorijoje. Adresas: Akademijos g. 4, LT-04812, Vilnius, Lietuva, tel. +370 5 210 9300; el. paštas: info@mii.vu.lt.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekose ir Vilniaus universiteto interneto svetainėje adresu:
`https://www.vu.lt/naujienos/ivykiu-kalendorius`

Abstract

According to the Global Marine Insurance Annual Report, marine casualties remain a major issue between human and non-human risk factors. Vessel collisions and anomalies at sea are among these factors. By analysing the massive historical data obtained from automatic identification systems (AIS), intelligent transport systems are being developed to address the challenges of predicting vessel trajectories. Most often, attempts are made to improve the accuracy of regression predictions by examining historical vessel behaviour, movement patterns, and similarities. Clustering techniques are also frequently adopted, as situational awareness of maritime traffic is a critical factor in maritime transport safety. Nevertheless, no globally agreed-upon solution has yet been proposed.

Situational awareness, which involves observing and interpreting the environment and making predictions about future dynamics, is perceived as the highest level of awareness (level 3). This dissertation investigates deep learning architectures and their hyperparameters, performs feature engineering on the derived data, and develops a recursive model that generates the most accurate prediction. In particular, this thesis investigates deep recurrent neural network architectures and builds models based on semi-structured AIS big data to extrapolate further geographic coordinates by regression. It has been empirically determined that the architecture with the highest accuracy is the multi-step multivariate Long short-term memory (LSTM) autoencoder, which includes vessel and meteorological observations. The thesis proposes solutions to improve the initial prediction and develop a generalised model. One suggestion would be to use different coordinate systems to determine vessel position locations. A second proposal is calculating the delta vector difference instead of absolute coordinates, and to recursively reconstruct sequence positions by adding those deltas to the last known point. In addition, a method is proposed to integrate categorical vessel-type data into a common dataset. Finally, an assessment of the uncertainty in prediction accuracy is made by combining statistical estimates such as ellipsoidal prediction regions (EPRs), conformal

prediction regions (CPRs) and prediction intervals (PIs), which allow us to estimate the boundaries of probabilistically overlapping regions and, therefore, to identify possible vessel collisions.

The dissertation examines one of the most recent vessel collisions in 2021 near Bornholm Island in the Baltic Sea as one of the verifications of the accident investigations. Using deep recurrent neural networks, the developed models can predict the subsequent multi-step trajectory of the vessel, which is combined with probabilities and statistics to form regions of prediction intervals with a confidence level of 95% to calculate collision risk estimates. The integration of confidence indicators shows that a non-parametric approach with conformal regions can detect most of the possible collision scenarios. The results of the real-world accident case studies confirm that deep learning models with advanced predictive capabilities can effectively improve navigational decisions by contributing to the prevention of maritime incidents.

Acknowledgements

I am profoundly thankful to my supervisor, Prof. Dr. Povilas Treigys, for his expert guidance and encouragement throughout this work. His keen academic insight, candid feedback, and generously shared wisdom have sharpened my thinking and elevated the quality of this dissertation. I would also like to thank my former mentor and now colleague, Dr. Julius Venskus, whose early guidance as my Master's thesis supervisor inspired me to pursue doctoral studies in the first place. His thoughtful critique and steadfast belief in my potential have been invaluable.

A special word of thanks goes to Dr. Jurgita Markevičiūtė for her meticulous consultations and significant contributions to the preparation of the scientific article. Finally, I am indebted to the lecturers of the Institute of Data Science and Digital Technologies and my colleagues at Klaipėda University (ISK) for their insightful advice and collaborative spirit, which have continually fueled my motivation and intellectual growth.

The author is grateful for the high-performance computing resources provided by the Information Technology Open Access Centre at the Faculty of Mathematics and Informatics of Vilnius University, as well as the Information Technology Research Centre.

Last but not least, I am deeply grateful to my family for their unwavering support, which has given me the time and space to focus on my studies and research.

Contents

List of Figures

List of Tables

List of Abbreviations

AE	**A**uto**E**ncoder
AIS	**A**utomatic **I**dentification **S**ystem
ARIMA	**A**uto**R**egressive **I**ntegrated **M**oving **A**verage
ARPA	**A**utomatic **R**adar **P**lotting **A**id
CI	**C**onfidence **I**ntervals
CPA	**C**losest **P**oint of **A**pproach
CPR	**C**onformal **P**rediction **R**egion
DBSCAN	**D**ensity-**B**ased **S**patial **C**lustering of **A**pplications with **N**oise
DL	**D**eep **L**earning
EPR	**E**lipsoidial **P**rediction **R**egion
GRU	**G**ated **R**ecurrent **U**nit
HELCOM	Baltic Marine Environment Protection Commission (**Hel**sinki **Com**mission)
IMO	**I**nternational **M**aritime **O**rganization
IUMI	**I**nternational **U**nion of **M**arine **I**nsurance
LSTM	**L**ong **S**hort-**T**erm **M**emory
ML	**M**achine **L**earning
MAEH	**M**ean **A**bsolute **E**rror of **H**aversine
MAIC	**M**arine **A**ccident and **I**ncident **I**nvestigation **C**ommittee
MMSI	**M**aritime **M**obile **S**ervice **I**dentity
MSE	**M**ean **S**quared **E**rror
OOW	**O**fficer **o**f the **W**atch
SOG	**S**peed **O**ver **G**round
SOM	**S**elf-**O**rganizing **M**ap
TCPA	**T**ime to the **C**losest **P**oint of **A**pproach
UTM	**U**niversal **T**ransverse **M**ercator
PI	**P**rediction **I**ntervals
RNN	**R**ecurrent **N**eural **N**etwork
VTS	**V**essel **T**raffic **S**ervices
WGS	**W**orld **G**eodetic **S**ystem

List of Symbols

α_t	Optimised scaling parameter for nonconformity scores
Accuracy	Percentage of time steps where actual points fall within predicted zones
C_i	Label-encoded category (vessel type) for sample i
C_j	One-hot encoded matrix for categorical vessel data
E	Embedding matrix for categorical vessel data
$E_i^{(t)}$	Embedding vector for vessel type at time step t
$I(a_i, z_i)$	Indicator function for evaluating coverage probability
j	Number of unique categories (vessel types)
Λ_i	Difference between actual and predicted coordinates at time step i
MAE	Mean Absolute Error
$MAEH$	Mean Absolute Error for Haversine Distance (trajectory error, km/m)
$MAPE$	Mean Absolute Percentage Error
MSE	Mean Squared Error
n	Total number of data points in a sequence
$P(\text{collision})$	Probability of collision between vessels
MAE	Mean Absolute Error
$P(\text{MAEH})$	MAEH is a custom regression metric designed to evaluate prediction accuracy in geospatial forecasting tasks. Unlike standard MAE, which operates on scalar values, MAEH measures the average absolute geodesic (Haversine) distance between the predicted and true geographic coordinates (i.e. latitude, longitude) at each time step in the forecasted trajectory.

	The error is computed between corresponding points of the predicted and actual sequences for each future step (from 0 to 19), and the final MAEH score is calculated as the average error over all time steps and across all evaluated trajectories. This metric captures spatial accuracy and is particularly suited for trajectory prediction tasks where positional accuracy on the Earth's surface is critical.
$RMSE$	Root Mean Squared Error
$\bar{R}_t$	Average nonconformity score at time step t
SD	Standard deviation of the data sample
Score	Sum of correct predictions within the predicted zone
t	Time step in a sequence
t_α	t-distribution critical value for confidence level α
δ	Desired confidence level for CPR
Δlat, $\Delta long$	Differences in latitude and longitude between consecutive steps
V_A, V_B	Areas of predicted zones for vessels A and B
$V_{A \cap B}$	Intersection area between two predicted zones
W_{CI}	Width of the Confidence Interval
W_{PI}	Width of the Prediction Interval
X_i	Input feature vector for vessel i (e.g., speed, position)
$X_{\Delta t}$	Time difference between consecutive steps (difference of vectors)
$X_{\delta Lat}$	Rate of change in latitude over time
$X_{\delta Lon}$	Rate of change in longitude over time
$x_i^{(t,k)}$	Feature k of vessel i at time step t
Y_i	Actual (true) multivariate coordinate output for vessel i
$\hat{Y}_i$	Predicted multivariate coordinate output for vessel i

1 Introduction

1.1 Research Relevance and the Need for Enhanced Maritime Monitoring

Shipping is vital for global trade, but growing vessel traffic increases collision, security, and environmental risks. Information on vessel locations, trajectories, and destinations underpins effective maritime safety and environmental protection. According to the International Union of Marine Insurance (IUMI), about 10% of losses in 2019 stemmed from collisions, with additional losses from crew errors, equipment failures, and weather conditions [81]. Despite intense activity, shipping volumes continue to rise: in 2020, marine underwriting premiums reached USD 30.0 billion, a 6.1% increase from 2019, with cargo vessels accounting for 57.2% of premiums [80]. This surge underscores the need for enhanced monitoring to mitigate safety and security threats.

Rising vessel activity produces vast amounts of Automatic Identification System (AIS) data that exceed the capacity of human operators and traditional machine learning (ML) methods to process in real time. Effective situational awareness has three levels: initial levels focus on assessing the surrounding environment, while Level 3 involves perception and prediction of future states [63]. Achieving Level 3 requires advanced techniques such as deep learning (DL) to handle large, noisy time series. Recent incidents highlight this need: in March 2025, the oil tanker MC Stena Immaculate collided under fog off the Humber Estuary; in late 2018, the frigate Helge Ingstad struck a tanker in Norway; and the Nord Stream pipeline explosion disrupted Baltic Sea traffic. Such events show how quickly congested or strategic waterways can become hazardous, underscoring the urgency of developing reliable multi-step trajectory prediction models and real-time monitoring tools.

Sonar and radar help vessels detect nearby obstacles, but radar can miss smaller objects behind larger ones, and both systems are limited in low visibility or cluttered settings. Many ships also carry AIS transponders, but real-time AIS

updates can be delayed by minutes and cannot by themselves forecast future movements. In contrast, historical AIS records combined with meteorological and categorical vessel type data are well-suited for DL methods. Recurrent neural networks (RNNs), especially multi-step multivariate models, can learn from these time series to predict a vessel's trajectory and flag potential collisions or abnormal manoeuvres. By leveraging AIS signals in DL models, RNN-based forecasting can improve situational awareness and support proactive safety measures in increasingly congested maritime environments.

1.2 Research Problem

There is a growing need for accurate, long-term vessel trajectory forecasts to support maritime safety and traffic management. Surveys of ML and DL in land-based traffic systems (e.g., [4]) show that 45% of efforts target congestion prediction and 30% focus on flow management approaches that assume homogeneous vehicles and stable data. In contrast, maritime AIS streams are irregular, vessel types vary widely, and environmental factors (e.g., wind, currents) strongly influence movement. Large, loaded ships may require 20–25 minutes to stop (completely), so precise forecasts over that horizon are essential for timely risk mitigation [38].

Traditional monitoring tools, such as radar, sonar, and AIS, provide only current state awareness and can be misleading in congested or low visibility conditions: radar can miss smaller targets behind larger objects, and AIS transmissions can lag or drop, leaving gaps in coverage. Historical AIS data provide rich trajectory records, but their irregular sampling intervals and noise make simple extrapolation unreliable beyond a few minutes, and AIS data do not provide an awareness analysis.

Many existing maritime models rely on short-term or linear predictions, apply Universal Transverse Mercator (UTM) transformations without justifying their impact, or build separate predictors for each vessel type, requiring extensive pre-processing and failing to capture multivariate dependencies such as speed, heading, and environmental inputs. As a result, there is no generalised approach to handle all vessel type behaviours. Deep RNNs can learn

long-term patterns from noisy AIS time series; however, prior work has not fully integrated categorical vessel-type labels and meteorological features into a single model, nor addressed the compounding error in multi-step forecasts.

Nowadays, collision-risk assessment typically relies on deterministic metrics such as the Closest Point of Approach (CPA) and Time to CPA (TCPA), which calculate the single future point at which two vessels will be closest to each other within a fixed guard-zone radius (e.g., 0.25 NM or 15 minutes). While CPA/TCPA are widely used for real-time collision assessment, they ignore uncertainty in each vessel's predicted path, cannot handle overlapping trajectories among multiple ships, and make it difficult to assign a meaningful probability to a collision. To address these limitations, the proposed methodology evaluates clusters of uncertainty and extends beyond CPA's fixed-point analysis, offering a broader assessment of future vessel trajectories. By incorporating a 20-minute prediction window aligned with the stopping times of large vessels and forecasting dynamic trajectory boundaries, this approach replaces static proximity and time thresholds with probabilistic overlap metrics that better capture the range of possible interactions.

This dissertation develops a unified framework that leverages deep RNNs to produce accurate, multi-step vessel trajectories from historical semi-structured AIS data; integrates vessel-type and meteorological features into a single model to improve generalisation across all vessel classes; and replaces single-point CPA/TCPA rules with probabilistic risk zones that account for model and data uncertainty. By combining these elements, the proposed approach enhances real-time maritime situational awareness and provides actionable risk estimates when they are needed most.

1.3 Research Object

AIS-based multivariate, multi-step vessel trajectory data for prediction and collision-risk assessment. RNN methods are the application tool to study the object.

1.4 Aims and Objectives of the Research

The aim of this research is to propose and investigate deep neural network-based algorithms for multi-vessel trajectory prediction and evaluate collision risk in maritime navigation.

To achieve this aim, the following objectives are established:

1. Conduct a literature review on multi-step multivariate vessel trajectory prediction methodologies, examining different modelling approaches, data representations, and their applications in maritime situational awareness and safety.
2. Develop and compare RNN architectures for long-term vessel trajectory prediction, evaluating their robustness, hyperparameter sensitivity, and optimisation criteria to determine the most effective model configurations for vessel trajectory prediction.
3. Assess the impact of vessel trajectory prediction accuracy on categorical, meteorological, and spatial data, including coordinate system transformations, and identify the techniques for processing categorical data.
4. Evaluate and identify effective proposed techniques for assessing vessel collision risks by using model uncertainty quantification to measure forecast reliability, and applying deterministic statistical approaches to detect and quantify potential trajectory overlaps indicative of collision scenarios.
5. Validate collision detection techniques through an empirical study using unseen vessel trajectory data, having actual historical sea incidents.

1.5 Scientific Novelty and Practical Value

This research introduces several novel methods that together advance vessel-trajectory forecasting and collision-risk assessment in maritime navigation. Its main contributions are:

1. Identification of the LSTM autoencoder (AE) as the most effective supervised RNN for multi-step vessel-movement extrapolation and discovery of a critical cell-count threshold. Comparing various RNN architectures on time-series regression shows that the LSTM AE best captures long-term dependencies. Using between 75 and 300 LSTM cells optimises accuracy, while additional cells yield negligible gains yet dramatically increase computational cost, leading to an efficient model design.

2. Unified integration of vessel-type and meteorological data alongside AIS minimises pre-processing and improves generalisation across all ship classes. Embedding categorical vessel types together with environmental factors (e.g., wind speed, currents) in a single dataset ensures that forecasts account for real-world conditions, boosting accuracy and robustness in diverse operating environments.

3. Instead of predicting absolute coordinates directly, the model forecasts position vector differences between consecutive time steps. These predicted deltas are recursively added to the current known position to reconstruct the trajectory. This approach, applied in both World Geodetic System (WGS84) and UTM coordinate systems, helps reduce cumulative error, improves spatial continuity, and maintains consistency over long-term forecasts without requiring complex drift correction mechanisms.

4. Introduced non-parametric conformal prediction regions (CPR) for vessel collision detection by constructing probabilistic risk zones from an ensemble of LSTM AE forecasts. Sampling varied hyperparameter configurations generates a non-Gaussian distribution of possible positions, and CPR envelops these predictions without assuming any specific error distribution. When two vessels' CPR zones overlap, a collision risk score is computed, offering a region-based metric that captures uncertainty across multiple forecasts rather than a single-point CPA/TCPA estimate. This novel application of CPR directly leverages model variability and noisy data to provide more informative collision-risk assessments.

1.6 Claims to be Defended

The following research-based claims are defended in the thesis:

1. When predicting a vessel's trajectory with LSTM in a neural network architecture, selecting the number of cells from a certain threshold no longer affects, or has a very minimal effect on, the accuracy of the prediction, but increasing the number of cells significantly slows down the model training.
2. In multi-step multivariate vessel trajectory extrapolation, recursive recalculation of the prediction time series from the previous coordinate allows a more accurate determination of the next prediction point, especially in the initial prediction steps.
3. The inclusion of vessel types and meteorological information in the common training dataset of the LSTM recurrent multi-step multivariate neural network improves the accuracy of vessel trajectory prediction when using an embedded encoding approach.
4. CPR allows the detection of vessel collision boundaries with the highest statistical probability at the 95th confidence level when the data are multivariate and do not contain a normal distribution.

The structure of the dissertation is described and presented in the following way. The introduction chapter presents the general situation of maritime traffic, the topic's relevance, the study's aim, the issues, the thesis statements and a workflow diagram of the whole research. Chapter 2 conducts a literature review of the subject area and examines DL recurrent network architectures, categorical data encoding techniques and prediction intervals for uncertainty detection. Chapter 3 presents a recurrent LSTM AE approach, which incorporates categorical and meteorological data and creates prediction regions for vessel collision detection. Chapter 4 describes the research data, its processing and the stages of model development. This chapter also describes the course of the empirical experiment. Chapter 5 presents the results, and the thesis is finalised with the general conclusions.

The thesis consists of 158 pages, with the summary in Lithuanian starting from page 129, 7 tables, 40 figures, 7 tables, 5 appendices, and 122 references.

The research process can be divided into the main parts (see Fig. 1.1): 1. Introducing the source and structure of the data; 2. The raw data pre-processing step, where the data are cleaned, filtered, normalised and structured into fixed sequences (vectors) of equal length suitable for RNNs; 3. Presenting methods that empirically attempt to improve the prediction of vessel trajectory through so-called feature engineering, where additional information is created or manipulated from existing features; 4. Adapting recurrent network architectures, providing network hyperparameters (network, training, architecture) and including a hidden layer of categorical and meteorological data with vessel types; 5. The developed models are evaluated by a test sample using classical regression metrics, with the derivative of the mean absolute value of the Haversian (MAEH) distance used to determine the error; 6. The developed models are practically applied to find prediction regions and calculate the probability uncertainties of vessel collisions when the experiment evaluates the coverage probability with the sequences of the test sample and validates on a previously unseen real marine incident.

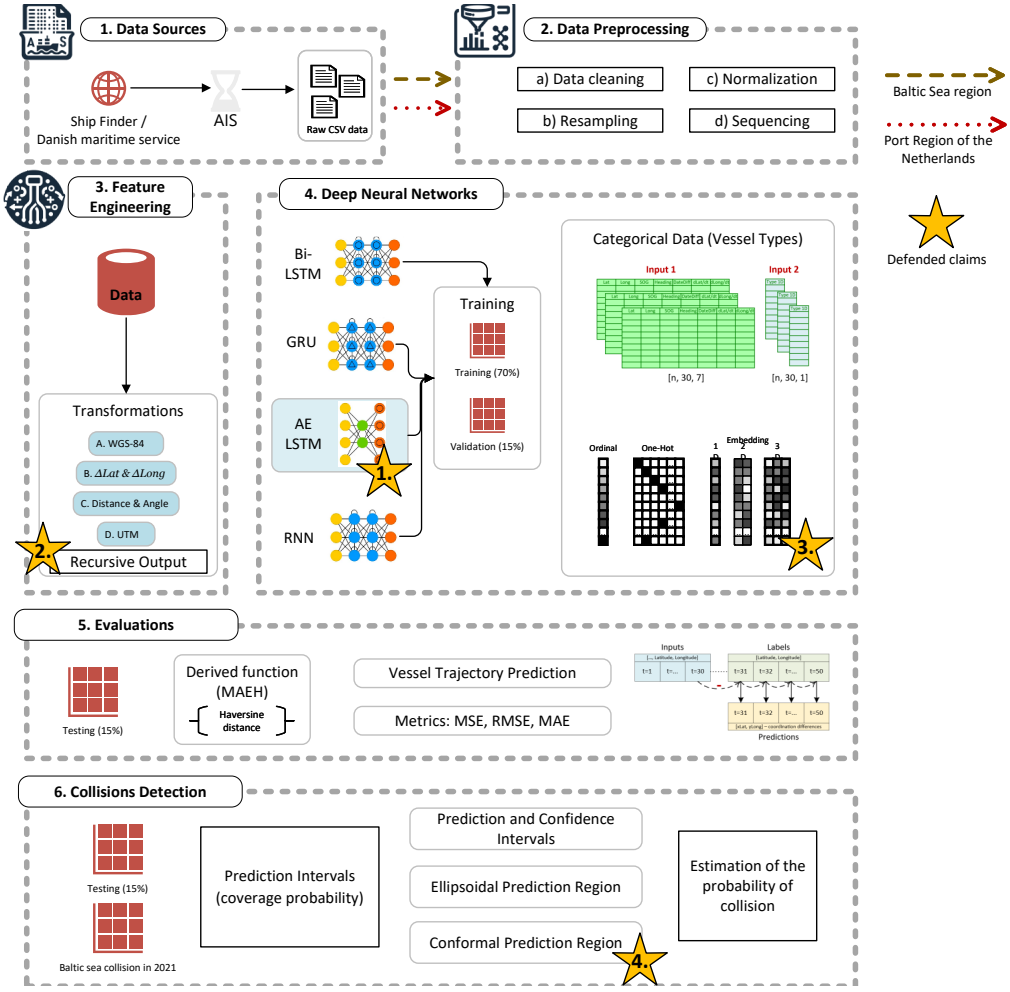


Figure 1.1: Research workflow diagram.

2 Related Studies on Maritime Traffic Prediction and DL Approaches

This chapter reviews scientific literature on vessel trajectory prediction and maritime situational awareness, including related thesis work. It then examines algorithmic approaches, particularly deep RNNs, and their integration into maritime situational awareness. The section also explores various RNN architectures, including basic LSTM, bidirectional LSTM, simple RNNs, GRU, stacked LSTMs, and LSTM AEs, explaining their theoretical foundations and practical uses. Additionally, it covers categorical data encoding techniques such as ordinal encoding, one-hot encoding, and multidimensional embedding, which are essential for integrating multi-vessel types and meteorological data into prediction models. The methods for determining prediction boundaries are also discussed, which are important for assessing the uncertainty and reliability of model predictions. The findings of this chapter were presented at scientific conferences [D.1–D.5].

2.1 Incidents, Accidents, and Risk in Maritime Transportation

Within the safety and risk management framework, particularly pertinent to maritime transportation, workplace safety, and emergency services, the terms "incident" and "accident" are frequently invoked. While these terms are interrelated and often used interchangeably in colloquial discourse, they possess distinct technical meanings. An incident is characterised as an event or circumstance with the potential for causing harm or disruption, irrespective of the materialisation of actual harm. In contrast, an accident is an event that has resulted in detrimental outcomes, which may include injuries, property damage, or loss. The diligent management of incidents is critical as it serves as the first line of defence in preventing their escalation into accidents and mitigating the risk potential. Incidents, particularly those involving vessel

movements, can escalate into collisions if not addressed promptly. The Marine Accident and Incident Investigation Committee (MAIC) categorises accidents based on severity, from very serious marine casualties to serious and less serious ones, down to marine incidents [60]. Such categorisations highlight the importance of early detection and intervention, where accurate vessel trajectory predictions and collision risk assessments play a pivotal role. By anticipating vessel movements and identifying potential anomalies or collision risks, decision-makers can act proactively to prevent incidents from escalating into accidents. These classifications consider the extent of damage inflicted upon the vessel, the environmental repercussions, and the effects on the crew. The precise delineation of these terms establishes a foundation for regulatory practices and guides the systematic investigation and reporting protocols. This taxonomy is instrumental in shaping the strategies for accident prevention and fostering a culture of safety within the maritime industry.

Despite all the information systems, specific rules at sea should be considered and followed in certain situations to avoid collisions. For example, in open waters, it is an accepted rule that vessels should maintain a distance of 1 to 1.5 nautical miles apart CPA to avoid creating a hazard when passing each other. These rules are reminiscent of road traffic rules outlined in the COLREG (Convention on the International Regulations for Preventing Collisions at Sea, 1972). The guidelines also specify how vessels should act to avoid collisions if at least one vessel makes a slight course change. For instance, Rule 14 head-on situation says, "When two power-driven vessels are meeting on reciprocal or nearly reciprocal courses to involve risk of collision, each shall alter her course to starboard so that each shall pass on the port side of the other". If the captain has not determined whether there is a danger and a real threat, the situation is considered hazardous, and all possible actions must be taken to avoid a disaster. In addition to Rule 5 (look-out): "Every vessel shall at all times maintain a proper look-out by sight and hearing as well as by all available means appropriate in the prevailing circumstances and conditions to make a full appraisal of the situation and or the risk of collision". Even so, there are situations where human errors occur and rules are broken, which is why additional systems are needed. Accurate predictions made 20–25 minutes ahead are particularly critical in such cases, as they allow enough time for vessels to

assess risks and take corrective action. This is especially important considering that large, loaded vessels, such as tankers, may require up to half an hour to come to a complete stop, highlighting the need for timely trajectory forecasts to prevent collisions [38].

The consequences of collisions can be severe, particularly regarding loss of life, including crew members, and environmental pollution from fuel or cargo spills. If pollutants are spilt into the waters (whether local or open), the vessel's captain must inform local authorities immediately. In the event of a collision resulting in hull damage, the time and cost required for repairs are substantial. Even minor defects can take one to six months or longer to repair. A damaged vessel loses its offshore permit and, after restoration, must pass all inspections to regain it. Larger shipping companies are better equipped to handle these financial challenges, while smaller companies may face the risk of bankruptcy. Compared to other means of transport, such as cars or aeroplanes, the processes involved are significantly more complex, underscoring the importance of minimising such incidents.

In a comprehensive examination of marine safety over the past three decades, Polish researcher Magdalena Bogalecka et al. [9] provided an interesting analysis of ship accidents in the Baltic Sea. Their work meticulously documents a series of specifically chosen accidents, including those involving the vessels *Dan Trimmer*, *Eagon W*, *Breant*, and *Victoria Seaways*, among others. These events have been uniformly classified as serious within the parameters set forth by the MEPC.3/Circ.3 Convention [36]. Notably, the predominant causative factor identified in nearly all these cases was severe meteorological and climatic conditions. The Baltic Sea, which is the focus of this research, remains one of the most active yet environmentally sensitive maritime regions. More about potential accidents and the situation in the Baltic Sea are investigated in the article [90].

The Baltic Sea Environment Protection Commission's (HELCOM) 2020 report provides a pertinent overview of maritime casualties within the region, indicating that cargo vessels were predominantly involved in such incidents. In 2020 alone, they accounted for 51% of all reported maritime casualties, amounting to a total of 125 incidents. Additionally, passenger

vessels accounted for 29% of the casualties, with 74 reports logged [102]. These statistics demonstrate that cargo and passenger vessels represent the bulk of maritime traffic in the Baltic Sea. A significant observation is that most collisions occur in areas of port activity, or when vessels navigate their approach or within the confines of the ports themselves, contributing to 30% of all maritime accidents. Ports represent highly congested and dynamic environments where frequent vessel manoeuvres such as stopping, turning, and docking are necessary, increasing the risk of collisions. Accurate trajectory predictions in such regions are critical for collision avoidance, as vessels must account for limited manoeuvring space, interactions with other vessels, and static obstacles. However, despite the high collision rate in ports, maritime incidents are not limited to these areas. Offshore collisions also pose substantial risks, as demonstrated by HELCOM reports and real-world events, such as the 2021 collision near Bornholm. Therefore, this study incorporates both port-adjacent (Netherlands region) and offshore (Baltic Sea) scenarios to ensure the applicability of the proposed methods across diverse maritime environments. Collisions have been further categorised by the nature of the entity they involve, which includes collisions with other moving vessels, static objects such as bridges, docks, or breakwaters, and a range of other causes such as being adrift, on fire, containment loss, or flooding. This stratification of collision types offers a nuanced perspective on maritime accident dynamics and is integral to developing targeted safety measures and regulations.

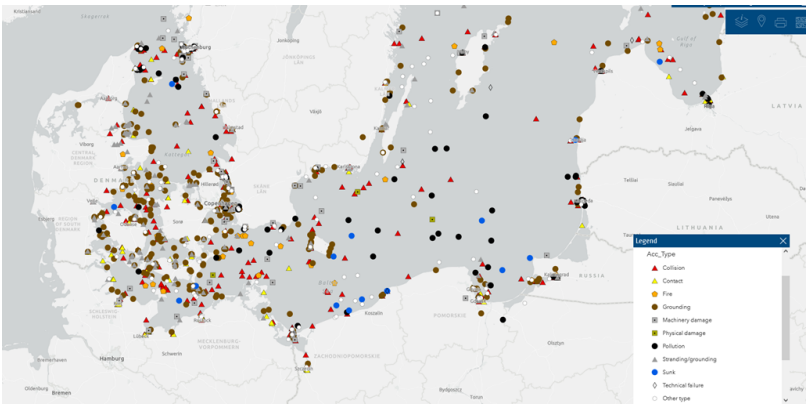


Figure 2.1: HELCOM maritime incidents map in the Baltic Sea.

HELCOM provides an interactive map (see Fig. 2.1) that shows different incidents in the Baltic Sea, including collisions, various contacts, fires, groundings, pollution, etc. For example, on December 13, 2021, a significant maritime incident occurred in the Baltic Sea, where two cargo ships collided near the Danish island of Bornholm, off the coast of the southern Swedish town of Ystad. By forecasting and predicting the trajectory of a particular vessel, it would have been possible to assess the risk of a collision or to detect abnormal behaviour in ongoing traffic. Although increasingly complex constructions of deep neural networks are emerging, LSTM networks are still being analysed and applied to solve the problem of maritime traffic forecasting [16, 29, 31]. This research aims to investigate integrating DL algorithms in improving maritime navigational decisions and the progression of risk management techniques to enhance maritime safety.

2.2 Review of Algorithms and Related Research in Maritime Awareness

Traditional approaches to vessel trajectory prediction and maritime awareness rely on classical algorithms, such as regression models (linear regression, polynomial regression, ARIMA), clustering techniques (k-means, DBSCAN, hierarchical clustering, SOM), and physics-based simulations (Kalman filters, Particle filters, Monte Carlo simulations). These methods have been widely used to estimate vessel movement, identify navigational patterns, and assess collision risks. While interpretable and computationally efficient, these models struggle to handle large-scale AIS data, complex vessel interactions, and highly dynamic maritime environments. As maritime traffic continues to increase, the need for more scalable and adaptive predictive models becomes evident.

DL is a subset of ML and AI, leveraging multi-layer neural networks to model and understand complex data patterns. Modern DL algorithms can process vast amounts of data, making them essential for automating vessel movement prediction in intensive shipping environments. Utilising AIS satellite data, these algorithms provide crucial support for monitoring and managing the

increasing maritime traffic and its participants, thereby enhancing maritime safety and operational efficiency.

This thesis applies DL architectures, specifically RNNs, to predict future vessel movements. The architectural model is sequence-based, capturing how data change over time. The investigation focuses on identifying the optimal recurrent network architecture and hyperparameters to achieve accurate prediction results. Various recurrent network architectures, including fully connected (simple) RNNs, basic (vanilla) RNNs, bidirectional LSTM networks, stacked LSTM networks, LSTM AE, and gated recurrent units (GRU), are tested for their inherent ability to model temporal dependencies and retain long-term memory. Recurrent networks excel at recognising and memorising longer patterns in sequential data, compared to algorithms like a multi-layer perceptron or similar, which is valuable for vessel trajectory prediction, where historical movement plays a significant role in forecasting future positions. Prediction accuracy for each architecture is evaluated by varying the number of cells in the hidden layer, incorporating meteorological data, and applying various coordinate system transformations to enhance prediction quality. While transformer models with attention mechanisms have shown great success in fields like natural language processing, preliminary experiments conducted during this study indicated no significant improvement over RNNs for the AIS dataset used. Due to their higher computational demands when processing extensive time series data, transformers were not explored in greater depth; however, further detailed analysis could be undertaken as part of future research. Meanwhile, transformers often require larger training datasets and substantial fine-tuning to achieve comparable performance for long-term time-series predictions. One of the articles [85] presents results where the attention mechanism did not improve the trajectory prediction. In contrast, recurrent networks are well-suited for time-dependent tasks and provide a balance between computational efficiency and prediction accuracy, making them more appropriate for the specific requirements of vessel trajectory forecasting in this thesis.

Additionally, categorical vessel type data and meteorological data are incorporated into the models to enhance trajectory predictions. Different

categorical encoding techniques, such as ordinal, one-hot, and multidimensional embedding, are employed to assess the impact on prediction accuracy. Long-short memory models are used for recursive multi-step forecasting across these encoding techniques.

Two datasets are utilised in this research. The first dataset, from the Netherlands (North Sea) coastal region, tests different RNN architectures and determines the most accurate prediction model. Results show that the bidirectional LSTM and LSTM AE network architectures provide the most precise vessel trajectory predictions, with minimal variance even with the most minor cell selection. The AE network architecture demonstrates reliable performance, depending on the appropriate cell size selection, particularly with increases in 100 and 150 cells. Experiments, arguments and discussions from all the studies are presented in the results section of this thesis.

The second dataset focuses on the Baltic Sea region, which was selected due to its high maritime activity and complexity. The previously identified top three RNN architectures - LSTM AE, bidirectional LSTM, and GRU - were rigorously re-evaluated and tested with the new AIS dataset. These experiments involved deeper investigations, including extensive hyperparameter fine-tuning, comparative analyses of various spatial coordinate system transformations (WGS84 and UTM projections), integration of meteorological data, and the application of advanced vessel collision detection techniques. Moreover, the models were validated using real-case maritime incidents, notably the 2021 collision between the cargo ships *Scot Carrier* and *Karin Hoej*, demonstrating their practical applicability and reinforcing their effectiveness in real-world maritime safety scenarios.

To ensure the reliability of the predictions, deep RNN models are combined with statistical uncertainty quantification techniques to produce reliable collision risk assessments. Bounds are constructed and compared based on prediction and confidence intervals (PIs, CIs), ellipsoidal prediction regions (EPRs), and conformal prediction regions (CPRs). These techniques are applied to simulated test scenarios and actual case studies to validate the effectiveness of approaches.

The findings indicate that DL models with advanced predictive capabilities and integrated reliability indicators can significantly enhance navigational

decisions and pre-emptive maritime safety strategies. The integration of categorical vessel-type data, meteorological data, and predictive trajectory bounds ensures accurate and reliable vessel trajectory predictions. This approach encourages a shift toward more proactive, AI/ML-enhanced maritime risk management protocols, ultimately contributing to the prevention of incidents and promoting safer navigation.

Related Thesis Work

To demonstrate relevance in maritime awareness at the doctoral level, the following PhD dissertations in this field will be examined.

One of the earliest doctoral theses in this field, presented by Philipp Last [45] in 2016 at Jacobs University Bremen, Germany, primarily addressed short-range predictions. These predictions visualised potential vessel movements within a short time frame, helping mariners identify collision risks. The system aimed to enhance decision-making for collision avoidance by indicating the radius of potential vessel movements with uncertainty bands, allowing mariners to take preventive actions. The thesis did not employ explicit dynamic models such as collision risk probability metrics or ship domain infringement evaluations. Instead, it relied on visual analysis tools and a background model derived from AIS data to flag abnormal behaviours, such as vessels deviating from traffic directions or leaving restricted waterways. These features were intended to assist vessel traffic services (VTS) operators in identifying and addressing risky situations, indirectly contributing to collision detection.

The study primarily focused on Class A AIS systems used by professional vessels, including cargo ships, tankers, and passenger ships. Its approach did not involve learning from data or adopting pattern-based models (e.g., supervised learning, clustering, or neural networks). Instead, it utilised fixed mathematical rules derived from the vessel's physical state and dynamic AIS data, classifying it as rule-based or physics-based modelling rather than ML. The prediction algorithm was a motion model based on dynamic AIS data, such as speed over ground (SOG), course over ground (COG), and rate of turn (ROT). Linear extrapolation with adjustments for dynamic parameters formed the basis of this model. However, the accuracy of predictions depended heavily on the

consistency of AIS reporting intervals and the availability of dynamic data fields (e.g., ROT and heading). Missing or inconsistent data increased uncertainty and led to prediction drift.

In 2021, Lei Du [25] defended a doctoral thesis at Aalto University, Finland, introducing a novel framework for maritime traffic risk assessment, emphasising near-miss detection based on ship manoeuvres. The research integrated the non-linear velocity obstacle (NLVO) algorithm to model the spatiotemporal relationships between vessels, projecting collision risks into velocity domains. This method identified conflicting velocities, termed the velocity obstacle zone, to determine whether a vessel's current trajectory posed a collision risk. The framework used AIS data to calculate multiple risk indicators, such as perceived navigator risk, action quality, and compliance with COLREGs, tailored for both ship pairs and multi-vessel encounters. The study highlighted that collision risks were concentrated in high-density traffic, such as the Gulf of Finland and waterways near Stockholm and Turku. It also incorporated vessel attributes, such as type, size, and manoeuvrability, to customise risk assessments. For instance, passenger ships were found to adopt earlier evasive manoeuvres than cargo ships and tankers, reflecting different risk perceptions and operational strategies.

However, the thesis assumed that vessels primarily change course, not speed, to avoid collisions, which may not fully reflect real-world behaviour, particularly in close-quarters scenarios. Environmental factors, such as wind and currents, were excluded from the risk modelling, potentially limiting the accuracy in dynamic maritime environments. While the results aligned with earlier studies in the Baltic Sea, broader validation across different regions and environmental contexts remains unexplored. This work underscores the importance of combining AIS data with manoeuvre-based analysis for maritime safety while highlighting the need for expanded environmental considerations and real-world validations to achieve more robust results.

In 2021, Brian Murray [63] defended his doctoral thesis at the Arctic University of Norway. The research focused on enhancing navigational safety by emulating human-like situational awareness through historical AIS data. It explored three levels of situational awareness: perception of the environment,

comprehension of the situation, and projection of future dynamics. By leveraging clustering techniques such as Gaussian Mixture models and hierarchical density-based clustering (HDBSCAN), the author categorised ship behaviour into distinct patterns and matched observed trajectories to these clusters. Predictive models were developed to project future ship behaviours within a 30-minute window, supporting proactive collision avoidance. Additionally, the research introduced methods for live trajectory predictions, aiming to improve computational efficiency and accuracy in maritime traffic regions.

Despite these contributions, the study had notable limitations. It did not incorporate meteorological data or parameters, which are critical for capturing real-world conditions that affect ship behaviour. Collision risk assessments relied on conventional techniques such as CPA and ship domain evaluation. However, the research did not address multi-vessel encounters or scenarios involving proactive collision avoidance by both vessels. Furthermore, while the methods aimed to enhance situational awareness for navigators, the effectiveness of long-range predictions in operational settings was not validated, leaving a gap in assessing their practical utility. The study emphasised the need for further exploration of classification techniques, hyperparameter tuning, and the integration of long-range and short-range prediction models to achieve more comprehensive solutions.

Finally, in 2024, Lubna Mohamed Eljabu [27] defended a doctoral thesis at Dalhousie University, Canada. The thesis contributed to Maritime Situational Awareness (MSA) through advanced data-driven frameworks for destination port prediction. It focused on detecting maritime routes, identifying anomalies, and predicting vessel destination ports by clustering trajectory segments and employing similarity measures such as Discrete Fréchet Distance (DFD) and Dynamic Time Warping (DTW). The author effectively segmented trajectories and created graph representations of maritime routes, enhancing the understanding of traffic flows and facilitating predictive modelling. Nonetheless, the study did not extend its analysis to evaluate potential collision risks, limiting its application in safety-critical scenarios. Reliance on similarity-based clustering posed challenges in offshore environments with less structured trajectory patterns. Overlapping clusters also complicated the classification

tasks, occasionally resulting in misclassification.

The thesis did not address the evaluation of prediction uncertainty, proposing it as a direction for future work. Monte Carlo simulations and Polynomial Chaos Expansions (PCE) were suggested as potential methods for assessing prediction uncertainties. While these techniques offer analytical rigour, they were not implemented or evaluated in the research, leaving uncertainty quantification as a theoretical proposition. These limitations highlight an area for future research to strengthen the robustness and reliability of vessel destination predictions.

These dissertations show the relevance of and interest in the topic internationally because, as will be seen later, the number of individual articles in scientific journals is growing. Two recent theses by Lithuanian researchers also explore ML applications in maritime traffic.

The first thesis, titled "Machine learning-based prediction of the behaviour of marine traffic participants and discovering non-standard marine traffic situations" by Andrius Daranda [22], explores the application of ML methods for manoeuvre modelling and threat assessment to enhance maritime safety. Daranda's work primarily uses clustering techniques like DBSCAN/OPTICS to segment historical marine traffic data and ML algorithms to predict turning points and vessel routes, aiming to support safe navigation planning. The study also introduces a contextual knowledge method to evaluate threats in manoeuvring situations.

While the study successfully utilises clustering methods to reduce data size, its focus is more on route planning than precise trajectory prediction, limiting its application in real-time collision avoidance. The prediction model primarily forecasts the next turning point based on cluster centres, with limited consideration of continuous movement dynamics such as vessel speed and heading changes. Although neural networks, including RNNs and LSTMs, are briefly mentioned, their potential for handling sequential data is not fully utilised. The reliance on multi-layer perceptron (MLP) single-step predictions instead of leveraging recursive strategies overlooks the benefits of temporal dependency modelling, which is critical for accurate multi-step trajectory forecasting and collision risk assessment. Furthermore, while clustering is used

to detect anomalies, there is limited discussion on how these anomalies correlate with collision risks, especially in the context of real-time AIS data. The choice of parameters for clustering methods and the dependence on geographical location are also important considerations. The study's application of SVMs for predicting manoeuvre points is based on static cluster results, lacking the dynamic adaptability necessary for real-time safety assessments. As a result, clusters that are "outliers" are considered an anomaly simply because their route is not historically intensive. Integrating more advanced techniques, such as recursive forecasting with RNNs and region-based methods like EPR and CPR for assessing overlapping boundaries, could enhance the approaches' applicability for predicting vessel trajectories and evaluating collision risks, aligning more closely with the research focus on maritime safety.

The second thesis, "Semi-supervised and Unsupervised Machine Learning Methods for Sea Traffic Anomaly Detection" by Julius Venskus [95], explores marine vessel traffic anomaly detection as an extension of maritime situational awareness. This work suggests combining AIS data with meteorological data, utilising six key parameters (wind direction, speed, wave direction, height, day/night cycle, and tide level) to predict vessel movements under various traffic and weather conditions. A notable strength of this research is the use of a multi-layer LSTM architecture to fill in missing data, enhancing the dataset's quality for downstream tasks. Additionally, the thesis introduces upper and lower bounds models to calculate prediction regions, supporting uncertainty estimation. The study also presents a modified self-organising map (SOM) algorithm for classifying marine vessel movement data into normal and abnormal classes, retraining strategies for SOM methods, a vessel type prediction method using LSTM, and two LSTM-based methods for unsupervised detection of abnormal marine vessel trajectories. These methods aim to improve maritime situational awareness for smaller ports with moderate traffic and enhance the detection of anomalous trajectories in larger areas with substantial traffic.

Some observations can be made regarding the two theses. The focus is primarily on data preparation, missing data imputation, and anomaly detection, with less emphasis on trajectory prediction or collision risk assessment. While

the usage of prediction boundary methods (upper and lower interval bounds) is beneficial, these methods are more suited for data following a normal distribution, which is often not the case in maritime traffic. Different vessel types were also separated into distinct data samples, requiring additional pre-processing and data handling. This segmentation approach contrasts with methods integrating categorical data, such as embedding layers or one-hot encoding, to create a unified dataset for diverse vessel types. Despite these observations, the thesis offers valuable insights into anomaly detection and advances the application of unsupervised and semi-supervised learning techniques in maritime situational awareness.

