VILNIAUS UNIVERSITETO
KAUNO FAKULTETAS

SOCIALINIŲ MOKSLŲ IR TAIKOMOSIOS INFORMATIKOS INSTITUTAS

Finansų technologijų studijų programa
Kodas 6211BX022

KOTRYNA MAČERNYTĖ

MAGISTRO BAIGIAMASIS DARBAS

FINANSINIŲ PASLAUGŲ GINČŲ BAIGTIES PROGNOZAVIMO MODELIS

Kaunas 2024

VILNIAUS UNIVERSITETO
KAUNO FAKULTETAS

SOCIALINIŲ MOKSLŲ IR TAIKOMOSIOS INFORMATIKOS INSTITUTAS

KOTRYNA MAČERNYTĖ

MAGISTRO BAIGIAMASIS DARBAS

FINANSINIŲ PASLAUGŲ GINČŲ BAIGTIES PROGNOZAVIMO MODELIS

Magistrantas _____
(parašas)

Darbo vadovas _____
(parašas)

_____
(darbo vadovo mokslo laipsnis, mokslo pedagoginis vardas, vardas ir pavardė)

Darbo įteikimo data
Registracijos Nr.

Kaunas 2024

TURINYS

SANTRUMPŲ SĄRAŠAS

EŽŽT – Europos Žmogaus Teisių Teismas

NLP – Natūralios kalbos apdorojimo metodai (angl. Natural language processing)

AVM – Atraminių vektorių mašinos

MM – Mašininis mokymasis

BERT – Dvikryptis transformacinio kodavimo būdas (angl. Bidirectional Encoder Representations from Transformers)

TF-IDF – Terminų dažnumas – atvirkštinis dokumentų dažnumas (angl. term frequency inverse document frequency)

SMOTE – arba sintetinės mažumos perteklinės atrankos metodas

DI – dirbtinis intelektas

PAVEIKSLŲ SĄRAŠAS

LENTELIŲ SĄRAŠAS

Mačernytė, Kotryna (2024). Finansinių paslaugų ginčų baigties prognozavimo modelis. Kaunas: Vilniaus universitetas, Kauno fakultetas, Socialinių mokslų ir taikomosios informatikos institutas, Finansų technologijos. 52 p.

SANTRAUKA

Šio magistro baigiamojo darbo tikslas – sukurti modelį, skirtą ginčų dėl finansinių paslaugų baigties prognozavimui, pasitelkiant pažangias duomenų analizės technologijas ir dirbtinio intelekto metodus. Pagrindinis uždavinys – sujungti teisinių procesų analizę su moderniomis mašininio mokymosi technologijomis, siekiant padidinti prognozavimo tikslumą ir sumažinti teisinių procesų kaštus.

Darbo metu atlikta mokslinės literatūros apžvalga, kurioje nagrinėtos esamos ginčų sprendimo metodikos, teisinės praktikos iššūkiai bei teisinių technologijų (angl. LegalTech) panaudojimo galimybės. Didelis dėmesys skirtas giluminio mokymosi modeliams, ypač *BERT* algoritmui, kuris išsiskiria gebėjimu analizuoti tekstinius duomenis ir atpažinti sudėtingas semantines struktūras.

Tyrime panaudoti viešųjų teismų sprendimų duomenų bazių duomenys, kurie buvo kruopščiai apdoroti ir parengti mašininio mokymosi algoritmų taikymui. Sukurtas modelis vertinamas pagal tikslumą prognozuojant ginčų baigtį, atsižvelgiant į bylos tipą, šalių teisinius argumentus bei kitus reikšmingus veiksnius. Tyrimo rezultatai atskleidė, kad *BERT* modelis pasiekė 81,7 % tikslumą, rodydamas pakankamai aukštą efektyvumą prognozuojant finansinių ginčų baigtį. Šis rezultatas yra konkurencingas tarptautiniame kontekste, lyginant su kitose šalyse atliktų tyrimų rezultatais.

Sukurtas modelis gali būti pritaikytas ne tik finansinių ginčų analizei, bet ir kitoms teisės sritims, tokioms kaip darbo teisė ar šeimos teisė, kur tiksli ginčų baigties prognozė padėtų racionalizuoti teisinių procesų eigą. Praktinis modelio pritaikymas leistų teisės specialistams greitai įvertinti galimą ginčo baigtį, sumažinti procesų trukmę ir kaštus. Tyrimo rezultatai rodo, kad giluminio mokymosi technologijų integracija į teisinį sektorių yra perspektyvi kryptis, galinti reikšmingai pagerinti teisinių procesų efektyvumą.

Mačernytė, Kotryna (2024). A Model for Predicting the Outcome of Financial Services Disputes. MBA Graduation Paper. Kaunas: Vilnius University, Kaunas Faculty, Institute of Social Sciences and Applied Informatics. 52 p.