2.3 Short-and Long-Term Predictions Using Multivariate Data

Time series forecasts are widely used in the transport sector for various purposes. Models have been developed to make short-term predictions, such as one-step ahead forecasts [57, 85], where only the immediate next position is predicted. Additionally, some approaches utilise rolling forecasting [22], a form of iterative multi-step prediction where forecasts are continuously updated one step at a time. The model is retrained using the latest observations and prior predictions at each step. However, such forecasts may become increasingly distorted over longer horizons as they rely on previously predicted values rather than actual data, causing error accumulation. In maritime prediction, multi-step multivariate prediction refers to the ability of a model to forecast multiple future positions of a vessel over time using various input variables. Instead of predicting the next immediate position, a multi-step model can simultaneously extrapolate the vessel's coordinates for several future time steps, allowing for more comprehensive trajectory predictions.

Multivariate prediction involves the use of multiple features to inform these forecasts. This can include data on the vessel's current speed, heading, latitude, longitude, vessel type, and relevant meteorological information in maritime scenarios. By incorporating these diverse variables, the model can capture the complex interactions and dependencies that influence a vessel's movement.

Thus, a multi-step multivariate model in maritime traffic prediction can provide detailed and extended forecasts of a vessel's future positions, leveraging a rich set of input data to enhance the accuracy and reliability of the trajectory predictions. Two WGS parameters are sufficient for the extrapolation of the trajectory prediction data: longitude and latitude, which are time-varying and can be geographically determined.

Predicting vessel trajectories holds increasing significance in maritime safety, environmental conservation, and the optimisation of port operations. The intricacies of maritime activities, encompassing diverse elements such as weather conditions, vessel attributes, and human behaviour, pose significant challenges to accurately predicting vessel trajectories. Typically, forecasts rely on historical numerical data collected from AIS stations [17, 18, 74, 104, 113]. However, these predictions often overlook the distinct traffic characteristics of different vessel types.

Situational awareness of maritime traffic is a key factor in maritime transport safety. Modern research is primarily concerned with the control of autonomous vessels, collision risks [1] and anomalies [41]. Movement trajectories are a vessel's forward or backwards passage, which is analysed using vessel thrusters [44] to detect maritime traffic incidents, but often does not focus on specific vessels, as shown even in the presented thesis described in Chapter 1. At sea, as on land, different types of transport with established routes and traffic flows can provide an additional context for maritime mobility.

2.4 Exploring Predictive Models for Vessel Trajectory Navigation

Maritime transport data encompass the gross weight of goods, passenger movements, and vessel traffic across waterways. The action at sea remains intense, with maritime flows continuing to grow. However, the sector faces persistent challenges, particularly regarding container vessels, severe weather conditions (e.g., strong winds, waves, and floods), collisions, and other risks, whether natural or man-made. To mitigate risks such as collisions, scientists are actively researching and applying various methods to monitor and predict

vessel movements. While considered a lower-level safety component of the broader marine environment, accidents pose significant threats to human life, economic activity, and environmental security. A complete representation of the marine security matrix, detailing the relationships between actors and components of maritime safety, is provided by Bueger [10]. Monitoring and analysing naval shipping can help to ensure safety at sea. Such surveys are made possible by information systems, such as AIS, which provide real-time and historical vessel tracking data. AIS data, collected via terrestrial receivers and satellites, enable researchers to develop and apply prediction methods to anticipate vessel trajectories [46], identify anomalies, and assess collision risks. Such data include the geographical position of the vessels, the directions of movement, and many other characteristics, often not only about ships but also about their environment. Large amounts of data allow researchers to use DL methods, which provide superior information that can be useful in obtaining more accurate research results. Various methods have been explored for vessel trajectory prediction in recent years, ranging from traditional motion models to advanced ML and DL techniques. These methods include statistical approaches, rule-based systems, and modern data-driven models such as RNNs, LSTM networks, and transformer architectures.

One of the tools that can help ensure safety is monitoring the movement of ships and creating intelligent systems to improve the control of routes. Historical ship routes can refine habits and learn usual patterns from previous movements, as in research [7, 65]. According to the publication by Rong [76], maritime traffic behaviour is based on AIS data, consisting of three main steps: grouping ship trajectories, waypoints, route legs identification, and ship behaviour characterisation. Patterns are used to identify whether the vehicle is behaving normally and not deviating from the course.

Moreover, an essential criterion of accuracy emerges. Trajectory deviations can be categorised and addressed by clustering or regression tasks and their combinations. Unsupervised ML techniques can effectively address trajectory deviations and abnormal vessel movements. For example, Venskus [99], with co-authors, applied LSTM networks to detect real-time abnormal vessel movements by adapting unsupervised learning techniques. Additionally,

the same author implemented a SOM with a virtual pheromone model to classify maritime vessel movements into normal and abnormal categories [97]. Further insights were provided in another experiment, where the SOM network retraining achieved a significant reduction in computation time, nearly 50% faster without compromising accuracy, even when tested across different datasets [98].

For regression tasks, the focus is on predicting continuous trajectory values, enabling the precise forecasting of vessel movement. Supervised learning techniques, particularly neural network architectures, have been widely applied. For instance, RNNs and LSTM networks are often combined with data reduction techniques, such as convolutional neural networks (CNNs), to forecast vessel traffic based on sparse trajectory (GPS) data [8, 23, 73]. Similarly, spatiotemporal methods, such as those proposed by Wang [105], divide regional flows into hexagonal grids for improved prediction accuracy. In comparative studies, LSTM networks have demonstrated significant advantages in forecasting cargo traffic, often outperforming classical methods like ARIMA, even with incomplete time series values [51, 110]. LSTM networks are particularly effective in automatically capturing spatial and temporal dependencies in traffic flow data, as highlighted by Zhang [118]. Over recent years, AEs have gained prominence for their ability to compress high-dimensional input data into a latent space and reconstruct it, making them valuable for generative modelling. These techniques have been successfully applied in maritime traffic predictions, as demonstrated in studies by Capobianco [12] and others using dual-linear AE approaches [64].

Geographical location is critical in positioning a vessel's trajectory to solve prediction problems. Across various architectures, geographic coordinates (longitude and latitude) are widely used as inputs for predicting multi-step vessel trajectories [32, 87]. These coordinates are sometimes transformed into alternative map projections to simplify distance measurements and improve computational efficiency. For instance, cartography's standard WGS can be converted into the UTM projection. In this transformation, geographic coordinates are mapped to a Cartesian system, expressed as x (easting) and y (northing) coordinates, as demonstrated by Keller [42]. Two-dimensional

coordinate systems like UTM have been successfully integrated into the training of deep neural networks for trajectory prediction [3, 66, 71]. However, despite adopting such transformations, the choice of a particular coordinate system is often not explicitly justified, and the benefits of these transformations remain unclear. This thesis addresses this gap by systematically comparing multiple positioning transformations across different coordinate systems and evaluating their impact on trajectory prediction accuracy with various RNN architectures to identify the most effective approach for accurate vessel trajectory forecasting.

2.5 Deep RNN Architectures

The characteristics of RNNs are defined by their cyclic pathways formed through synaptic connections, allowing the network to retain and process sequential information over time [75]. While standard feed-forward neural networks process data in a single pass, typical RNNs transmit information recursively from one block to another, enabling the modelling of temporal dependencies in data sequences. The research investigates simple RNN and improved DL recurrent networks, LSTM and GRU, with different architectural combinations. These improved architectures can solve the problems of vanishing gradient and long-term dependency [100] because it has feedback links and a unique memory management structure (cells). A typical LSTM cell comprises input, output, and forget gates that regulate the flow of information, allowing the network to retain or discard information over long sequences selectively. In contrast, the GRU simplifies this mechanism by using only update and reset gates, offering a more computationally efficient alternative while managing dependencies effectively. These memory cells enable the processing of sequential time-step data, where the gates dynamically adjust the information flow between units. Unlike traditional methods, LSTM-based models better predict vessel behaviour, offering higher precision, better adaptability, and faster prediction speeds [88]. These advantages make recurrent networks highly effective for maritime trajectory forecasting, particularly LSTM and GRU models. The following text provides a detailed overview of the architectures explored in this study.

Basic LSTM

The standard LSTM network architecture consists of a single hidden LSTM cell layer. The general network design includes an input layer, hidden LSTM and dropout layers, and a dense output layer (see Fig. 2.2). The hidden layer comprises blocks of LSTM cells connected in a unidirectional sequence, allowing information to flow through the network. These LSTM cells are interconnected by a main status signal, regulated by the cell's gates to control how information is stored, updated, or discarded over time.

Each LSTM cell contains specialised gates, input, forget, and output gates, that manage the flow of information and enable the network to retain important long-term dependencies while filtering out less relevant data. This gating mechanism allows the network to address common issues, such as the vanishing gradient problem, making it well-suited for modelling sequential data like vessel movements. The dropout layer between the LSTM and output layers is a regularisation technique, preventing overfitting by randomly deactivating certain neurons during training. The dense output layer transforms the processed information into final predictions, such as future vessel positions over multiple time steps.

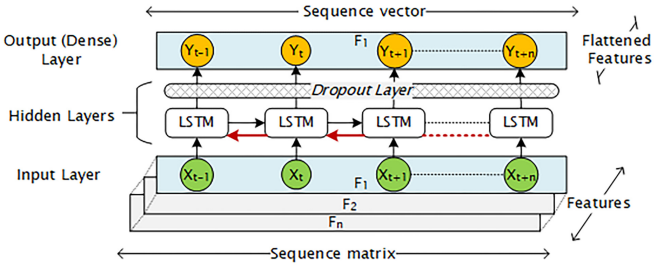


Figure 2.2: Basic unidirectional and bidirectional LSTM architectures. Note: bidirectional marked with a red line.

Bidirectional LSTM

Bidirectional LSTMs enhance the standard LSTM architecture by processing sequence data in both forward and backwards directions. This allows the network to capture dependencies from past and future contexts simultaneously

[112]. In this case, forward and backwards ways are constructed in the same single LSTM hidden layer (see a note in Fig. 2.2). The rest of the architecture remains similar to the standard LSTM, with a dropout layer applied to the hidden layer to reduce overfitting by randomly deactivating input units during training. The key difference lies in the dual flow of information signals moving in both directions, enabling the network to learn more comprehensive sequence representations. Literature [53] suggests that bidirectional LSTM architectures often perform better in sequence prediction tasks than their standard counterparts.

Simple RNN

Simple RNN processes sequential data by using a loop to iterate over each time step in the sequence while maintaining an internal state that encodes information from previous time steps. At any given point, the hidden state is computed solely based on the current input and the prior hidden state without additional mechanisms for managing long-term memory. This structure allows the Simple RNN to model short-term dependencies effectively. Still, it makes it less capable of capturing long-term patterns due to the absence of specialised gates for memory control.

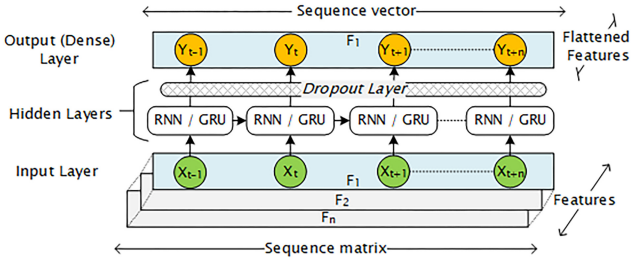


Figure 2.3: Simple RNN and GRU architectures.

Unlike more advanced recurrent architectures such as LSTM and GRU, the simple RNN does not incorporate input, forget, or output gates to regulate information flow. As a result, it is more prone to issues such as the vanishing gradient problem, which limits its ability to learn long-range dependencies in sequential data. The overall architecture of the simple RNN is similar to that of

other recurrent models, consisting of an input layer, hidden RNN layers, and a dense output layer. The primary difference lies in the internal cell structure of the hidden layers, which lack the gating mechanisms seen in the LSTM and GRU networks (see Fig. 2.3 with RNN cells).

Gated recurrent unit

The GRU is an advanced type of RNN that simplifies the memory management process by combining the functionalities of LSTM's multiple gates into two key components: the update gate and reset gate. These gates control the flow of information within the network, determining which data should be retained and which should be discarded, thereby addressing the vanishing gradient problem and enabling the model to capture longer-term dependencies more effectively than a Simple RNN.

The update gate regulates how much past information should be carried forward to future time steps, functioning similarly to the LSTM's input and forget gates combined. On the other hand, the reset gate determines how much of the previous information to forget when computing the current state, allowing the model to reset memory as needed adaptively. The overall architecture of the GRU network is similar to that of a basic LSTM or RNN, consisting of an input layer, hidden layers, and a dense output layer. The primary difference lies in the cell structure within the hidden layers, where GRU cells replace LSTM or standard RNN cells (see Fig. 2.3 with GRU cells).

Stacked LSTM

An extension of the standard LSTM architecture is the Stacked LSTM, which consists of multiple LSTM layers stacked on top of one another to create a deeper network [50]. Unlike the basic LSTM, which has a single hidden layer, the stacked architecture allows the model to learn more complex and abstract temporal patterns by passing information through multiple processing layers. In this configuration (see Fig.2.4), the output from each LSTM hidden layer serves as the input for the next hidden layer, enabling hierarchical feature extraction across the sequence data.

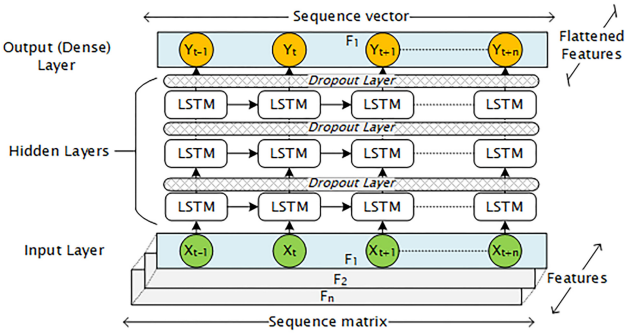


Figure 2.4: Stacked LSTM architecture.

A dropout layer is incorporated under each hidden LSTM layer to improve generalisation and prevent overfitting. This dropout layer randomly deactivates a portion of neurons during training, encouraging the model to learn more generalised patterns. Based on findings in the literature [96], which compared network accuracy across different layer depths, the stacked LSTM architecture in this study was designed with three hidden LSTM layers.

LSTM Autoencoder

The LSTM AE is a specialised neural network designed to learn efficient data representations by compressing input data into a lower-dimensional space and then reconstructing it back to its original form [64]. This architecture is particularly effective for sequence data, enabling the network to capture complex temporal dependencies and patterns. The LSTM AE consists of three main components (see Fig. 2.5): the encoder, the latent space vector, and the decoder.

The encoder compresses the input sequence into a compact latent representation. This is achieved by processing the sequence through LSTM layers that capture the temporal structure of the data and condense it into a single vector that summarises the entire sequence. The resulting latent vector is repeated n times to prepare it for decoding, where n is the number of time steps in the output sequence. The decoder takes this repeated latent vector and attempts to reconstruct the original input sequence. Using LSTM layers, the decoder transforms the compressed information into the target sequence,

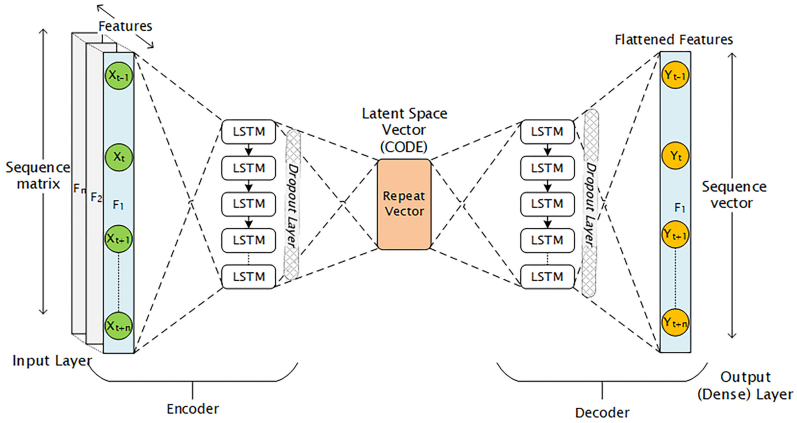


Figure 2.5: LSTM AE architecture.

learning to replicate the input as closely as possible. Dimensionality reduction in the LSTM AE is achieved by adjusting the number of LSTM cells in the encoder. By limiting the number of units, the model is forced to compress the input data, extracting only the most relevant features for reconstruction.

2.6 Impact by Vessel Type Category

Incidents do not often change the behaviour of traffic flows. The Nord Stream gas pipeline explosion in the Baltic Sea resulted in a ban on traffic, transit, anchoring, diving, use of underwater vehicles, geophysical mapping and other activities in the area due to safety and investigations [86]. In addition, additional traffic of military and reconnaissance vessels was intensified. Changes in traffic flows and routing of conventional vessels increase the risk of collisions between different types of vessels. The behaviour of vessels at sea is inherently diverse and contingent on their respective types. The predominantly quantitative data are invaluable for probing intricate phenomena, behaviours, and tendencies. The primary challenge lies in effectively integrating this wealth of information into the calculations performed by DL models.

Vessels can be categorised based on their intended purpose, such as transportation, fishing, auxiliary, or technical roles, and later they were classified according to the geographical areas they navigate, including offshore, raiding,

inland, or mixed routes [72]. For instance, offshore vessels sailing more than 200 miles from port cover long distances. The extensive array of ship types and their classifications is explored comprehensively in the article [19]. The collective impact of these multifaceted factors significantly influences overall maritime traffic, allowing for the re-purposing of historical data in system modelling. When applied in analysing AIS data, the recursive method has demonstrated notable advantages in various experimental contexts [56, 76, 120]. This method has proven effective in tasks such as anomaly detection, computation of the width of categorical routes derived from maritime traffic data [48], and the development of optimised Seq2Seq models tailored for short-term ship trajectory prediction [107, 116].

The integration of categorical data into ML algorithms to enhance performance has garnered attention in diverse domains, particularly transportation and medicine, as highlighted in the authors' paper [70], which focuses on disease prediction. Delving into the nuanced processing of categorical data using various encoding techniques, Dahouda and Joe explore the comparative efficacy of one-hot encoding and DL embedding in their article [21]. While the work predominantly contrasts these encoding methods, it duly acknowledges limitations stemming from dataset constraints and the specific binary classification problem under consideration.

In a separate investigation, Polish researcher Sebastian Gnat [33] performed an experiment evaluating diverse encoding techniques within regression models. Gnat's findings indicate that one-hot encoding exhibits greater accuracy in classical regression algorithms. Despite potential challenges associated with high cardinality due to dummy variables, the article [14] proposes solutions for complex problems without compromising accuracy.

Conversely, author Changro's study [47] asserts the viability of employing DL embedding encoding even with very high cardinality. The advantages of embedded encoding are not isolated, as corroborated by findings in other publications [15, 24, 43, 78, 101]. This collective body of research underscores the multifaceted landscape of categorical data processing within ML paradigms and offers insights into the diverse encoding strategies that can be applied to address specific challenges.

After the analytical exploration, it was identified that only a limited number of studies have systematically analysed and compared categorical data encoding techniques, particularly in the context of recursive multi-step trajectory prediction. Previous works have often explicitly selected specific vessel types for individual analysis or prepared separate datasets for distinct vessel categories. This thesis, therefore, tries to integrate multiple vessel types into a single unified dataset and evaluate the effectiveness of different encoding methods in preserving inter-category relationships.

2.7 Trajectory Prediction Boundaries for Collision Risk and Uncertainty Estimation

This subsection reviews the scientific literature on vessel trajectory prediction methods and their integration with collision risk assessment frameworks. Traditional maritime safety systems commonly rely on deterministic collision detection measures, such as CPA and TCPA. However, these methods provide limited insight when vessel behaviour dynamically changes due to environmental conditions, human factors, or unforeseen circumstances. To address this limitation, recent studies increasingly incorporate DL techniques, notably RNNs, to enhance the precision of trajectory forecasts and quantify associated uncertainties. Unlike deterministic approaches, advanced probabilistic methods utilise uncertainty estimation, such as confidence intervals, prediction intervals, and conformal prediction regions, to offer a more comprehensive assessment of potential collision scenarios. This review emphasizes integrating accurate forecasts with uncertainty quantification to strengthen maritime situational awareness and safety by examining vessel trajectory prediction methodologies alongside collision detection frameworks.

2.7.1 Vessel Trajectory Prediction Methods and Uncertainty Estimation

In maritime navigation, the accurate prediction of vessel trajectories is paramount to avoiding collisions and ensuring safety at sea. A body of research has been dedicated to addressing this challenge, utilising a variety of sophisticated

methodologies. For instance, one study by Biao Zhang [117] and colleagues focuses on refining the precision of ship motion predictions, especially within complex environments characterised by non-stationary, non-linear, and stochastic variables. This research introduces the IWOA-TCN-Attention model, a novel predictive framework that integrates DL networks, time-sequential data, and an attention mechanism, substantially improving the prediction accuracy.

Complementing this, another study by Rong [77] approaches maritime traffic prediction using a dual lens, contemplating both ship destination and route. This study distinguishes itself by its methodology of extracting vessel motion patterns from archival data, deploying multinomial logistic regression alongside Gaussian Process regression models to construct probabilistic forecasts. However, it observes that the hourly forecast increases the error and notes that additional features may hypothetically reduce this error.

The research landscape reveals a diversity of approaches for assessing predictions of vessel trajectories, notably focusing on their utility in identifying anomalous marine traffic patterns. The detection and analysis of abnormal vessel trajectories are carried out in a study by Kristoffer et al. [69], who propose a kinematic similarity measure and even provide access to labelled data to support further research on abnormal trajectory detection.

LSTM neural networks have also played a significant role in enhancing the trajectory prediction accuracy. For example, Mehri [61] employs LSTM models with a context-aware approach that integrates spatial data, such as vessel type and meteorological conditions, for data-driven movement predictions. Additional research explores the broader application of LSTM networks in ship trajectory prediction, showcasing their versatility in handling maritime data [39, 89, 103, 119, 121].

Uncertainty quantification methods have been introduced to enhance the reliability of trajectory predictions. Venskus [99] proposed unsupervised wild bootstrapping to evaluate prediction reliability, demonstrating its effectiveness in detecting a wide array of abnormal marine traffic behaviours. Additionally, abnormal vessel trajectories have been tested by creating multivariate cases of prediction intervals, assuming that the trajectory entering a defined region is considered normal.

Graphical methods for region detection, such as those proposed by Carnerero et al. [13], utilise convex optimisation techniques to delineate prediction regions. The reliability of predictions for multidimensional data is further examined using boundary estimation techniques that define regions with predefined nominal coverage rates. Golestaneh et al. [34] have contributed to this field by constructing and evaluating multivariate EPRs, delineating the uncertainty inherent in multidimensional stochastic processes. While not directly applied to ship trajectory prediction, their findings offer insights into minimising the conservativeness of prediction regions.

Forecasting models often rely on statistical approaches to evaluate prediction reliability. Yin et al. [114] analyse bus travel time forecasts using the construction of PIs, and Noma and Lucagbo [58, 68] discuss multidimensional prediction and CIs in broader contexts. However, maritime forecasting data frequently lack a clearly defined distribution, necessitating non-parametric techniques such as CPRs. This approach, initially introduced by Shafer [83], has been applied in diverse fields, including pedestrian localisation scenarios [20], and, more recently, time-series prediction interval detection [2, 91].

2.7.2 Collision Risk Assessment and Predictive Frameworks

While trajectory prediction provides the foundation for maritime safety, collision risk assessment is equally critical to ensuring safe navigation. Ship collisions pose significant threats to maritime operations due to their potential to cause substantial harm. Addressing this, Ryan Wen Liu [55] lays out a comprehensive framework for assessing and analysing ship collision risks. This framework incorporates advanced methodologies, including the quaternion ship domain and kernel density estimation. It integrates a ConvLSTM model for spatial-temporal risk prediction, marking a substantial stride in maritime collision avoidance strategies.

In expanding on collision avoidance strategies, several studies propose advanced decision-making frameworks. For instance, Xie et al. [108] introduce a deep reinforcement learning (DRL) approach to multi-vessel collision avoidance that adheres to COLREGs. Their model integrates a collision risk index (CRI) into its reward function, optimising vessel behaviour in various encounter

scenarios. Similarly, Seo et al. [82] develop a CRI-based A* algorithm that balances economic route optimisation with safety considerations, effectively addressing collision avoidance routing.

Yoshioka et al. [115] present a decision-making algorithm that utilises collision risk maps to visually represent potential risks and guide route planning, focusing on enhancing explainability for seafarers. Zhou et al. [122] focus on determining collision avoidance timing using an ML framework that incorporates dynamic and static factors to guide officers on Watch (OOW). Additionally, Liu et al. [54] propose QSD-LSTM, a novel trajectory prediction model that integrates a quaternion ship domain to enhance the prediction of vessel interactions in complex maritime environments.

Probabilistic models for collision risk assessment have also been explored in previous research. Mujeeb et al. [62] assess collision risks between vessels and offshore platforms using a model based on traffic density, causation probabilities, and mitigation measures. However, this approach lacks adaptability to real-time dynamic conditions. Yim et al. [111] apply multiple linear regression to analyse the perceived collision risks based on separation distance, though their reliance on subjective evaluations limits the model's real-time applicability.

In 2020, Du [26] developed a COLREG-compliant alert system for stand-on vessels that classifies conflicts into four encounter stages and nine severity levels to improve collision risk detection under dynamic conditions. However, uncertainties regarding ship domain boundaries remain a challenge in this framework. In 2024, Lin [52] introduced an encoder-decoder LSTM model for regional collision risk prediction, achieving high accuracy but facing limitations due to aggregation density and clustering uncertainties. Likewise, Gao [30] (2024) employs ST-ENAGCN (spatiotemporal edge-node attention graph convolutional network) to analyse multi-ship collision avoidance scenarios. Yet, this approach may not fully address ship domain boundary uncertainties, which are crucial for reliable collision risk assessment.

Additional research has focused on risk evaluation frameworks. Studies such as those by Weng [106], Tritsarolis [92, 93], and Jia [37] delve into collision risk assessment using probabilistic and statistical methods, further

highlighting the complexity of accurately predicting and mitigating maritime collision risks.

The thesis addresses the limitations of prior work by using a DL based trajectory forecasting framework enhanced with advanced uncertainty quantification techniques, including EPR, CPR, CI and PI. These approaches estimate the probability distribution around RNN forecasts within defined trajectory boundaries and compute collision risk scores from their overlaps, enabling more reliable and proactive real-time maritime risk assessments.

By applying these predictive boundaries to real-world scenarios, such as the 2021 collision between *Scot Carrier* and *Karin Hoej*, we demonstrate the effectiveness of DL models in improving maritime safety with a 95% confidence level. Our approach also integrates reliability indicators, particularly emphasising the strength of conformal prediction in accurately predicting potential collision scenarios. As the reviewed articles indicate, one specific method or its derivatives is typically used to define prediction boundaries. Therefore, this study applied various methods to specify the domain region of vessels based on predictions from DL models while assessing the probability score of collision risk for vessels in actual historical accidents.

2.8 Summary of the Chapter

This chapter reviewed the literature on advanced methodologies for multi-step multivariate vessel trajectory prediction, emphasising their importance in improving maritime situational awareness and safety. A particular focus was placed on RNNs, especially advanced architectures like LSTM and GRU, for their ability to process sequential maritime data. With their feedback links and specialised memory structures, these networks effectively address challenges such as the vanishing gradient problem and long-term dependencies, offering high precision and adaptability in trajectory prediction. Given that large vessels, such as tankers, require up to half an hour to come to a complete stop, accurate long-term forecasts, mainly those made 20–25 minutes ahead, are crucial for preventing collisions. Since AIS signal data are semi-structured and often incomplete, extracting meaningful vessel behaviour patterns requires analysing

spatiotemporal characteristics across multiple variables. As vessel movement depends on various interrelated factors such as latitude, longitude, speed, vessel type, and environmental conditions, trajectory prediction is inherently a multivariate problem that requires models capable of capturing complex dependencies within large-scale maritime data.

The review also highlighted the critical role of maritime awareness, addressing the increasing maritime traffic, diverse vessel classifications, and ongoing challenges like collisions and adverse weather conditions by applying various methods. While some studies focus on anomaly detection using unsupervised learning techniques such as data clustering, others employ regression-based supervised learning to predict continuous vessel trajectories to forecast vessel movement based on sparse GPS and AIS data. Moreover, the impact of different coordinate system transformations has been explored in previous research, demonstrating that the choice of representation, whether WGS84 or UTM, can affect prediction accuracy. This is attributed to advantages such as improved spatial distance calculations, uniform data scaling, and enhanced clustering of ship movements. The transformation ensures consistent distance metrics using Euclidean calculations rather than geodesic formulas, and simplifies model computations. Furthermore, UTM projections provide a structured view of AIS data, preventing distortions caused by latitude-longitude scaling differences. These findings reinforce the need to investigate further coordinate transformations in vessel trajectory forecasting and expand the research focus beyond anomaly detection, allowing for long-term trajectory predictions to be analysed later to identify deviations, unusual behaviour, or other maritime risks.

Additionally, the chapter explored categorical data encoding studies for integrating vessel-type information into predictive models alongside various methods for defining prediction intervals and boundaries to assess the reliability of model predictions. Some studies analysed vessel types separately by splitting them into different datasets. In contrast, others focused only on a limited number of vessel categories, potentially overlooking the relationships between various types of ships. This highlights the need to investigate methods for integrating multiple vessel types into a single prediction framework while preserving their

inherent dependencies. When predicting vessel trajectories, an important aspect is estimating the uncertainty of predictions, particularly in scenarios involving multiple vessels in close proximity. Many studies focus on forecasting a single vessel's movement without explicitly considering how surrounding vessels influence the overall risk of collision. Some researchers suggest that uncertainty estimation could improve collision risk assessment, but few have implemented it in practical applications. Existing approaches often assume that trajectory errors follow a specific distribution (e.g., Gaussian), which may not accurately reflect real vessel movement. As a result, non-parametric methods, which do not rely on strong distributional assumptions, have been explored as alternative solutions. Techniques such as Monte Carlo simulations have been proposed in the previously analysed thesis. Still, other PI-based methods, including PI, CI, EPR, and CPR, may provide a more flexible way to estimate uncertainty and assess the probability of vessel interactions. Investigating these approaches is essential for improving trajectory accuracy and the ability to detect potential collision risks based on surrounding vessel movements.

3 Guidelines on Methods

This chapter provides an in-depth analysis of the methods employed in the thesis, highlighting new insights and advancements in vessel trajectory prediction and maritime safety. One of the key innovations of this research is the development of a recursive recalculation logic for forecasting outputs. Rather than building models directly with specific (absolute) coordinate sequences, the study identifies the differences between vector sequences. These differences are then predicted and recursively extrapolated to determine the vessel's future positions based on the last known location. This approach enhances the accuracy and reliability of trajectory predictions.

The chapter begins by exploring multi-step recursive models and explaining this novel approach's theoretical foundations and practical applications. It then delves into various categorical data encoding techniques, including ordinal encoding, one-hot encoding, and embeddings, which are important for integrating categorical vessel-type data into DL models. These encoding methods are compared to assess their impact on prediction accuracy.

Subsequently, the chapter addresses prediction intervals and boundaries, which are essential for evaluating the reliability of predictive models. Detailed discussions on PI, CI, EPR, and CPR are provided. These techniques offer a probabilistic understanding of potential future vessel positions, enabling dynamic assessments of collision risks and strategic planning to mitigate such risks. The study demonstrates how DL models can effectively identify high-risk scenarios and enhance maritime safety by integrating confidence levels and prediction boundaries. These methods were presented in the main papers by Jurkus et al. [A.1, A.2, B.1, B.2].

3.1 Multivariate, Multi-step and Recursive Models

The paradigm of multi-step forecasting extends beyond traditional single-step predictions by forecasting a series of future values rather than only the immediate next step. This approach is particularly valuable in maritime scenarios, where

predicting vessel positions several time steps ahead is critical for timely risk assessment and collision prevention. In multi-step forecasting, prediction models are trained using sequences of historical AIS data containing various features, such as latitude, longitude, speed, and heading, to output sequences of future predicted values. Longitude and latitude determine an object's geographical location according to the globally accepted WGS84, which uses polar coordinates. Latitude specifies the north-south position relative to the equator, while longitude measures the east-west position relative to the prime meridian. Thus, each sequence represents a chronological vessel movement from one geographical point to another at determined time intervals. The neural network uses such historical positions as input, while the output sequence represents future positions that the model learns by optimising the selected architecture's loss function. A simple visual example is presented in Figure 3.1.

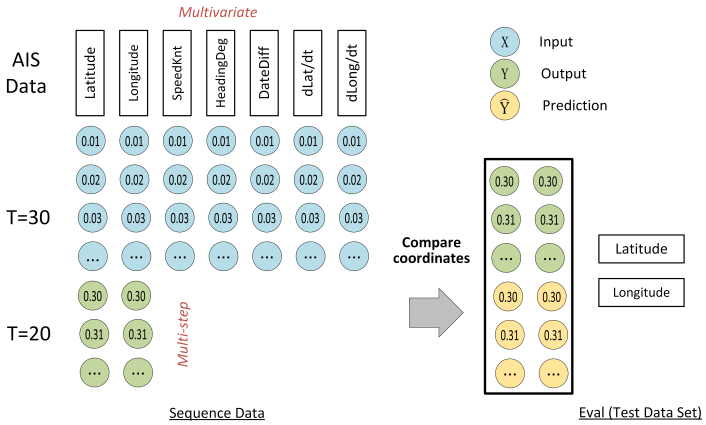


Figure 3.1: Multivariate multi-step structure of RNNs.

This multi-step forecasting approach is a supervised learning technique because it relies on labelled historical data for training. The labels in this context are multivariate, specifically consisting of actual geographic coordinates (latitude and longitude) representing vessel positions over time. During the model evaluation, these labelled sequences allow predictions to be directly compared against actual vessel movements, thus enabling the objective assessment of predictive accuracy and reliability. Given the inherent spatio-temporal complexity of vessel movements, multivariate input data are essential to capture

and accurately predict future trajectories. The prediction outputs are also multivariate since each forecasted point consists of multiple features, typically latitude and longitude coordinates. These models iteratively reuse previous predictions as inputs, progressively extending the forecasting horizon.

Several transformations were applied in this study. Instead of directly using geographic coordinates, the Haversine distance and azimuth angle were calculated, from which geographic positions were recalculated. Additionally, the Cartesian coordinate transformation using the UTM projection was employed. In these experiments, the original structure remains the same, but in the output of the network, for example, longitude and latitude are replaced by distance and turning angle. A more detailed transformation experiment is described in the paper [40].

The approach applied in this study involves transforming the geographic coordinates (latitude and longitude) into two new spatial features: distance and turning angle. The distance between two consecutive vessel positions is calculated using the Haversine formula, which measures the shortest distance between two geographic points along the Earth's surface. This formula calculates distances based on the Earth's curvature and a mean radius of approximately 6,371 kilometres. While accurate for short distances, the error may slightly increase over long distances due to the Earth's irregular shape.

The second feature involves calculating the azimuth angle, which indicates the directional bearing from one vessel position to the following relative to the geographic north. Azimuth angles range from 0° to 360°, measured clockwise, with 0° corresponding to true north. The networks capture more meaningful spatial and directional relationships by training the prediction models using these derived distance and angle features. Subsequently, the predicted distance and angle values are combined with the last known vessel position to reconstruct a series of geographic coordinates, effectively providing an accurate representation of future vessel trajectories.

The UTM projection was another transformation applied in this study. Unlike geographic coordinates, UTM is a global coordinate system defined in meters, allowing locations to be identified with high precision. The UTM grid divides the Earth's surface into 60 longitudinal zones, each spanning 6

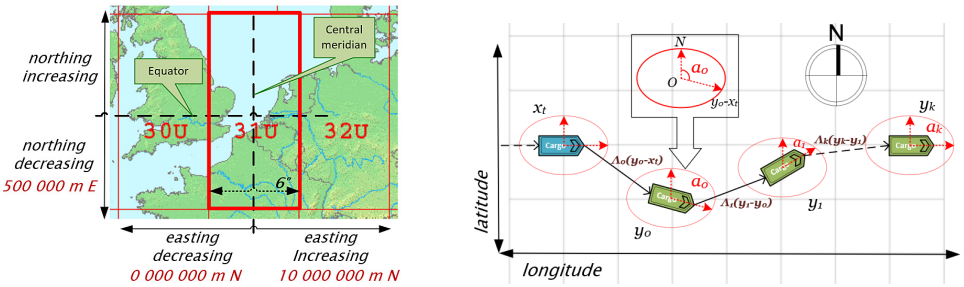


Figure 3.2: Coordinate transformations: (**a**) UTM projection in the Netherlands region; (**b**) distance and angle calculation projection (input is blue, output is green).

degrees. Within these zones, positions are expressed using two coordinates, easting and northing, analogous to x and y coordinates in a Cartesian coordinate system. Each zone has a central meridian with an assigned false easting value of 500,000 meters to avoid negative coordinates. Specifically, the Netherlands region, which serves as part of this study's dataset, falls into UTM zone 31.

The transformation from longitude and latitude to UTM coordinates employs trigonometric functions, resulting in x coordinates increasing eastward and decreasing westward relative to the central meridian and y coordinates increasing northward and decreasing southward. By converting geographic positions into two-dimensional Cartesian coordinates, the scale of maritime traffic routes can be better represented and analysed. In this study, the recurrent neural network architectures are trained using these UTM coordinates and formatted into flattened output vectors. Predictions made by these models are subsequently reconverted from UTM back to WGS84 geographic coordinates, ensuring compatibility and allowing comprehensive evaluation of the predicted vessel trajectories against actual routes. The concepts for both transformations are given in Figure 3.2.

Nevertheless, forecasting over extended horizons introduces inherent challenges. Accumulating errors over successive predictions can compromise forecast accuracy over longer prediction intervals. The study applied recursive data extrapolation techniques to address this, particularly with transformations

like Haversine distances, azimuth angles, and UTM projections. This recursive approach allows improved accuracy by reducing cumulative error propagation, thereby ensuring the adaptability and effectiveness of the model in managing the complexities associated with multi-step predictions.

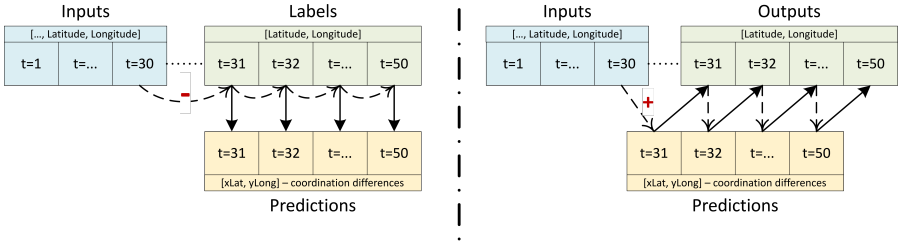


Figure 3.3: Recursive multi-step trajectory prediction model.

Vessel trajectory prediction constitutes a regression ML task falling under the domain of supervised training, necessitating the availability of labelled data. Within the RNN architectures, the input comprises sliced sequences, each mirroring the aforementioned matrices' length, with vessel features as the key inputs. Concurrently, the output matrix typically represents the continuous input coordinates of the vessel, effectively capturing the vessel's evolving trajectory. In this specific context, the sequence output deviates (Δlat, $\Delta long$) from direct geographical coordinates, instead presenting the disparity between their respective vectors. To generate labels for this sequence output, the actual (real) coordinates are computed by incorporating the last input timeline and calculating the coordinate differences relative to it. During the predictive phase, wherein the model anticipates these differences, retrieving actual coordinates necessitates an additional step: vector addition (see Figure 3.3). Ingrained in the architecture, this intricate process exemplifies the meticulous orchestration required to transform predictions into meaningful geographical coordinates within vessel trajectory prediction.

The sequence defined in Equation (3.1) can be illustrated through the following simplified example. In this scenario, the input encapsulates vessel features at distinct time points. For simplicity, assume the first two columns denote longitude and latitude coordinates, and the output exclusively features these coordinates. As highlighted in the predictions line, the neural network

isn't trained on precise geographical coordinates (or using other transformation features as UTM coordinates) but learns disparities within these delta coordinates. This concept is elucidated further in the comprehensive discussion in the article [40]. The neural network yields learned differences as part of its output. Recursive derivation of the original geographical coordinates from the last input point unveils a more accurate prediction, particularly in the short term. This demonstration underscores how the network's training on coordinate differences facilitates a more precise prediction of geographical coordinates, enhancing the model's efficacy.

$$\begin{aligned}
\text{Original Input and Output:} &\quad \Big[[1,1,\ldots] \quad [2,2,\ldots] \quad [3,3,\ldots] \quad \ldots \quad [4,4] \quad [5,5] \quad [6,6]\Big] \\
\text{Output (differences):} &\quad \Big[(4-3,4-3) \quad (5-4,5-4) \quad (6-5,6-5)\Big] \\
\text{Predictions:} &\quad \Big[[1,1] \quad [1,1] \quad [1,1]\Big] \\
\text{Input and Output:} &\quad \Big[[1,1,\ldots] \quad [2,2,\ldots] \quad [3,3,\ldots] \quad \ldots \quad [1,1] \quad [1,1] \quad [1,1]\Big] \\
\text{Output (based on the last point):} &\quad \Big[[3,3,\ldots] \quad \ldots \quad [4,4] \quad [5,5] \quad [6,6]\Big]
\end{aligned} \tag{3.1}$$

Recursive forecasting encompasses creating lagged features from the target series and training ML models on delta features rather than absolute positions. Past predictions are recursively utilised to generate new lagged features as the prediction extends into the future. Precisely, models predict incremental differences (deltas) between consecutive points rather than directly predicting absolute positions. These predicted deltas are subsequently added to the most recent absolute position, reconstructing future positions in the trajectory. For instance, as illustrated in the equation above, the original absolute positions are $[1,1],[2,2],[3,3]...[4,4][5,5][6,6]$ are transformed into differences $(4-3,4-3),(5-4,5-4),(6-5,6-5)$. Predictions are then made on these differences, and future absolute positions are reconstructed recursively by adding predicted increments to the previously known absolute position. This recursive recalculation reduces cumulative forecasting errors, thereby enhancing the accuracy and reliability of multi-step trajectory predictions.

Estimation of errors

The accuracy of the neural networks is initially assessed using rigorous regression metrics in the Netherlands region dataset. These metrics, including the Mean Absolute Error of the Haversine distance (MAEH), Mean Squared Error (MSE, (3.2)), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), provide a detailed evaluation of the model's performance in predicting coordinate values. A specific metric, MAEH, is derived from the Haversine distance function based on Figure 3.4. This metric calculates the predicted and actual values of geographic coordinates and measures the average sequence error across multi-step points. The Haversine function facilitates distance calculations in SI system units like meters or kilometres, offering a practical assessment of predictive accuracy. This meticulous evaluation framework ensures a precise and factual understanding of the neural networks' performance in trajectory prediction, considering diverse aspects of the predictive model. Due to its direct spatial interpretability, MAEH serves as the primary criterion for subsequent model evaluation and comparison. Estimates are calculated using the test sample dataset:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3.2}$$

where:

- n is the number of samples,
- y_i is the true value,
- $\hat{y}_i$ is the predicted value.

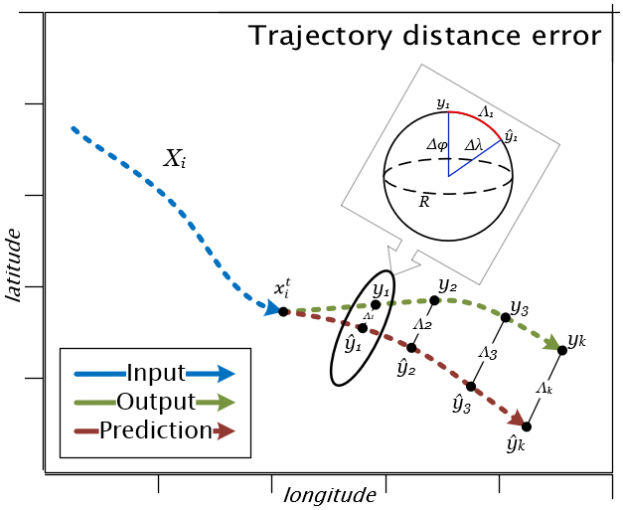


Figure 3.4: Haversine distance error calculation in a trajectory.

Here, X denotes the matrix of vessel characteristics, which is the network's input ($t = 30$ time steps). Knowing the actual y and predicted direction of $\hat{y}$ movement ($k = 20$ time steps), it is possible to measure the distance between each point in the time series $\Lambda_k = \hat{y_k} - y_k$. The shortest path between two points of a geographical position is calculated by the Haversine function (Λ), where $\Delta\phi$ is latitude, $\Delta\lambda$ is longitude, R is the Earth's radius (mean radius = 6,371 km), and n is the output sequence in the test dataset. Distance is measured in kilometres. Furthermore, the total trajectory error is determined by the mean absolute Haversine error MAEH (3.3). This principle determines how much the predicted range differs from reality. By this logic, all architectures were evaluated in the entire sample of test data.

$$MAEH = \frac{1}{n}\sum_{k=1}^{n}\left|\Lambda_k\left(y_k - \hat{y}_k\right)\right| \tag{3.3}$$

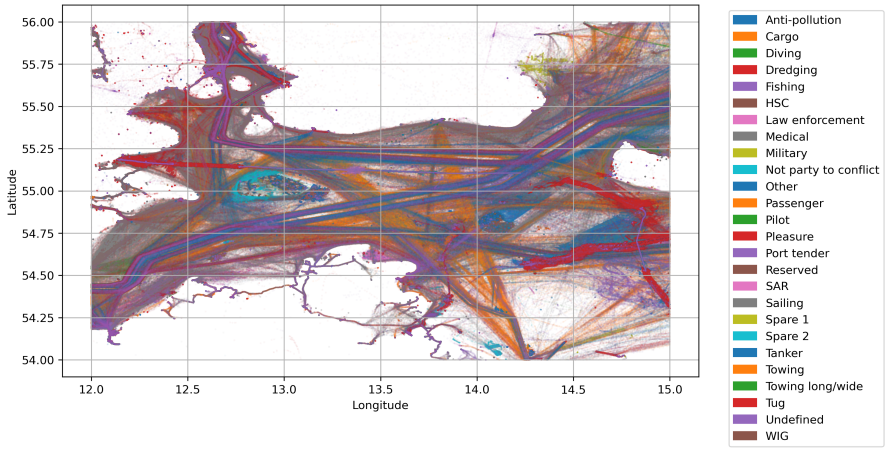


Figure 3.5: AIS data trajectories by vessel type in the Baltic Sea region.

3.2 Encoding Techniques

Within the satellite data sample, vessel types are inherently non-numeric, presented textually when transmitted by the vessel (see Fig. 3.5). Categorical data, encapsulating various discrete types or labels of vessels, have over 20 unique classifications within the designated research region. Developing a sophisticated deep ML model necessitates transforming categorical data into a format compatible with the recurrent network algorithm. Some standard methods for encoding categorical data are ordinal (similar to label encoding), one-hot, and embedding techniques.

3.2.1 Ordinal Encoding

This encoding involves the conversion of each categorical value into an integer representation. This technique is particularly relevant when the order of categories holds significance, and preserving this order is imperative. Within the entire region under consideration, 26 distinct vessel types exist (see Table 3.1). In the process of ordinal encoding, values are systematically assigned and transformed in increments ranging from 0 to 26. Notably, this assignment is carried out randomly, acknowledging the unknown relationships between various types of ships traversing the region. These vessel types span diverse

Table 3.1: Ship type classification.

Ship Type	ID	Ship Type	ID
Anti-pollution	1	Pleasure	14
Cargo	2	Port tender	15
Diving	3	Reserved	16
Dredging	4	SAR	17
Fishing	5	Sailing	18
HSC	6	Spare 1	19
Law enforcement	7	Spare 2	20
Medical	8	Tanker	21
Military	9	Towing	22
Not party to conflict	10	Towing long/wide	23
Other	11	Tug	24
Passenger	12	Undefined	25
Pilot	13	WIG	26

categories such as cargo, fishing, military, passenger, and tug. In both label and ordinal encoding, each non-numeric category undergoes substitution with a unique integer value. Subsequently, this encoded representation is introduced as an additional feature, manifesting as an extra column in the input matrix at each time step. For instance, cargo ships within the entire dataset are consistently assigned a specific value, such as 1, facilitating the integration of ordinal encoding as a meaningful component of the broader encoding strategy.

$$X^{(i)}_{\text{Ordinal}} = \left[\begin{array}{cccc|c} x_i^{(1,1)} & x_i^{(1,2)} & \dots & x_i^{(1,n)} & \mathrm{C}_i^{(1,n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_i^{(t,1)} & x_i^{(t,2)} & \dots & x_i^{(t,n)} & \mathrm{C}_i^{(t,n+1)} \end{array}\right] \left| \begin{bmatrix} \mathrm{C}_i \\ 0 \\ 1 \\ \vdots \\ 26 \end{bmatrix} \right. \tag{3.4}$$

In this input matrix (3.4), the index i represents different samples of sequences, where x is the vessel feature (speed, coordinates, etc.), and each sample has its label-encoded values C for the ship type (category number)

feature at each time step t. The input of the recurrent neural network consists of a three-dimensional matrix made up of the axes: the samples/sequences, the time steps, and the vessel features.

3.2.2 One-hot Encoding

This encoding involves the transformation of categories into a binary system, generating new attributes by allocating a value of one to the specific category of interest and zero to all other categories. This technique adopts a binary representation, efficiently capturing the presence or absence of each category within the dataset. However, it is essential to acknowledge a notable drawback associated with one-hot encoding, where each distinct category creates a separate feature. Consequently, when dealing with a substantial number of elements, the data matrix expands significantly, influencing the overall shape of the data (with a subsequent increase in network parameters). For instance, if a cargo ship is present in the dataset, the corresponding column in the encoded matrix is set to 1. In contrast, all other categorical columns are uniformly assigned a value of 0. This distinctive binary representation forms a fundamental aspect of the one-hot encoding technique, albeit with considerations for its impact on data dimensionality.

$$X^{(i)}_{\text{One-hot}} = \begin{bmatrix} x_i^{(1,1)} & \dots & x_i^{(1,n)} & \mathrm{C}_i^{(1,n+1)} & \dots & \mathrm{C}_i^{(1,n+j)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_i^{(t,1)} & \dots & x_i^{(t,n)} & \mathrm{C}_i^{(t,n+1)} & \dots & \mathrm{C}_i^{(t,n+j)} \end{bmatrix} \Bigg| \begin{bmatrix} \mathrm{C}_i^1 & \mathrm{C}_i^2 & \dots & \mathrm{C}_i^j \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \tag{3.5}$$

For instance, input matrix (3.5), for each different vessel type, it is required to create an additional feature in the dataset $[i,t,n+C_j]$, where i is the sequence, t is the time series, and n is the features combined with the one-hot encoding matrix C_j (binary values). Here, each vessel type is transformed into a binary vector j.

3.2.3 Embeddings

Embeddings is a process that reduces the dimensionality of data and preserves the relationships between categories. In this technique, the embedding layer changes the ordinarily assigned values for each vessel type and turns positive integers (indices) into fixed-sized dense vectors. This study tests three dimensions of the embedding matrices: 1D, 2D, and 3D. Embedding involves representing categorical data as continuous vectors in a lower-dimensional space (3.6), allowing the model to learn meaningful relationships between categories.

$$E = \begin{bmatrix} e_{1,1} & e_{1,2} & \dots & e_{1,d} \\ e_{2,1} & e_{2,2} & \dots & e_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ e_{j,1} & e_{j,2} & \dots & e_{j,d} \end{bmatrix} \tag{3.6}$$

The model maps j unique categories in the dataset to an embedding layer with d dimensions. Each type is tokenised, and a dictionary is created. The dictionary lists unique ship types assigned a positive integer, as in ordinal encoding. This prepared data passes as input through the neural network's embedding layer and outputs a compressed embedding matrix. During training, the model learns the values of the embedding vectors. The embedding vectors start with random values and are updated through gradient-based optimisation to minimise a specific loss function. Subsequently, the dictionary converts each sequence with a ship type to the appropriate index. The embedding layer looks up the dense embedding vector for each category. The encoding layer facilitates the development of a model capable of learning distinct vessel trajectories corresponding to each categorical feature (vessel type). This means that the LSTM AEs' prediction trajectory will be designed for the relevant vessel type context to include the context of ship types.

In this matrix (3.7), E represents the embedding matrix, where each row corresponds to a unique category, and each column represents a dimension in the embedding space. In other words, an embedding matrix is concatenated to the numerical features of the vessel at each timestamp, which is reshaped to match the sizes of the two matrices, as can be seen in the architecture of the LSTM AE.

$$X_{\text{Embed}}^{(i)} = \begin{bmatrix} x_i^{(1,1)} & x_i^{(1,2)} & \dots & x_i^{(1,n)} & E_i^{(1)}) \\ \dots & \dots & \dots & \dots & \dots \\ x_i^{(t,1)} & x_i^{(t,2)} & \dots & x_i^{(t,n)} & E_i^{(t)}) \end{bmatrix} \tag{3.7}$$

Embedding is particularly useful when dealing with high-cardinality categorical features and allows the model to learn encoded representations. The key idea is that ship types with similar meanings or contexts tend to have similar embedding vectors, allowing the model to capture relationships between categories. This technique is also often used in DL natural language processing for word encoding. It helps reduce dimensionality over one-hot encoding since the number of features can be controlled. It can be extended by adding categorical data without fundamentally changing the architectural structure.

3.3 Prediction Intervals

This thesis investigates vessel trajectory prediction by integrating DL algorithms with advanced statistical methods for trajectory forecasting, boundary estimation, and collision risk assessment. DL models serve as the basis for generating probabilistic forecasts of vessels' future positions, explicitly accounting for uncertainties inherent in maritime navigation. These trajectory predictions are further analysed using statistical approaches such as PI, CI, EPR, and CPR, defining boundaries that represent potential collision zones. Collision risk is evaluated by comparing these probabilistically determined boundaries against actual vessel positions, providing a dynamic and comprehensive assessment of potential maritime collisions closely linked to the accuracy of trajectory predictions.

3.3.1 Confidence and Prediction Intervals

CIs and PIs are statistical measures that are used to quantify uncertainty in predictions, particularly for models that output numerical values, such as those predicting vessel trajectories. Both intervals often rely on assumptions about data distribution, for example, a normal distribution, and are typically derived

from theoretical distributions. CIs estimate the range within which a population parameter (e.g., mean) is expected to lie with a certain confidence level [59]. PIs provide a range within which we can expect future data points to fall with a certain confidence level, α, typically 95%. W_{PI} is wider than W_{CI} because it accounts for the individual variability of each data point around the predicted mean, not just the variability of the mean itself [49]. In our scenario, the intervals for each time step are computed individually by aggregating point data from various models that forecast the same sequence. This approach ensures that the intervals accurately reflect the range of predictions and the associated uncertainty at each specific moment, thereby accommodating the variability inherent in the models' outputs. A set of points at a given time containing the predicted coordinates, the CI (3.9) and the PI (3.8) at the confidence level α are calculated using the equations, where the upper and lower bounds are subsequently found by adding/subtracting the sums from the sample mean and formula expressions based on [11, 28, 84]:

$$W_{\text{PI}} = t_{\left(\frac{1+\alpha}{2}, n-1\right)} \cdot SD \cdot \sqrt{1 + \frac{1}{n}}, \tag{3.8}$$

$$W_{\text{CI}} = t_{\left(\frac{1+\alpha}{2}, n-1\right)} \cdot \frac{SD}{\sqrt{n}}, \tag{3.9}$$

where the following applies:

- $t_{\frac{1+\alpha}{2}, n-1}$ is the critical value of the Student's distribution (t-score) corresponding to the desired confidence level, α, and $n-1$ degrees of freedom, where n is the sample size.
- SD is the sample's standard deviation, representing the spread of the data points.
- $\frac{SD}{\sqrt{n}}$ is the standard error of the estimate, which adjusts the standard deviation for the size of the sample.
- The term $\sqrt{1 + \frac{1}{n}}$ accounts for the added uncertainty when predicting a single future observation rather than estimating the population's mean.

In standard univariate linear regression models, statistical intervals are delineated and visualised in a relatively uncomplicated manner, typically with time series data plotted along the x-axis and any other variable of interest on the y-axis. However, in our scenario, the forecast data are multivariate, encompassing longitude, latitude, and time series. For bivariate coordinate data, these intervals are often represented as an ellipse around the mean of the data points, capturing the uncertainty in both dimensions. Here, the intervals serve as the ellipse's radii. This approach is utilised to define the domain of the ship's trajectory, thereby accommodating the multidimensional nature of the data.

3.3.2 Ellipsoidal Prediction Regions

Unlike univariate data, which contains a single variable, multivariate data contains multiple variables that can be related to each other. For example, in geographic information systems (GISs), coordinates such as latitude and longitude are often analysed together. In vessel trajectory prediction, quantifying the uncertainty of the prediction is as important as the prediction itself. For this purpose, the EPR is constructed to encapsulate the potential future position of the vessel within a confidence range. EPR is mathematically formulated (3.10) by considering the predicted trajectory points' variance and spatial distribution based on the author's formulation [34]. The axes of the ellipsoid correspond to the principal directions of variability in the data, and their lengths are proportional to the standard deviations of the data along these directions. For the expression of data points with mean vector μ and covariance matrix Σ, an EPR that contains a $100(\alpha)\%$ portion of the distribution can be defined as

$$EPR = \{x : (x-\mu)^T \Sigma^{-1} (x-\mu) \leq \chi^2_{p,\alpha}\}, \tag{3.10}$$

where x is a vector in the multivariate space, $\chi^2_{p,\alpha}$ is the critical value of the chi-squared distribution with p degrees of freedom corresponding to the desired confidence level α, and p is the number of variables. Given a set of predicted points for a vessel's trajectory and a centre point, the EPR can be calculated using the provided pseudo-code, which the authors of the article provide.

Algorithm 1: EPR Calculation

Input: $points$ - array of latitude and longitude points
Input: $center$ - centre point (mean) of the EPR
Output: epr - the calculated EPR as a polygon

$hull \leftarrow \text{ConvexHull}(points)$;
$hullVertices \leftarrow points[hull.vertices]$;
$covMatrix \leftarrow \text{Covariance}(hullVertices)$;
$[eigValues, eigVectors] \leftarrow \text{Eigen}(covMatrix)$;
$confidenceLevel \leftarrow 0.95$;
$radii \leftarrow \sqrt{\text{ChiSquaredInverse}(confidenceLevel, 2)} * \sqrt{eigValues}$;
$angleArray \leftarrow$ array of angles from 0 to 2π;
$ellipsoidPoints \leftarrow$ EmptyArray(size: 100, dimensions: 2);
for $i \leftarrow 1$ **to** 100 **do**
    $ellipsoidPoints[i][1] \leftarrow radii[1] * \cos(angleArray[i])$;
    $ellipsoidPoints[i][2] \leftarrow radii[2] * \sin(angleArray[i])$;
end
$rotatedPoints \leftarrow eigVectors \cdot ellipsoidPoints$;
$translatedPoints \leftarrow rotatedPoints + center$;
$epr \leftarrow \text{CreatePolygon}(translatedPoints)$;
return epr

First, the convex hull surrounding all points is determined to find the outer boundary. The covariance matrix is then derived, and eigenvalue decomposition is performed to extract eigenvalues and eigenvectors. The chosen confidence level (represented by α) is set to 0.95 (corresponding to 95%). The radius of the EPR ellipsoid is calculated using the chi-square distribution, considering α and the degrees of freedom, which, for a 2D point, are equal to 2, resulting in the radius $\sqrt{\chi^2_{2,\alpha}} \cdot \sqrt{\lambda_i}$. Angles ranging from 0 to 2π are then generated to represent the ellipsoid parametrically. The surface points of the ellipsoid are calculated using these radii and then rotated by the eigenvectors to align the ellipsoid with the data distribution. Finally, an ellipsoid is formed from the centre point, creating the EPR (see Algorithm 1).

EPRs have basic geometric characteristics and are classified as parametric methods. This classification arises from the reliance on distribution parameters

such as shape, orientation, and size to define the domain. Nonetheless, these domains of the predictions are particularly valuable for the analysis of multidimensional ship orbital collision data. Their usefulness becomes apparent when the study focuses on understanding the ships' geographical position and spatial extent.

3.3.3 Conformal Prediction Regions

CPRs aim to provide a range (or a region, in the case of multidimensional predictions) within which future observations are expected to fall, with a predefined level of confidence or significance, using past data [6, 109]. CPR leverages the past distribution of data to define prediction regions guaranteed to contain the true value of new observations with a specified probability, assuming that future data will resemble the past data.

For CPR, a calibration set C, consisting of data instances with known true outcomes, is employed to calculate nonconformity measures, denoted as N. This calibration set is distinct and non-overlapping with the model's training and test sets, ensuring the utilisation of validation samples previously employed to ascertain the model's accuracy during its training phase.

We use Y_t to denote the actual multivariate coordinates at time step t and $\hat{Y}_t$ to denote the predicted coordinates. The nonconformity measure for each time step is quantified using the Euclidean distance (3.11):

$$N_t(Y_t, \hat{Y}_t) = \|Y_t - \hat{Y}_t\|_2 \tag{3.11}$$

where $t \in \{1, \ldots, T\}$ indexes the time steps within the sequence. The Euclidean distance is a natural choice here and is suitable for determining the width of a region. The nonconformity measure should account for errors in all dimensions of the multivariate output, including latitude and longitude in the case of spatial coordinates.

Considering the tendency for prediction errors to increase over longer forecast horizons, the nonconformity scores are not aggregated over the entire sequence. Instead, they are averaged at each time step, resulting in a distinct threshold, τ_t, for each time step. This threshold is subsequently applied to new

prediction sequences from the test sample to determine if each predicted point is within the expected conformal region. To ensure an unbiased evaluation, the calibration dataset comprises 15% of the full dataset, specifically selected from the validation set. This dataset is distinct from the training data used for model building and the test data used for evaluating the final model performance. The model generates predictions for each instance using this calibration set, and the corresponding nonconformity scores are calculated. In the multi-model approach, these scores represent an average of the distances across all models at each time step. The problem of calibrating CPR is approached by calculating the average nonconformity scores across all models at each time step in the calibration set. For the i-th instance in the calibration set and time step t, the nonconformity score $R_t^{(i)}$ is defined as the Euclidean distance between the actual value, $Y_t^{(i)}$, and the predicted value, $\hat{Y}_t^{(i)}$.

The average nonconformity score at time step t across all instances is obtained using (3.12):

$$\bar{R}_t = \frac{1}{n}\sum_{i=1}^{n} R_t^{(i)}, \tag{3.12}$$

where n denotes the number of instances in the calibration set. Consequently, we seek to optimise the following problem (3.13):

$$\begin{aligned} \text{minimise} \quad & \text{Quantile}\left(\{\bar{R}_1,\ldots,\bar{R}_T\}, 1-\delta\right) \\ \text{subject to} \quad & \sum_{t=1}^{T} \alpha_t \bar{R}_t = 1 \\ & \alpha_t > 0, \quad t = 1,\ldots,T \end{aligned} \tag{3.13}$$

Here, α_t represents the parameters to be optimised, which scale the nonconformity scores at each time step, and δ represents the desired confidence level for CPR. The objective function minimises the quantile of the averaged nonconformity scores, aligning with the confidence level to construct valid PIs.

Figure 3.6 illustrates the process from using the calibration data to obtaining the new trajectory prediction with uncertainty estimation. The calibration data sequences, which contain both the input and output parts (as part of the supervised learning process), are used as input for the trained models

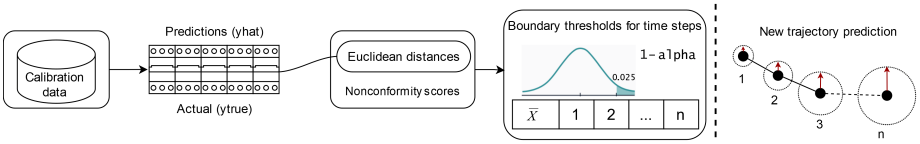


Figure 3.6: Illustration of boundary width determination in CPR using nonconformity scores.

to generate predictions ($\hat{y}$) over multiple time steps. These predictions are then compared to the corresponding ground truth values (y_{true}) from the calibration set output to calculate nonconformity scores based on the Euclidean distances between the predicted and actual values. These nonconformity scores are then used to compute boundary thresholds for each time step, determined by the confidence level $1 - \alpha$. These thresholds define the prediction region width, representing the outlying distance radius, illustrated as circles around each future predicted position in the new trajectory. This ensures that the true position lies within the defined regions with a given confidence level. The region's width for new predictions, which is consistent with the desired confidence level, is established by this quantile, allowing for an assessment of the accuracy of the forecasts. This method is more empirical, constructed from the data without strong parametric assumptions, making it widely applicable.

3.4 Summary of the Chapter

This chapter provided a detailed overview of the algorithms employed to enhance vessel trajectory prediction and improve maritime safety. The discussion began with analysing multi-step recursive models, which predict the differences (Δlat, $\Delta long$) between consecutive vectors rather than absolute values. By recursively recalculating forecast outputs, this approach effectively reduces cumulative errors over multiple time steps, resulting in more accurate long-term trajectory predictions. The performance of these models was evaluated using standard regression metrics such as MSE and MAE, alongside the introduced MAEH.

The chapter also explored various categorical data encoding techniques for integrating vessel-type information into DL models. Methods such as ordinal encoding, one-hot encoding, and embedding layers were evaluated for

their effectiveness in representing categorical data. Embedding techniques, in particular, were highlighted for their ability to capture relationships between vessel types and minimise the MAEH metric, ensuring more precise trajectory forecasts.

A critical focus was placed on using probabilistic measures to define prediction intervals and assess the uncertainty of model predictions. Several approaches were evaluated, including PI, CI, EPR, and CPR. These probabilistic interval-based methods are essential for quantifying the uncertainty in trajectory predictions, enabling the identification of safe navigation zones and even assessing collision risks.

To evaluate the reliability of these prediction models, key probabilistic metrics such as coverage probability and collision risk scores were incorporated. Coverage probability assesses how often actual vessel trajectories fall within the predicted regions, directly reflecting the model's accuracy in uncertainty estimation or implying an anomaly. The collision risk score, derived from the Jaccard index, measures the degree of overlap between predicted trajectories of vessels, providing a probabilistic assessment of potential collision risks. This probabilistic evaluation method offers a significant advantage over traditional deterministic methods like CPA and TCPA by considering the entire prediction boundary rather than a single-point estimate.

4 Data and Experiment Workflow Setup

This chapter details the empirical experiments to develop previously discussed methods and models, enhancing vessel trajectory prediction and maritime safety. The experiments utilise AIS data and meteorological information from platforms such as Ship Finder [G.2] and Danish Maritime Authority [G.1] databases. These datasets provide comprehensive information on vessel movements and environmental conditions, which is necessary for accurate trajectory prediction.

A key characteristic of big data is its variety, indicating that data often comes from heterogeneous sources and is typically semi-structured rather than fully structured. Therefore, the collected information must be processed before it can be effectively applied. The initial step involves processing the AIS and meteorological data, collected in comma-separated values (CSV) format, to ensure they are appropriately structured for input into the RNN algorithms. This process includes generating data sequences and augmenting the dataset with relevant features, such as delta coordinate parameters and related feature engineering transformations. These preparations aim to enhance the model's ability to capture complex patterns and dependencies in the data. Moreover, the focus is on data frequency adjustment, where the irregular time step intervals in the Netherlands region are modified using logarithmic normalisation. Additionally, time-frequency resampling methods are employed in the Baltic Sea region to standardise the frequency and reduce redundant observations, ensuring a more consistent dataset.

Subsequently, the chapter explores the embedding of vessel types into the dataset. The categorical data are transformed into a format that the DL models can effectively utilise by employing various encoding techniques. This step is essential for improving the prediction accuracy by providing the models with richer and more informative inputs.

The core of the experimental work involves creating and training RNN models, including the LSTM AE architecture, to predict vessel movements.

These models are evaluated on their ability to accurately forecast the future positions of vessels, taking into account both the temporal sequences and the diverse set of input features. The predicted trajectories estimate collision probabilities by assessing the overlap of prediction regions and identifying potential high-risk encounters. This integration of trajectory prediction with probabilistic collision risk assessment enhances the model's practical application. These experiments were presented in the main papers by Jurkus et al. [A.1, A.2, B.1, B.2].

4.1 AIS and Meteorological Data

AIS units are designed to automatically communicate with one another, sharing data about course, speed, and intended routes. This allows mariners to be aware of other ships well before they are visible, ensuring timely and appropriate actions can be taken. AIS data can be categorised into three types:

- static;
- dynamic;
- voyage-related.

Static information, entered during the system's installation on the vessel, includes details such as MMSI, IMO number, name, vessel length and beam, and vessel type. Dynamic information is updated automatically and includes parameters like speed, heading, geographical coordinates, and position timestamp. Voyage-related information, which must be input from the dashboard, covers details such as the type of cargo, voyage itinerary, and destination. This information can be presented as text (alphanumeric) or through other suitable equipment, such as radar with a graphical representation. AIS can contribute to an international maritime information system, aiding voyage planning and monitoring. However, there are certain limitations to be aware of. The OOW needs to remember that not all vessels, especially leisure craft, fishing boats, warships, and some coastal shore stations, including VTS centres, are equipped with AIS. Additionally, very small vessels are not mandated to carry AIS, and AIS can be switched off. Incorrect information about one ship displayed on the

bridge of another can lead to dangerous confusion. All information, such as AIS or automatic radar plotting aid (ARPA) data, is only as accurate as the information from sources such as the gyro and speed log, excluding any equipment error. This has a direct impact on the reliability of the data and information. When using AIS information to assist in collision avoidance decision-making, several cautionary points should be considered:

- AIS serves as an additional source of navigational information, complementing but not replacing systems such as radar, ARPA, and VTS.
- The use of VTS does not relieve the OOW from the responsibility of always adhering to the COLREG.
- Users should not rely solely on AIS for information; instead, they should utilise all available safety information.

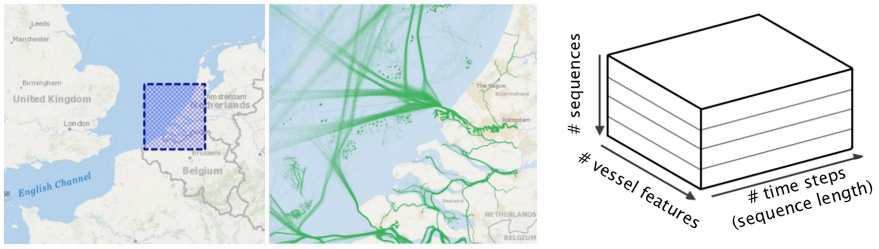
Meanwhile, the radar can only display all surrounding objects' positions, calculated courses, and speeds. However, it should be noted that the radar's range is not as extensive as AIS, yet it is the primary instrument for collision prevention.

4.1.1 AIS from Ship Finder API

At the United Nations Conference on Trade and Development, the Review of Maritime Transport 2023 highlighted that the Netherlands, particularly the Port of Rotterdam, continues to dominate European maritime trade. In 2022, the port recorded a container throughput of approximately 14 million twenty-foot equivalent units (TEUs), solidifying its position as Europe's busiest container port [94]. Due to this significant traffic density and strategic importance, the Netherlands was selected as the primary region for this study on vessel trajectory prediction. Traffic data for this region was obtained from Ship Finder, a provider of vessel tracking and maritime intelligence services, which relies on coastal AIS network stations (historical sea traffic data can be downloaded at MarineCadastre[1] or purchased from MarineTraffic[2]). Although the collected dataset exhibited specific gaps and occasional signal disruptions

[1] https://marinecadastre.gov/ais

[2] https://www.marinetraffic.com/en/p/ais-historical-data



(a) **(b)**

Figure 4.1: Collected data views: (**a**) visualisation of AIS traffic data in the research region; (**b**) three-dimensional view of the vessel data structure.

inherent to AIS-based tracking systems, the data quality was sufficient for conducting the intended trajectory prediction experiments. A graphical image was generated to visualise the available data (see Fig. 4.1). Densely spaced data indicate that traffic flow is significantly more intense. The primary dataset is three-dimensional; the dimensions can be viewed by axis: sequences, vessel features, and time steps.

Cargo ships are one of the most common industrial vessels in maritime traffic, transporting goods and materials in containers between ports. Due to their abundant AIS data availability, cargo ships were explicitly selected for the initial research stage, which involved comparative analyses of different RNN architectures. In subsequent stages conducted in the Baltic Sea region, the research was expanded to incorporate all vessel types by integrating categorical data into the neural networks, allowing for comprehensive multi-type vessel trajectory forecasting. In the Netherlands, 5 months of cargo data were collected between September 2018 and February 2019 (nearly 21 million records), distributed in the territory of the North Sea. The research region's latitude range varies from 51'635 ° to 52'12 °, while the longitude value ranges between 3′10 and 4′50.

Raw vessel records are stored in multiple CSV files, segmented by date-time periods. The collected AIS data contains inconsistencies such as duplicate records, missing vessel characteristics, and undefined data types. To ensure data quality, pre-processing was applied (subsection 4.2) to filter incomplete records and standardise key vessel attributes. The dataset consists of

essential vessel features that characterise traffic flow within the selected region, enabling a structured approach to trajectory prediction. The dataset consists of the main vessel features, which represent traffic flow in a particular region:

- VesselId (MMSI): maritime mobile service identity (Id used only for sequence generation);
- Latitude: geographic latitude coordinate;
- Longitude: geographic longitude coordinate;
- SpeedKnt: vessel speed measured in knots (Speed Over the Ground, SOG);
- HeadingDeg: vessel sailing direction;
- DateDiff: the difference between two time steps in the trajectory (4.1), measured in minutes;
- Δ Latitude: latitude difference of two time steps in time (4.2);
- Δ Longitude: longitude difference of two time steps in time (4.3).

Changes in coordinates over time between two points can be derived through feature engineering techniques [99]. These values represent the vessel's movement speed between consecutive time steps. The derived features are integrated with existing parameters, combining temporal and spatial information to enhance the predictive capability of the model:

$$X_{\Delta t} = t_s - t_{s-1} \tag{4.1}$$

$$X_{\delta Lat} = \frac{X_{Lat,s} - X_{Lat,s-1}}{X_{\Delta t}} \tag{4.2}$$

$$X_{\delta Lon} = \frac{X_{Lon,s} - X_{Lon,s-1}}{X_{\Delta t}} \tag{4.3}$$

where:

- s is the time in different measurement intervals,
- $X_{\Delta t}$ is the time difference value of the previous step,

- $X_{\delta Lat}$ is the latitude coordinate (variation),
- $X_{\delta Lon}$ is the longitude coordinate (variation).

This dataset has several limitations: firstly, the amount of data is insufficient (compared to the Danish historical dataset, which represents only 10% of the sample); secondly, it lacks frequent time series observations and additional relevant features. A key challenge is the inconsistency in trajectory sequences due to missing vessel motion observations, leading to disproportionate changes in time, space, and location. While such issues are often addressed by interpolating missing records through resampling techniques, this study did not apply such methods for data imputation. The dataset originates from AIS tracking, where vessel observations may be delayed or intermittently lost, resulting in data gaps. Instead of filling in missing values, resampling was applied in the Baltic Sea region dataset to standardise the time intervals between observations, ensuring structured and uniform data for analysis.

4.1.2 AIS from the Danish Maritime Authority

The Danish shore-based AIS system, which gathers and transmits data, is under the ownership of the Danish Maritime Authority. While an external supplier handles a portion of the system's operation and maintenance, the responsibility for data storage lies with the Agency for Governmental IT Services (Statens IT). The Danish Maritime Authority stores the Danish AIS data and provides real-time and historical access.

The validity of the initial results for the Netherlands region was verified by experiments on a dataset from the Baltic Sea region close to Bornholm using identical features. The AIS satellite data were obtained from a Danish maritime service[1]. The location of the selected region is in the Baltic Sea, and it borders Sweden, Poland, and Denmark (southeast of Copenhagen). This region boundary box coordinates (see Fig. 4.2): west = 12'00 °, east = 15'00 °, north = 56'00 °, south = 54'00 °. The data cover a six-month observation period for selecting vessels (June 2021 to December 2021) belonging to the cargo vessel type, including others. The dataset consists of the main vessel

[1] http://web.ais.dk/aisdata/

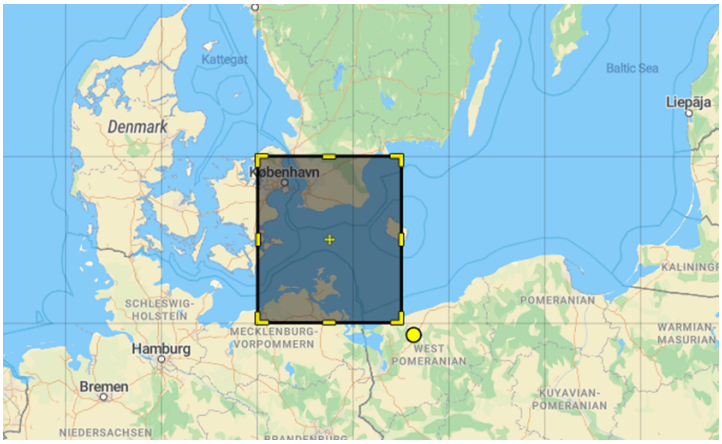


Figure 4.2: Research area in the Baltic Sea off the island of Bornholm.

features: geographic coordinates (latitude, longitude), speed measured in knots, sailing direction, and time series - the difference between two time steps in the trajectory (minutes). Nearly 100 million observations with data points are available.

Regardless of the features chosen for the research, the Danish Maritime Authority also provides other information, which can be static, dynamic or voyage. The data are stored in CSV format, without needing an additional API to retrieve them, and include vessel trajectories and identifiers (Maritime Mobile Service Identity, MMSI). Vessel trajectories are recorded with various attributes, including timestamps, coordinates, speed, course, vessel type, and other navigational parameters. AIS information comes from the vessels' instruments. The recorded information may contain errors or technical faults that can misrepresent a vessel's speed or position. Therefore, data pre-processing and handling steps are necessary as part of the data mining stage to ensure accuracy.

A frequency graph was created to visualise the distribution of data points within a defined geographical region (see Fig. 4.3(a)). Further data exploration, illustrated in Figure 4.3(b), reveals that almost half of the original dataset comprises observations of vessels with a speed of 0 knots, indicating that these vessels are stationary or moored. Since these stationary data points introduce noise in vessel trajectory predictions, they were excluded from

subsequent analysis in both regions examined in this thesis. A representation of the frequency distribution of the data after filtering is given in Appendix A.

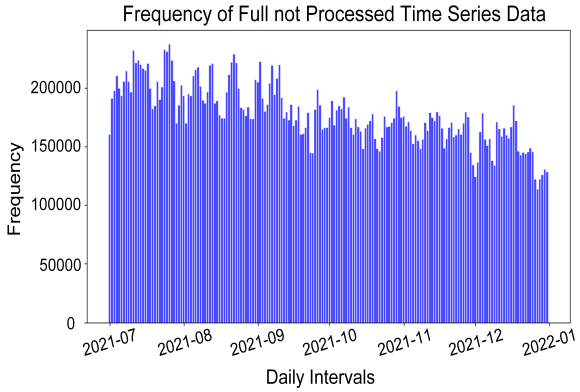
4.1.3 Weather Data

One of the important factors influencing the movement of vessels is weather. For instance, wind can affect aerodynamic force, depending on whether the vessel is moving windward or leeward. Depth is another important aspect, measured as the distance from the sea surface to the seabed, which can vary due to tides caused by wind or moon phases, potentially leading to grounding incidents. Seafarers must often consult nautical charts to navigate these changes. Historical weather data are provided by various information systems, both in real time and historically. One such provider is the Weatherbit.io API [G.3], which offers high-quality weather forecasts, observations, and historical weather data. This API integrates historical weather data from the NOAA Integrated Surface Database, MADIS, and GHCN datasets. Additionally, Weatherbit.io enhances its historical datasets by sourcing data from alternative sources, including historical satellite data, rainfall radar, and reanalysis projects, to build the most comprehensive historical weather data record possible.

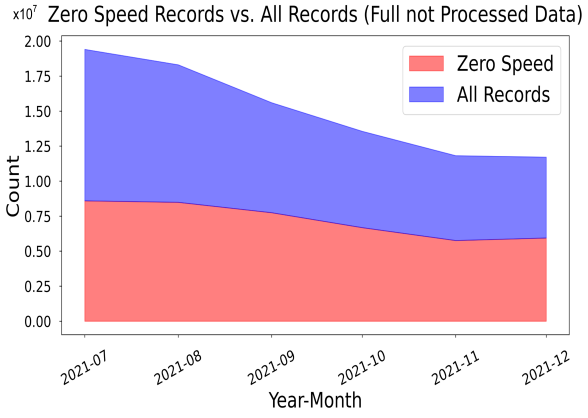
In Figure 4.4, a grid has been constructed to extract weather data for a given position, covering the Baltic Sea region in the second half of 2021, ensuring that AIS and weather data overlap. The dataset contains hourly aggregated observational data, including air temperature, wind direction and speed, wave height and direction, swell height and direction, cloud cover, visibility, water temperature, atmospheric pressure, and additional marine weather parameters. The weather data grid was spatially matched to AIS data points by identifying the grid point closest to each AIS position based on latitude and longitude. Each AIS observation was thus assigned the corresponding weather conditions from its nearest grid location.

4.2 Data Pre-processing and Sequence Generation

As the research is based on the principles of supervised training, where input variables X and output variables Y are available, DL architectures can learn the



(a)



(b)

Figure 4.3: Raw data visualisation: (**a**) frequency of not processed time series data; (**b**) proportion between stationary and moving vessels.

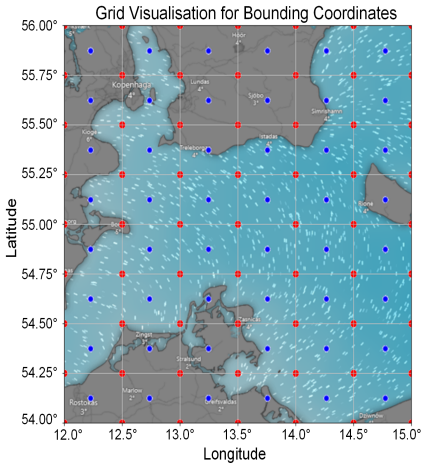


Figure 4.4: Weather region and API coordinate grid points.

interdependencies of these variables, allowing for the reconstruction of further vessel trajectories from new data. The trajectory forecasting by time steps can be divided into one-time step prediction and multiple time-step prediction. The forecasting problem is being analysed by Xu Liu et al. [57] in their work to create a bike-sharing opportunity and predict their availability. The authors argued that a multi-time-step output LSTM is much better than the standard uni-time-step output. Based on their research and the gathered AIS data, the researchers developed a multivariate multi-step data structure.

The Netherlands region's raw data must be analysed, structured and prepared for use in the recurrent networks. The main dataset is three-dimensional (see Fig. 4.1(b)), the dimensions of which can be viewed by axis: sequences, vessel features, and time-steps. Analysing the data from the Netherlands region, it was observed that only about $1/4$ of the vessels were actively moving during the studied period. A vessel is considered stationary or drifting when its speed attribute value is zero. After several experiments, results showed that the dataset with stationary vessel data contains noise that can offset net weights for forecasting progress. Therefore, these records were eliminated, leaving only vessels in motion (more than 5 million records).

The Baltic region's raw data were processed: consolidated, cleared out,

and structured into sequences. Observations with a vessel having zero speed were removed from the data as noise. Also, time series data resampling was performed, allowing the frequency of time step observations to be changed. In the original dataset, observations were made several times per minute per vessel. It was decided to select a time series every minute, applying the k-nearest neighbour (k-NN) algorithm and thus eliminating widespread observations. Here, the time difference feature is calculated in seconds. The other processes shown in the workflow diagram in the thesis's first chapter are identical in both regions. The vessel flow was sorted by time and MMSI and cut into equal-length sequences of 50 time steps. Normalisation was performed, and the sequences were divided into training, validation, and testing samples with already applied transformations described in this article. The total observations are ~11,450,000 (records) and ~460,000 generated sequences, of which 30 observations are input, and 20 observations are output (forecast). Considering standard maritime operational practices, a sequence of 30 historical AIS observations provides a sufficient representation of vessel trajectory, while forecasting 20 future observations align closely with the estimated stopping time and reaction intervals required to effectively assess collision risk and perform preventive manoeuvres. The distribution of time steps between observations is, on average, about 1 minute. The average movement of the vessel is 340 meters, so here, in one sequence, the vessel moves on average about 16 km (this is almost 10 km at the input and 6 km at the output). Statistical estimates are shown in Table 4.1.

Table 4.1: Sequence data characteristics.

Feature	**Value**
Total sequence length	50 time steps
Input trajectory length	30 time steps
Output trajectory length	20 time steps
Average time step change in time	~60 s
Average distance change in space	~320 m
Average length of one trajectory	~16 km
Number of generated sequences	943,584

The decomposition process of sequencing is illustrated in Figure 4.5, where vessel features (F_n) are segmented into fixed-length sequences (i) using

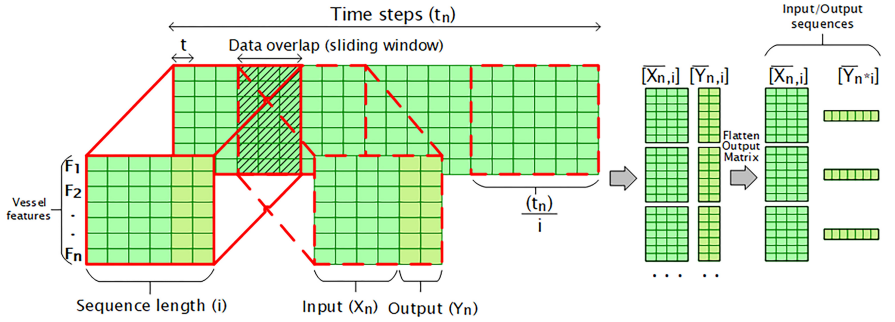
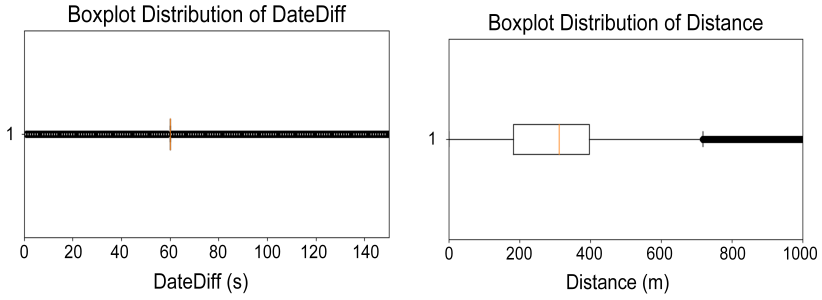


Figure 4.5: Vessel data sequencing.

a sliding window approach with overlapping time steps (t_n). This method ensures that each new sequence retains past observations, which is essential for time series forecasting in recurrent networks. The overlap strategy helps maintain temporal continuity by dividing each sequence in half, allowing the model to capture dependencies over time. Each sequence is further split into input (X_n, F_n) and output (Y_n, F_n) matrices, representing feature-label pairs in a supervised learning regression framework. The input features are passed into the network, which predicts the output features. Since all deep recurrent architectures in this study employ a two-dimensional output structure, the output matrix is flattened before training to ensure compatibility with the model's final prediction layer. Figure 4.7 illustrates an example of the generated vessel movement sequences in the Baltic Sea region, visually demonstrating the extracted trajectories used in the study.

Specific criteria were applied for sequence generation to ensure consistency and reliability across datasets. Each sequence is classified based on vessel type (e.g., cargo or other categories), and no sequence contains data from multiple MMSI identifiers. In the Netherlands dataset, a 150-minute filter was applied to smooth the time steps and eliminate large gaps caused by inconsistent AIS observations. This ensures that no time step interval exceeds 150 minutes, maintaining a stable temporal relationship between consecutive data points. The filter threshold was determined based on the frequency distribution of time step differences, with the highest data concentration occurring within a 150-minute range.



(a) **(b)**

Figure 4.6: Raw data filtering distributions: (**a**) distribution of time series between $t+1$ records; (**b**) distribution of distance between $t+1$ records.

In contrast, the Baltic Sea dataset required different filtering parameters due to variations in regional traffic density and AIS data sources. Here, an 800-meter spatial filter and a 120-minute time difference filter were used to account for the higher frequency of recorded observations (see Fig. 4.6). These adjustments reflect the distinct characteristics of AIS data collection in different maritime regions, optimising sequence continuity and preserving the integrity of vessel trajectory predictions. The complete data source and filtering are provided in Appendix B.

4.3 Data Frequency Adjustment

4.3.1 Irregular Time Step Intervals in the Netherlands Region

One of the key areas for improving the Netherlands dataset is compressing the wide range of sequence values into a narrow one between time steps, because the data contains missed vessel motion observations. This means there may be a disproportionate trajectory change in time, space, and location. Tasks of this nature are often addressed by filling in the gaps with new data that mimic the lack of records (resampling), but this method has not been applied in this region. The actual data are from AIS, which means that the tracking of the vessel's observations may be interrupted (or delayed) in the system, and data gaps may appear. One such trajectory is shown in Figure 4.8(a). This visualisation

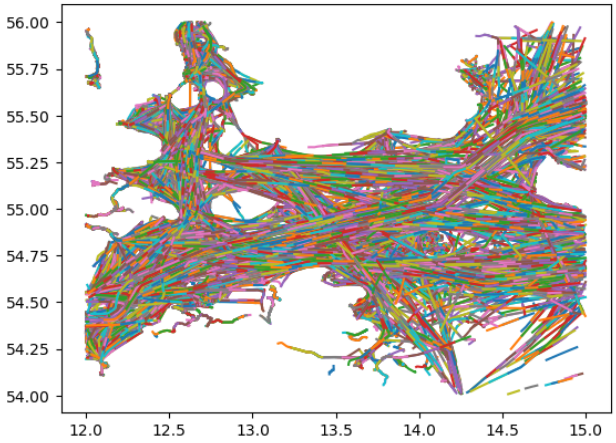
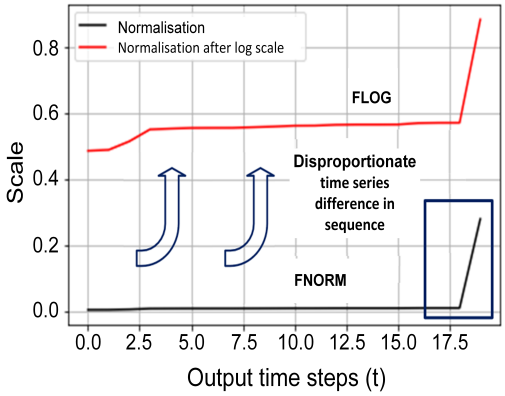
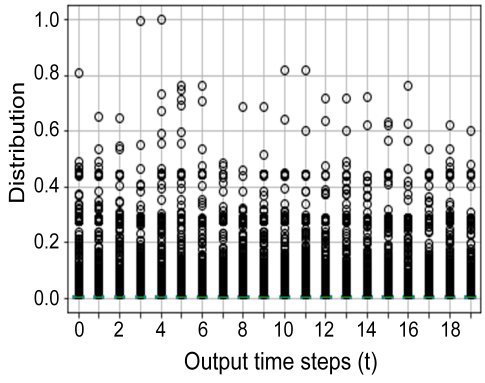


Figure 4.7: Generated vessel movement sequences in the Baltic Sea region, where each colour represents the trajectory of an individual vessel over time.

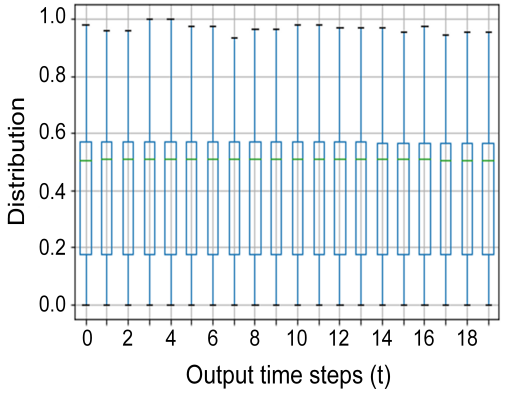
demonstrates how the longitude coordinate of consecutive points changes with the vessel moving further due to larger and irregular time steps toward the end of the trajectory sequence. Applying normalisation to such a trajectory (4.5) introduces another challenge: time series with higher data density become excessively flat, with values approaching zero. Moreover, extensive data gaps are rising critically and are becoming close to one on the y-axis. First, this problem is minimised by applying a logarithmic scale (4.4), which reduces and stabilises data gap spikes of the original data. Only then does normalisation (4.5) apply to the entire set. This requires one additional step later to invert the exponential function, returning the normalised and logarithmic values to the initial coordinate system. The image shows that the trajectory is better distributed (normalised) between 0 and 1 after the additional scale. Figure 4.8(b) shows how the time steps (date difference feature) are distributed over the entire data sample with normalisation only, and Figure 4.8(c) shows how the normalisation covers the whole sample from 0 to 1 with the $FLOG$ function. The logarithmic scale makes it possible to create more favourable conditions for the activation function to work better and perform a better dispersion of normalised values between the minimum and maximum range. The given functions, where X denotes the vessel's features, are general.



(a)



(b)



(c)

Figure 4.8: Dataset features transformation using normalisation and logarithmic scaling: (**a**) a single sequence longitude normalisation, where the x-axis represents time steps and the y-axis contains normalised values [0,1]; (**b**) normalised date difference feature distribution over the output time step; (**c**) date difference feature distribution after FLOG scale and normalisation.

$$FLOG = \log(x) \tag{4.4}$$

4.3.2 Time Frequency Resampling in the Baltic Sea Region

A resampling method (see Fig. 4.9) is applied to address challenges in AIS data, such as signal gaps, variability in reporting frequencies across different vessels, and overlapping observations. In the Baltic Sea region, where maritime data analysis requires consistent time intervals for accurate modelling, the k-NN method is utilised to standardise the time series. This approach ensures a uniform one-minute time step, mitigating inconsistencies in the dataset while preserving essential patterns and trends.

The resampling process begins by identifying periods with multiple AIS observations within each one-minute interval. The k-nearest data points are then selected and averaged, effectively reducing redundancy while maintaining the structural integrity of vessel trajectories. This method is particularly useful when multiple observations are recorded rapidly, preventing excessive data density and making the dataset more manageable for trajectory prediction models. By enforcing a uniform time interval, the k-NN resampling process eliminates irregularities while preserving the vessel's movement characteristics, ultimately enhancing the accuracy and reliability of predictive models for maritime monitoring and collision risk assessment.

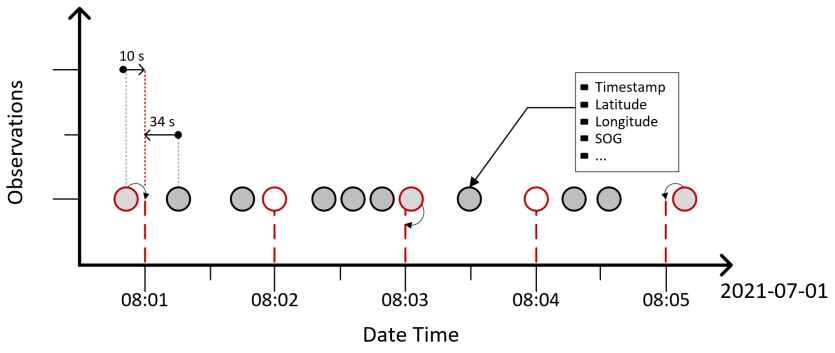


Figure 4.9: AIS data resampling based on time series. Grey dots represent all AIS observations, while the red boundary circles indicate points selected using the k-NN method for standardising steps.

4.4 Embedding Vessel Types and Permutations

The vessel features and categorical data encoding techniques are represented in the LSTM AE neural network architecture (see Fig. 4.10). The figure shows that the numerical characteristics of the vessel do not change in the experiments and that the vessel type features are connected by different encoding in each empirical test. The input matrix into the RNN is distinctly disparate from the encoding process, introducing an additional layer of complexity to the algorithm, as demonstrated through empirical experimentation. The figure shows how the data shapes are combined, how it is transmitted across different layers and the output. For example, the shape (None, 30, 7) indicates that the data are multidimensional, consisting of 30 time steps with 7 AIS features per input sequence. Meanwhile, the shape of the output sequence (None, 40) indicates a stack of two attributes (longitude and latitude or delta coordinates) for 20 time steps when the shape is flattened. Different encoding techniques add additional features to the numerical input data.

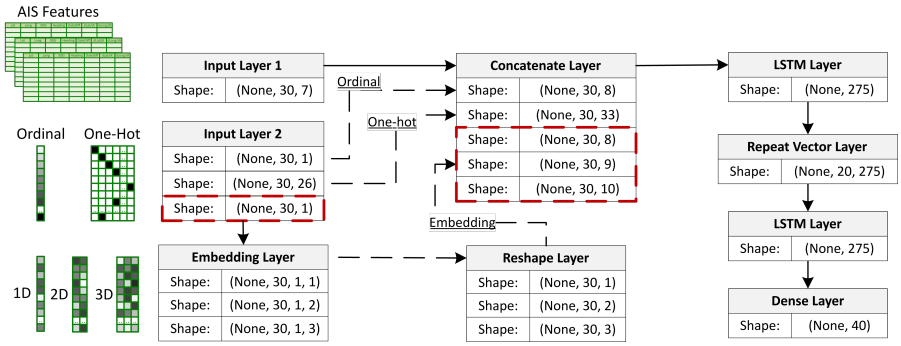


Figure 4.10: LSTM AE architecture with different vessel type encoding techniques.

A vessel observation captures a vessel type identifier, of which there are more than 20 in the region and over the period under study (see Fig. 3.5). The area is characterised by the emergence of maritime routes and maritime highways, where historical traffic is most intense. The more intense traffic is concentrated in certain ship types: cargo, passenger, tankers, and other ship types that are distributed throughout the region. Still, different ships have

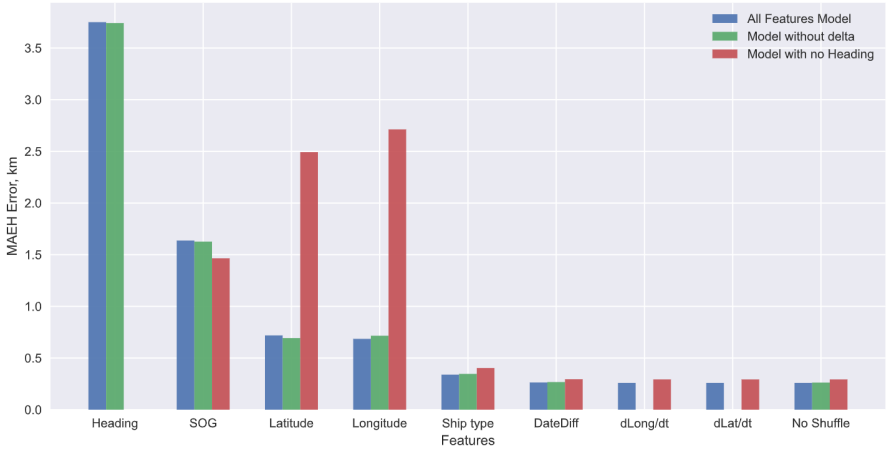


Figure 4.11: Permutations of dataset features.

specific routes and movements, such as pollution clean-up, harbour maintenance, spare, ground effect vehicle (wing-in-ground-effect, WIG), etc.

In addition, feature permutations are added to calculate the permutations (see Fig. 4.11), which show how each feature individually affects the ship's forecast. The permutations are calculated by shuffling the positions of the features relative to each other when the accuracy of the network model is evaluated. The image shows that all the selected features are essential, but the most remarkable are direction, speed, and coordinates. It is evident that removing the heading feature significantly increases coordinate dependence. More details on feature permutations are given in [67].

4.5 Recurrent LSTM AE

In collision prediction within maritime navigation, a DL model utilising an LSTM AE architecture is detailed in related publications [40]. The model's strength lies in its ability to encode temporal sequences of vessel movements into a lower-dimensional latent space and reconstruct them, capturing the essential features for predicting future trajectories. The architecture comprises three key components: an encoder, a latent space, and a decoder. The encoder processes input sequences, reducing their dimensionality, while the decoder reconstructs

the output from the latent representation. This specific implementation involves a sequence-to-sequence approach, where input data are flattened to transform the time steps into a singular vector that encapsulates the temporal features essential for predicting vessel paths. This architecture has already been selected as the best forecasting architecture following comparisons with other RNNs.

The ensuing stage witnesses the deployment of a sophisticated deep RNN, meticulously trained on the experimental dataset, featuring a dedicated categorical data layer for ship types within the scope of training and validation samples. Acknowledging the non-deterministic nature of neural networks, a series of iterative tests is conducted to enhance the reliability and robustness of the results obtained. The general concept of the data mining workflow can be seen in Figure 4.12.

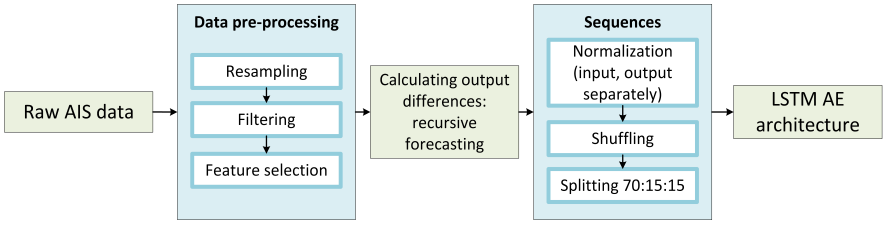


Figure 4.12: General workflow and permutations of dataset features.

During the research, the sequences are constructed from 50 time steps, split into 30 steps for the input and 20 for the output, with an average time interval of 1 minute. The data used for the network comprises features such as encoded vessel type, geographical coordinates, speed, heading, and differences in latitude and longitude between consecutive time steps. The output sequence consists only of the coordinate features, while the input sequence consists of the aforementioned vessel characteristics. These features are valuable for understanding vessel behaviour and predicting further points along the trajectory. The parameters used in the architecture are given in Table (4.2).

The dataset, comprising nearly one million sequences, was divided using a standard 70:15:15 split for training, validation, and testing, respectively, a practice commonly adopted in scientific research [5, 79] to ensure balanced evaluation and generalisation. Before training, the generated and normalised

Table 4.2: LSTM AE configuration.

Parameter	Value	Note
Layers	3	Total number of encoder and decoder parts
Seq. Len.	50	30 input, 20 output
Epochs	100	—
Optimiser	0.001	Adam (with learning rate)
Regularisation	0.01	Dropout layers
Number of Units	275	Cells in each LSTM layer
Batch Size	128	Examples utilised in one iteration
Models Size	20	Models trained on the same data
Loss Function	MSE	Measures prediction quality
Activation Function	ReLU	Used in LSTM gates and Dense layers

sequence matrices were shuffled to minimise sequence dependency, reduce variance, and improve gradient variability. Mixing the matrices ensures that each new sequence in the model will be independent of the previous sequence, reducing the variance and making the gradient more variable. This randomisation enhances the model's exposure to diverse scenarios, contributing to better generalisation. Normalisation (4.5) was applied to all numerical features, scaling values between 0 and 1 to ensure uniformity across features and accelerate the learning process. The training subset was used to fit the models, the validation subset was employed to monitor the training progress and prevent overfitting, and the test subset was used to evaluate final performance. Notably, the study employed not a single model but an ensemble of models trained on the same dataset. Each model, influenced by factors like regularisation and initialisation, produces slightly different predictions, enabling a more comprehensive uncertainty assessment and supporting the robustness of the applied collision detection techniques. The cross-validation results shown in Appendix C further illustrate the models' convergence and stability:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4.5}$$

where:

- X is the feature of the vessel,
- X_{norm} is the normalised value,
- X_{max} and X_{min} are the maximum and minimum value of the feature.

The experiments were performed by changing the architectures and the number of cells/units. The same network parameters and cell size interval keep the same conditions for all architectures, and the same data order for training. In the literature, RNNs commonly utilise cell sizes ranging from tens to several hundred. Initially, the research followed this practice, evaluating cell sizes incrementally from 25 to 300 cells, increasing by increments of 25. However, to rigorously test the effect of cell size and verify whether the prediction accuracy improves or stabilises beyond conventional sizes, further experiments extended the range up to 5000 cells, employing larger increment steps. Each cell-size configuration across different architectures, including experiments with embedding techniques, was trained 10 times to ensure reliability and stability in predictions. Subsequent experimental phases involving coordinate system transformations, uncertainty quantification, and collision detection required training additional models, resulting in an extensive set of trained configurations. This comprehensive experimental approach enabled robust validation of model performance across various conditions.

Various hyperparameter modifications were tested during the experiment, such as the changed activation functions RELU and TANH. As a prevention for regression models' overfitting, L2 and dropout regularisation techniques were tested. However, the most suitable parameters were obtained after practical testing with the components listed in Table 4.2.

4.6 Collision Probability Estimation

According to one of the primary objectives, this thesis aims to enhance maritime collision risk assessment by integrating prediction PIs, CIs, EPRs, and CPRs. These interval-based methods provide a probabilistic understanding of vessels' future positions, accounting for uncertainties inherent in maritime navigation,

and allow for the definition of potential collision boundaries, highlighting areas based on LSTM AE recursive predictions. The LSTM architecture is based on real data from the Baltic Sea region using AIS. By combining confidence levels with these predictive methods, the models dynamically assess collision risks and identify high-risk scenarios from nearby vessels' perspectives, enabling pre-emptive safety measures.

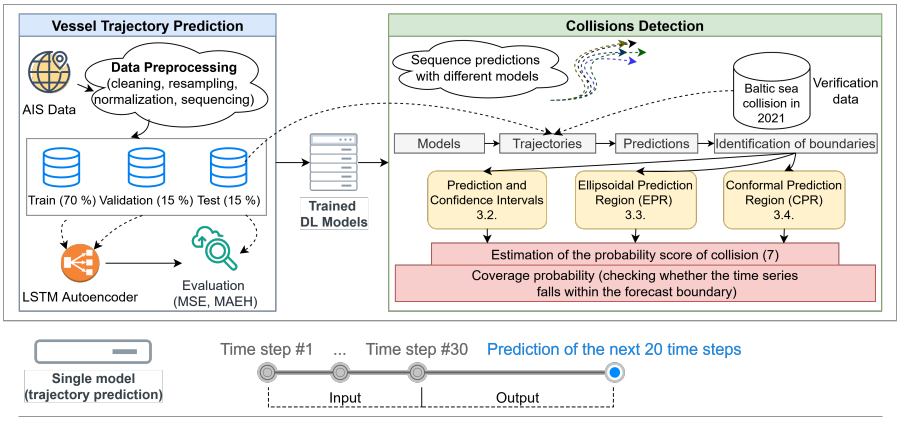


Figure 4.13: Diagram of Vessel Path Prediction, Boundary Mapping, and Collision Detection Workflow.

A comprehensive workflow has been developed to provide a clear study overview (see Fig. 4.13). The diagram illustrates a general flowchart for vessel trajectory prediction, which builds upon the results from the previous studies. The primary focus of this study is on utilising model predictions to define the boundaries of predicted trajectories, assessing overlapping regions with collision risk scores in real-case scenarios, and evaluating prediction accuracy using coverage probabilities. The study employs multiple trained models, each capable of producing different predictions for the same trajectory due to inherent uncertainties, allowing for the generation of diverse samples for probabilistic estimates. These methods are discussed in detail throughout the thesis. Notably, each model accepts an input sequence of 30 time steps (equivalent to 30 min) and predicts the next 20 time steps (equivalent to 20 min). The process involves evaluating the predicted trajectories using a test data sample, comparing the real-time steps of the vessel's movement against the prediction bounds, and

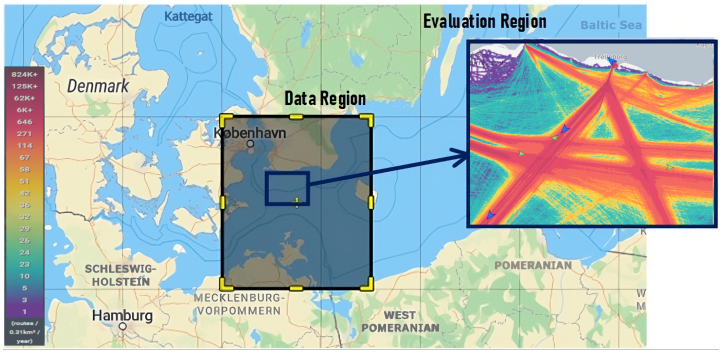


Figure 4.14: AIS study region with highlighted evaluation area.

calculating how many actual positions fall within these predicted regions. The study is validated using a historical maritime accident in the Baltic Sea in late 2021. This case study demonstrates how different methods generate distinct boundaries and how these overlapping thresholds with collision risk scores can be evaluated from the perspective of multiple vessels, highlighting the practical application of the framework in real-world collision detection scenarios.

The geographical scope of the data collection is shown in Figure 4.14. The regions used for simulated test case predictions are delineated in smaller subsets of the data, called evaluation regions, in the figure. This subset is chosen due to its high traffic density and proximity to ports. The research methods are evaluated in this particular sample, although predictions could also be made using other sequences over a wider range of data regions.

Predictions must be made for each sequence using a different model to construct and evaluate the prediction regions. Figure 4.15 shows randomly selected sequences and their trajectory predictions generated by the 20 models. Each subplot in the figure represents a density estimate or distribution of the predicted points at a given time. Areas with higher density around the geometrical centre are shown in warmer colours (yellow tones), and areas with lower probability are shown in cooler colours (blue tones). Actual positions are marked with red dots corresponding to the 20-minute further predictions. The plot indicates that the data do not exhibit a perfectly normal distribution, as evidenced by the absence of a single distinct peak. This characteristic poses

limitations for using parametric methods such as PI and CI, which rely on assumptions of normality. However, forecast points are generally centralised, especially early in the forecast period. This tendency to focus on the early stages of forecasting is common in predictive analytics.

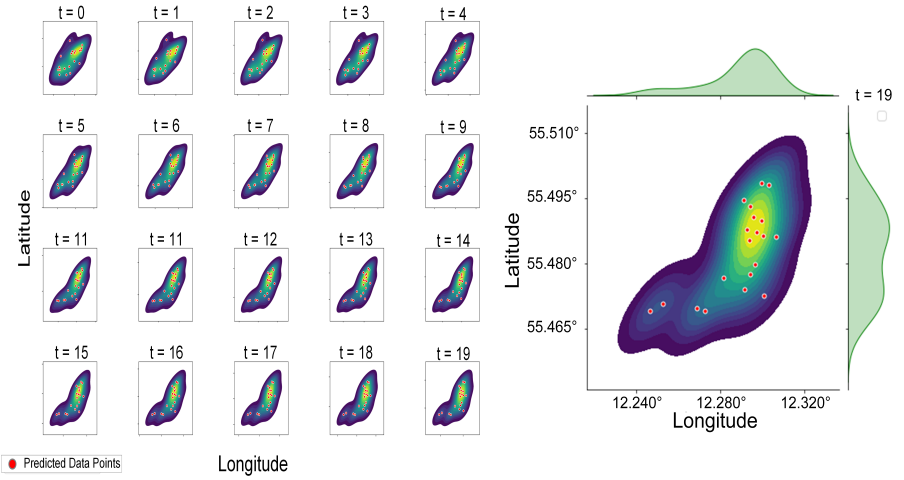


Figure 4.15: Distribution of data points for selected sequence predictions.

The estimated area of the width of a region is used for the standard assessment. The evaluation of the sequences within the sub-region, totalling 8,984, and their corresponding prediction regions hinges on two principal criteria: (a) an examination of whether the actual trajectory of the vessel at each time step resides within the predicted region, a method known as coverage probability; and (b) the verification of the methods' effectiveness in a real marine traffic incident through the calculation of probabilistic risk score of collision. All the methods used correspond to a 95% confidence level.

Suppose $A = \{a_1, a_2, \ldots, a_n\}$ represent the set of actual points, where a_i is the actual coordinate at time step i, and let $Z = \{z_1, z_2, \ldots, z_n\}$ represent the set of predicted zones, where z_i is the predicted zone at time step i. We define the indicator function (4.6) I for each time step i as follows:

$$I(a_i, z_i) = \begin{cases} 1 & \text{if point } a_i \text{ is contained within zone } z_i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.6}$$

The score for the full set of sequences, assessing the predicted zones' suitability for the actual points, is the sum of the individual scores over all time steps (4.7), where n is the total number of time steps or points. The accuracy of the predicted zones can then be expressed as a percentage (4.8). This accuracy metric shows the percentage of time steps where the actual coordinates fall within the predicted intervals.

$$\text{Score} = \sum_{i=1}^{n} I(a_i, z_i), \tag{4.7}$$

$$\text{Accuracy} = \left(\frac{\text{Score}}{n} \right) \times 100\%. \tag{4.8}$$

To calculate the probability of intersection, for example, between the predicted trajectories of ships (which are represented as ellipses/circles within prediction zones), we use the following formula (E):

$$P(\text{collision}) = \frac{V_{A \cap B}}{V_A + V_B - V_{A \cap B}}, \tag{4.9}$$

where V_A and V_B are the areas of the individual regions, and $V_{A \cap B}$ is the area of their intersection. The formula evaluates the areas of both zones and the overlapping area. If the two regions do not intersect, the overlapping area is zero, resulting in a zero probability of intersection.

The idea is related to a classical probability formula, where the probability of an event C happening is the number of outcomes that result in C divided by the total number of possible outcomes. We assess the likelihood of ships intersecting based on their projected paths within their respective prediction zones.

4.7 Summary of the Chapter

This chapter outlines the preparation of empirical experiments designed to validate the proposed methods and enhance vessel-trajectory prediction using AIS data from two areas: the Netherlands and the Baltic Sea region around Bornholm Island, Denmark. The AIS data from the Netherlands, extracted

from Ship Finder, provided information on cargo ships, a prevalent type of industrial ship, comprising more than 5 million records after filtering. Despite the abundance of data, the dataset had limitations due to insufficient size and noticeable gaps in the time series. Simple normalisation overly compresses dense sequences, flattening values near zero. Applying logarithmic scaling effectively redistributed values across the entire range, ensuring an even and consistent representation of the date difference feature from 0 to 1.

Conversely, the Danish AIS dataset, owned by the Danish Maritime Authority, was significantly larger, comprising approximately 11,450,000 records from all vessel types. After filtering and resampling to adjust the frequency of time-step observations, around 460,000 vessel movement sequences were generated. By applying the k-NN algorithm, observations were resampled to a consistent one-minute interval. Feature permutations were also introduced to evaluate how individual features impact the ship's forecast by shuffling their positions relative to each other during model evaluation. The data frequency adjustment addressed irregular time step intervals in the Netherlands region through logarithmic normalisation and applied time frequency resampling in the Baltic Sea region to reduce observation redundancy.

The chapter also describes the implementation of a DL model using an LSTM AE architecture for collision prediction within maritime navigation. The model's performance was optimised by evaluating various hyperparameters, resulting in an optimal configuration of 275 LSTM cells per layer, Adam optimiser with a learning rate of 0.001, regularisation using dropout layers at a rate of 0.01, MSE as the loss function, and Rectified Linear Unit (ReLU) activation. The accuracy of the neural networks was assessed using rigorous regression metrics, including MAEH, MSE, MAE, RMSE, and MAPE, to evaluate the model's performance in predicting coordinate values. The objective was to minimise the prediction errors as much as possible.

Additionally, the weather dataset used in the experiments provided comprehensive information such as wind direction, wind speed, temperature (nighttime and daytime averages), mean sea level pressure, humidity, visibility, cloud cover, moon phase, and other similar observational data. These data were presented in a daily time series, recording the aggregated weather changes

every hour. The permutation experiments were conducted by shuffling feature positions relative to each other and assessing their impact on model accuracy, revealing that the most critical AIS features for trajectory prediction are heading, speed (SOG), latitude, and longitude.

Furthermore, the collision risk score was assessed using interval-based methods described in Chapter 3, including EPR, CPR, PI and CI. These techniques were applied to estimate the areas of predicted regions and evaluate the coverage probability, ensuring that the actual trajectory of the vessel at each time step falls within the predicted region. The effectiveness of these methods was also verified by calculating the probabilistic risk of collision in real marine traffic incidents. All methods employed corresponded to a 95% confidence level, ensuring robust and reliable predictions.

The DL models are the foundation for predicting vessel trajectories, which are then used to assess collision risks. Each model generates a probabilistic forecast of the vessel's future positions, accounting for uncertainties inherent in maritime navigation. The predicted trajectories are further analysed using statistical techniques such as prediction intervals, EPR, and CPR to establish boundaries for possible collision zones. The likelihood of a collision is evaluated by comparing the predicted boundaries with the actual positions of vessels. This integrated approach ensures that the collision risk assessment is closely tied to the accuracy of trajectory predictions, providing a dynamic and probabilistic evaluation of potential collision scenarios.

5 Results

This chapter presents the results of the experiments conducted to improve vessel trajectory prediction and assess collision risks. The study begins with a comparative evaluation of various DL models, focusing on RNNs using the Netherlands dataset. This initial analysis identifies the most effective architectures for trajectory forecasting by assessing prediction accuracy and the impact of hyperparameters, such as cell size.

Building on these findings, the top-performing RNN models, LSTM AE, bidirectional LSTM, and GRU, are further evaluated in both the Netherlands and the Baltic Sea regions. The experiments incorporate multiple data transformations, including coordinate system adjustments and recursive forecasting strategies, to determine their influence on prediction accuracy. Following this, the most promising model - the LSTM AE - was selected for further refinement. Additional experiments examine the integration of categorical vessel-type data using various encoding techniques, as well as the incorporation of meteorological information to assess their contributions to prediction performance.

Finally, the chapter presents the results of collision risk assessment using real-case maritime incidents. The developed models are applied to historical vessel collisions, where calculated trajectory boundaries and probabilistic estimations are used to quantify potential risks. This practical validation demonstrates the applicability of the proposed methods in real-world maritime safety operations. These results were presented in the main papers by Jurkus et al. [A.1, A.2, B.1, B.2].

5.1 DL Models Evaluation

The models were evaluated using test samples after training six deep RNN architectures. The MSE loss was computed as the average across all experiments, normalised by the number of trained models. The results, presented in Table 5.1, provide a comparative analysis of each architecture's prediction accuracy. A

Table 5.1: MSE loss results with different cell hyperparameter sets.

Cells	RNN $\times 10^{-4}$	Basic LSTM $\times 10^{-4}$	LSTM Stack $\times 10^{-4}$	GRU $\times 10^{-4}$	AE $\times 10^{-4}$	BiLSTM $\times 10^{-4}$
25	6.910	5.181	4.434	5.563	4.581	6.582
50	5.334	4.381	4.361	4.620	3.944	3.939
75	4.904	4.211	4.146	4.290	4.010	4.177
100	4.819	3.947	4.150	4.021	4.031	3.697
125	4.702	4.039	4.075	3.914	**3.688**	3.721
150	4.739	**3.875**	4.240	3.916	3.996	3.683
175	4.635	3.920	4.067	3.914	3.789	3.822
200	**4.628**	3.999	4.564	3.826	3.706	3.900
225	4.679	4.024	4.482	3.819	3.987	3.636
250	4.630	4.085	4.133	3.742	3.946	3.736
275	4.660	3.998	4.099	3.827	3.724	3.645
300	4.681	4.086	**3.956**	**3.710**	3.782	**3.618**

lower MSE value indicates higher prediction accuracy, with the most effective model achieving the smallest error in vessel trajectory forecasting. Regression prediction is considered accurate when the error value is closest to zero.

Table 5.1 highlights each architecture's lowest error values (marked in bold) at the optimal cell size. The results indicate that the Bidirectional LSTM achieved the lowest error on the test dataset, followed closely by the AE and GRU models. A comparative ranking of all architectures, based on their best-performing cell sizes, is visualised in Figure 5.1.

Descriptive statistics are illustrated through box plots (see Fig. 5.2), which depict the distribution and skewness of the distance loss data for the recursive multi-step prediction strategy. The comparison evaluates different architectures by varying the number of cells and computing the average distance errors across experimental trials. The results indicate that the AE and bidirectional LSTM exhibit the most stable error distributions, while the latter also shows the most remarkable variation between minimum and maximum error values. Across all architectures, using more than 100 cells consistently reduces error, demonstrating the impact of network capacity on prediction accuracy. For a detailed empirical analysis of how varying the LSTM cell count affects prediction accuracy, see Appendix D. Empirical results demonstrate that MAEH

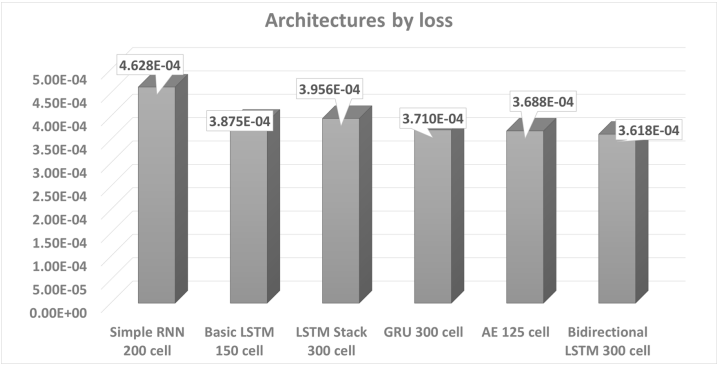


Figure 5.1: Architectures with the lowest loss with a combination of specific cell sizes.

is consistently observed for models employing between 75 and 300 cells, with errors clustered around 0.18–0.20 km. Beyond 300 cells, the prediction accuracy stabilises, and further increments offer negligible improvements. Notably, increasing the cell count significantly elevates computational demands, especially evident in extreme cases (e.g., 5,000 cells), where training time is markedly extended without proportional gains in accuracy, thus highlighting the critical importance of optimal hyperparameter selection.

Figure 5.3 illustrates the results with the lowest average distance errors across different architectures, independent of the selected number of cells. The graph demonstrates that vessel trajectory predictions can be enhanced beyond polar coordinate systems. Specifically, training RNNs using subtracted distances and rotation angles improves accuracy, whether working with direct AIS data or augmented delta features. However, the greatest improvement is observed when employing the UTM projection, which converts geographic coordinates into a two-dimensional space. Among all architectures, AE benefits the most from this transformation, achieving a minimum error of 1.141 km in datasets with irregular time steps, representing an improvement of nearly 30% compared to AIS-based features without delta latitude and longitude adjustments. However, the polar coordinate approach (distance and angle) and the UTM transformation relied on a recursive conversion process, where predicted delta coordinates were interpolated from the last known input point. This recursive recalculation likely

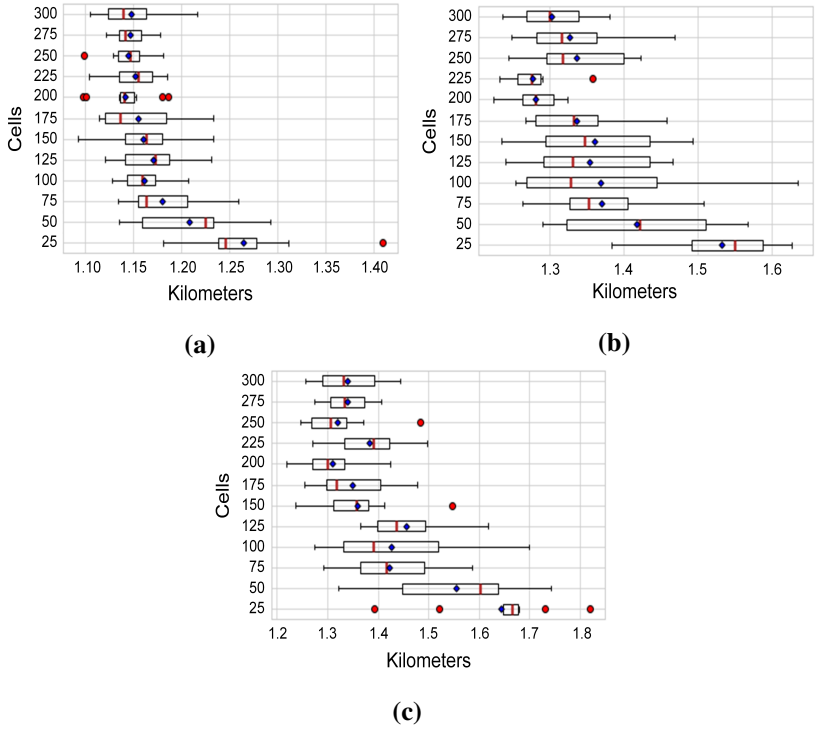


Figure 5.2: Standardised distribution of recursive models error: (**a**) AE MAEH distance; (**b**) bidirectional LSTM MAEH distance; (**c**) GRU MAEH distance.

impacted accuracy more than the UTM transformation itself. This assumption was confirmed through experiments conducted with geographic coordinates, where output sequences were constructed using delta values. The results demonstrated that recursively updating predictions from delta coordinates led to more accurate forecasts than directly predicting absolute position values in the sequence.

5.2 Assessing the Accuracy of Vessel Predictions

Figure 5.4 illustrates a comparative analysis of different coordinate transformations (WGS84 and UTM) and their recursive prediction counterparts based on vector differences. The MAEH values indicate that recursively recalculating predicted positions using delta coordinates significantly enhances prediction

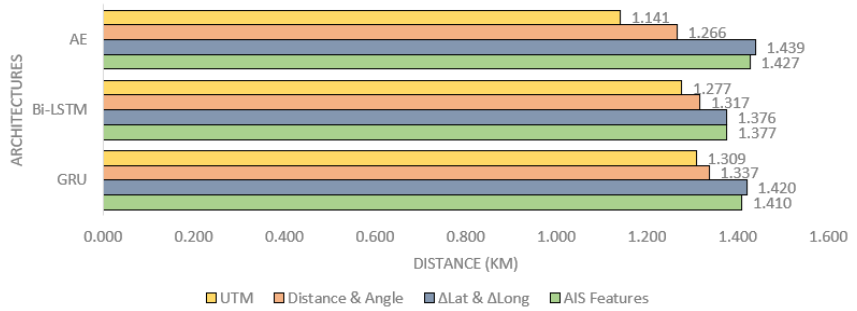


Figure 5.3: Lowest experiment distance loss results in the Baltic region among cargo vessels.

accuracy across both coordinate systems. Specifically, the recursive WGS84 approach achieved the lowest prediction error (0.3537 km), demonstrating superior accuracy compared to standard predictions without recursive recalculations. This emphasises the advantage of employing recursive delta methods in trajectory forecasting over absolute values. Furthermore, incorporating meteorological data, such as wind, wave, and temperature conditions, further reduced MAEH to 0.3203 km, highlighting the added predictive value of environmental context in vessel trajectory modelling. Predictions of the trajectories of random vessels are given in Appendix E.

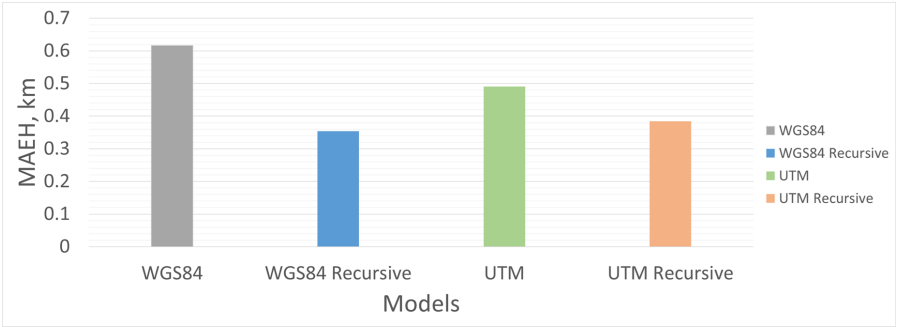


Figure 5.4: Lowest experiment distance loss results in the generalised model between different coordinate systems.

A randomly selected sequence was used for the final network evaluation,

with the results presented in Figure 5.5. The visualisation depicts the best-performing model with optimal cell sizes. The blue line represents the input sequence (30 time steps), while the green line marks the actual vessel trajectory (20 time steps). The predicted vessel movement for the same period is illustrated in yellow, and the grey lines indicate surrounding traffic intensity from other sequences. The results show that the short-term forecast (~10 minutes) aligns closely with the actual trajectory. In contrast, the long-term prediction exhibits a slight decline in accuracy, which is expected in multi-step forecasting. Additionally, the distance error in metres is displayed in red.

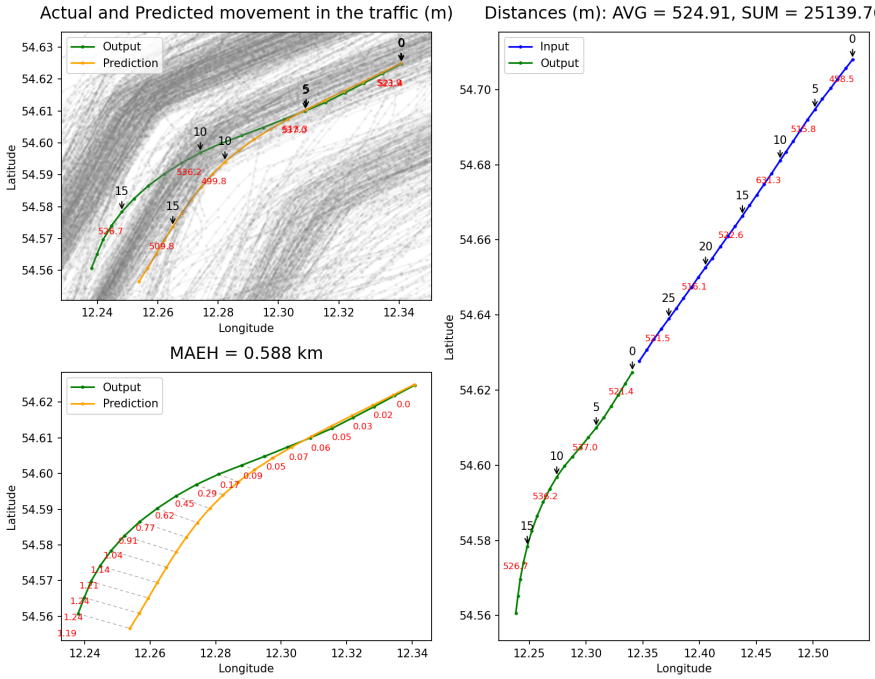


Figure 5.5: Actual and predicted random vessel movement in sea traffic.

The feature permutation analysis identifies the most impactful features for vessel trajectory prediction accuracy (see Table 5.2). Heading and SOG were consistently ranked as the top two most critical features across all models, emphasising their substantial influence on trajectory prediction. Geographical coordinates (latitude and longitude) and their derivatives (dLat/dt and dLong/dt)

also showed considerable importance. Among meteorological data, wave height, wind wave height, water temperature, air temperature, and wind speed were identified as essential features. The last row in the table provides overall prediction accuracy (MAEH error) for each model without feature permutation, serving as a baseline for evaluating feature significance. A larger error resulting from permutation indicates greater importance of that feature in determining overall prediction accuracy.

Table 5.2: Feature importance determined by permutation for three different trained models.

Rank	Feature (Model 1)	Value	Feature (Model 2)	Value	Feature (Model 3)	Value
1	Heading	3.7129	Heading	3.6744	Heading	3.6657
2	SOG	1.7340	SOG	1.7125	SOG	1.7171
3	Longitude	0.4614	Latitude	0.5019	Latitude	0.5209
4	Latitude	0.4535	Longitude	0.4808	Longitude	0.4993
5	dLong/dt	0.4088	windWaveHeight	0.4058	waveHeight	0.3892
6	waveHeight	0.4086	waveHeight	0.4002	windWaveHeight	0.3881
7	windWaveHeight	0.4054	dLong/dt	0.3817	waterTemperature	0.3767
8	dLat/dt	0.4054	airTemperature	0.3792	airTemperature	0.3766
9	windWavePeriod	0.4015	waterTemperature	0.3789	dLong/dt	0.3761
10	waterTemperature	0.4010	dLat/dt	0.3786	dLat/dt	0.3745
11	airTemperature	0.4003	windWavePeriod	0.3765	windWavePeriod	0.3737
12	windSpeed	0.3987	windSpeed	0.3745	windSpeed	0.3732
Overall	**0.3963**		**0.3710**		**0.3693**	

Additional meteorological parameters, including secondary swell height, swell direction, secondary swell period, swell height, secondary swell direction, visibility, pressure, wind direction, swell period, cloud cover, wind wave direction, and wave direction, were also evaluated. Although these did not rank among the top features, their inclusion provides a comprehensive environmental context that potentially enhances trajectory prediction accuracy. These meteorological factors influence vessel dynamics and significantly impact navigation conditions and trajectory movement. Consequently, integrating these top-ranked and additional meteorological features ensures the model's predictions benefit from detailed and context-rich environmental information, capturing the external influences on vessel movements more accurately. To maintain a balanced feature set, seven top-ranked meteorological variables (wave height, wind wave height, water temperature, air temperature, wind speed,

wind wave period, and wind wave direction) were selected, ensuring that the number of meteorological features did not exceed the primary AIS vessel features.

5.3 Impact of Categorical Data on Vessel Trajectory Prediction

The accuracy of the neural networks, specifically the LSTM AE trained on the Baltic Sea dataset, is assessed using rigorous regression metrics on the test dataset. These metrics, including MAEH, MSE, MAE, RMSE, and MAPE, comprehensively evaluate the model's performance in predicting coordinate values and assessing trajectory accuracy with transformations (see Appendix D).

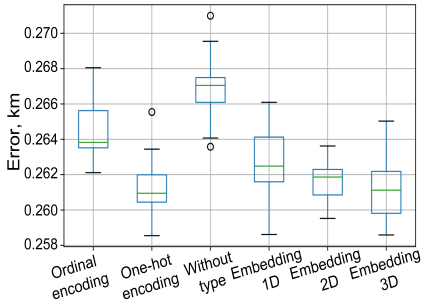


Figure 5.6: Repeated experiment results of the MAEH metric using different encoding techniques for categorical data.

Table 5.3: Mean error metrics of multidimensional embedding.

Metric	Embed 1D	Embed 2D	Embed 3D
MAEH	0.26252	0.26162	0.26124
MSE	2.51E-05	2.49E-05	2.47E-05
RMSE	0.00488	0.00485	0.00483
MAE	0.002046	0.002039	0.002036
MAPE	0.000110	0.000110	0.000109

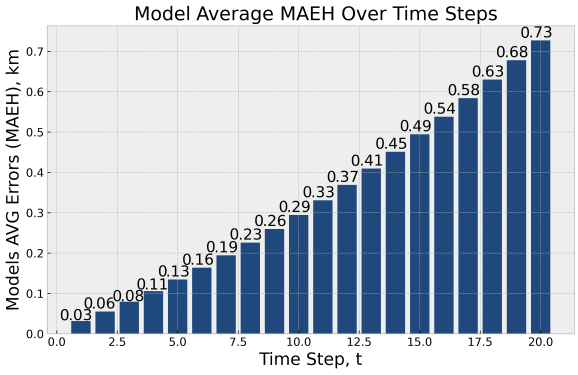
The experimental findings on trajectory prediction accuracy are illustrated in Figure 5.6. Incorporating vessel type as a feature in the dataset contributes to more stable predictions by reducing error variance. Both one-hot and embedded encoding techniques enhance accuracy, with embedded encoding, particularly in a two-dimensional structure, demonstrating the lowest variance. Table 5.3 presents detailed results of the embedded encoding experiments. On average, the inclusion of vessel type encoding improves prediction accuracy by approximately 0.0027 kilometres compared to models without categorical

vessel type information. These findings highlight the importance of encoding techniques in enhancing model performance and overall trajectory prediction reliability.

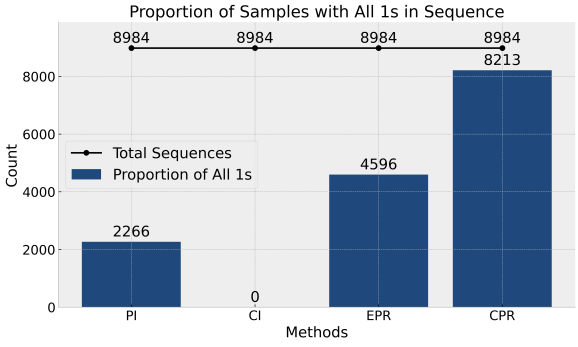
5.4 Analysis of Prediction Accuracy and Risk Estimation

This subsection evaluates the model's performance using predictions from the Baltic Sea test dataset. The analysis includes the average prediction error by time steps, offering insight into the accuracy of the model's forecasts. The coverage probability is also assessed to determine how frequently the actual vessel positions fall within the predicted collision regions. Finally, the proposed methods are verified through a real-world case study involving the cargo ships *Scot Carrier* and *Karin Hoej*. This analysis calculates the predicted ship domain boundaries based on future vessel positions using the various methods presented in the research. The probability of a potential collision is assessed by evaluating the overlap between these boundaries, offering a practical evaluation of the model's effectiveness in forecasting maritime safety risks.

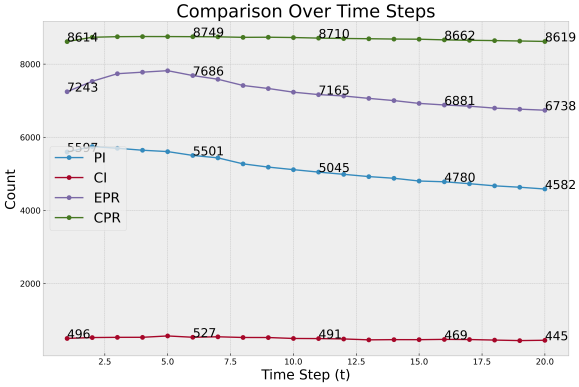
The model's accuracy utilising a derived MAEH function, which aligns conceptually with MAE but calculates the discrepancy based on the distance between the time series of the prediction and the actual sequence rather than direct coordinate comparisons [40]. This approach quantifies the distance error in the International System of Units (SI) as meters or kilometres. Figure 5.7(a) depicts the cumulative average prediction error across all models. The error fluctuates between 30 to 730 meters for forecasts extending up to 20 minutes, increasing as the forecast timeline extends. This metric's evaluation was conducted on a designated test sample set (15% of all data sequences).



(a)



(b)



(c)

Figure 5.7: Graphs of accuracy and evaluation results: (**a**) an average model error by time series of trajectory prediction (km); (**b**) sequences with all time steps within the calculated boundary widths (coverage probability) estimation; (**c**) coverage probability count across individual time steps.

Figure 5.7(b) assesses the subset sequences following the scoring methodology outlined previously (4.7), with an exception: this graph displays the aggregate count of sequences accurately predicted by the methods, specifically those instances where all-time series points within a sequence were contained within the predicted region. It reveals that most predictions aligned with the CPR areas, indicating high accuracy. Conversely, the confidence interval, though set at a 95% confidence level, was markedly narrow, resulting in the absence of sequences fully encompassed within these intervals. Figure 5.7(c) illustrates the distribution of sequence entries in the forecast region across time series and the applied method. It shows that as the forecast horizon increases, the accuracy of EPR and the prediction and confidence intervals tend to decrease. In contrast, CPR areas are established and calculated using the calibration sample, featuring a broader radius. This characteristic increases the likelihood of points falling within the CPR, enhancing its inclusiveness. However, it is important to note that this method relies on empirical data from past observations rather than a theoretical probabilistic assumption.

Visualisation of the regions as circular or elliptical vessel guard zones (domains) is feasible if the semi-major and semi-minor axes are known or, at the very least, the radius areas are known. In the case of EPR, both axes are defined, allowing the domain to be represented as an ellipse. Conversely, in CPR, only the radius is specified, resulting in a circular safe zone representing the maximum Euclidean distance a vessel may deviate from the centre of the predicted position. Figure 5.8 exhibits the marine incident case previously discussed, showcasing various implementations of these methods (without resorting to additional interpolations purely for visualisation purposes).

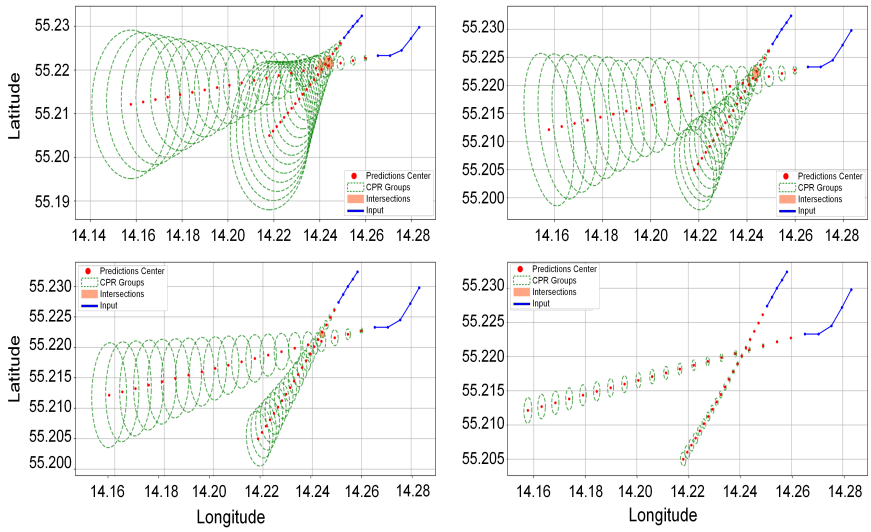


Figure 5.8: Comparison of method boundary widths in a marine accident.

5.5 Verification: the *Scot Carrier* and *Karin Hoej* Cargo Ships collision

On December 13, 2021, a critical maritime incident occurred in the Baltic Sea involving a collision between two cargo ships. The event unfolded off the coast of the southern Swedish city of Ystad and near the Danish island of Bornholm, within Swedish territorial waters. This collision involved a Danish-flagged vessel, *Karin Hoej* (MMSI: 219021240), and a British cargo ship, identified as *Scot Carrier* (MMSI: 232018267). In the aftermath of the collision, the Danish ship capsized, leading to an immediate and urgent search and rescue operation for the crew members aboard. The search efforts, intensified by the participation of *Scot Carrier*, focused on locating at least two people initially reported missing. Tragically, the situation took a sombre turn when one of the missing crew members from the Danish vessel was found dead in the hull, highlighting the severe consequences of the incident. The collision, amidst fog and darkness, has prompted a comprehensive investigation into the circumstances leading to this unfortunate event, examining factors such as maritime traffic negligence and the adherence to safety protocols in one of Europe's busiest maritime corridors.

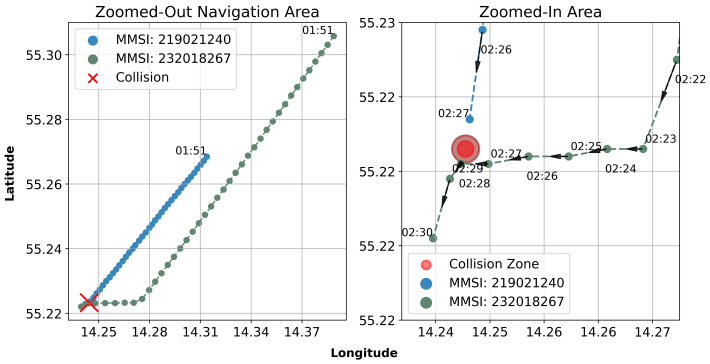


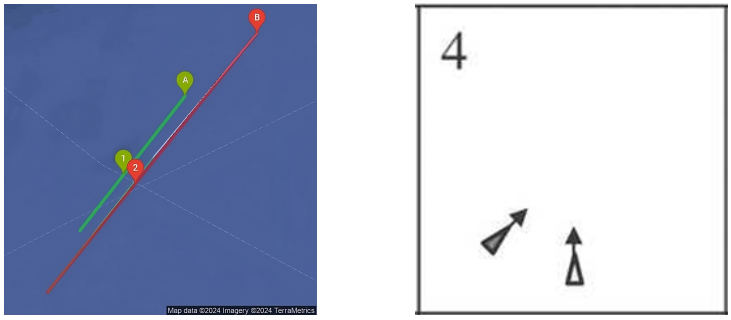
Figure 5.9: Spatiotemporal analysis of maritime collision accident.

Figure 5.9 presents two scatter plots that delineate the spatiotemporal trajectories of two vessels before and immediately following a collision in the Baltic Sea. The data, extracted from the Danish Maritime Authority's AIS, offers a high-resolution depiction of the incident, which occurred around 2:27 a.m. on December 13, 2021. The "Zoomed-Out Area" on the left illustrates the direction paths of the parallel moving vessels identified by MMSI 219021240 (blue) and MMSI 232018267 (green) before their collision (timestamp 1:51 a.m.). The adjacent "Zoomed-In Area" on the right provides a detailed view of the vessel movements at one-minute intervals from 02:22 to 02:30 a.m., showing the collision zone and directions of movement. This granular temporal resolution reveals the immediate navigational circumstances that led to the collision. Following the collision, the AIS signal transmission from *Karin Hoej* was disrupted, cutting off data flow to the information system.

This visual analysis is crucial for the forensic examination of the causes and consequences of the collision. It precisely demonstrates the trajectories and relative positions of the vessels, highlighting the importance of AIS data in maritime safety investigations. The figure not only contributes to elucidating this particular incident but also exemplifies the significance of AIS data from maritime authorities for enhancing navigational safety protocols and collision avoidance strategies.

According to the COLREGs, the captain of a vessel on the open sea should maintain a minimum distance of 1.5 to 2 nautical miles from other vessels to

avoid dangerous situations. One of the tools used in maritime and aviation contexts is CPA, which represents the shortest distance between two moving objects if they continue on their current course and speed without any changes. Meanwhile, TCPA measures the time until two moving objects reach the CPA to each other. In this study, these methods are applied to the collision verification to utilise the existing tools. Measurements are made based on the position of the vessels at 01:51 a.m. timestamp, as shown in the figures. The geographical position (longitude and latitude), speed in knots, bearing, and angles in degrees are used for the CPA calculation. The resulting CPA is 0.53 nautical miles, indicating an increased risk since the distance is less than 1 nautical mile. TCPA is calculated as 34.80 minutes to the calculated CPA points. Figure 5.10a illustrates the results of the calculations, indicating the starting point of the calculation with letters, the lines representing direction, and the projected position points with numbers. In maritime navigation, collision avoidance scenarios can be analysed using established theoretical models such as the Imazu problem, which provides a framework for evaluating vessel manoeuvring strategies to minimise collision risks. Our real-case scenario closely aligns with the Imazu problem case 4 (see Fig. 5.10b), where the relative positions and courses of two vessels necessitate a careful assessment of collision avoidance measures [35].



(a) **(b)**

Figure 5.10: Graphs of collision risk assessment: (**a**) calculation of CPA and TCPA for cargo vessels half an hour before collision; (**b**) Imazu problem case 4 schematic representation [35].

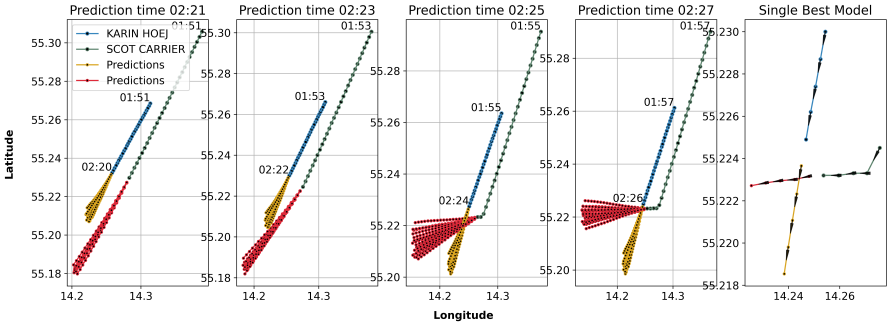


Figure 5.11: Comparative trajectory predictions from DL models. Triangles indicate the vessels' moving direction.

Figure 5.11 displays a series of scatter plots juxtaposing actual and predicted vessel trajectories leading up to the maritime collision near Bornholm, as DL models forecasted. Each panel represents a different prediction start time, ranging from 02:21 to 02:27 a.m., with a 30-step input sequence utilised for predicting the future positions of the vessels. The plots depict the actual paths of *KARIN HOEJ* (blue) and *SCOT CARRIER* (green), overlaying these with the predicted trajectories (yellow and red) generated by our trained models at each interval. The architecture of the DL AE, with its inherent complexities and the incorporation of regularisation techniques such as Dropout layers, results in the model producing a range of predictions with subtle variations. The rightmost graph, titled "Single Best Model", showcases the most precise model forecasted trajectory, identified by its minimal error value. Additionally, all depicted paths include directional indicators.

This ensemble of predictions illustrates the capabilities of our DL models to anticipate vessel movements and their potential for application in real-world maritime navigation and safety systems. It also underscores the value of employing a suite of models to capture the variability and uncertainty inherent in such dynamic prediction scenarios.

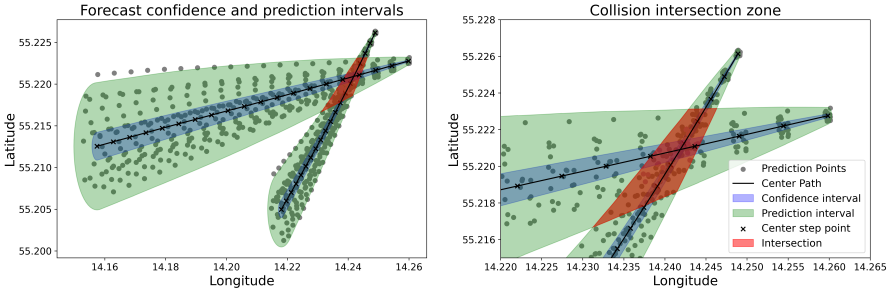


Figure 5.12: Visualisation of prediction and confidence intersection zones for bivariate spatial data.

Predictive calculations for the ship trajectories are initiated from the 2:25 a.m. timestamp, coinciding with observing the ships' turning movement changes. However, the starting point for predictions can be chosen at any time. Figure 5.12 comprises two panels illustrating the use of CIs in predicting bivariate spatial data for maritime trajectory forecasting. The left panel, "Forecast Confidence and Prediction Intervals", depicts the calculated CIs for a 20-minute (steps) prediction horizon determined by our trained model. For each step, the data are represented by a dot, aligning within the shaded confidence interval bands, demonstrating the range of potential vessel locations over time. The centre of the predicted path is marked by cross symbols, which indicate the model's central forecast for the vessel's position at each time step. The red zone marks the intersection area, representing the forecasted convergence of the two vessel paths. The graphical representation of the forecasts is made by interpolating the data between the time series, as the prediction regions are normally represented as circular, elliptical or sector-based (hybrid) ship domains.

The right panel, named the "Collision Intersection Zone", zooms in on the area where the CIs of the two predicted trajectories overlap. This intersection signifies a heightened risk of collision as indicated by the narrowing intervals in the initial prediction steps. The darkest shaded area denotes the most immediate risk zone, based on the closest predictions where the paths are forecasted to intersect. The reliability of both intervals is 95%. The desired level of

confidence and the range of the interval decrease (become narrower), and as the confidence level increases, the range of the interval increases (becomes wider).

Together, these panels convey the dynamic nature of predictive modelling in maritime navigation, emphasising the importance of calculating and visualising the predictions and associated CIs to assess the risk of collision more effectively. The models demonstrated a high degree of accuracy in forecasting the collision, which occurred.

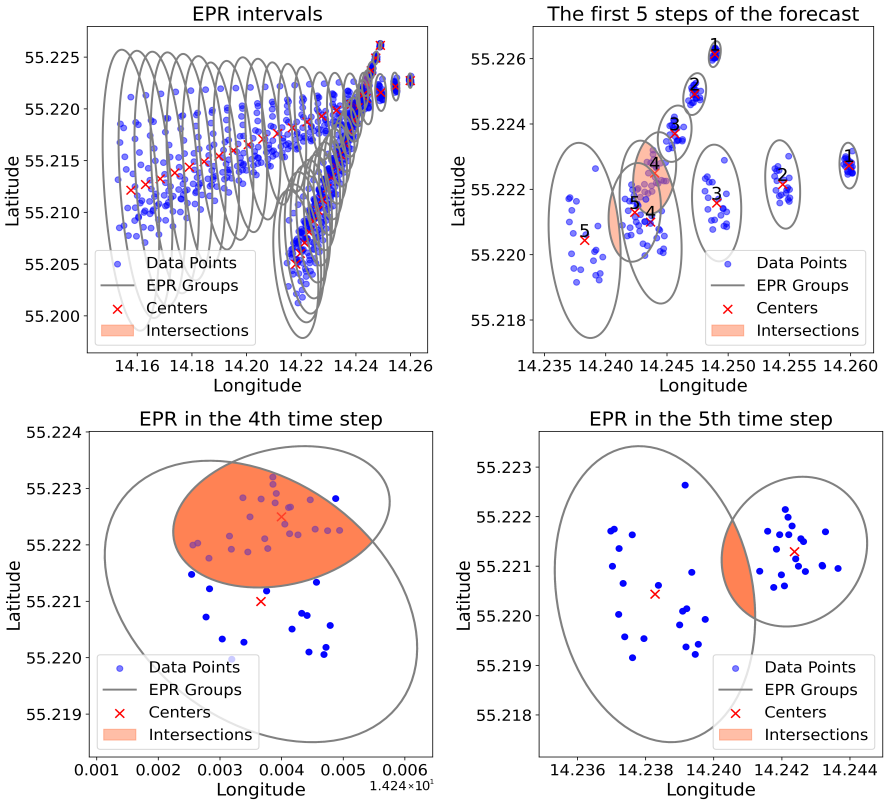


Figure 5.13: Ellipsoidal prediction regions. Numbers next to the red crosses represent the corresponding prediction time steps.

Figure 5.13 presents a graphical analysis of vessel trajectory forecasts using the EPR method. It consists of four distinct panels, each illustrating the application of EPRs at different stages of the prediction process.

In the first panel, labelled “EPR Intervals“, a series of ellipses represents the EPRs for multiple forecasted time steps of the two vessels. The data points (blue and red dots) indicate the actual positions of the vessels. At the same time, the ellipses encapsulate the regions where the vessels are predicted to be with a certain level of confidence. The ellipses oscillate, demonstrating the dynamic nature of the prediction over time, with overlapping areas suggesting potential collision risks.

The second panel, “The first 5 steps of the forecast“, zooms in on the initial predictions. It shows a closer view of the individual EPR ellipses for the first five time steps, where intersections of the ellipses (black outlines) indicate moments where the prediction models suggest a higher likelihood of a collision. The blue and red dots continue to represent the actual positions of the vessels, and the “X“ marks denote the centres of the EPR ellipses, which are the model’s best guess for the vessels’ future positions.

In the third and fourth panels, “EPR in the 4th time step“ and “EPR in the 5th time step“, we see the isolated EPR snapshots at specific moments. These individual ellipses provide a clear visual representation of the predicted movement areas for each vessel at given time steps. The overlap between the ellipses is shaded, highlighting the critical areas where the vessels’ paths are predicted to converge, indicating moments of heightened collision risk.

Once an area is demarcated, its possible overlap with adjacent areas can be analysed by predicting ship trajectories. To make it easier to understand the area covered by each zone, they are converted to a UTM projection, which allows the actual area of these zones to be calculated in square meters.

Applying the formula (E) can obtain a probability estimate of the collision risk score. The data presented in Table 5.4 show that using the CPR method results in a risk of collision between the two ships within a 5-minute time frame of nearly 40%, as predicted at 02:24 a.m. Conversely, even if the risk of a collision is 1%, i.e., an incident is likely to happen, it is much less likely. Zone A is *Karin Hoej*, and Zone B is *Scot Carrier.* The intersection areas between the trajectories of the two ships at a specific point in the time series, for instance, at step 1, are examined by comparing them with the forecast trajectory of the other ship at the same step, and this process continues sequentially. The centres

Table 5.4: Collision risk score comparison over time steps (highest probability marked in bold).

Time Step	EPR			
	Area A, m^2	Area B, m^2	$A \cap B$, m^2	Prob (%)
1	11,425.5	32,978.4	0.0	0.00
2	35,960.2	117,517.0	0.0	0.00
3	72,216.4	220,926.0	0.0	0.00
4	132,867.0	368,967.0	96,986.7	**23.96**
5	198,320.0	536,512.0	22,066.5	**3.10**
6	274,578.0	762,444.0	0.0	0.00

Time Step	CPR			
	Area A, m^2	Area B, m^2	$A \cap B$, m^2	Prob (%)
1	20,812.7	20,808.5	0.0	0.00
2	64,937.4	64,927.8	0.0	0.00
3	153,537.9	153,523.8	0.0	0.00
4	302,596.1	302,584.9	171,576.5	**39.57**
5	524,334.6	524,344.3	89,920.4	**9.38**
6	826,551.7	826,611.9	0.0	0.00

Time Step	PI			
	Area A, m^2	Area B, m^2	$A \cap B$, m^2	Prob (%)
1	4,717.5	14,151.0	0.0	0.00
2	14,568.8	53,199.1	0.0	0.00
3	31,642.8	112,766.8	0.0	0.00
4	53,492.5	190,868.5	8,420.9	**3.57**
5	79,293.0	288,027.5	0.0	**0.00**
6	107,950.1	405,907.8	0.0	0.00

Time Step	CI			
	Area A, m^2	Area B, m^2	$A \cap B$, m^2	Prob (%)
1	224.6	673.9	0.0	0.00
2	693.8	2,533.3	0.0	0.00
3	1,506.8	5,369.8	0.0	0.00
4	2,547.3	9,089.0	0.0	**0.00**
5	3,775.9	13,715.6	0.0	**0.00**
6	5,140.5	19,328.9	0.0	0.00

of the predicted trajectories for all models are marked with red dots. This methodical approach ensures a thorough evaluation of the potential overlaps or intersections among the predicted paths of various vessels, providing insight into the likelihood of close encounters or collisions at each step of the forecast.

By leveraging a multi-model approach, we effectively quantify prediction variability, where uncertainty increases as the forecast horizon extends. This is directly reflected in the size of the ellipsoidal boundaries, as longer-term predictions exhibit greater divergence due to the accumulating error. This characteristic is inherent in stochastic trajectory forecasting, where uncertainty propagation is a fundamental aspect of risk assessment. The probabilistic nature of EPRs allows for a more robust evaluation of vessel movement safety compared to single deterministic methods such as CPA, which focus on immediate proximity rather than future trajectory regions.

5.6 Summary of the Chapter

This chapter presents the results of the experiments conducted to enhance vessel trajectory prediction and maritime safety. The research involved analysing and developing recurrent networks based on time sequences managed by specific cell structures. The study found that different cell sizes have a direct impact on the results, depending on the type of architecture. Six different architectures were tested: simple RNN, basic LSTM, LSTM stack, GRU, AE, and basic bidirectional LSTM. The bidirectional LSTM and AE architectures achieved the most accurate predictions and the lowest MSE loss, with optimal cell sizes crucial for their performance.

DL models were evaluated in multiple stages, with the North Sea region dataset initially used to compare different recurrent architectures. The LSTM AE achieved the most precise trajectory predictions, with UTM-based transformations producing the lowest MAEH error of 1.141 km, outperforming other transformations such as distance and angle (1.266 km), delta latitude and longitude (1.439 km) and raw AIS features (1.427 km).

Further studies in the Baltic Sea region introduced recursive delta transformations with WGS84 and UTM coordinates. The results demonstrated

that WGS84 delta-based predictions achieved the lowest MAEH error of 0.3537 km, showing comparable or superior performance to UTM-based transformations (0.3839 km). Additionally, experiments assessing categorical vessel-type encoding methods revealed that 3D embedding yielded the lowest MAEH error (0.2636 km), outperforming one-hot encoding (0.2655 km) and ordinal encoding (0.2681 km), while predictions without vessel-type information resulted in a slightly higher error (0.2710 km). These findings highlight the critical role of coordinate transformations, recursive delta-based predictions, and vessel-type encoding in optimising vessel trajectory forecasts, demonstrating that refined data representations and feature engineering significantly enhance prediction accuracy. Incorporating key meteorological features such as wave period, wind speed, wind wave period, air temperature, water temperature, wind wave height, and wave height further enhanced the prediction accuracy, demonstrating that environmental conditions play a crucial role in improving vessel trajectory forecasts, reducing the MAEH to 0.3203 km.

The chapter also explored the use of prediction boundaries, including prediction and confidence intervals, ellipsoidal prediction regions, and conformal prediction regions, to estimate potential future positions of vessels. These methods provided a probabilistic assessment of collision risks, with conformal prediction regions proving particularly effective. Coverage probability analysis across time steps demonstrated that CPR had the highest inclusion rate at 86.7%, followed by EPR at 71.4%, PI at 51.0%, and CI at 4.9%, reinforcing their reliability in maritime collision assessment.

A notable application was the real-case collision analysis between *Scot Carrier* and *Karin Hoej* cargo ships, where the conformal methodology predicted the accident with a 39.6% probability. These findings highlight the advantage of probabilistic approaches over deterministic methods such as CPA and TCPA, which rely on fixed threshold values rather than dynamic trajectory uncertainties. The framework proposed by the study extends beyond traditional methods by evaluating trajectory boundary overlaps dynamically, integrating multiple models without bootstrapping, and aligning predictions with real-world stopping times of large vessels.

The study emphasised the importance of integrating multiple models and

employing advanced prediction techniques to estimate forecast intervals and dynamically observe variability. Methods provided a probabilistic framework for defining trajectory boundaries and assessing collision risks. Compared to traditional approaches such as CPA and TCPA, these techniques offer a more robust assessment of vessel movement uncertainty by evaluating clusters of probabilistic boundaries rather than relying on a single deterministic point.

A key advantage of this approach is its ability to move beyond predefined static thresholds, such as CPA's fixed proximity and response time conditions, by forecasting dynamic trajectory boundaries and evaluating their overlap probabilistically. This method enhances long-term collision risk assessment, aligning with real-world maritime conditions where vessel behaviour is influenced by external factors such as weather, navigational constraints, and operational decisions.

However, despite the advantages, challenges remain. The reliance on multiple DL models increases computational demands, particularly for real-time applications. Additionally, the effectiveness of these methods depends on the accuracy of the underlying trajectory predictions, which are influenced by AIS data quality, including potential gaps, delays, and inconsistencies. Future research could further refine data pre-processing techniques, integrate physical vessel parameters, and explore comparative studies with reinforcement learning-based approaches. By addressing these limitations, the proposed framework offers a promising step forward in maritime collision detection and risk assessment, contributing to safer and more reliable vessel navigation systems.

6 General Conclusions

This thesis proposed and investigated DL approaches for multi-vessel trajectory forecasting and probabilistic collision-risk assessment in maritime navigation. Specifically, the thesis (a) reviewed existing trajectory-prediction methods, highlighting RNNs ability to perform long-term forecasts rather than relying on linear extrapolation; (b) designed, trained, and tuned various RNN architectures to identify those yielding the lowest prediction error in vessel-movement forecasting; (c) examined input representations, such as coordinate transformations, recursive delta recalculation, vessel-type embeddings, and meteorological data to enhance prediction accuracy; (d) implemented and compared uncertainty quantification techniques (EPR, CPR, PI, and CI) for collision-risk estimation; and (e) applied these models to a previously unseen real-world collision incident to demonstrate practical applicability. The key conclusions are:

1. RNNs (such as LSTM/GRU) are necessary for 20–25 minute vessel-trajectory forecasts. Clustering or linear extrapolation cannot produce continuous predictions over a 20–25 minute horizon, critical for large-vessel stopping times. In the North Sea study, six architectures (simple RNN, basic LSTM, stacked LSTM, GRU, LSTM AE, bidirectional LSTM) were compared: bidirectional LSTM (300 cells, MSE = 3.618×10^{-4}) and LSTM AE (125 cells, MSE = 3.688×10^{-4}) achieved the lowest errors with a specific cell combination. While bidirectional LSTM maintained stable accuracy across configurations, AE gets plateaued beyond a threshold cell count, which suggested that architecture and hyperparameter tuning dramatically affect long-term accuracy.

2. Careful tuning of LSTM cell counts is critical: larger architectures do not guarantee better accuracy. In the Baltic Sea experiments, determined models with 75–300 LSTM cells achieved the lowest MAEH (around 0.18–0.20 km). Beyond 300 cells, error reduction plateaued, and increasing cells past $1,000$ caused a sharp rise in training time with negligible gains. Thus, selecting an appropriate cell count is essential; excessively large architectures impair efficiency without improving accuracy.

3. Enhanced input representations, coordinate transformations, recursive delta extrapolation, vessel-type embeddings, and meteorological features substantially reduce prediction error.

- Introduced recursive delta recalculation (feeding each predicted position into the next step), cut MAEH from 0.6171 km (raw WGS84) to 0.3537 km (WGS84-delta) and from 0.4903 km (UTM) to 0.3839 km (UTM-delta), preventing error accumulation, especially in early steps.
- In North Sea tests with the proposed LSTM AE, projecting raw AIS into UTM reduced MAEH to 1.141 km (versus 1.427 km for raw WGS84), with distance-angle transformations at 1.266 km and delta-lat/lon at 1.439 km.
- In Baltic Sea experiments with the proposed LSTM AE, integrating a 3D embedding for vessel type reduced MAEH to 0.2636 km (versus 0.2655 km for one-hot, 0.2681 km for ordinal, and 0.2710 km without any vessel-type).
- Integrating meteorological data with AIS inputs to capture the surrounding environmental impact reduced MAEH from 0.3537 km to 0.3203 km, confirming that richer, recursively updated, and locally projected inputs yield more accurate multi-step forecasts.

4. Proposed probabilistic interval-based methods (CPR, EPR, PI, CI) outperform CPA/TCPA. Forecasting trajectory-boundary overlap via CI, EPR, and CPR provides a more complete uncertainty assessment than point-based CPA/TCPA. CPR achieved 86.7% coverage (versus 71.4% for EPR, 51.0% for PI, and 4.9% for CI) by avoiding normal distribution assumptions and evaluating uncertainty clusters over a 20-minute horizon. This replaces fixed guard zone rules with dynamic region overlaps for a more adaptive, statistically sound collision-risk framework.

5. Real-world validation (*Scot Carrier–Karin Hoej*, 2021) confirms probabilistic detection but highlights AIS limitations. Overlapping CPRs from multiple LSTM AE forecasts yielded a larger 39.6% collision risk score at 95% confidence, matching the actual incident. However, AIS gaps and delays still constrain reliability; resampling reduced some noise, but incorporating vessel-specific physical parameters (e.g., turning radii, inertia) could further improve robustness.

Bibliography

[1] M. Abebe, Y. Noh, Y.-J. Kang, C. Seo, D. Kim, and J. Seo. "Ship Trajectory Planning for Collision Avoidance Using Hybrid ARIMA–LSTM Models". In: *Ocean Engineering* 256 (2022), p. 111527. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2022.111527.

[2] N. Ajroldi, J. Diquigiovanni, M. Fontana, and S. Vantini. "Conformal Prediction Bands for Two-Dimensional Functional Time Series". In: *Computational Statistics and Data Analysis* 187 (2023), p. 107821. ISSN: 01679473. DOI: 10.1016/j.csda.2023.107821. arXiv: 2207.13656.

[3] D. Alizadeh, A. A. Alesheikh, and M. Sharif. "Prediction of Vessels Locations and Maritime Traffic Using Similarity Measurement of Trajectory". In: *Annals of GIS* 27.2 (2021), pp. 151–162. ISSN: 19475691. DOI: 10.1080/19475683.2020.1840434.

[4] H. Almukhalfi, A. Noor, and T. H. Noor. *Traffic Management Approaches Using Machine Learning and Deep Learning Techniques: A Survey*. July 2024. DOI: 10.1016/j.engappai.2024.108147.

[5] M. S. Almutairi, K. Almutairi, and H. Chiroma. "Hybrid of Deep Recurrent Network and Long Short-Term Memory for Rear-End Collision Detection in Fog-Based Internet of Vehicles". In: *Expert Systems with Applications* 213 (2023), p. 119033. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2022.119033.

[6] A. N. Angelopoulos and S. Bates. *A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification*. 2022. arXiv: 2107.07511 [cs.LG]. URL: https://arxiv.org/abs/2107.07511.

[7] S. Arasteh, M. A. Tayebi, Z. Zohrevand, U. Glässer, A. Y. Shahir, P. Saeedi, and H. Wehn. "Fishing Vessels Activity Detection from Longitudinal AIS Data". In: *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems* December (2020), pp. 347–356. DOI: 10.1145/3397536.3422267.

[8] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx. "A Graph CNN-LSTM Neural Network for Short and Long-Term Traffic Forecasting Based on Trajectory Data". In: *Transportation Research Part C: Emerging Technologies* 112.December 2019 (2020), pp. 62–77. ISSN: 0968090X. DOI: 10.1016/j.trc.2020.01.010.

[9] M. Bogalecka, E. Jakusik, and K. Kołowrocki. "Ship Accidents of Last 30 Years Ship Accidents at the Baltic Sea Area". In: *Journal of Polish Safety and Reliability Association, Summer Safety and Reliability Seminars* 8.4 (2017), pp. 147–154.

[10] C. Bueger. "What is Maritime Security?" In: *Marine Policy* 53 (2015), pp. 159–164. ISSN: 0308597X. DOI: 10.1016/j.marpol.2014.12.005.

[11] G. W. Burruss and T. M. Bray. "Confidence Intervals". In: *Encyclopedia of Social Measurement*. Ed. by K. Kempf-Leonard. New York: Elsevier, 2005, pp. 455–462. ISBN: 978-0-12-369398-3. DOI: 10.1016/B0-12-369398-5/00060-8.

[12] S. Capobianco, N. Forti, L. M. Millefiori, P. Braca, and P. Willett. "Uncertainty-Aware Recurrent Encoder-Decoder Networks for Vessel Trajectory Prediction". In: *Proceedings of 2021 IEEE 24th International Conference on Information Fusion, FUSION 2021* (2021), pp. 1–10. ISSN: 15579603. DOI: 10.23919/fusion49465.2021.9626839.

[13] A. D. Carnerero, D. R. Ramirez, S. Lucia, and T. Alamo. "Prediction Regions Based on Dissimilarity Functions". In: *ISA Transactions* 139 (2023), pp. 49–59. ISSN: 00190578. DOI: 10.1016/j.isatra.2023.03.048.

[14] E. Carrizosa, M. Galvis Restrepo, and D. Romero Morales. "On Clustering Categories of Categorical Predictors in Generalized Linear Models". In: *Expert Systems with Applications* 182.February (2021), p. 115245. ISSN: 09574174. DOI: 10.1016/j.eswa.2021.115245. arXiv: 2110.10059.

[15] P. Cerda and G. Varoquaux. "Encoding High-Cardinality String Categorical Variables". In: *IEEE TRANSACTIONS ON KNOWLEDGE*

AND DATA ENGINEERING 34.3 (Mar. 2022), pp. 1164–1176. ISSN: 1041-4347. DOI: 10.1109/TKDE.2020.2992529.

[16] C.-W. Chen, C. Harrison, and H.-H. Huang. *The Unsupervised Method of Vessel Movement Trajectory Prediction*. 2020. arXiv: 2007.13712 [cs.CV].

[17] J. Chen, H. Chen, Y. Zhao, and X. Li. "FB-BiGRU: A Deep Learning Model for AIS-Based Vessel Trajectory Curve Fitting and Analysis". In: *Ocean Engineering* 266 (2022), p. 112898. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2022.112898.

[18] W. Chen, J. Chen, J. Geng, J. Ye, T. Yan, J. Shi, and J. Xu. "Monitoring and Evaluation of Ship Operation Congestion Status at Container Ports Based on AIS Data". In: *Ocean & Coastal Management* 245 (2023), p. 106836. ISSN: 0964-5691. DOI: 10.1016/j.ocecoaman.2023.106836.

[19] X. Chen, Y. Yang, S. Wang, H. Wu, J. Tang, J. Zhao, and Z. Wang. "Ship Type Recognition via a Coarse-to-Fine Cascaded Convolution Neural Network". In: *The Journal of Navigation* 73.4 (2020), pp. 813–832. DOI: 10.1017/S0373463319000900.

[20] M. Cleaveland, I. Lee, G. J. Pappas, and L. Lindemann. "Conformal Prediction Regions for Time Series using Linear Complementarity Programming". In: (2023), pp. 1–16. arXiv: 2304.01075. URL: http://arxiv.org/abs/2304.01075.

[21] M. K. Dahouda and I. Joe. "A Deep-Learned Embedding Technique for Categorical Features Encoding". In: *IEEE Access* 9 (2021), pp. 114381–114391. DOI: 10.1109/ACCESS.2021.3104357.

[22] A. Daranda. "Mašininiu mokymusi grindžiamas laivybos eismo dalyvių elgsenos prognozavimas bei nestandartinių laivybos srauto situacijų atradimas". In: (2021). Prieiga per eLABa – nacionalinė Lietuvos akademinė elektroninė biblioteka. DOI: 10.15388/vu.thesis.178.

[23] A. Daranda and G. Dzemyda. "Reinforcement Learning Strategies for Vessel Navigation". In: *Integrated Computer-Aided Engineering* 30.1 (2023), pp. 53–66. ISSN: 1069-2509. DOI: 10.3233/ICA-220688.

[24] H. De Meulemeester and B. De Moor. “Unsupervised Embeddings for Categorical Variables”. In: *2020 INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN)*. IEEE International Joint Conference on Neural Networks (IJCNN). International Joint Conference on Neural Networks (IJCNN) held as part of the IEEE World Congress on Computational Intelligence (IEEE WCCI), ELECTR NETWORK, JUL 19-24, 2020. IEEE; IEEE Computat Intelligence Soc; Int Neural Network Soc. 2020. ISBN: 978-1-7281-6926-2. DOI: 10.1109/ijcnn48605.2020.9207703.

[25] L. Du. “Maritime Traffic Risk Analysis in the Northern Baltic Sea from AIS Data”. Doctoral dissertation, Department of Mechanical Engineering, School of Engineering. Defence held online via Zoom on 2021-10-01. Supervisors: Prof. Pentti Kujala, Prof. Floris Goerlandt, Prof. Osiris A. Valdez Banda. PhD thesis. Finland: Aalto University, Oct. 2021. ISBN: 978-952-64-0500-1 (electronic), 978-952-64-0499-8 (printed). URL: https://aaltodoc.aalto.fi/handle/123456789/109976.

[26] L. Du, O. A. Valdez Banda, F. Goerlandt, Y. Huang, and P. Kujala. “A COLREG-Compliant Ship Collision Alert System for Stand-On Vessels”. In: *Ocean Engineering* 218 (Dec. 2020). ISSN: 00298018. DOI: 10.1016/j.oceaneng.2020.107866.

[27] L. M. Eljabu. “Novel Maritime Traffic Analysis Techniques to Enhance Maritime Situational Awareness”. Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Faculty of Computer Science. PhD thesis. Halifax, Nova Scotia: Dalhousie University, Aug. 2024. URL: http://hdl.handle.net/10222/84534.

[28] R. N. Forthofer, E. S. Lee, and M. Hernandez. “7 - Interval Estimation”. In: *Biostatistics (Second Edition)*. Ed. by R. N. Forthofer, E. S. Lee, and M. Hernandez. Second Edition. San Diego: Academic Press, 2007, pp. 169–212. ISBN: 978-0-12-369492-8. DOI: 10.1016/B978-0-12-369492-8.50012-1.

[29] N. Forti, L. M. Millefiori, P. Braca, and P. Willett. “Prediction of Vessel Trajectories From AIS Data Via Sequence-To-Sequence Recurrent Neural Networks”. In: *ICASSP 2020 - 2020 IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain, 2020, pp. 8936–8940. ISBN: 978-1-5090-6631-5. DOI: 10.1109/ICASSP40776.2020.9054421.

[30] M. Gao, M. Liang, A. Zhang, Y. Hu, and J. Zhu. "Multi-Ship Encounter Situation Graph Structure Learning for Ship Collision Avoidance Based on AIS Big Data with Spatio-Temporal Edge and Node Attention Graph Convolutional Networks". In: *Ocean Engineering* 301 (June 2024). ISSN: 00298018. DOI: 10.1016/j.oceaneng.2024.117605.

[31] D.-w. Gao, Y.-s. Zhu, J.-f. Zhang, Y.-k. He, K. Yan, and B.-r. Yan. "A Novel MP-LSTM Method for Ship Trajectory Prediction Based on AIS Data". In: *Ocean Engineering* 228 (2021), p. 108956. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2021.108956.

[32] W. D. Gao, S. Y. Zhu, F. J. Zhang, K. Y. He, K. Yan, and B. Yan. "A Novel MP-LSTM Method for Ship Trajectory Prediction Based on AIS Data". In: *Ocean Engineering* 228.October 2020 (2021), p. 108956. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2021.108956.

[33] S. Gnat. "Impact of Categorical Variables Encoding on Property Mass Valuation". In: *Procedia Computer Science* 192 (2021), pp. 3542–3550. ISSN: 18770509. DOI: 10.1016/j.procs.2021.09.127.

[34] F. Golestaneh, P. Pinson, R. Azizipanah-Abarghooee, and H. B. Gooi. "Ellipsoidal Prediction Regions for Multivariate Uncertainty Characterization". In: *IEEE Transactions on Power Systems* 33.4 (2018), pp. 4519–4530. ISSN: 08858950. DOI: 10.1109/TPWRS.2018.2791975. arXiv: 1708.02703.

[35] T. Higaki, H. Hashimoto, and H. Yoshioka. "Investigation and Imitation of Human Captains' Maneuver Using Inverse Reinforcement Learning". In: *Journal of the Japan Society of Naval Architects and Ocean Engineers* 36 (2022), pp. 137–148. ISSN: 1342-3105. DOI: 10.2534/jjasnaoe.36.137.

[36] International Maritime Organization. *CASUALTY-RELATED MATTERS: Reports on Marine Casualties and Incidents*. MSC-MEPC.3/Circ.3 T1/12.01. Revised harmonized reporting procedures – Reports required under SOLAS regulation I/21 and MARPOL, articles 8 and 12. London, UK: International Maritime Organization, Dec. 2008.

[37] C. Jia, J. Ma, X. Yang, and X. Lv. "RAGAN: A Generative Adversarial Network for Risk-Aware Trajectory Prediction in Multi-Ship Encounter Situations". In: *Ocean Engineering* 289.P1 (2023), p. 116188. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2023.116188.

[38] Jithin. *Stopping Distance, Turning Circle, Ships Manoeuvring*. Marine studies blog, *Knowledge of Sea*. Founder: Jithin, sailing on tankers since 2005. Jan. 2020. URL: https://knowledgeofsea.com/stopping-distance-turning-circle-ships-manoeuvring/.

[39] R. Jurkus, P. Treigys, and J. Venskus. "Investigation of Recurrent Neural Network Architectures for Prediction of Vessel Trajectory". In: *Information and Software Technologies*. Cham: Springer International Publishing, 2021, pp. 194–208. ISBN: 978-3-030-88304-1. DOI: 10.1007/978-3-030-88304-1_16.

[40] R. Jurkus, J. Venskus, and P. Treigys. "Application of Coordinate Systems for Vessel Trajectory Prediction Improvement Using Recurrent Neural Networks". In: *Engineering Applications of Artificial Intelligence* 123.July 2022 (2023), p. 106448. ISSN: 09521976. DOI: 10.1016/j.engappai.2023.106448.

[41] G. B. Karataş, P. Karagoz, and O. Ayran. "Trajectory Pattern Extraction and Anomaly Detection for Maritime Vessels". In: *Internet of Things* 16 (2021), p. 100436. ISSN: 2542-6605. DOI: 10.1016/j.iot.2021.100436.

[42] C. P. Keller. "Map Projections: A Reference Manual". In: *Cartographic Perspectives* 23 (1996), pp. 25–26. ISSN: 1048-9053. DOI: 10.14714/cp23.770.

[43] Khedkar, Sanskruti, Lambor, Shilpa, Narule, Yogita, and Berad, Prathamesh. "Categorical Embeddings for Tabular Data using PyTorch". In: *ITM Web Conf.* 56 (2023), p. 02002. DOI: 10.1051/itmconf/20235602002.

[44] Klaipėdos universiteto Lietuvių kalbos katedra. *Internetinis jūreivystės terminų žodynas*. 2015. URL: https://my.ku.lt/jtvd.

[45] P. Last. "Analysis of Automatic Identification System Data for Maritime Safety". Doctoral thesis, Department of Computer Science & Electrical

Engineering. Date of Defense: 31.10.2016. Dissertation Committee: Prof. Dr. Lars Linsen, Prof. Dr. Martin Hering-Bertram, Prof. Dr. Andreas Birk. PhD thesis. Bremen, Germany: Jacobs University Bremen, Oct. 2016.

[46] P. Last, C. Bahlke, M. Hering-Bertram, and L. Linsen. "Comprehensive Analysis of Automatic Identification System (AIS) Data in Regard to Vessel Movement Prediction". In: *Journal of Navigation* 67 (2014), pp. 791–809. URL: https://api.semanticscholar.org/CorpusID:129028002.

[47] C. Lee. "How Can We Use Neural Network With Entity Embedding for Product Valuations? A Case Study for the Car Industry". In: *International Journal of Information Management Data Insights* 3.2 (2023), p. 100187. ISSN: 26670968. DOI: 10.1016/j.jjimei.2023.100187.

[48] J.-S. Lee and Y.-U. Yu. "Calculation of Categorical Route Width According to Maritime Traffic Flow Data in the Republic of Korea". In: *JOURNAL OF MARINE ENGINEERING AND TECHNOLOGY* 22.5, SI (Sept. 2023), pp. 222–232. ISSN: 2046-4177. DOI: 10.1080/20464177.2023.2223396.

[49] B. Li, Z. Shi, Z. Wu, and Z. Yang. "LSTM-Based Distributed Photovoltaic Output Interval Prediction Method for Source-Load Collaboration". In: *2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)*. 2023, pp. 366–369. DOI: 10.1109/ICSP58490.2023.10248839.

[50] Y. Li and H. Cao. "Prediction for Tourism Flow based on LSTM Neural Network". In: *Procedia Computer Science* 129 (2018). 2017 INTERNATIONAL CONFERENCE ON IDENTIFICATION,INFORMATION AND KNOWLEDGEIN THE INTERNET OF THINGS, pp. 277–283. ISSN: 1877-0509. DOI: 10.1016/j.procs.2018.03.076.

[51] S. Lim, S. J. Kim, Y. Park, and N. Kwon. "A Deep Learning-Based Time Series Model With Missing Value Handling Techniques to Predict Various Types of Liquid Cargo Traffic". In: *Expert Systems with Applications* 184 (2021), p. 115532. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2021.115532.

[52] C. Lin, R. Zhen, Y. Tong, S. Yang, and S. Chen. "Regional Ship Collision Risk Prediction: An Approach Based on Encoder–Decoder LSTM Neural Network Model". In: *Ocean Engineering* 296 (Mar. 2024). ISSN: 00298018.

[53] C. Liu, Y. Li, R. Jiang, Q. Lu, and Z. Guo. "Trajectory-Based Data Delivery Algorithm in Maritime Vessel Networks Based on Bi-LSTM". In: Qingdao, China, Sept. 2020, pp. 298–308. ISBN: 978-3-030-59015-4. DOI: 10.1007/978-3-030-59016-1_25.

[54] R. W. Liu, K. Hu, M. Liang, Y. Li, X. Liu, and D. Yang. "QSD-LSTM: Vessel Trajectory Prediction Using Long Short-Term Memory with Quaternion Ship Domain". In: *Applied Ocean Research* 136 (2023), p. 103592. ISSN: 0141-1187. DOI: 10.1016/j.apor.2023.103592.

[55] R. W. Liu, X. Huo, M. Liang, and K. Wang. "Ship Collision Risk Analysis: Modeling, Visualization, and Prediction". In: *Ocean Engineering* 266.P4 (2022), p. 112895. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2022.112895.

[56] R. W. Liu, M. Liang, J. Nie, W. Y. B. Lim, Y. Zhang, and M. Guizani. "Deep Learning-Powered Vessel Trajectory Prediction for Improving Smart Traffic Services in Maritime Internet of Things". In: *IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING* 9.5 (Sept. 2022), pp. 3080–3094. ISSN: 2327-4697. DOI: 10.1109/TNSE.2022.3140529.

[57] X. Liu, A. Gherbi, W. Li, and M. Cheriet. "Multi Features and Multi-time steps LSTM Based Methodology for Bike Sharing Availability Prediction". In: *Procedia Computer Science* 155 (2019). The 16th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2019),The 14th International Conference on Future Networks and Communications (FNC-2019),The 9th International Conference on Sustainable Energy Information Technology, pp. 394–401. ISSN: 1877-0509. DOI: 10.1016/j.procs.2019.08.055.

[58] M. D. Lucagbo and T. Mathew. "Rectangular Confidence Regions and Prediction Regions in Multivariate Calibration". In: *Journal of the*

Indian Society for Probability and Statistics 23.1 (2022), pp. 155–171. ISSN: 23649569. DOI: 10.1007/s41096-022-00116-7.

[59] D. Ma, Y. Gao, D. Shi, Y. Shen, Y. Zhou, and Q. Yang. "Nonparametric Bootstrap Estimation of Confidence Interval in Base Station Test". In: *2009 5th Asia-Pacific Conference on Environmental Electromagnetics*. 2009, pp. 390–394. DOI: 10.1109/CEEM.2009.5303672.

[60] Marine Accident and Incident Investigation Committee. *Definitions: Casualty Investigation Code*. http://www.maic.gov.cy/mcw/dms/maic/maic.nsf/page02_en/page02_en?OpenDocument. Accessed: 2024-03-20. Republic of Cyprus, 2014-2024.

[61] S. Mehri, A. A. Alesheikh, and A. Basiri. "A Context-Aware Approach for Vessels' Trajectory Prediction". In: *Ocean Engineering* 282.April (2023), p. 114916. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2023.114916.

[62] M. P. Mujeeb-Ahmed, J. K. Seo, and J. K. Paik. "Probabilistic Approach for Collision Risk Analysis of Powered Vessel With Offshore Platforms". In: *Ocean Engineering* 151 (Mar. 2018), pp. 206–221. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2018.01.008.

[63] B. Murray. "Machine Learning for Enhanced Maritime Situation Awareness: Leveraging Historical AIS Data for Ship Trajectory Prediction". Doctoral thesis, supported by MARKOM2020 and the Norwegian Ministry of Education and Research in cooperation with the Norwegian Ministry of Trade, Industry and Fisheries. PhD thesis. Tromsø, Norway: UiT The Arctic University of Norway, May 2021. ISBN: 978-82-8236-433-1. URL: https://hdl.handle.net/10037/20984.

[64] B. Murray and L. P. Perera. "A Dual Linear Autoencoder Approach for Vessel Trajectory Prediction Using Historical AIS Data". In: *Ocean Engineering* 209 (2020), p. 107478. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2020.107478.

[65] B. Murray and L. P. Perera. "An AIS-based Deep Learning Framework for Regional Ship Behavior Prediction". In: *Reliability Engineering*

and System Safety 215 (2021), p. 107819. ISSN: 09518320. DOI: 10.1016/j.ress.2021.107819.

[66] B. Murray and L. P. Perera. "Ship Behavior Prediction via Trajectory Extraction-Based Clustering for Maritime Situation Awareness". In: *Journal of Ocean Engineering and Science* (2021), pp. 0–21. ISSN: 24680133. DOI: 10.1016/j.joes.2021.03.001.

[67] S. Nirmalraj, S. M. A. Antony, P. Sridevipon malar, S. A. Oliver, J. K. Velmurugan, V. Elanangai, and G. Nagarajan. "Permutation Feature Importance–Based Fusion Techniques for Diabetes Prediction". In: *SOFT COMPUTING* (2023 APR 24 2023). ISSN: 1432-7643. DOI: 10.1007/s00500-023-08041-y.

[68] H. Noma, T. Shinozaki, K. Iba, S. Teramukai, and T. A. Furukawa. "Confidence intervals of prediction accuracy measures for multivariable prediction models based on the bootstrap-based optimism correction methods". In: *Statistics in Medicine* 40.26 (2021), pp. 5691–5701. ISSN: 10970258. DOI: 10.1002/sim.9148. arXiv: 2005.01457.

[69] K. V. Olesen, A. Boubekki, M. C. Kampffmeyer, R. Jenssen, A. N. Christensen, S. Hørlück, and L. H. Clemmensen. "A Contextually Supported Abnormality Detector for Maritime Trajectories". In: *Journal of Marine Science and Engineering* 11.11 (2023). ISSN: 2077-1312. DOI: 10.3390/jmse11112085.

[70] C. Pan, A. Poddar, R. Mukherjee, and A. K. Ray. "Impact of Categorical and Numerical Features in Ensemble Machine Learning Frameworks for Heart Disease Prediction". In: *Biomedical Signal Processing and Control* 76.January (2022), p. 103666. ISSN: 17468108. DOI: 10.1016/j.bspc.2022.103666.

[71] J. Park, J. Jeong, and Y. Park. "Ship Trajectory Prediction Based on Bi-LSTM Using Spectral-Clustered AIS Data". In: *Journal of Marine Science and Engineering* 9.9 (2021), p. 1037. ISSN: 20771312. DOI: 10.3390/jmse9091037.

[72] V. Paulauskas, B. Plačienė, A. Paulauskienė, R. Maksimavičius, V. Lukauskas, R. Barzdžiukas, D. Paulauskas, A. Kaulickis, and J. Banaitis. "Jūrinė technologija, laivo sandara I dalis: mokymo medžiaga – vadovėlis jūreiviui". In: (2015).

[73] N. C. Petersen, F. Rodrigues, and F. Camara Pereira. "Multi-Output Bus Travel Time Prediction with Convolutional LSTM Neural Network". In: *Expert Systems with Applications* 120 (2019), pp. 426–435. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2018.11.028.

[74] J. Qu, R. W. Liu, Y. Guo, Y. Lu, J. Su, and P. Li. "Improving Maritime Traffic Surveillance in Inland Waterways Using the Robust Fusion of AIS and Visual Data". In: *Ocean Engineering* 275 (2023), p. 114198. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2023.114198.

[75] A. A. Rizal, S. Soraya, and M. Tajuddin. "Sequence to Sequence Analysis with Long Short Term Memory for Tourist Arrivals Prediction". In: *Journal of Physics: Conference Series* 1211.1 (Apr. 2019), p. 012024. DOI: 10.1088/1742-6596/1211/1/012024.

[76] H. Rong, A. P. Teixeira, and C. Guedes Soares. "Data Mining Approach to Shipping Route Characterization and Anomaly Detection Based on AIS Data". In: *Ocean Engineering* 198.January (2020), p. 106936. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2020.106936.

[77] H. Rong, A. P. Teixeira, and C. G. Soares. "Maritime Traffic Probabilistic Prediction Based on Ship Motion Pattern Extraction". In: *Reliability Engineering and System Safety* 217.April 2021 (2022), p. 108061. ISSN: 0951-8320. DOI: 10.1016/j.ress.2021.108061.

[78] Y. Russac, O. Caelen, and L. He-Guelton. "Embeddings of Categorical Variables for Sequential Data in Fraud Context". In: *The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2018)*. Ed. by A. E. Hassanien, M. F. Tolba, M. Elhoseny, and M. Mostafa. Cham: Springer International Publishing, 2018, pp. 542–552. ISBN: 978-3-319-74690-6.

[79] J. dos Santos Sá, E. do Nascimento da Silva, L. Nunes Gonçalves, C. Mateus Machado Cardoso, A. Antloga do Nascimento, G. Protásio dos Santos Cavalcante, M. Emília de Lima Tostes, J. Priscyla Leite de Araújo, F. José Brito Barros, and F. de Souza Farias. "Multimodal Urban Mobility Solutions for a Smart Campus Using Artificial Neural Networks for Route Determination and an Algorithm for Arrival Time Prediction". In: *Engineering Applications of Artificial Intelligence* 137

(2024), p. 109074. ISSN: 0952-1976. DOI: 10.1016/j.engappai.2024.109074.

[80] A. Seltmann. *IUMI Stats Report*. Seoul, South Korea, 2021.

[81] A. Seltmann. *Marine Insurance – Casualty trends*. Oslo, Norway, 2019.

[82] C. Seo, Y. Noh, M. Abebe, Y.-J. Kang, S. Park, and C. Kwon. "Ship Collision Avoidance Route Planning Using a CRI-Based A* Algorithm". In: *International Journal of Naval Architecture and Ocean Engineering* 15 (2023), p. 100551. ISSN: 2092-6782. DOI: 10.1016/j.ijnaoe.2023.100551.

[83] G. Shafer and V. Vovk. "A Tutorial on Conformal Prediction". In: *Journal of Machine Learning Research* 9 (2008), pp. 371–421. ISSN: 15324435. arXiv: 0706.3188.

[84] A. F. Siegel. "Chapter 9 - Confidence Intervals: Admitting That Estimates Are Not Exact". In: *Practical Business Statistics (Sixth Edition)*. Ed. by A. F. Siegel. Sixth Edition. Boston: Academic Press, 2012, pp. 219–247. ISBN: 978-0-12-385208-3. DOI: 10.1016/B978-0-12-385208-3.00009-2.

[85] K. A. Sørensen, P. Heiselberg, and H. Heiselberg. "Probabilistic Maritime Trajectory Prediction in Complex Scenarios Using Deep Learning". In: *Sensors* 22 (5 Mar. 2022). ISSN: 14248220. DOI: 10.3390/s22052058.

[86] N. Stream. *Incident on the Nord Stream Pipeline: Secure Gas Supply for Europe*. Press release. 2022. URL: https://www.nord-stream.com/press-info/press-releases/incident-on-the-nord-stream-pipeline-updated-02112022-529/.

[87] Y. Sun, X. Chen, L. Jun, J. Zhao, Q. Hu, X. Fang, and Y. Yan. "Ship Trajectory Cleansing and Prediction With Historical AIS Data Using an Ensemble ANN Framework". In: *International Journal of Innovative Computing, Information and Control* 17.2 (2021), pp. 443–459. ISSN: 13494198. DOI: 10.24507/ijicic.17.02.443.

[88] Y. Sun, X. Peng, Z. Ding, and J. Zhao. "An Approach to Ship Behavior Prediction Based on AIS and RNN Optimization Model". In: *International Journal of Transportation Engineering and Technology* 6.1

(2020), pp. 16–21. ISSN: 2575-1743. DOI: 10.11648/j.ijtet.20200601.13.

[89] M. A. B. Syed and I. Ahmed. "A CNN-LSTM Architecture for Marine Vessel Track Association Using Automatic Identification System (AIS) Data". In: *Sensors* 23.14 (2023). ISSN: 14248220. DOI: 10.3390/s23146400. arXiv: 2303.14068.

[90] T. Szubrycht. "Marine Accidents as Potential Crisis Situations on the Baltic Sea". In: *Archives of Transport* 54.2 (2020), pp. 125–135. ISSN: 23008830. DOI: 10.5604/01.3001.0014.2972.

[91] S. Tajmouati, B. EL Wahbi, and M. Dakkon. "Applying Regression Conformal Prediction with Nearest Neighbors to Time Series Data". In: *Communications in Statistics: Simulation and Computation* (2022), pp. 1–13. ISSN: 15324141. DOI: 10.1080/03610918.2022.2057538. arXiv: 2110.13031.

[92] A. Tritsarolis, E. Chondrodima, N. Pelekis, and Y. Theodoridis. "Vessel Collision Risk Assessment using AIS Data: A Machine Learning Approach". In: *Proceedings - IEEE International Conference on Mobile Data Management* 2022-June.Mdm (2022), pp. 425–430. ISSN: 15516245. DOI: 10.1109/MDM55031.2022.00093.

[93] A. Tritsarolis, B. Murray, N. Pelekis, and Y. Theodoridis. "Collision Risk Assessment and Forecasting on Maritime Data". In: (2023), pp. 1–10. DOI: 10.1145/3589132.3625573.

[94] United Nations Conference on Trade and Development. *Review of Maritime Transport 2023*. Accessed: 2025-01-12. Geneva: United Nations Publications, 2023. URL: https://unctadstat.unctad.org/CountryProfile/MaritimeProfile/en-GB/528/MaritimeProfile528.pdf.

[95] J. Venskus. "Semi-supervised and Unsupervised Machine Learning Methods for Sea Traffic Anomaly Detection". Prieiga per eLABa – nacionalinė Lietuvos akademinė elektroninė biblioteka. PhD thesis. Vilniaus universitetas, 2021. DOI: 10.15388/vu.thesis.179.

[96] J. Venskus and P. Treigys. "Preparation of Training Data by Filling in Missing Vessel Type Data Using Deep Multi-Stacked LSTM Neural Network for Abnormal Marine Transport Evaluation". In: TISE 2019 :

International Conference on Time Series and Forecasting : proceedings of abstracts (Sept. 2019), p. 38.

[97] J. Venskus, P. Treigys, J. Bernatavičienė, V. Medvedev, M. Voznak, M. Kurmis, and V. Bulbenkienė. "Integration of a Self-Organizing Map and a Virtual Pheromone for Real-Time Abnormal Movement Detection in Marine Traffic". In: *Informatica* 28.2 (2017), pp. 359–374. ISSN: 08684952. DOI: 10.15388/Informatica.2017.133.

[98] J. Venskus, P. Treigys, J. Bernatavičienė, G. Tamulevičius, and V. Medvedev. "Real-Time Maritime Traffic Anomaly Detection Based on Sensors and History Data Embedding". In: *Sensors* 19.17 (2019). ISSN: 14248220. DOI: 10.3390/s19173782.

[99] J. Venskus, P. Treigys, and J. Markevičiūtė. "Unsupervised Marine Vessel Trajectory Prediction Using LSTM Network and Wild Bootstrapping Techniques". In: *Nonlinear Analysis: Modelling and Control* 26.4 (2021), pp. 718–737. ISSN: 23358963. DOI: 10.15388/namc.2021.26.23056.

[100] R. Vinayakumar, K. P. Soman, and P. Poornachandran. "Applying Deep Learning Approaches for Network Traffic Prediction". In: *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. Amrita Vishwa Vidyapeetham, India, 2017, pp. 2353–2358. ISBN: 978-1-5090-6367-3. DOI: 10.1109/ICACCI.2017.8126198.

[101] N. Wahid, T. Nur Adi, H. Bae, and Y. Choi. "Predictive Business Process Monitoring – Remaining Time Prediction using Deep Neural Network with Entity Embedding". In: *Procedia Computer Science* 161 (Jan. 2019), pp. 1080–1088. DOI: 10.1016/j.procs.2019.11.219.

[102] N. Waltteri, N. Florent, H. Markus, and L. Meski. "Shipping Accidents in the Baltic Sea 2020". In: *Baltic Marine Environment Protection Commission* (2021), p. 35.

[103] S. Wang, Y. Li, and H. Xing. "A Novel Method for Ship Trajectory Prediction in Complex Scenarios Based on Spatio-Temporal Features Extraction of AIS Data". In: *Ocean Engineering* 281.April (2023), p. 114846. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2023.114846.

[104] X. Wang and Y. Xiao. "A Deep Learning Model for Ship Trajectory Prediction Using Automatic Identification System (AIS) Data". In: *INFORMATION* 14.4 (Apr. 2023). DOI: 10.3390/info14040212.

[105] X. Wang, J. Li, and T. Zhang. "A Machine-Learning Model for Zonal Ship Flow Prediction Using AIS Data: A Case Study in the South Atlantic States Region". In: *Journal of Marine Science and Engineering* 7.12 (2019). ISSN: 20771312. DOI: 10.3390/JMSE7120463.

[106] J. Weng, J. Du, K. Shi, and S. Liao. "Effects of Ship Domain Shapes on Ship Collision Risk Estimates Considering Collision Frequency and Severity". In: *Ocean Engineering* 283.April (2023), p. 115070. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2023.115070.

[107] W. Wu, P. Chen, L. Chen, and J. Mou. "Ship Trajectory Prediction: An Integrated Approach Using ConvLSTM-Based Sequence-to-Sequence Model". In: *JOURNAL OF MARINE SCIENCE AND ENGINEERING* 11.8 (Aug. 2023). DOI: 10.3390/jmse11081484.

[108] W. Xie, L. Gang, M. Zhang, T. Liu, and Z. Lan. "Optimizing Multi-Vessel Collision Avoidance Decision Making for Autonomous Surface Vessels: A COLREGs-Compliant Deep Reinforcement Learning Approach". In: *Journal of Marine Science and Engineering* 12.3 (2024). ISSN: 2077-1312. DOI: 10.3390/jmse12030372.

[109] C. Xu, H. Jiang, and Y. Xie. *Conformal Prediction for Multi-Dimensional Time Series by Ellipsoidal Sets*. 2024. arXiv: 2403.03850 [stat.ML]. URL: https://arxiv.org/abs/2403.03850.

[110] C. H. Yang, C. H. Wu, J. C. Shao, Y. C. Wang, and C. M. Hsieh. "AIS-Based Intelligent Vessel Trajectory Prediction Using Bi-LSTM". In: *IEEE Access* 10 (2022), pp. 24302–24315. ISSN: 21693536. DOI: 10.1109/ACCESS.2022.3154812.

[111] J. B. Yim, D. S. Kim, and D. J. Park. "Modeling Perceived Collision Risk in Vessel Encounter Situations". In: *Ocean Engineering* 166 (Oct. 2018), pp. 64–75. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2018.08.003.

[112] C. Yin and W. Zhang. "Concise Deep Reinforcement Learning Obstacle Avoidance for Underactuated Unmanned Marine Vessels".

In: *Neurocomputing* 272 (2018), pp. 63–73. ISSN: 0925-2312. DOI: 10.1016/j.neucom.2017.06.066.

[113] Z. Yin, D. Yang, and X. Bai. "Vessel Destination Prediction: A Stacking Approach". In: *Transportation Research Part C: Emerging Technologies* 145 (2022), p. 103951. ISSN: 0968-090X. DOI: 10.1016/j.trc.2022.103951.

[114] Z. Yin and B. Zhang. "Construction of Personalized Bus Travel Time Prediction Intervals Based on Hierarchical Clustering and the Bootstrap Method". In: *Electronics (Switzerland)* 12.8 (2023). ISSN: 20799292. DOI: 10.3390/electronics12081917.

[115] H. Yoshioka, H. Hashimoto, and H. Makino. "Decision-making Algorithm for Ship Collision Avoidance with Collision Risk Map". In: *Ocean Engineering* 286 (2023), p. 115705. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2023.115705.

[116] L. You, S. Xiao, Q. Peng, C. Claramunt, X. Han, Z. Guan, and J. Zhang. "ST-Seq2Seq: A Spatio-Temporal Feature-Optimized Seq2Seq Model for Short-Term Vessel Trajectory Prediction". In: *IEEE ACCESS* 8 (2020), pp. 218565–218574. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3041762.

[117] B. Zhang, S. Wang, L. Deng, M. Jia, and J. Xu. "Ship Motion Attitude Prediction Model Based on IWOA-TCN-Attention". In: *Ocean Engineering* 272.January (2023), p. 113911. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2023.113911.

[118] Q. Zhang, C. Yin, Y. Chen, and F. Su. "IGCRRN: Improved Graph Convolution Res-Recurrent Network for Spatio-Temporal Dependence Capturing and Traffic Flow Prediction". In: *Engineering Applications of Artificial Intelligence* 114.July (2022), p. 105179. ISSN: 0952-1976. DOI: 10.1016/j.engappai.2022.105179.

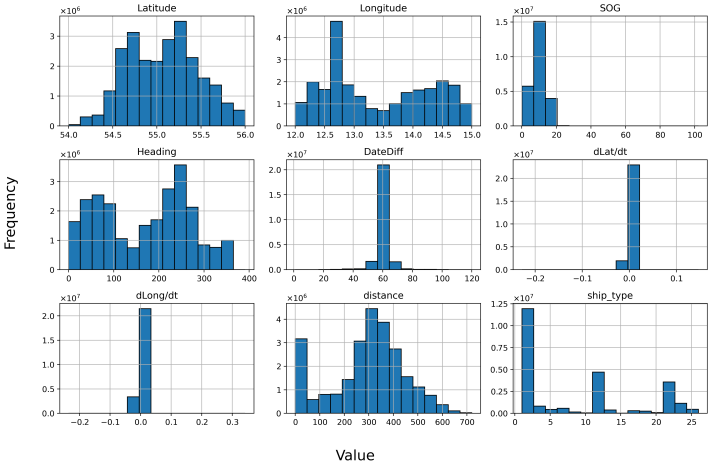
[119] S. Zhang, L. Wang, M. Zhu, S. Chen, H. Zhang, and Z. Zeng. "A Bi-directional LSTM Ship Trajectory Prediction Method based on Attention Mechanism". In: *IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* 2021 (2021), pp. 1987–1993. ISSN: 26896621. DOI: 10.1109/IAEAC50856.2021.9391059.

[120] X. Zhang, J. Liu, P. Gong, C. Chen, B. Han, and Z. Wu. "Trajectory Prediction of Seagoing Ships in Dynamic Traffic Scenes via a Gated Spatio-Temporal Graph Aggregation Network". In: *Ocean Engineering* 287 (2023), p. 115886. ISSN: 0029-8018. DOI: 10.1016/j.oceaneng.2023.115886.

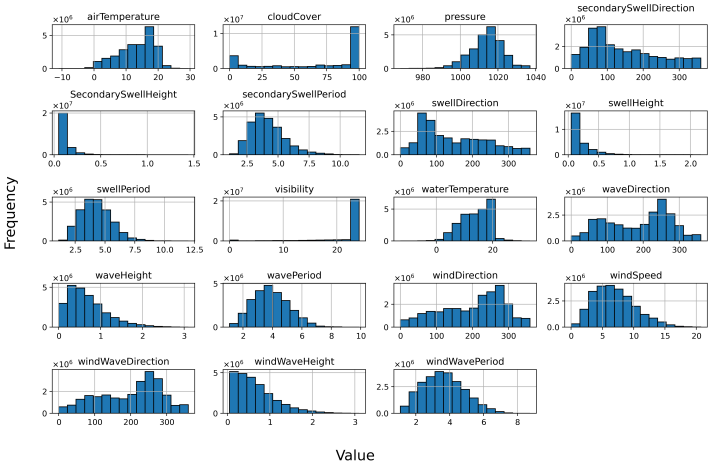
[121] J. Zhao, Z. Yan, Z. Z. Zhou, X. Chen, B. Wu, and S. Wang. "A Ship Trajectory Prediction Method Based on GAT and LSTM". In: *Ocean Engineering* 289.P1 (2023), p. 116159. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2023.116159.

[122] Y. Zhou, W. Du, J. Liu, H. Li, M. Grifoll, W. Song, and P. Zheng. "Determination of Ship Collision Avoidance Timing Using Machine Learning Method". In: *Sustainability* 16.11 (2024). ISSN: 2071-1050. DOI: 10.3390/su16114626.

A Appendix - Numerical data



(a)



(b)

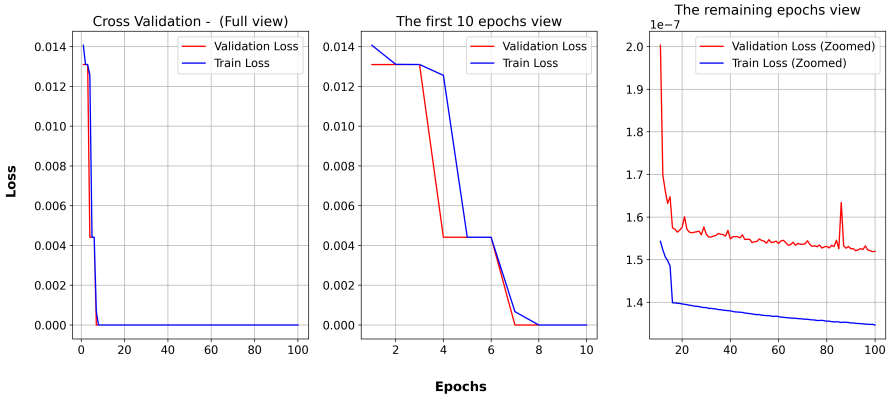
Figure A.1: Baltic Sea region representation of: (**a**) AIS dataset distribution; (**b**) weather dataset distribution.

B Appendix - AIS data & filtering

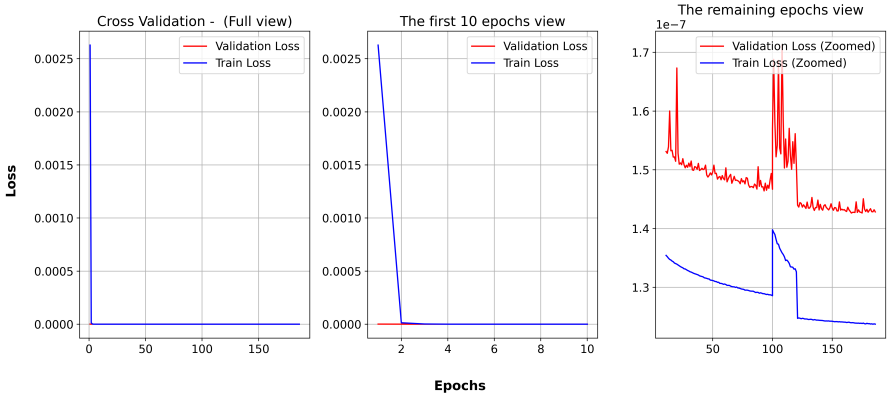
Table B.1: Dataset meta information before and after processing.

Raw Data (Before Processing)	**Processed Data (After Filtering)**
Total Entries: 93,822,057	**Total Entries:** 26,441,425
Features: `# Timestamp`: Object `MMSI`: Float64 `Latitude`: Float64 `Longitude`: Float64 `Navigational status`: Object `ROT`: Float64 `SOG`: Float64 `COG`: Float64 `Heading`: Float64 `IMO`: Object `Callsign`: Object `Name`: Object `Ship type`: Object `Cargo type`: Object `Width`: Float64 `Length`: Float64 `Draught`: Float64 `Destination`: Object `ETA`: Object	**Features:** `id`: Int64 `DateTime`: Datetime64 `# Timestamp`: Object `Latitude`: Float64 `Longitude`: Float64 `SOG`: Float64 `Heading`: Float64 `DateDiff`: Float64 `dLat/dt`: Float64 `dLong/dt`: Float64 `distance`: Float64 `Ship type`: Object
Applied Filters: None	**Applied Filters:** `SOG` $\neq 0$ `DateDiff` > 0 `DateDiff` ≤ 120 `Distance` ≤ 800 Resampling (60s) Drop NA Drop Duplicates

C Appendix - Model learning process



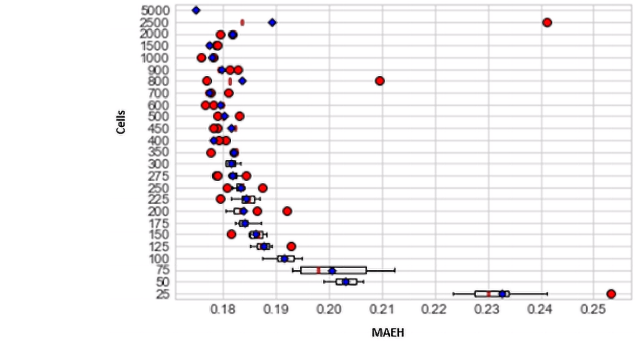
(a)



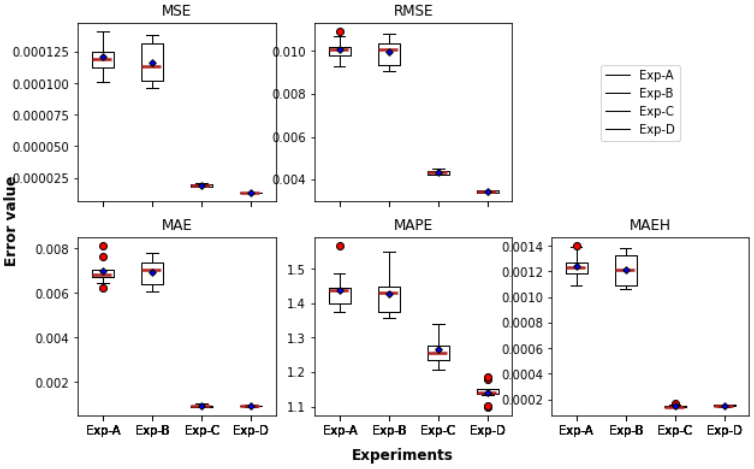
(b)

Figure C.1: Cross-validation results for different models: (**a**) cross-validation plot for a random model; (**b**) cross-validation plot for the best model.

D Appendix - Cell Count Impact on Prediction Accuracy



(a)



(b)

Figure D.1: Analysis of configurations and prediction accuracy: (**a**) impact of LSTM cell count on MAEH prediction error; (**b**) comparative evaluation of coordinate transformation experiments, where Exp-A uses raw AIS features, Exp-B includes delta latitude and longitude, Exp-C employs distance and azimuth angle transformations, and Exp-D utilises UTM projection coordinates.

E Appendix - Random routes

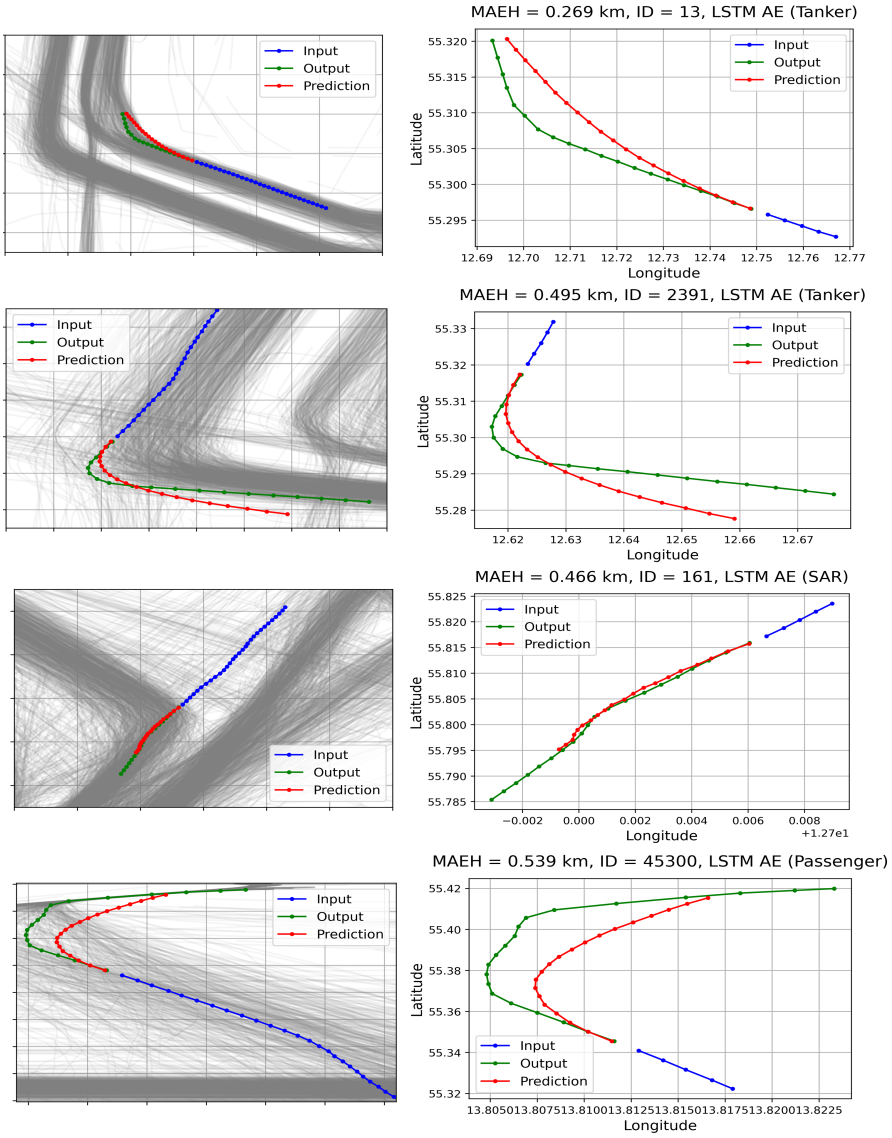


Figure E.1: Prediction examples (a).

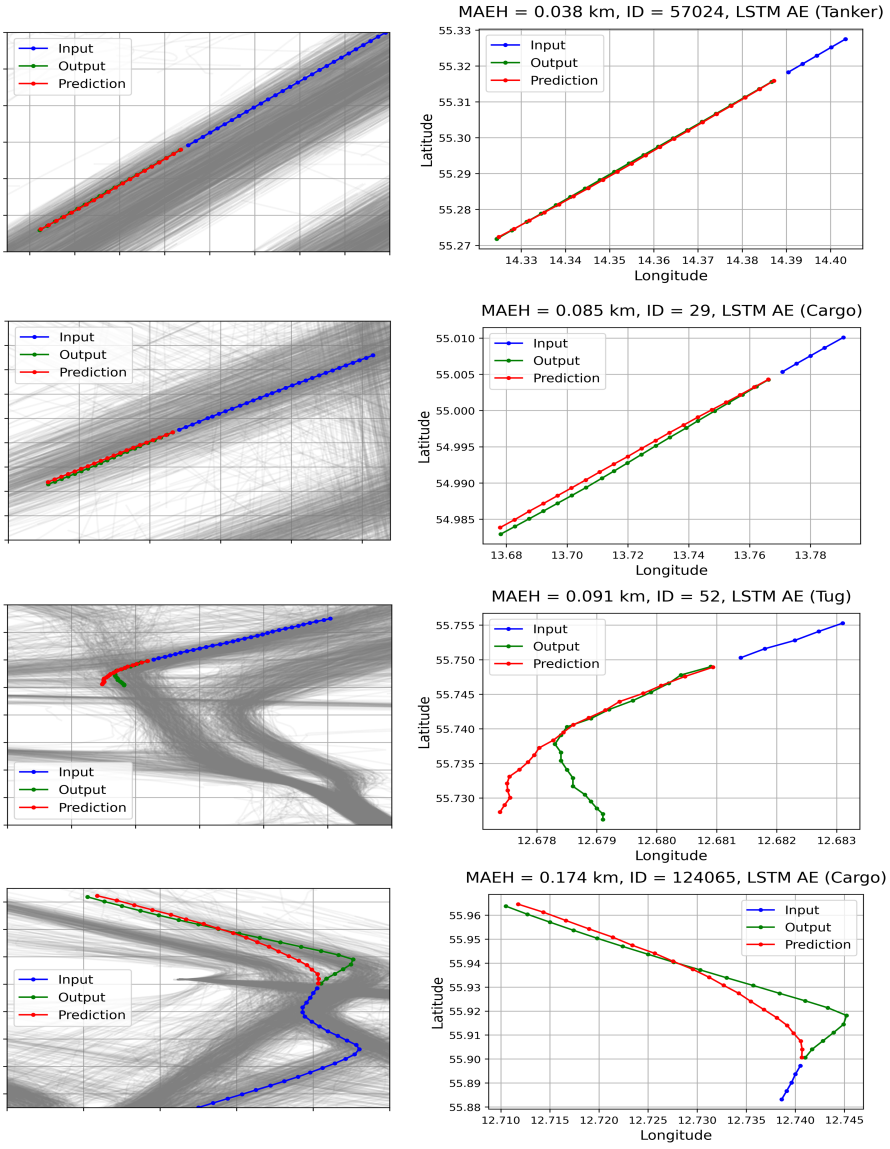


Figure E.2: Prediction examples (b). Baltic sea trajectories for different vessel types: grey shows traffic intensity, red is the prediction of the current state, blue is the model input, and green is the actual movement.

List of Publications

The results obtained in this dissertation have been published in five scientific publications: two research articles in periodical journals indexed by Web of Science, two articles at peer-reviewed scientific conferences, and one national non-peer-reviewed publication of conference proceedings. The results were presented at two international and four national scientific conferences. The list of publications and conference presentations is given below.

Articles in periodic scientific journals indexed in The Web of Science:

[A.1] **Robertas Jurkus**, Julius Venskus, and Povilas Treigys, "Application of coordinate systems for vessel trajectory prediction improvement using a recurrent neural networks". In *Engineering applications of artificial intelligence*, Oxford : Pergamon-Elsevier Science Ltd. ISSN 0952-1976. eISSN 1873-6769. 2023, vol. 123, part C, art. no. 106448, p. [1-10]. DOI: 10.1016/j.engappai.2023.106448.

[A.2] **Robertas Jurkus**, Julius Venskus, Jurgita Markevičiūtė, and Povilas Treigys, "Enhancing Maritime Safety: Estimating Collision Probabilities with Trajectory Prediction Boundaries Using Deep Learning Models". In: *Sensors*, vol. 25, issue 5, article 1365, 2025. DOI: 10.3390/s25051365.

Articles in peer-reviewed scientific conference proceedings:

[B.1] **Robertas Jurkus**, Povilas Treigys, and Julius Venskus, "Investigation of Recurrent Neural Network Architectures for Prediction of Vessel Trajectory". In: *Information and Software Technologies*, Springer, Chapter No: 16, vol 1486. Chapter DOI:10.1007/978-3-030-88304-1_16, KTU, Lithuania, 2021.

[B.2] **Robertas Jurkus**, Julius Venskus, and Povilas Treigys, "Extended Research on Categorical Data Encoding Techniques for Recursive Multi-Step Prediction of Vessel Trajectory". In: *Time Series Analysis and Forecasting*. ITISE 2023. Contributions to Statistics. Springer, Cham, 16 April 2025. eBook ISBN: 978-3-031-69750-0. DOI:10.1007/978-3-031-69750-0_16.

ARTICLES IN NATIONAL NON-PEER-REVIEWED PUBLICATION:

[C.1] **Robertas Jurkus**, Julius Venskus, and Povilas Treigys, "Laivo tipų įtaka prognozuojant laivo judėjimo trajektoriją naudojant giliuosius rekurentinius neuroninius tinklus". In: *Technologijų ir verslo aktualijos – 2022*, ISSN 2538-8045, KTU, 22-oji studentų mokslinė konferencija, Panevėžys, Lithuania, November 25, 2022.

PRESENTATIONS AT SCIENTIFIC CONFERENCES:

[D.1] **Robertas Jurkus**, Povilas Treigys, and Julius Venskus, "Investigation of Recurrent Neural Network Architectures for Prediction of Vessel Trajectory" in *International Conference on Information and Software Technologies (ICIST) 2021*, Kauno technologijos universitetas, Kaunas, Lithuania, September 14-16, 2021.

[D.2] **Robertas Jurkus**, Julius Venskus, and Povilas Treigys, "Prediction of vessels trajectory using different coordinate systems" in *Data Analysis Methods for Software Systems (DAMSS 2021)*, DAMSS: 12th conference on data analysis methods for software systems, Druskininkai, Lithuania, December 2–4, 2021.

[D.3] **Robertas Jurkus**, Julius Venskus, and Povilas Treigys, "Laivo tipų įtaka prognozuojant laivo judėjimo trajektoriją naudojant giliuosius rekurentinius neuroninius tinklus" in *Technologijų ir verslo aktualijos*, TVA 2022 22-oji studentų mokslinė konferencija, KTU, Panevėžio

technologijų ir verslo fakultetas, Panevėžys, Lietuva, lapkričio 25 d., 2022.

[D.4] **Robertas Jurkus**, Julius Venskus, and Povilas Treigys, "Categorical Data Encoding Techniques for Recursive Multi-Step Prediction of Vessel Trajectory" in *International Conference on Time Series and Forecasting (ITISE 2023)*, Gran Canaria, Spain, July 12-14, 2023.

[D.5] Julius Venskus, **Robertas Jurkus**, and Povilas Treigys, "Confidence And Prediction Intervals Usage In Maritime Traffic Awareness Evaluation Using LSTM Deep Neural Networks" in *Data Analysis Methods for Software Systems (DAMSS 2023)*, DAMSS: 14th conference on data analysis methods for software systems, Druskininkai, Lithuania, November 30 - December 2, 2023.

[D.6] Julius Venskus, **Robertas Jurkus**, and Povilas Treigys, "A complex network approach to marine traffic interaction modeling: feature extraction for real-time traffic awareness" in *Data Analysis Methods for Software Systems (DAMSS 2024)*, DAMSS: 15th conference on data analysis methods for software systems, Druskininkai, Lithuania, November 28 - November 30, 2024.

PRESENTATIONS AT NATIONAL SCIENTIFIC INSTITUTIONS:

[E.1] **Robertas Jurkus**, "Giliųjų rekurentinių tinklų ir AIS duomenų tyrimas laivo trajektorijos prognozavimui". Systems analysis seminar. Institute of Data Science and Digital Technologies. Vilnius University, May 31, 2021.

[E.2] **Robertas Jurkus**, "Polinės ir Dekarto koordinačių sistemų taikymas laivo trajektorijos prognozavimui naudojant rekurentinius neuroninius tinklus". Systems analysis seminar. Institute of Data Science and Digital Technologies. Vilnius University, March 7, 2022.

[E.3] **Robertas Jurkus**, Presentation at the event "AI Lithuania Meetup Klaipėda", title of the presentation "Application of polar and Cartesian coordinate systems for trajectory prediction using recurrent neural networks". Klaipėda University, May 20, 2022.

[E.4] **Robertas Jurkus**, “Laivo tipų įtaka prognozuojant laivo judėjimo trajektoriją naudojant giliuosius rekurentinius neuroninius tinklus bei jų modifikacijas“. Systems analysis seminar. Institute of Data Science and Digital Technologies. Vilnius University, January 23, 2023.

[E.5] **Robertas Jurkus**, “Pasikliovimo intervalų taikymas vertinant laivų trajektorijos prognozę naudojant LSTM giliuosius neuroninius tinklus“. Systems analysis seminar. Institute of Data Science and Digital Technologies. Vilnius University, November 6, 2023.

RELATED AWARDS:

[F.1] Winner of the second prize of the “Competition of Student Final Theses on Transport Topics” organized by the Ministry of Transport and Communications, on the topic of water transport, titled “LSTM Deep Neural Network Research for Prediction of Vessel Movement Using Big Traffic Data”, 2022, Vilnius, Lithuania.

[F.2] Klaipėda Industrialists Association Award for the Master’s thesis “LSTM Deep Neural Network Research for Prediction of Vessel Movement Using Big Traffic Data”, 2020, Klaipėda, Lithuania.

DATASETS:

[G.1] **Danish maritime service**, Safety of Navigation, National Waters, Caspar Brands Plads 9, 4220 Korsør, Denmark. [Online]. Available: `http://web.ais.dk/aisdata/`. Baltic Sea region, accessed: June 03, 2024.

[G.2] **Ship Finder**, “The Live Marine Traffic Tracking App”. Available: `https://shipfinder.co/`. Port Region of the Netherlands, accessed: June 03, 2024.

[G.3] **Weatherbit**, Provider of the highest quality weather forecasts, observations and historical weather data. Available: `https://www.weatherbit.io`. Baltic Sea region, accessed: June 22, 2024.

Santrauka

Remiantis Pasauline jūrų draudimo metine ataskaita, nelaimės jūroje išlieka svarbi problema, susijusi tiek su žmogiškaisiais, tiek su nežmogiškaisiais rizikos veiksniais. Tarp jų – laivų susidūrimai ir eismo anomalijos. Analizuojant AIS istorinius duomenis, kuriami intelektualūs sprendimai laivų trajektorijoms prognozuoti. Dažniausiai siekiama gerinti regresinių prognozių tikslumą pasitelkiant istorinių judėjimo modelių analizę, elgsenos panašumus ar klasterizavimo metodus, nes situacinis jūrų eismo suvokimas yra ypač svarbus transporto saugai. Vis dėlto, visuotinio sprendimo šioje srityje vis dar nėra pasiūlyta.

Situacinis suvokimas, apimantis aplinkos stebėseną, interpretavimą ir būsimų pokyčių prognozavimą, laikomas aukščiausiu (3-iuoju) suvokimo lygiu. Disertacijoje nagrinėjamos giliojo mokymosi rekurentinės tinklų architektūros, jų hiperparametrai bei atliekama požymių inžinerija naudojant pusiau struktūrizuotus AIS duomenis. Sukurtas rekursinis transformacijos modelis leidžia tiksliau ekstrapoliuoti kelių žingsnių laivų trajektorijas. Empiriniai rezultatai parodė, kad tiksliausias modelis – daugiamačių duomenų ilgos–trumpalaikės atminties autoenkoderis (angl. *long short-term memory autoencoder*, LSTM AE), kuriame integruojami laivo tipų ir meteorologiniai duomenys kaip generalizuotas sprendimas. Tikslumui gerinti siūloma taikyti skirtingas koordinačių sistemas, apskaičiuoti delta erdvinių koordinacinių taškų (vektorių) skirtumus – pokytį tarp ankstesnio bei einamojo taško ir rekursyviai atkurti sekos pozicijas pridedant šias deltas prie paskutinio žinomo taško, nenaudojant absoliučių reikšmių. Kategoriniai laivų tipų duomenys integruojami į bendrą mokymo rinkinį naudojant skirtingas kodavimo technikas. Prognozių patikimumui ir neapibrėžtumui įvertinti taikomi statistiniai metodai, elipsiniai ir konforminiai prognozės regionai (sritys) bei intervalai, leidžiantys apskaičiuoti tikimybinį trajektorijų persidengimą, ir nustatyti galimą greta esančių laivų susidūrimo riziką.

Disertacijoje analizuojamas realus avarijos tyrimo atvejis – 2021 m. laivų susidūrimas prie Bornholmo salos. Taikant giliuosius rekurentinius neuroninius tinklus, sukurti modeliai prognozuoja kelių žingsnių į priekį trajektorijas ir formuoja prognozės sritis 95 % pasikliovimo lygmeniu susidūrimo rizikai vertinti. Neparametrinis konforminių sričių metodas pasirodė veiksmingas

identifikuojant galimus susidūrimo scenarijus. Rezultatai rodo, kad giliojo mokymosi modeliai, turintys trumpalaikio ir ilgalaikio prognozavimo galimybes, gali reikšmingai pagerinti navigacijos sprendimus ir prisidėti prie incidentų prevencijos.

I. Įvadas

Jūrų transportas yra gyvybiškai svarbus pasaulinei prekybai, tačiau didėjantis laivybos intensyvumas kelia susidūrimų, saugumo ir aplinkosaugos rizikų. Informacija apie laivų padėtį, trajektorijas ir paskirties vietas yra būtina siekiant užtikrinti jūrų saugumą, aplinkos apsaugą ir ekonomikos stabilumą. Tarptautinė jūrų draudimo sąjunga (IUMI) praneša, kad 2019 m. apie 10 % nuostolių kilo dėl susidūrimų, o likusi dalis – dėl įgulos klaidų, įrangos gedimų ir nepalankių oro sąlygų [81]. Nepaisant jau didelės apkrovos, laivybos apimtys toliau auga: 2020 m. IUMI duomenimis jūrų draudimo įmokos siekė 30 mlrd. USD, tai 6,1 % daugiau nei 2019 m., o krovinių laivai užėmė 57,2 % visos draudimo įmokų sumos [80]. Tokie augantys srautai didina jūrų saugumo rizikas.

Didėjantis laivybos intensyvumas generuoja didžiulius automatinės identifikavimo sistemos (angl. *automatic identification system*, AIS) duomenų kiekius, kuriuos realiuoju laiku apdoroti ne tik sunku, bet dažnai ir neįmanoma pasitelkiant tik žmogiškuosius išteklius ar įprastus mašininio mokymosi (angl. *machine learning*, ML) metodus. Veiksmingas situacijos suvokimas turi tris lygius: pradiniai lygmenys orientuojasi į aplinkos vertinimą, o III lygis apima ir būsimos būklės suvokimą bei prognozavimą [63]. Norint pasiekti III lygį, reikia pažangių metodų, tokių kaip gilusis mokymasis (angl. *deep learning*, DL), kuriais galima analizuoti dideles, triukšmingas laiko eilučių duomenų apimtis. Neseni incidentai tai puikiai iliustruoja: 2025 m. kovo mėn. tanklaivis „MC Stena Immaculate“ rūke susidūrė su krovininiu laivu netoli Humberio žiočių; 2018 m. pabaigoje fregata „Helge Ingstad“ Norvegijoje susidūrė su naftą gabenančiu laivu, o Baltijos jūroje 2022 m. „Nord Stream“ dujotiekio sprogimai sutrikdė eismą. Tokie įvykiai rodo, kaip greitai perpildyti ar strateginiai vandens keliai gali tapti pavojingi, trikdyti įprastą eismą, todėl reikalingi papildomi sprendimai, pabrėžiant patikimų daugiažingsnių trajektorijų prognozavimo modelių ir stebėjimo realiuoju laiku priemonių svarbą.

Sonarai ir radarai padeda laivams aptikti artėjančias kliūtis, tačiau radaras tankaus eismo sąlygomis gali nepastebėti mažesnių objektų, užstojamų didesnių laivų, o abiejų sistemų patikimumas mažėja nepalankaus matomumo sąlygomis ar triukšmingoje aplinkoje. Daugelis laivų naudoja AIS atsakiklius, tačiau AIS duomenys realiuoju laiku gali vėluoti kelias minutes, be to, negali savaime prognozuoti būsimų judėjimų, o teikia tik esamos situacijos būsenas. Priešingai, istoriniai AIS įrašai, kartu su meteorologiniais ir kategoriniais laivo tipo duomenimis, puikiai tinka DL metodams ir prognozėms modeliuoti. Rekurentiniai neuroniniai tinklai (angl. *recurrent neural networks*, RNNs), ypač kelių žingsnių daugiamačiai modeliai, geba mokytis iš tokių laiko eilučių, prognozuoti laivo trajektoriją ir fiksuoti galimus susidūrimus ar neįprastus manevrus. Pasitelkus AIS signalus DL modeliuose, RNN pagrįstos prognozės prisideda prie geresnio situacijos suvokimo ir skatina proaktyvias saugumo priemones jūrų vandenyse, kur pasitaiko vis daugiau spūsčių.

Tyrimo problematika

Nors pažangūs mašininio ir giliojo mokymosi metodai vis plačiau taikomi sausumos transporto sistemose (pvz., apžvalginiame straipsnyje [4]), kur 45 % tyrimų skirti spūstims prognozuoti, o 30 % – srautams valdyti, jūrų transporto specifika beveik nenagrinėjama. AIS duomenų srautai jūroje yra netolygūs, laivų tipai – įvairesni, o aplinkos veiksniai (pvz., vėjas, srovės) stipriai veikia judėjimą. Masyviems, visiškai pakrautiems laivams, pvz., tanklaiviams, gali prireikti net 20–25 minučių visiškai sustoti, todėl reikalingos prognozės, apimančios ilgalaikius 20–25 min. išankstinius spėjimus, kad esant poreikiui būtų galima išvengti susidūrimų ar priimti kitus svarbius laivo valdymo sprendimus [38].

Įprastai laivams stebėti naudojami radarai, sonarai ir AIS, tačiau jie suteikia tik dabartinės padėties vaizdą ir gali klaidinti tankaus eismo ar prasto matomumo sąlygomis: radaras gali nepastebėti mažesnių objektų už didesnių laivų, o AIS signalai kartais vėluoja arba prarandami, todėl sprendimai realiuoju laiku tampa nepatikimi. Istoriniai AIS įrašai pateikia gausią trajektorijų informaciją, tačiau jų netolygūs matavimo intervalai ir duomenų triukšmas apsunkina trajektorijų

analizę bei panaudojimą prognozių ekstrapoliacijai, net trumpalaikiame (kelių minučių) laikotarpyje, o vien AIS duomenys neapima prognozavimo galimybių.

Daugelis iki šiol taikomų jūrų trajektorijų prognozavimo modelių pagrįsti trumpalaikiais arba tiesiniais spėjimais. Be to, neretai taiko universaliųjų skersinių Merkatoriaus (angl. *Universal Transverse Mercator*, UTM) koordinačių transformacijas be aiškaus paaiškinimo, kaip tai veikia prognozės tikslumą, arba kuria atskirus modelius skirtingiems laivų tipams, o tai reikalauja daug paruošiamojo darbo ir neįtraukia tokių kintamųjų kaip greitis, kursas ar aplinkos sąlygos. Todėl nėra vieningo sprendimo, galinčio vienodai gerai apdoroti visų laivų tipų elgseną. Gilieji rekurentiniai neuroniniai tinklai gali įsiminti ilgalaikes priklausomybes iš triukšmingų AIS laiko eilučių, tačiau ankstesniuose darbuose nėra ištirtas vienu metu integruotas kategorinių laivų tipų ir meteorologinių duomenų panaudojimas, taip pat nebuvo spręsta absoliučių koordinačių prognozavimo įtaka daugiažingsnėse prognozėse, kurią galima gerinti modeliuojant vektorių pokyčius tarp paskutinių laiko taškų.

Susidūrimų rizikos vertinimas šiuo metu dažniausiai grindžiamas deterministiniais rodikliais, tokiais kaip minimalus artimiausio priartėjimo taškas (angl. *closest point of approach*, CPA) ir laikas iki minimalaus priartėjimo (angl. *time to closest point of approach*, TCPA), kurie apskaičiuoja vieną būsimą tašką, kai du laivai bus arčiausiai vienas kito fiksuoto atstumo intervale (pvz., 0,25 jūrmylės arba 15 minučių). Nors CPA / TCPA plačiai naudojami sprendimams realiuoju laiku, juos sudėtinga interpretuoti, nes jie apskaičiuoja tik konkretų tašką, ignoruoja trajektorijos neapibrėžtumą, nesugeba tvarkyti kelių laivų trajektorijų persidengimų ir neleidžia priskirti prasmingos susidūrimo tikimybės. Siekiant pašalinti šiuos trūkumus ir pasiūlyti platesnį būsimų laivų trajektorijų vertinimą, pristatomuose metoduose vertinami neapibrėžtumo klasteriai ir plečiamas CPA fiksuoto taško analizės požiūris. Šis požiūris, apimantis 20 minučių prognozės intervalą, atitinkantį didelių laivų stabdymo laiką, ir numatantis dinamiškas trajektorijų ribas, pakeičia statinius atstumo ir laiko ribojimus tikimybiniais persidengimų rodikliais, kurie geriau atspindi galimų sąveikų įvairovę.

Šia disertacija kuriamas vieningas sprendimas, naudojant giliuosius RNN ir siekiant gauti tikslesnius daugiažingsnius laivų trajektorijų spėjimus

pagal istorinius, pusiau struktūrizuotus AIS duomenis; integruojami laivų tipų elgsenos ir meteorologiniai kintamieji į vieną modelį, gerinantį prognozavimo galimybes bendrai visoms laivų klasėms; ir pakeičiamos vieno taško CPA / TCPA taisyklės tikimybinėmis rizikos sritimis, atsižvelgiančiomis į modelio ir duomenų neapibrėžtumą. Dėl tokio derinio siūlomas požiūris sustiprina realaus laiko jūrų situacijos suvokimą ir pateikia veiksmingus rizikos vertinimus tada, kai jų labiausiai reikia.

Tyrimo objektas

Daugiamačiai, daugiažingsniai laivų trajektorijų duomenys, pagrįsti AIS, skirti susidūrimų rizikai prognozuoti ir vertinti. RNN metodai yra taikymo priemonė objektui tirti.

Tyrimo tikslai ir uždaviniai

Šio tyrimo tikslas – pasiūlyti ir ištirti giliųjų neuroninių tinklų pagrindu veikiančius algoritmus, skirtus kelių žingsnių (ilgalaikėms) laivų trajektorijų prognozėms ir susidūrimų rizikai vertinti jūrų navigacijoje.

Siekiant šio tikslo, keliami šie uždaviniai:

1. Atlikti kelių žingsnių daugiamačių laivų trajektorijų prognozavimo metodų literatūros analizę, nagrinėjant skirtingas modelių kūrimo strategijas, duomenų reprezentacijas ir jų taikymą jūriniam situacijos suvokimui bei saugai.
2. Išvystyti ir palyginti RNN architektūras ilgalaikiam laivų trajektorijų prognozavimui, vertinant jų patikimumą, jautrumą hiperparametrams ir optimizavimo kriterijus, siekiant nustatyti veiksmingiausias trajektorijų prognozavimo modelių konfigūracijas.
3. Įvertinti trajektorijų prognozavimo tikslumo priklausomybę nuo kategorinių, meteorologinių ir erdvinių duomenų, įskaitant koordinačių sistemų transformacijas, bei nustatyti tinkamiausias kategorinių duomenų apdorojimo metodikas.

4. Įvertinti ir identifikuoti veiksmingus siūlomus metodus laivų susidūrimų rizikai vertinti, taikant modelio prognozių neapibrėžtumo kiekybinį įvertinimą prognozės patikimumui nustatyti bei naudojant deterministinius statistinius metodus galimiems trajektorijų persidengimams, rodantiems susidūrimo scenarijus, aptikti ir kiekybiškai įvertinti.

5. Patikrinti susidūrimų aptikimo metodikas empirinio tyrimo metu, naudojant anksčiau netirtus laivų trajektorijų duomenis su realiais istoriniais jūriniais incidentais.

Mokslinis naujumas ir praktinė vertė

Šiame tyrime pristatomi keli naujoviški metodai, kurie kartu tobulina laivų trajektorijų prognozavimą ir susidūrimų rizikos vertinimą jūrų navigacijoje. Pagrindiniai indėliai yra:

1. Nustatyta, kad LSTM AE yra veiksmingiausia prižiūrimojo mokymo RNN architektūra daugiažingsniam laivų judėjimo ekstrapoliavimui, ir atrasta kritinė ląstelių skaičiaus riba. Palyginus įvairias RNN architektūras laiko eilučių regresijai, paaiškėjo, kad LSTM AE geriausiai įsisavina ilgalaikius priklausomybių modelius. Naudojant 75–300 LSTM ląstelių, pasiekiamas optimalus tikslumas. Nors didinant ląstelių skaičių, papildomos ląstelės suteikia menką naudą, tačiau smarkiai padidina skaičiavimo sąnaudas, o tai padeda projektuoti veiksmingus modelius.

2. Laivų tipų ir meteorologinių duomenų (pvz., vėjo greičio, srovių), ir AIS duomenų vieningo integravimo sprendimas sumažina duomenų paruošimo apimtį ir gerina įvairių laivų klasių generalizavimą. Kategorinių laivų tipų įterpimas (angl. *embedding*) drauge su aplinkos kintamaisiais vienoje duomenų imtyje užtikrina, kad prognozės atspindi realias sąlygas, taip didinant tikslumą ir patikimumą skirtingose aplinkose.

3. Užuot prognozavus absoliučias koordinates, modeliu prognozuojami gretimų laiko žingsnių padėties vektoriaus pokyčiai (delta). Šie prognozuoti delta vektoriai rekursyviai pridedami prie žinomos esamos padėties, taip atkuriant trajektoriją. Šis metodas, taikomas tiek pasaulinėje geodezinėje WGS84, tiek UTM koordinačių sistemose, padeda sumažinti

kaupiamąją paklaidą, gerina erdvinį vientisumą ir užtikrina prognozės nuoseklumą ilgesniam laikotarpiui.

4. Pasiūlyta naudoti neparametrines konforminių prognozės regionų (angl. *conformal prediction region*, CPR) metodikas laivų susidūrimams aptikti, konstruktyviai sudarant tikimybinės rizikos zonas iš kelių LSTM AE prognozių. Imant skirtingas hiperparametrų konfigūracijas, susidaro nežymių klaidų pasiskirstymas (stochastiškumas), o CPR apgaubia šias prognozes nenaudodama jokių prielaidų apie klaidų pasiskirstymą. Kai dviejų laivų CPR zonos persidengia, apskaičiuojamas susidūrimo rizikos balas – tai suteikia regionu pagrįstą rodiklį, atspindintį neapibrėžtumą keliuose prognozių rinkiniuose, o ne vieno taško kaip CPA / TCPA įvertį. Naujoviškai taikant CPR, tiesiogiai naudojamas modelio kintamumas ir triukšmingi duomenys, siekiant pateikti informatyvesnių susidūrimo rizikos įvertinimų.

Ginamieji teiginiai

Disertacijoje ginami šie tyrimu pagrįsti teiginiai:

1. Kai laivo trajektorija prognozuojama naudojant LSTM tipo neuroninio tinklo architektūrą, pasiekus tam tikrą LSTM ląstelių skaičių, prognozės tikslumas nebekinta arba kinta nežymiai, tačiau tinklo mokymo trukmė smarkiai pailgėja.

2. Kelių žingsnių daugiamačių laivų trajektorijos ekstrapoliacijai taikant rekursyvų prognozės sekos perskaičiavimą nuo ankstesnio taško, galima tiksliau nustatyti kitus prognozių taškus, ypač pradiniuose etapuose.

3. Laivų tipų ir meteorologinės informacijos įtraukimas į bendrą LSTM rekurentinio kelių žingsnių daugiamačio neuroninio tinklo mokymo duomenų rinkinį gerina laivų trajektorijų prognozavimo tikslumą, kai taikomas įterptasis kodavimo metodas.

4. Konforminis prognozavimo regionas (CPR) leidžia nustatyti laivų susidūrimo ribas su didžiausia statistine tikimybe (95 % pasikliovimo lygiu), kai duomenys yra daugialypiai ir neturi normaliojo skirstinio.

Disertacijos santraukoje pristatoma jūrų eismo situacija, tyrimo aktualumas, tikslas, problematika, ginamieji teiginiai ir bendra darbo eiga. Antroje dalyje apžvelgiami gilieji RNN, kategorinių duomenų kodavimo metodai ir prognozių neapibrėžtumo vertinimo būdai. Trečioje dalyje nagrinėjamas LSTM AE metodas su kategorinių ir meteorologinių duomenų integracija bei susidūrimų prognozavimo sritimis. Ketvirtoje dalyje aprašomi tyrimo duomenys, jų apdorojimas ir modelių kūrimo eiga. Penktoje dalyje pateikiami pagrindiniai eksperimentų rezultatai, o disertacija užbaigiama bendrosiomis išvadomis.

II. Literatūros analizė

Laivų trajektorijų prognozė jūroje reikalauja daugiakriterinio požiūrio, nes AIS duomenys yra nereguliariai atnaujinami ir sudėtingi. Daugiažingsnės (angl. *multi-step*) prognozės, kurių rezultatas – kelių laiko žingsnių iš eilės išskleistos koordinačių reikšmės, leidžia numatyti laivų judėjimą 20–25 minučių intervalu, būtinu norint atsižvelgti į didelių laivų stabdymo laiką. Tokio tipo modeliuose vienu metu naudojami požymiai: laivo greitis, kursas, geografinės koordinatės ir meteorologiniai rodikliai (pvz., vėjo greitis, bangų aukštis), nes vien tik absoliučios koordinačių vertės nebeužtikrina pakankamo tikslumo, ypač dinamiškoje aplinkoje, kurioje veikia tiek meteorologinės sąlygos, tiek įvairios povandeninės srovės ir pats jūrinis eismas. Tyrimuose pabrėžiama, kad daugiamačiai (angl. *multi-variate*) duomenys leidžia atskleisti sudėtingus priklausomybių sąryšius tarp laivo charakteristikų ir aplinkos veiksnių, gerinant prognozių patikimumą [57, 85].

RNN, ypač LSTM ir sklendžių rekurentinių vienetų (angl. *gated recurrent unit*, GRU) architektūros, tapo pagrindiniais įrankiais prognozuojant kelių žingsnių trajektorijas. LSTM sprendžia gradiento nykimo problemą ir geba išlaikyti ilgalaikes priklausomybes gana tiksliai apdorodamas sekų duomenis, o GRU supaprastina LSTM struktūrą, išlaikydamas aukštą našumą, bet mažesnį skaičiavimo laiką [18, 73]. Paprasti RNN modeliai dažnai neatlaiko ilgalaikių priklausomybių, todėl daugumoje tyrimų vertinamos ir lyginamos skirtingos RNN variacijos: paprastas RNN, LSTM, sluoksniuotas (angl. *stacked*) LSTM, dvikryptis LSTM ir LSTM AE [22, 63]. Mokslinėje literatūroje įrodyta, kad

LSTM autoenkoderiai efektyviai suspaudžia didelės apimties duomenis ir sumažina paklaidų kaupimąsi daugiažingsnėse prognozėse[12, 64].

Kategorinių duomenų, tokių kaip laivo tipas, integravimas į modelius taip pat yra svarbus tikslumo aspektas. Laivo tipai (transportinis, žvejybos, pagalbinis ir kt.) lemia skirtingą elgseną vandenyje, todėl duomenų rinkinį būtina papildyti kategorinėmis savybėmis. Dažniausiai taikomi kodavimo metodai yra fiktyvus kodavimas (angl. *one-hot*), etiketinis / ordinalus (angl. *label/ordinal*) ir įterptinis (angl. *embedding*) kodavimas [33, 47]. Tyrimai rodo, kad įterptinis kodavimas ypač naudingas, kai kategorijų skaičius didelis, nes leidžia sumažinti matmenų skaičių ir pagerinti modelio generalizavimą [24, 70]. Sistemingas skirtingų kodavimo metodų palyginimas laivų trajektorijų prognozėse dar nėra išsamiai atliktas, todėl šiame darbe nagrinėjamas ordinalus, fiktyvus ir įterptasis kodavimas, siekiant išsiaiškinti, kuris metodas geriausiai atspindi tarpkategorinius ryšius [21].

Susidūrimų rizikos vertinimas yra kitas būtinas jūrų saugos elementas. Įprastiniai artimiausio priartėjimo taško ir laiko iki artimiausio priartėjimo taško metodai nustato laivų priartėjimo taškus pagal pastovų saugų atstumą ar laiką, tačiau jie apskaičiuojami pagal esamą situaciją ir neįvertina trajektorijų prognozių neapibrėžtumo bei negali prisitaikyti prie dinamiškai kintančių situacijų [55, 108]. Dėl to plačiau analizuojamas tikimybinis vertinimas: prognozės intervalai (angl. *prediction interval*, PI), pasikliovimo intervalai (angl. *confidence intervals*, CI), elipsoidinės prognozės sritys (angl. *elipsoidal prediction region*, EPR) ir konforminės prognozės sritys (angl. *conformal prediction region*, CPR). Šios prognozių intervalais pagrįstos technikos suteikia galimybę apibūdinti galimų trajektorijų regionus, atsižvelgiant į modelio ir duomenų triukšmą [34, 83, 91]. EPR apibrėžia daugiamačius regionus, o CPR – neparametrines sritis, tinkamas negausiems duomenims, nes nereikalauja normalaus pasiskirstymo prielaidų [2, 13]. Šių sričių persidengimo analizė leidžia kiekybiškai įvertinti susidūrimo rizikos įvertį, kai kelios trajektorijos gali susikirsti vienu metu [20, 62].

Santraukos pabaigoje pažymėtina, kad nors daugiažingsnių daugiamačių modelių kūrimas ir tobulinimas gerokai progresavo, vis dar trūksta sistemingo kategorinių duomenų kodavimo ir skirtingų laivų tipų bei trajektorijų prognozių

neapibrėžtumo įtraukimo į generalizuotą modelį. Siekiant aukštesnio tikslumo, būtina toliau tirti skirtingų koordinačių transformacijų (pvz., WGS84 vs. UTM) poveikį prognozės stabilumui [3, 42], taip pat plėtoti tikimybinio susidūrimų rizikos vertinimo metodus, kurie veiktų realiuoju laiku ir prisitaikytų prie kintančių jūrų sąlygų. Šios kryptys atveria galimybes sukurti patikimesnes ir saugesnes navigacijos priemones jūrų transportui.

III. Metodų taikymas

Šiame skyriuje aprašomi pagrindiniai metodai, skirti sukurti laivų trajektorijų prognozavimo modeliams pagerinti gaunamiems pirminiams rezultatams ir pateikti susidūrimo rizikos vertinimo metodai. Pagrindiniai sprendimai yra:

- Rekursyvinė daugiažingsnė prognozė, paremta erdvinių koordinacinių taškų skirtumu – vektoriaus pokyčiu tarp ankstesnio ir einamojo taško (delta).
- Koordinačių transformacijos: WGS84, erdvinių koordinacinių taškų skirtumai Δ (lat / long), polinės (atstumas / kampas) ir Dekarto (UTM).
- Kategorinių duomenų (laivo tipo) kodavimo strategijos: etikečių / ordinalus, fiktyvus ir įterptinis.
- Prognozių neapibrėžtumo metodai: pasikliovimo intervalai (CI), prognozių intervalai (PI), elipsoidinės prognozės sritys (EPR) ir konforminės prognozės sritys (CPR).

Daugiažingsnė rekursyvinė prognozė

Daugiažingsnė prognozė leidžia išskleisti kelių taškų seką į priekį, o ne vien tik artimiausią poziciją. Pagrindinė idėja – prognozuoti vektorių skirtumus $(\Delta\text{lat}, \Delta\text{long})$ vietoje absoliučių koordinačių, o vėliau rekursyviai atkurti:

1. Tinklas mokomas pagal istorinius AIS duomenis: platuma, ilguma, greitis, kursas ir kt.

2. Išvestis – sekos vektorių skirtumai: $\Delta\text{lat}_t = \text{lat}_t - \text{lat}_{t-1}$, analogiškai $\Delta\text{long}_t = \text{long}_t - \text{long}_{t-1}$. Skirtumai gali būti skaičiuojami tiek UTM, tiek WGS84 projekcijose.
3. Rekursyviai sumuojant su paskutine faktine absoliučia padėtimi, atstatomos pirminės koordinatės.

Toks požiūris sumažina ilgalaikės prognozės klaidų kaupimąsi ir geriau atspindi trajektorijos nuokrypį [40].

Koordinačių transformacijos

Siekiant sumažinti geografinio mastelio iškraipymus ir pagerinti erdvinį aiškumą, bandyti keli transformavimo būdai:

- **WGS84 koordinatės** (AIS požymiai): absoliučios reikšmės (ilgumos, platumos duomenys).
- **Vektorių skirtumai** Δ (lat / long): erdvinių koordinacinių taškų skirtumai, kurie prognozuojami ir rekursyviai atstatomi į žinomą padėtį.
- **Atstumas ir kampas**: polinė koordinačių sistema, skaičiuojamas Haversinis atstumas d ir azimuto kampas θ nuo paskutinio taško.
- **UTM projekcija**: geografinės koordinatės konvertuojamos į Dekarto erdvę (x, y), leidžiančią tiesiogiai modeliuoti atstumus be geografinio mastelio iškraipymų.

Eksperimentai ir moksliniai tyrimai rodo, kad rekursyvus delta metodas ir UTM transformacija geriausiai sumažina MAEH ir kitas regresinių metrikų paklaidas [57, 85].

Prognozių klaidų vertinimo metrikos

Prognozių tikslumas vertinamas keliais regresijos metrikų rinkiniais, naudojant Šiaurės ir Baltijos jūros testavimo duomenų rinkinius. Čia pateikiamos trys pagrindinės metrikos:

- **Vidutinė kvadratinė paklaida** (angl. *mean squared error*, MSE):

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- **Vidutinė absoliutinė paklaida** (angl. *mean absolute error*, MAE):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

- **Vidutinė absoliutinė Haversinio atstumo paklaida** (angl. *mean absolute error of Haversine*, MAEH):

$$\mathrm{MAEH} = \frac{1}{n}\sum_{i=1}^{n}\left|\Lambda_i\left(y_i - \hat{y}_i\right)\right|$$

Formulėse n – imčių skaičius; y_i – tikroji reikšmė; $\hat{y}_i$ – prognozuojama reikšmė; Λ – Haversinė atstumo funkcija; i – daugiažingsnio laiko eilutė. Dėl erdvinio aiškumo MAEH pasirinkta kaip pagrindinė metrika, matuojanti vidutinį trajektorijos nuokrypį metrais arba kilometrais (SI sistemos vienetais).

Kategorinių duomenų kodavimo būdai

Dirbtiniai gilieji neuroniniai tinklai yra tam tikros struktūros matematinės funkcijos, galinčios apdoroti didžiulius duomenų kiekius, kurių reikšmės privalo būti diskrečios. Vis dėlto, daugumoje duomenų šaltinių dažnai susiduriama su dviejų tipų kintamaisiais:

- diskretūs kintamieji – dažniausiai tai yra sveikieji arba dešimtainiai skaičiai ir turi begalinį skaičių galimų reikšmių. Tokiomis reikšmėmis laikomi duomenys apie laivo požymius: greitis, judėjimo kryptis, ilgumos ir platumos koordinatės ir kt.;
- kategoriniai kintamieji – tai atskiri kintamieji, naudojami duomenims skaidyti pagal tam tikras charakteristikas. Šioms reikšmėms priskiriami skirtingi laivo tipai: krovininiai, keleiviniai, kariniai, žvejybos ir pan.

Kuriant giliojo mašininio mokymo modelius, reikalinga transformuoti kategorinius duomenis, kad būtų galima pritaikyti RNN tinklo algoritmus. Laivo tipai (transportinis, žvejybos, pagalbinis ir kt.) pateikiami ASCII tekstu. Kad būtų panaudoti neuroniniame tinkle, taikomi trys kodavimo būdai:

1. **Etikečių / ordinalus kodavimas**: tai procesas, kurio metu kiekvienam tipui priskiriamas unikalus sveikasis skaičius. Nesuteikia žinių apie tarpkategorinius panašumus, tačiau paprastas ir greitas. Ordinalus reiškia, kad svarbus kategorijų eiliškumas ir svarbu išlaikyti vieningą tvarką;

2. **Fiktyvus kodavimas**: tai procesas užkoduojantis kategorijas dvejetainėje sistemoje ir sukuriantis naujus požymius, priskiriant atitinkamai kategorijai vienetą, o kitoms nulį. Statistikoje, tai suprantama, kaip fiktyvių kintamųjų sukūrimas. Šios technikos trūkumas, kad kiekvienai kategorijai sukuriamas atskiras požymis, todėl turint labai daug elementų – labai išsiplečia duomenų struktūros matrica (padidėja ir tinklo parametrų skaičius). Kiekvienam laivo tipui sukuriama K–matė vektoriaus reprezentacija, kur K – kategorijų (laivo tipų) skaičius. Vienas elementas, atitinkantis laivo tipą, žymimas 1, kiti – 0. Tinkamas, kai kategorijų nedaug, nes kitaip didėja matmenys.

3. **Įterptinis kodavimas**: tai procesas, mažinantis duomenų matmenis ir išsaugantis semantinius kategorijų ryšius. Teigiamus sveikuosius skaičius paverčia fiksuoto dydžio vektoriais. Kiekvienam tipui priskiriamas vektorius d. Modelio mokymosi metu šie vektoriai optimizuojami, leidžiant atrasti kontekstinius panašumus tarp tipų. Dažniausiai pasirenkamas $d = 2$ arba 3, atsižvelgiant į duomenų kiekį.

Eksperimentai rodo, kad įterptinis kodavimas sumažina MAEH paklaidą ir geriau sugeba atspindėti tarpkategorinius ryšius, ypač kai laivo tipų įvairovė didelė [21, 33, 47]. Nors panašūs rezultatai gauti ir su fiktyviu kodavimu, tačiau esant dideliam kardinalumui jis yra mažiau tinkamas.

Prognozių neapibrėžtumo ir sričių nustatymo metodai

Norint patikimai įvertinti laivų susidūrimų riziką bei pačių prognozių tikslumą, būtina ne tik atlikti trajektorijos prognozę, bet ir nustatyti jos neapibrėžtumo

ribas. Pagrindiniai taikyti metodai:

- **Pasikliovimo intervalai (CI)**: $W_{\text{CI}} = t_{\frac{1+\alpha}{2},n-1} \cdot \frac{SD}{\sqrt{n}}$, kur SD – standartinis nuokrypis, n – imties dydis, t – Studento pasiskirstymo kritinė reikšmė (angl. *t-Score*), atitinkanti pageidaujamą pasikliautinąjį lygį α.
- **Prognozių intervalai (PI)**: $W_{\text{PI}} = t_{\frac{1+\alpha}{2},n-1} \cdot SD\sqrt{1+\frac{1}{n}}$, čia intervalas apima tiek modelio, tiek stebėjimo triukšmą, kur $n-1$ laisvės laipsnis, n – imties dydis. Išraiška $\sqrt{1+\frac{1}{n}}$ atsižvelgia į papildomą neapibrėžtumą, kai prognozuojamas vienas būsimas stebėjimas, o ne vertinamas populiacijos vidurkis.
- **Elipsoidinės prognozės sritys (EPR)**: matricinė forma, paremtas Mahalanobis atstumu (angl. *Mahalanobis distance*) nuo taško x iki vidurkio μ:

  $$EPR = \{x : (x-\mu)^{\mathsf{T}}\Sigma^{-1}(x-\mu) \leq \chi^2_{p,\alpha}\},$$

  kur μ – vektoriaus vidurkis, Σ – kovariacijos matrica, $\chi^2_{p,\alpha}$ – χ^2 – skirstinio slenkstis, $p = 2$ nurodo, kad turime dvimatę erdvę. EPR apibrėžia elipsę tolimesniems laiko žingsniams, įvertindama abiejų koordinačių tarpusavio koreliaciją [34]. Sąlyga $\leq \chi^2_{p,\alpha}$ nusako, ar taškas x patenka į prognozavimo sritį (elipsę).
- **Konforminės prognozės sritys (CPR)**: naudojama kalibravimo aibė C ir neatitikimo matas kiekviename laiko žingsnyje $N_t = \|y_t - \hat{y}_t\|_2$. , kur y_t yra faktinės kelių dimensijų koordinatės, $\hat{y}_t$ atitinkamos prognozės. Neatitikimas apskaičiuojamas kaip Euklido atstumas, apimantis visų dimensijų paklaidas. Atsižvelgiant į didėjančius prognozės neapibrėžtumus ilgesniuose horizontuose, neatitikimai neapjungiami visai sekai, bet vidurkinami kiekviename laiko žingsnyje, gaunant skirtingus slenksčius $\bar{R}_t$. Siekdami suformuoti galutinę prognozės sritį su pasirinktu pasikliovimo lygiu $1-\delta$, optimizuojami svoriai $\{\alpha_t\}$, pritaikytus kiekvieno žingsnio vidutiniams neatitikimams $\bar{R}_t$. Ši optimizacija leidžia apskaičiuoti bendrą slenkstį pagal kvantilio kriterijų, kuris apibrėžia kiekvieno taško atstumą nuo prognozės centro laikom momente t. Gautos sritys atitinka pasirinktą

pasikliovimo lygį be prielaidų apie tikslų duomenų pasiskirstymą [6, 83, 91].

Metodai leidžia apskaičiuoti dinamiškas prognozuojamo laivo trajektorijos ribas, atsižvelgiant į neapibrėžtumą kiekviename laiko žingsnyje. Kai dviejų laivų prognozės sritys (apskritimai arba elipsės) persidengia, galima įvertinti susidūrimo tikimybę. Tam naudojamas Žakardo indeksu (angl. *Jaccard index*) paremtas susidūrimo rizikos rodiklis (angl. *collision risk score*), kuris apskaičiuojamas pagal prognozės sričių sankirtos ir bendro ploto santykį:

$$P(\text{susidūrimas}) = \frac{V_{A\cap B}}{V_A + V_B - V_{A\cap B}},$$

kur V_A ir V_B yra atskirų regionų plotai, o $V_{A\cap B}$ yra jų susikirtimo plotas. Plotai apskaičiuojami naudojant UTM projekciją.

IV. Duomenų ir eksperimento darbo eiga

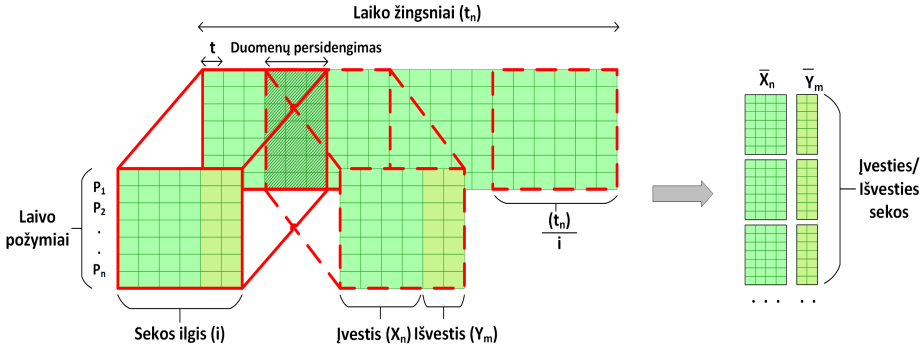
Eksperimentuose naudoti AIS ir meteorologiniai duomenys iš dviejų regionų:

- **Nyderlandų regionas (Šiaurės jūra)**: pirminis šaltinis – *„Ship Finder“* (istoriniai AIS duomenys), surinkti 2018 m. rugsėjo–2019 m. vasario laikotarpiu. Vėliausiai atliekamoms testinėms prognozėms naudota 5 mėn. įrašų, iš viso $\sim$21 mln. eilučių (naudoti tik krovininiai laivai). Duomenų valymo etapu pašalinti dubliai, įrašai su trūkstamais signalais, neteisingais MMSI ar greičio reikšmėmis. Dėl AIS stebėjimų nevienodumo laiko intervalų eilutėse taikytas duomenų dažnio pertvarkymas: pritaikytas logaritminis mastelio standartizavimas (FLOG), o tada įprastas $[0,1]$ normalizavimas, kad būtų sušvelninamas skirtumas tarp laiko eilučių neatitikimų (laiko žingsniai nuo 1 min iki kelių šimtų min.). Regiono koordinatės: platuma $51^\circ 39'$–$52^\circ 12'$, ilguma $3^\circ 10'$–$4^\circ 50'$.
- **Baltijos jūros regionas (prie Bornholmo)**: AIS duomenys iš Danijos jūrų administracijos apėmė 2021 m. birželio–gruodžio laikotarpį, eksportuojami CSV formatu per https://web.ais.dk/aisdata/. Pradinėse eksperimentinėse stadijose naudoti tik krovininių laivų duomenys, vėliau

įtraukti visi to regiono laivų tipai. Dėl gausybės įrašų (∼100 mln.) ir dažnų intervalų (<1 min.) taikyta laiko eilučių pakartotinė atranka (angl. *resampling*) kas 1 min.: kiekvienam laivui vidurkinami k artimiausi k–NN taškai, todėl kiekvieną minutę gaunamas vienas įrašas. Regiono koordinatės: ilguma 12°–15° E, platuma 54°–56° N.

- **Meteorologiniai duomenys**: gauti naudojant Weatherbit.io API (sujungiant NOAA ISD, MADIS, GHCN ir papildomus palydovinius bei klimatologinius šaltinius). Gauti duomenys yra Baltijos jūros regione, suformuoti stačiakampio tinklelio (angl. *grid*) schema, o kiekvienam AIS įrašui priskiriamas artimiausio meteorologinio tinklelio taško stebėjimas. Stebimų parametrų sąrašas: kasvalandinis oro temperatūros vidurkis, vėjo kryptis ir greitis, bangų aukštis ir kryptis, bangavimo bangos aukštis ir kryptis, debesuotumas, matomumas, vandens temperatūra, atmosferos slėgis ir kiti jūros meteorologiniai rodikliai.

Sekų generavimo tikslas – paversti AIS stebėjimus į atitinkamas, vienodo ilgio, laiko eilučių sekas, kurios naudojamos RNN tinklams. Laivų srautas buvo surūšiuotas pagal laiką ir MMSI bei suskirstytas į lygias 50 laiko žingsnių ilgio sekas. Atlikta normalizacija, o seka paskirstyta į mokymo, validavimo ir testavimo imtis, jau taikant straipsnyje aprašytas transformacijas [40]. Iš visų ~11450000 stebėjimų (eilučių) sugeneruota ~460 000 sekų, iš kurių 30 stebėjimų yra įvesčiai, o 20 – išvesčiai (prognozei). Atsižvelgiant į įprastą jūrų operacijų praktiką, 30 istorinių AIS stebėjimų seka gerai atspindi laivo trajektoriją, o 20 būsimų stebėjimų prognozė atitinka numatomą sustojimo ir reagavimo laiką, reikalingą veiksmingai vertinti susidūrimo riziką ir atlikti prevencinius manevrus. Vidutinis laiko žingsnių skirtumas tarp stebėjimų yra apie 1 min. Vidutinis laivo pajudamas atstumas – 340 metrų, todėl vienoje sekoje laivas įveikia vidutiniškai apie 16 km (iš jų ~10 km įvesties ir ~6 km išvesties). Statistiniai įverčiai pateikti 1 lentelėje.

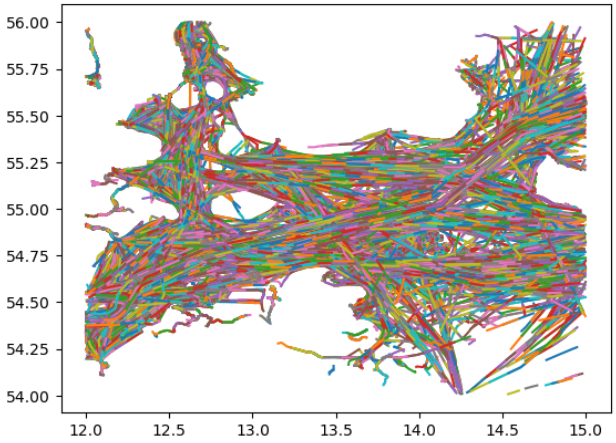


1 pav.: Laivų duomenų skaidymas į sekas.

1 lentelė: Sekų duomenų charakteristikos

Savybė	**Reikšmė**
Sekos bendras ilgis	50 laiko žingsnių
Įvesties trajektorijos ilgis	30 laiko žingsnių
Išvesties trajektorijos ilgis	20 laiko žingsnių
Vidutinis laiko žingsnio pokytis	~60 s
Vidutinis erdvinis pokytis	~320 m
Vidutinis trajektorijos ilgis	~16 km
Sugeneruotų sekų skaičius	943 584

Sekų generavimo procesas iliustruojamas 1 pav., kur laivo savybės (F_n) suskirstomos į fiksuoto ilgio sekas (i), naudojant slenkančio lango metodą su persidengiančiais laiko žingsniais (t_n). Šis metodas užtikrina, kad kiekviena nauja seka išsaugo ankstesnius stebėjimus, kurie yra būtini laiko eilutėms prognozuoti rekurentiniuose tinkluose. Persidengimo strategija padeda išlaikyti laiko tęstinumą, padalijant kiekvieną seką per pusę, leidžiant modeliui fiksuoti priklausomybes laiko atžvilgiu. Kiekviena seka dar labiau skaidoma į įvesties (X_n, F_n) ir išvesties (Y_n, F_n) matricas, kurios atitinka savybių ir etikečių poras prižiūrimojo regresinio mokymo kontekste. Įvesties požymiai perduodami į tinklą, kuris prognozuoja išvesties savybes. Visos giliųjų rekurentinių tinklų architektūros šiame tyrime naudoja dvimatę išvestį, todėl matrica prieš mokymą ištiesinama (angl. *flatten*), kad būtų suderinama su galutiniu prognozavimo sluoksniu. Pateiktame 2 pav. yra sugeneruotų laivo judėjimo sekų pavyzdys



2 pav.: Sukurtos laivų judėjimo Baltijos jūros regione sekos, kur kiekviena spalva vaizduoja atskiro laivo judėjimo trajektoriją per tam tikrą laiką.

Baltijos jūros regione, vizualiai demonstruojantis tyrime naudojamas gautas trajektorijas.

Tyrimo metu sekos buvo sudarytos iš 50 laiko žingsnių, padalytos į 30 įvesties žingsnių ir 20 išvesties žingsnių, vidutinio 1 minutės laiko tarpo sąlygomis. Tinkle naudojami duomenys apėmė tokias savybes kaip užkoduotas laivo tipas, geografinės koordinatės, greitis, kursas bei gretimų laiko žingsnių platumos ir ilgumos skirtumai. Išvesties seka susidėjo tik iš koordinačių savybių, o įvesties seka – iš minėtų laivo charakteristikų. Šios savybės padeda suprasti laivo elgseną ir prognozuoti tolimesnius trajektorijos taškus. Architektūroje naudoti parametrai pateikti 2 lentelėje.

Duomenų rinkinys, apimantis beveik milijoną sekų, padalytas 70:15:15 santykiu mokymui, validavimui ir testavimui, kas yra įprasta mokslo tyrimuose [5, 79], siekiant užtikrinti subalansuotą vertinimą ir generalizaciją. Prieš mokymą sugeneruotos ir normalizuotos sekų matricos buvo sumaišytos (angl. *shuffle*), kad būtų sumažinta sekų priklausomybė, dispersijos poveikis ir pagerintas gradientų kintamumas. Matricų maišymas užtikrina, kad kiekviena nauja seka modelyje būtų nepriklausoma nuo ankstesnės sekos, taip sumažinant dispersiją ir padidinant gradientų kintamumą. Ši atsitiktinė tvarka leidžia modeliui susidurti su įvairiomis situacijomis, prisidedant prie geresnės generalizacijos.

2 lentelė: LSTM AE konfigūracija.

Parametras	Reikšmė	Pastaba
Sluoksniai	3	Iš viso kodavimo ir dekodavimo dalių
Sekos ilgis	50	30 įvestyje, 20 išvestyje
Epochos	100	—
Optimizatorius	0,001	Adam (su nurodytu mokymosi greičiu)
Reguliarizavimas	0,01	Dropout sluoksniai
Ląstelių skaičius	275	Vienetų skaičius kiekviename LSTM sluoksnyje
Partijos dydis	128	Vienos iteracijos metu naudoti pavyzdžiai
Modelių kiekis	20	Pagal tuos pačius duomenis apmokyti modeliai
Nuostolių funkcija	MSE	Prognozavimo kokybės priemonės
Aktyvavimo funkcija	ReLU	Naudojama LSTM vartuose ir išvesties sluoksniuose

Normalizacija (E) taikyta visoms skaitinėms savybėms, transformuojant reikšmes tarp 0 ir 1, siekiant užtikrinti savybių vienodumą ir pagreitinti mokymosi procesą. Mokymo imtis (70 %) naudota modeliams apmokyti, validavimo imtis (15 %) – mokymosi pažangai stebėti ir persimokymo prevencijai, o testavimo imtis (15 %) – galutiniam modelių veikimui įvertinti. Svarbu paminėti, kad tyrime nebuvo naudojamas vienas modelis, o modelių rinkinys, apmokytas su tais pačiais duomenimis. Kiekvienas modelis, veikiamas tokių veiksnių kaip reguliarizacija ir žinių pradinių reikšmių parinkimas, generuoja šiek tiek skirtingas prognozes, leidžiančias išsamiau įvertinti neapibrėžtumą ir stiprinti taikomų susidūrimo aptikimo metodų patikimumą. Pateikiama normalizavimo funkcija:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

kur:

- X – yra laivo savybė,

- X_{norm} – yra normalizuota reikšmė,
- X_{max} ir X_{min} – atitinkamai didžiausia ir mažiausia tos savybės reikšmės.

Eksperimentuose buvo keičiamas architektūrų tipas ir ląstelių skaičius. Tie patys tinklo parametrai ir ląstelių dydžio intervalas užtikrino vienodas sąlygas visoms architektūroms ir vienodą duomenų tvarką mokymui. Analizuotoje literatūroje rekurentiniai neuroniniai tinklai dažniausiai naudoja ląstelių skaičių nuo keliolikos iki kelių šimtų. Iš pradžių šiame tyrime taikytas šis bandymas, vertinant ląstelių skaičių palaipsniui nuo 25 iki 300, didinant po 25 vienetus. Tačiau siekiant griežčiau patikrinti ląstelių skaičiaus įtaką ir įsitikinti, ar prognozės tikslumas gerėja ar stabilizuojasi už įprasto intervalo ribų, tolesniuose eksperimentuose išplėstas diapazonas iki 5000 ląstelių, taikant didesnius padidėjimo žingsnius. Kiekviena ląstelių skaičiaus konfigūracija skirtingose architektūrose, įskaitant eksperimentus su įterptojo kodavimo metodais, buvo apmokoma 10 kartų, siekiant užtikrinti prognozių patikimumą ir stabilumą. Vėlesniuose eksperimentų etapuose, apimančiuose koordinačių sistemų transformacijas, neapibrėžtumo kiekybinį įvertinimą ir susidūrimo aptikimą, prireikė įtraukti papildomų modelių apmokymą, todėl iš viso gautas platus apmokytų konfigūracijų rinkinys. Šis išsamus eksperimentinis metodas leido užtikrintai patvirtinti modelių jautrumą tinklo ląstelių hiperparametrams įvairiomis sąlygomis.

Be to, eksperimentų metu išbandyti įvairūs hiperparametrų pakeitimai: aktyvacijos funkcijų keitimas (ReLU vs. TanH), L2 ir dropout reguliarizacijos technikos, siekiant apsaugoti regresinius modelius nuo persimokinimo. Vis dėlto, geriausių rezultatų pasiekta praktiškai išbandžius su 2 lentelėje nurodytais parametrais.

Eksperimentuose naudotas meteorologinių duomenų rinkinys papildė AIS duomenis, įskaitant vėjo kryptį, vėjo greitį, temperatūrą (nakties ir dienos vidurkius), vidutinį jūros lygio slėgį, drėgmę, matomumą, debesuotumą, mėnulio fazę bei kitus panašius stebėjimo duomenis. Šie duomenys pateikti kaip kasdienės laiko eilutės, kuriose kiekvieną valandą fiksuoti agreguoti orų pokyčiai. Atlikti permutacijų eksperimentai, keičiant požymių eiliškumą tarpusavyje ir vertinant jų poveikį modelio tikslumui. Jie atskleidė, kad

svarbiausi AIS požymiai laivų trajektorijų prognozei yra kursas, greitis, platuma ir ilguma.

Laivų susidūrimo rizikos rodiklis vertinamas taikant intervalinius metodus: EPR, CPR, PI ir CI. Šie metodai naudojami prognozuojamų sričių plotams įvertinti ir aprėpties tikimybei apskaičiuoti, siekiant užtikrinti, kad tikroji laivo trajektorija kiekvienu laiko momentu patektų į prognozuotą sritį. Šių metodų veiksmingumas taip pat patvirtintas skaičiuojant susidūrimo tikimybę realių jūrinių incidentų kontekste. Visi taikyti metodai atitiko 95 % pasikliovimo lygį, užtikrinantį patikimas ir tvirtas prognozes.

Giliojo mokymosi modeliai sudarė pagrindą laivų trajektorijų prognozavimui, kurie vėliau naudojami susidūrimo rizikai vertinti (Žakardo indeksu). Kiekvienas modelis generavo tikimybinę laivo būsimų padėčių prognozę, įvertindamas su jūrine navigacija susijusį neapibrėžtumą. Prognozuotos trajektorijos toliau analizuojamos pasitelkiant statistinius metodus, tokius kaip prognozavimo intervalai, EPR ir CPR, siekiant nustatyti galimų susidūrimo zonų ribas. Susidūrimo tikimybė vertinama lyginant prognozuotas ribas su faktinėmis laivų padėtimis, kai sritys tam tikru momentu persidengia. Šis integruotas požiūris užtikrina, kad susidūrimo rizikos vertinimas būtų glaudžiai susijęs su trajektorijų prognozavimo tikslumu, suteikdamas dinamišką ir tikimybinį galimų susidūrimo scenarijų įvertinimą.

V. Rezultatai

Šiame skyriuje pristatomi eksperimentų rezultatai, kuriais siekiama pasiūlyti ir pagerinti laivų trajektorijų prognozę bei įvertinti susidūrimų rizikas. Iš pradžių, naudojant Nyderlandų regiono duomenis, palyginti keli RNN modeliai (LSTM AE, dvipusis LSTM, GRU) ir nustatytos optimalios LSTM ląstelių konfigūracijos. Toliau šie modeliai testuoti tiek su Nyderlandų, tiek su Baltijos jūros duomenimis, taikant koordinačių transformacijas ir rekursinį prognozavimo modelį. Galiausiai LSTM AE modelis išplėstas integruojant kategorinius laivo tipų ir meteorologinius duomenis, kurie panaudoti prognozių riboms nustatyti ir galimai susidūrimo rizikai vertinti.

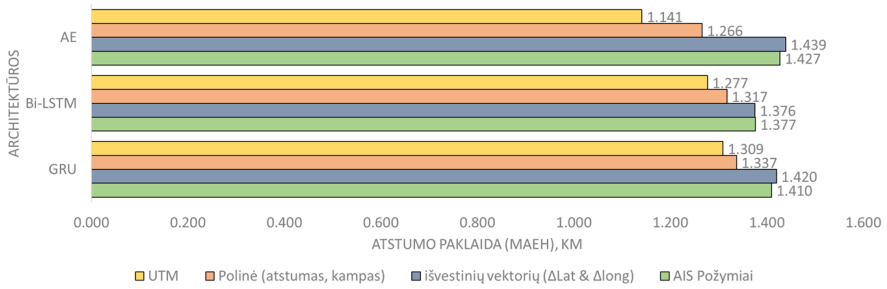
3 lentelė: MSE paklaidų įverčiai su skirtingais ląstelių hiperparametrų rinkiniais.

Ląstelės	RNN $\times 10^{-4}$	LSTM $\times 10^{-4}$	LSTM Stack $\times 10^{-4}$	GRU $\times 10^{-4}$	AE $\times 10^{-4}$	BiLSTM $\times 10^{-4}$
25	6,910	5,181	4,434	5,563	4,581	6,582
50	5,334	4,381	4,361	4,620	3,944	3,939
75	4,904	4,211	4,146	4,290	4,010	4,177
100	4,819	3,947	4,150	4,021	4,031	3,697
125	4,702	4,039	4,075	3,914	**3,688**	3,721
150	4,739	**3,875**	4,240	3,916	3,996	3,683
175	4,635	3,920	4,067	3,914	3,789	3,822
200	**4,628**	3,999	4,564	3,826	3,706	3,900
225	4,679	4,024	4,482	3,819	3,987	3,636
250	4,630	4,085	4,133	3,742	3,946	3,736
275	4,660	3,998	4,099	3,827	3,724	3,645
300	4,681	4,086	**3,956**	**3,710**	3,782	**3,618**

Skirtingų RNN architektūrų vertinimas Nyderlandų regione

Modeliai Nyderlandų regione su krovininių laivų tipais įvertinti testavimo imtimis po šešių giliųjų RNN architektūrų mokymosi. MSE nuostolių funkcija apskaičiuota kaip visų eksperimentų vidurkis ir normalizuota pagal modelių skaičių. Pateikta 3 lentelė atvaizduoja kiekvienos architektūros prognozių tikslumą, kur mažesnė MSE reikšmė reiškia didesnį tikslumą. Lentelėje paryškinti mažiausi klaidų lygiai (optimaliems ląstelių dydžiams): Nyderlandų regione geriausiai pasirodė dvipusis LSTM, jam nedaug nusileido AE ir GRU modeliai. AE ir dvipusis LSTM išlaikė stabiliausią klaidų pasiskirstymą, o dvipusis LSTM rodė didžiausią skirtumą tarp minimalių ir maksimalių klaidų. Visose architektūrose daugiau kaip 100 ląstelių nuosekliai mažino paklaidą, rodydamos, kad tinklo talpa daro didelę įtaką tikslumui. Tačiau padidinus ląstelių skaičių žymiai išauga skaičiavimo kaštai, ypač esant itin didelei ląstelių apimčiai (pvz., 5000 ląstelių): mokymo laikas gerokai pailgėja be proporcingo tikslumo pagerėjimo, todėl optimalių hiperparametrų parinkimas yra itin svarbus.

Modeliai po 10 kartų apmokyti kiekviena architektūra, ląstelių skaičiumi ir koordinačių transformacijos strategija, o pateikti rezultatai (žr. 3 pav.) – tai vidutinės atstumo paklaidos (MAEH) rodikliai. Taigi pateikiami trijų atrinktų



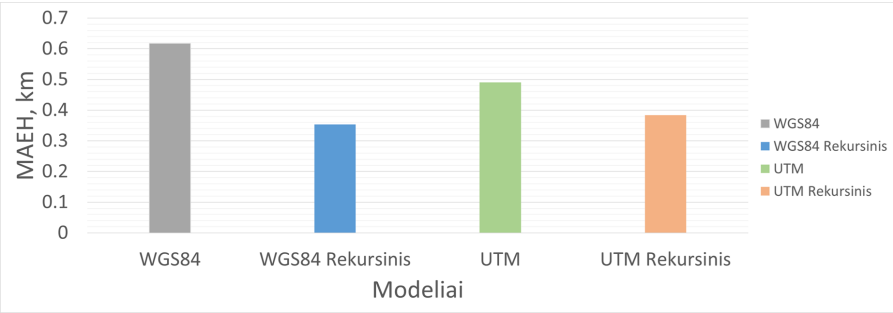
3 pav.: Mažiausi eksperimento atstumo paklaidos rezultatai.

modelių (LSTM AE, dvipusis LSTM ir GRU) vertinimai Nyderlandų regione, taikant UTM, polinių koordinačių (atstumas ir kampas), išvestinių vektorių skirtumo (δplatumos / δilgumos) bei įprastus AIS (WGS84 koordinačių) požymius. UTM transformacija suteikė didžiausią naudą – AE minimalus vidutinis paklaidos lygis buvo 1,141 km, tai beveik 30 % geriau nei prognozuojant pagal įprastus AIS duomenis. Atkreiptinas dėmesys, kad UTM šiuo atveju jau taikė rekursijos būdu perskaičiuojamus prognozių skirtumus. Polinių koordinačių metodas irgi pagerino tikslumą, tačiau rekursinis vektorių skirtumų perskaičiavimas, kai prognozuojamos delta vertės imamos iš paskutinio žinomo taško, labiausiai prisidėjo prie paklaidų mažinimo. Rekursinis atnaujinimas iš vektorių skirtumo davė tikslesnes prognozes nei absoliučių koordinačių tiesioginis spėjimas, o LSTM AE architektūra šiuo atveju buvo tiksliausia.

Laivų prognozių tikslumo vertinimas Baltijos jūros regione

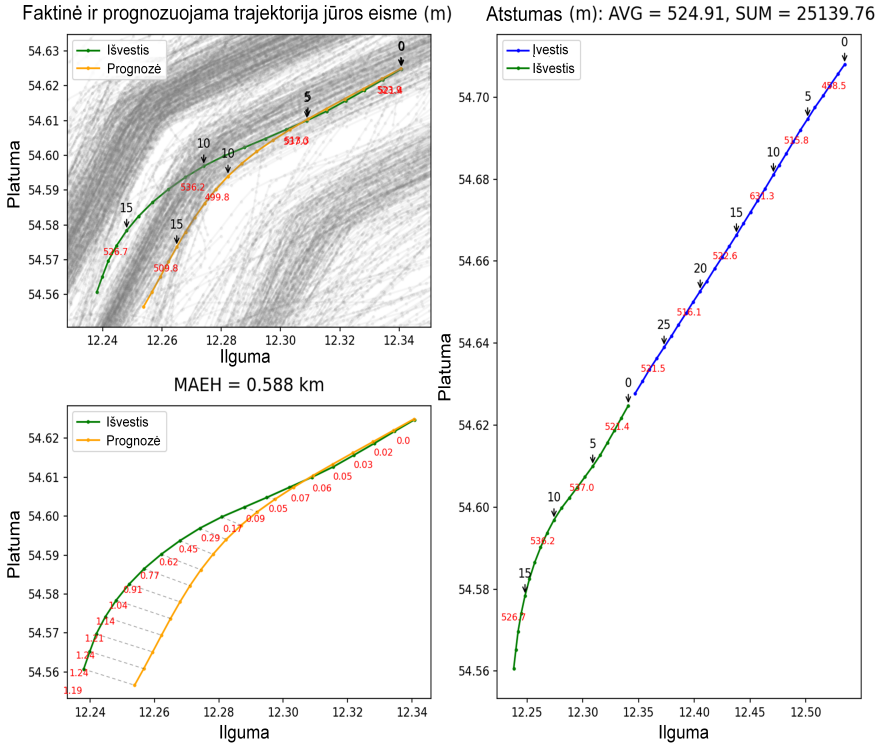
Pateiktas 4 paveikslas atvaizduoja skirtingų koordinačių transformavimo metodų (WGS84 ir UTM) ir jų rekursinių prognozių, pagrįstų vektorių skirtumais, rezultatus. MAEH reikšmės rodo, kad prognozuotas padėties rekursyvus perskaičiavimas pagal delta koordinates reikšmingai pagerina prognozių tikslumus abiejose koordinačių sistemose. Konkrečiai, rekursinis WGS84 metodas pasiekė mažiausią prognozės klaidą (0,3537 km), demonstruodamas geresnį tikslumą, palyginti su standartiniais prognozavimo būdais be rekursinio perskaičiavimo. Tai išryškina delta metodo pranašumą trajektorijų prognozėse prieš absoliučių reikšmių naudojimą. Be to, meteorologinių duomenų (pvz., vėjo, bangų, temperatūros sąlygų) integravimas dar labiau sumažino MAEH iki

0,3203 km, pabrėžiant aplinkos sąlygų pridėtinę prognozavimo vertę laivų trajektorijų modeliavimo procese.



4 pav.: Mažiausi eksperimento atstumo (MAEH) paklaidos rezultatai.

Atsitiktinai parinkta seka panaudota galutiniam modelio įvertinimui, o rezultatai pateikti 5 paveiksle. Vizualizacijoje pavaizduotas geriausiai veikiantis modelis su optimaliu ląstelės dydžiu. Mėlyna linija atspindi įvesties seką (30 laiko žingsnių), žalioji linija žymi faktinę laivo trajektoriją (20 laiko žingsnių). Geltona linija iliustruoja prognozuotą laivo judėjimą tame pačiame laikotarpyje, o pilkos linijos žymi aplinkinį eismo intensyvumą iš kitų sekų. Rezultatai rodo, kad trumpojo laikotarpio prognozė (apie 10 minučių) glaudžiai atitinka faktinę trajektoriją. O ilgojo laikotarpio prognozei tikslumas šiek tiek sumažėja, tačiau tai būdinga ilgalaikėms prognozėms. Be to, atstumo paklaida metrais pateikiama raudona spalva.

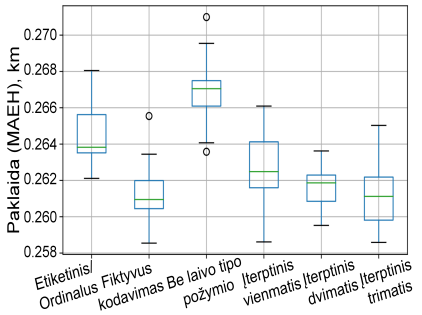


5 pav.: Faktinis ir prognozuojamas laivo judėjimas atsitiktinio eismo sąlygomis.

Kategorinių laivo tipų integracijos rezultatai

Eksperimento rezultatai, susiję su trajektorijų prognozavimo tikslumu, pavaizduoti 6 paveiksle. Į duomenų rinkinį įtraukus skirtingus laivų tipus ir juos užkodavus ordinaliu (etiketiniu), fiktyviu, įterptuoju sluoksniu bei rezultatus palyginus nenaudojant laivo tipų kodavimo, prognozės tapo stabilesnės dėl sumažėjusios klaidų dispersijos. Tiek fiktyvus kodavimas, tiek įterptasis kodavimas padidino tikslumą. Įterptojo kodavimo sluoksnis su dvimate struktūra pasižymėjo mažiausia dispersija. 7 lentelėje pateikti išsamūs šių eksperimentų MAEH rezultatai. Vidutiniškai laivo tipų kodavimo įtraukimas pagerino prognozavimo tikslumą maždaug 0,0027 km, palyginti su modeliu be kategorinės laivo tipo informacijos. Šie duomenys atskleidžia, kad tinkamos kodavimo metodikos yra svarbios gerinant modelio ir trajektorijų prognozių

patikimumą. Kodavimo būdas leidžia sukurti generalizuotą modelį skirtingiems laivo tipams.



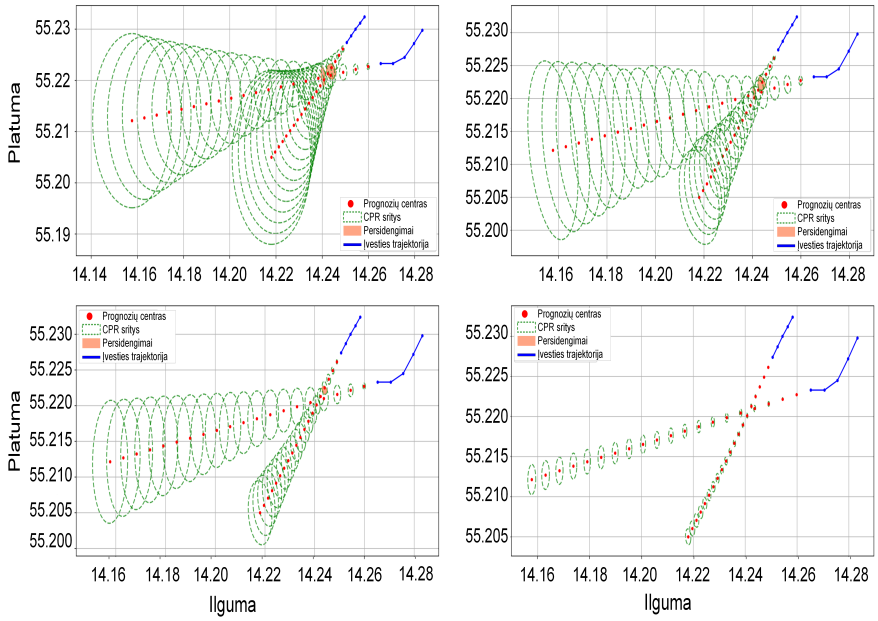
6 pav.: MAEH metrikos pakartotinio eksperimento rezultatai naudojant skirtingus kategorinių duomenų kodavimo būdus.

3 lentelė: Daugiamačio įterptojo kodavimo būdo vidutinės paklaidos.

Metrika	Vienmatis	Dvimatis	Trimatis
MAEH	0,26252	0,26162	0,26124
MSE	2,51E-05	2,49E-05	2,47E-05
RMSE	0,00488	0,00485	0,00483
MAE	0,002046	0,002039	0,002036
MAPE	0,000110	0,000110	0,000109

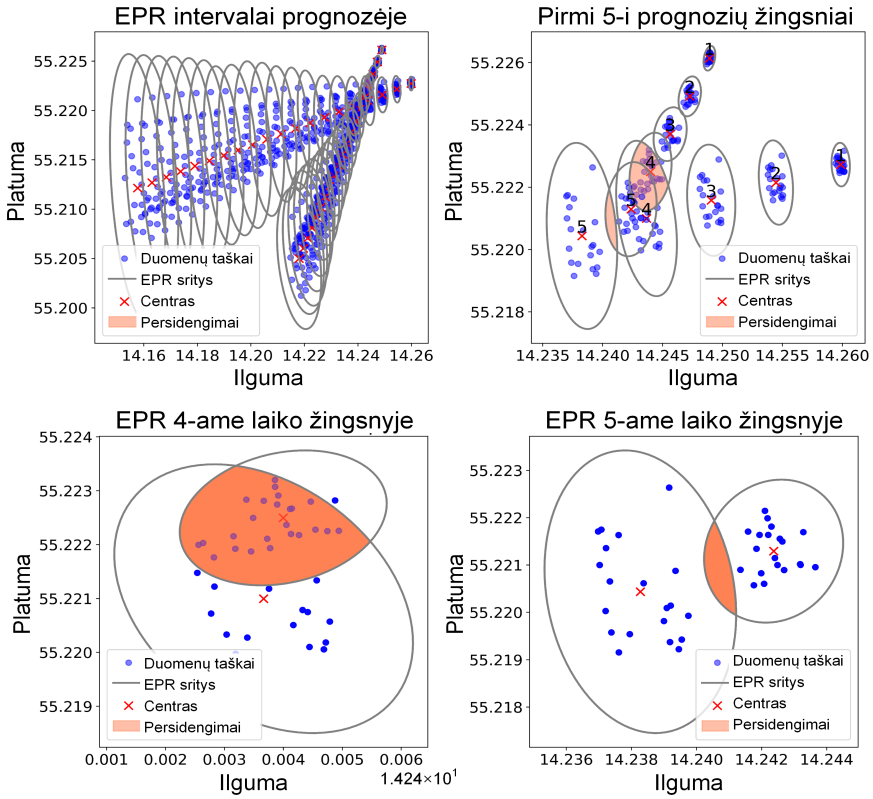
Susidūrimų vertinimo analizė

Vertinant modelio veikimą Baltijos jūros regione, LSTM AE apmokytas kelioms modelių kopijoms, o testavimo prognozės atliktos su pasirinkta duomenų pogrupio imtimi, į kurią įtrauktas ir realus „Scot Carrier“ ir „Karin Höj“ susidūrimo atvejis. Gautų prognozių neapibrėžtumui vertinti taikyti EPR, CPR, PI ir CI metodai, kiekvienam prognozuotam laiko žingsniui nustatomos galimos laivo vietos sritys. EPR zonos vaizduojamos elipsėmis, o CPR (bei PI, CI) zonos, apskritimais, kai nurodomas tik spindulys (žr. 8 pav.). Šių zonų persidengimai tarp skirtingų laivų leidžia kiekybiškai įvertinti susidūrimo rizikos balą.



8 pav.: Skirtingų metodų ir jų ribų pločio palyginimas jūrų avarijos atveju.

Pavaizduotame 9 paveiksle kaip vienas iš pavyzdžių pateikiamas EPR metodas. Pirmoje dalyje elipsės žymi EPR ribas keliems prognozuotiems žingsniams – mėlyni ir raudoni taškai rodo faktines laivų padėtis, o elipsės – prognozuotas zonas. Antroje dalyje („Pirmieji 5 žingsniai“) matyti elipsių susikirtimai su didesne susidūrimo tikimybe. Trečioje ir ketvirtoje dalyse atskiri EPR vaizdai konkrečiu laiko žingsniu parodo laivų prognozuojamas judėjimo sritis. UTM projekcija leidžia apskaičiuoti elipsių plotus kvadratiniais metrais ir paskaičiuoti regionų plotų persidengimus.



9 pav.: EPR formos prognozavimo sritys. Prie raudonų kryželių esantys skaičiai žymi atitinkamus prognozavimo laiko žingsnius.

„Scot Carrier“ ir „Karin Höj“ susidūrimo atvejis pagal CPR metodiką prognozuotas su didžiausiu 39 % tikimybės įverčiu, parodant, kad tikimybiniai sprendimai (skaičiuojant trajektorijų ribų persidengimus) pranašesni už CPA / TCPA, kurie remiasi fiksuotomis slenkstinėmis vieno taško reikšmėmis. Siūlomi metodai prognozuoja dinamiškas trajektorijų ribas ir vertina jų persidengimą, o ne vieną tašką, leidžiant įtraukti kelių modelių prognozes. Pagrindinis privalumas – gebėjimas pereiti nuo statinių slenksčių prie tikimybinio įverčio, tačiau kelių DL modelių naudojimas didina skaičiavimo kaštus, o veiksmingumas priklauso nuo AIS duomenų kokybės (spragos, vėlavimai). Ateityje verta tobulinti duomenų paruošimą, integruoti fizinius laivo parametrus ir nagrinėti stiprinamojo mokymosi sprendimus.

VI. Bendrosios išvados

Šiame darbe pasiūlytos ir ištirtos giliojo mokymosi metodikos, skirtos laivų trajektorijoms prognozuoti ir susidūrimų tikimybės rizikai vertinti jūrų navigacijoje. Konkretūs indėliai apima: (a) esamų trajektorijų prognozavimo metodų analizę, atkreipiant dėmesį į rekurentinių neuroninių tinklų (RNN) gebėjimą atlikti ilgalaikes prognozes, o ne remtis tiesine ekstrapoliacija; (b) įvairių RNN architektūrų projektavimą, mokymą ir derinimą, siekiant identifikuoti tas, kurios lemia mažiausią prognozavimo klaidą modeliuojant laivų judėjimą; (c) įvesties reprezentacijų, tokių kaip koordinačių transformacijos, rekursinį delta padėties perskaičiavimą, laivų tipų įterptąjį kodavimą ir meteorologinių duomenų integraciją, siekiant pagerinti prognozavimo tikslumą; (d) kiekybinio neapibrėžtumo metodų (EPR, CPR, PI ir CI) realizavimą ir palyginimą susidūrimų rizikai vertinti; (e) šių modelių pritaikymą anksčiau netirtam realiam susidūrimo atvejui, siekiant parodyti praktinį pritaikomumą. Pagrindinės išvados yra:

1. Rekurentiniai neuroniniai tinklai (tokie kaip LSTM ar GRU) yra būtini 20–25 minučių laivų trajektorijų prognozėms. Klasterizavimas arba tiesinė ekstrapoliacija negali sukurti nenutrūkstamų prognozių per 20–25 minučių intervalą, o tai yra būtina didelių laivų stabdymo laikui. Šiaurės jūros tyrime palygintos šešios architektūros (paprastas RNN, standartinis LSTM, sluoksniuotas LSTM, GRU, LSTM autoenkoderis, dvipusis LSTM): dvipusis LSTM (300 ląstelių, MSE = $3,618 \times 10^{-4}$) ir LSTM autoenkoderis (125 ląstelių, MSE = $3,688 \times 10^{-4}$) pasiekė mažiausias paklaidas su tam tikra ląstelių (dar žinoma kaip celės) kombinacija. Dvipusis LSTM išlaikė pastovų tikslumą skirtingose konfigūracijose, o autoenkoderio modelio tikslumas pradėjo nebegerėti (arba labai nežymiai), pasiekus atitinkamą LSTM ląstelių kiekį, taip parodant, kad architektūros ir hiperparametrų derinimas turi lemiamą įtaką ilgalaikiam tikslumui.

2. Atitinkamas LSTM ląstelių skaičiaus parinkimas yra itin svarbus, nes didesnis jų kiekis negarantuoja geresnio tikslumo. Baltijos jūros eksperimente LSTM modeliai su 75–300 ląstelių pasiekė mažiausią MAEH (apie 0,18–0,20 km). Viršijus 300 ląstelių, klaidos mažėjimas stabilizavosi, o didinant ląstelių skaičių virš 1000, labai išaugo mokymo laikas be reikšmingos naudos tikslumo atžvilgiu. Pasirinkus per mažą

kiekį (pvz., iki 50), tikslumas nėra tenkinamas, dėl to būtina parinkti optimalų ląstelių skaičių.

3. Patobulinti modelio įvesties atvaizdavimai: koordinačių transformacijos, rekursinė delta ekstrapoliacija, laivų tipų ir meteorologinių duomenų integracija, kurie padeda sumažinti prognozių paklaidas ir generalizuoti modelį.

- Įdiegus rekursinį delta perskaičiavimą (kai kiekviena prognozuota padėtis naudojama kitam žingsniui), MAEH sumažėjo nuo 0,6171 km (WGS84) iki 0,3537 km (WGS84–delta) ir nuo 0,4903 km (UTM) iki 0,3839 km (UTM–delta), ypač ankstyvuosiuose žingsniuose užkertant kelią klaidų kaupimuisi.
- Šiaurės jūros bandymuose su pasiūlytu LSTM autoenkoderiu, transformavus pradinius WGS84 AIS duomenis į UTM koordinates, MAEH sumažėjo iki 1,141 km (prieš tai WGS84: 1,427 km). Taikant atstumo ir kampo transformacijas, MAEH buvo 1,266 km, o Δplatumos / Δilgumos transformacija – 1,439 km.
- Baltijos jūros eksperimentuose su siūlomu LSTM autoenkoderiu integravus trimačio įterptinį kodavimo sluoksnį su skirtingais laivų tipais, MAEH sumažėjo iki 0,2636 km (prieš tai – 0,2655 km su fiktyviu kodavimu, 0,2681 km su etiketiniu kodavimu ir 0,2710 km be laivų tipo duomenų).
- Integravus meteorologinius duomenis su AIS įvestimis ir siekiant atspindėti aplinkos poveikį, MAEH sumažėjo nuo 0,3537 km iki 0,3203 km, patvirtinant, kad prognozėse tikslinga atsižvelgti ne tik į laivo charakteristikas, bet ir į supančią aplinką, kuri leidžia išgauti tikslesnes kelių žingsnių (ilgalaikes) prognozes.

4. Pasiūlytos tikimybinės intervalų technikos (CPR, EPR, PI, CI) pranoksta standartinius CPA / TCPA metodus. Iš kelių skirtingų modelių prognozuojamos trajektorijos ir vertinamas jų neapibrėžtumas, sudarant trajektorijų ribas (angl. *trajectory boundary*) PI, CI, EPR ir CPR būdais. Šios ribos vaizduoja galimą laivo padėties zoną ir suteikia prasmingesnės

informacijos nei taškiniu pagrindu veikiantys CPA / TCPA metodai. CPR apima 86,7 % (priešingai nei 71,4 % EPR, 51,0 % PI ir 4,9 % CI) padengimą, lyginant prognozuojamų ribų patekimą į faktinius laivo duomenis. CPR išvengia normaliojo pasiskirstymo prielaidų ir vertina neapibrėžtumo klasterius per 20 minučių laikotarpį. Tokiu būdu regionų ribos tampa dinamiškesnės, o jų persidengimas su kitų laivų ribomis gali būti laikomas tikėtina susidūrimo rizika.

5. Realaus incidento atvejis („Scot Carrier“ ir „Karin Höj“ susidūrimas, 2021 m.) patvirtina tikimybinį aptikimą, tačiau pabrėžia AIS duomenų trūkumus ir prognozių ribotumus. CPR metodas parodė 39,6 % susidūrimo riziką (95 % pasikliovimo lygiu), atitinkančią faktinį įvykį. Tačiau norint pritaikyti šį metodą, būtina pakankamai tiksli laivų trajektorijų prognozė, o AIS duomenų spragos gali riboti modelio tikslumą. Be to, metodai neatsižvelgia į laivų fizinius parametrus (pvz., posūkio spindulį, inerciją, laivo išmatavimus), kurie gali pagerinti saugios zonos vertinimą – tai galėtų būti tolesnių tyrimų kryptis.

Notes

Curriculum Vitae

Robertas Jurkus was born in Plungė, Lithuania. His first steps in education were in Plungė, where he left Saulė High School. In 2016, he earned a B.Sc. in Informatics (Specialist in Informatics) from Klaipėda State University of Applied Sciences, and in 2020, he obtained an M.Sc. in Technical Information Systems Engineering from Klaipėda University. Since 2020, he has been pursuing a Ph.D. in Informatics Engineering at Vilnius University.

From 2016 to 2024, Jurkus worked as a programmer for the IT department of an international food industry company, where he developed data analytics tools, built software solutions, and supported various enterprise systems. In 2020, he joined Klaipėda University as an assistant, and since 2023, he has served as a lecturer. He also participates as a committee member in the review and evaluation of final theses at Klaipėda University and other Lithuanian higher education institutions. Jurkus is actively involved in scientific research projects.

Gyvenimo aprašymas

Robertas Jurkus gimė Plungėje, Lietuvoje. Savo pirmuosius mokslus jis pradėjo Plungėje, čia baigė Saulės vidurinę mokyklą. 2016 m. įgijo informatikos, specialisto bakalauro laipsnį Klaipėdos valstybinėje kolegijoje, o 2020 m. – techninių informacinių sistemų inžinerijos magistro laipsnį Klaipėdos universitete. Nuo 2020 m. studijavo informatikos inžinerijos mokslo krypties doktorantūroje Vilniaus universitete.

Nuo 2016 m. R. Jurkus aštuonerius metus dirbo tarptautinės maisto pramonės įmonės IT skyriuje programuotoju, kur kūrė duomenų analizės įrankius, programinę įrangą ir palaikė įvairias įmonės sistemas. 2020 m. jis prisijungė prie Klaipėdos universiteto kaip asistentas, o nuo 2023 m. dirba lektoriumi. Taip pat dalyvauja Klaipėdos universiteto ir kitų Lietuvos aukštųjų mokyklų baigiamųjų darbų recenzavimuose ir vertinimuose kaip komisijos narys. Aktyviai dalyvauja moksliniuose projektuose.

Robertas Jurkus

Maritime Traffic Awareness Evaluation Using Deep Neural Networks
Doctoral Dissertation
Technological Sciences
Informatics Engineering (T 007)
Thesis Editor: Zuzana Šiušaitė

Jūrų eismo informuotumo vertinimas naudojant giliuosius neuroninius tinklus
Daktaro disertacija
Technologijos mokslai
Informatikos inžinerija (T 007)
Santraukos redaktorė: Jorūnė Rimeisytė-Nekrašienė

Vilnius University Press
9 Saulėtekio Ave., Building III, LT-10222 Vilnius
El. p. info@leidykla.vu.lt, www.leidykla.vu.lt
bookshop.vu.lt, journals.vu.lt
Print run of 20 copies