SUMMARY

The aim of this master's thesis is to develop a model for predicting the outcome of financial services disputes using advanced data analytics and artificial intelligence techniques. The main objective is to combine the analysis of legal processes with modern machine learning technologies to increase the accuracy of the prediction and to reduce the costs of legal processes.

The thesis includes a review of scientific literature, which examines existing dispute resolution methodologies, the challenges of legal practice and the opportunities for the use of legal technology (LegalTech). A great deal of attention was paid to deep learning models, particularly the *BERT* algorithm, which is distinguished by its ability to analyze textual data and recognize complex semantic structures.

The study used data from public court decision databases, which were carefully processed and prepared for the application of machine learning algorithms. The developed model was evaluated in terms of its accuracy in predicting the outcome of disputes, considering the type of case, the legal arguments of the parties and other relevant factors. The results showed that the *BERT* model achieved an accuracy of 81.7%, indicating high efficiency in predicting the outcome of financial disputes. This result is competitive in an international context when compared to the results of studies carried out in other countries.

The model developed can be applied not only to the analysis of financial disputes, but also to other areas of law, such as labor law or family law, where an accurate prediction of the outcome of disputes would help to streamline the course of legal proceedings. The practical application of the model would allow legal professionals to quickly assess the likely outcome of a dispute, reducing the length and cost of proceedings. The results of the study show that the integration of deep learning technologies into the legal sector is a promising direction that can significantly improve the efficiency of legal processes.

ĮVADAS

Finansinių paslaugų sektoriuje ginčai tarp klientų ir finansinių institucijų yra dažnas reiškinys, kuris gali turėti rimtų pasekmių tiek pačioms institucijoms, tiek jų klientams. Šie ginčai gali kilti dėl įvairių priežasčių, pradedant vartotojų teisių pažeidimais ir baigiant netinkamu finansinių produktų teikimu. Atsižvelgiant į šių ginčų mastą ir sudėtingumą, labai svarbu turėti patikimus ir efektyvius metodus, galinčius padėti prognozuoti jų baigtį. Tai gali prisidėti prie spartesnio ginčų sprendimo, sumažinti teismo išlaidas ir padidinti pasitikėjimą finansinėmis institucijomis.

Vienas iš galimų tokių metodų yra ginčų dėl finansinių paslaugų baigties prognozavimo modelis. Šis modelis naudotų pažangias duomenų analizės ir mašininio mokymosi technologijas, siekiant išanalizuoti istorinius ginčų duomenis ir nustatyti galimus sprendimo scenarijus. Naudojant didelių duomenų analitiką ir mašininio mokymosi algoritmus, modelis turėtų apdoroti didžiulius duomenų kiekius ir identifikuoti dėsningumus, kurie yra nepastebimi tradiciniais analizės metodais.

Temos aktualumas. Šiandienos visuomenėje, kurioje vyksta nuolatiniai pokyčiai, vis dažniau susiduriame su žmogaus ir dirbtinio intelekto bei naujų technologijų sąveika. Teisininko darbas taip pat sparčiai keičiasi – atsiranda robotų, padedančių rengti sutartis ar kitus dokumentus, dirbtinio intelekto sistemos, galinčios atpažinti kelių eismo taisyklių pažeidimus, bei nauji sprendimai, palengvinantys vykdymo procesą. Tai tik keli pavyzdžiai, rodantys, kad teisininko profesija prisitaiko prie modernėjančio pasaulio ir technologijų pažangos.

Kadangi finansinių paslaugų sektorius yra dinamiškas ir nepastovus, tai dažnai kelia ginčų tarp finansinių institucijų ir jų klientų. Šie ginčai gali turėti įvairių priežasčių, todėl būtina turėti efektyvius įrankius, kurie padėtų numatyti ginčų baigtį ir taikyti tinkamas strategijas jų sprendimui. Technologijos neabejotinai veikia teisę tad modelis, palengvinantis tam tikrus procesus būtų puiki pagalba šiam sektoriui.

Problemos ištyrimo lygis. Teisiniai ginčai finansų paslaugų srityje yra sudėtingi ir dažnai kyla dėl nesutarimų, informacijos trūkumo ar netinkamo elgesio. Siekiant efektyviai prognozuoti tokių ginčų baigtį, būtina suprasti finansinių paslaugų teisinę bazę ir naudoti duomenų analizės bei dirbtinio intelekto technologijas (Salmerón-Manzano, 2021, Dubois, 2021, Frolova ir Ermakova, 2021, Rodionov, 2023).

Teisminių bylų rezultatų prognozavimas padeda suprasti teisminių sprendimų priėmimo procesą. Shaikh, Sahu ir Anand (2020) nustatė svarbius veiksnius, darančius įtaką bylų baigtims, naudodami mašininio mokymosi algoritmus. Mahfouz ir Kandil (2011) sukūrė automatizuotą teisminių ginčų rezultatų prognozavimo metodą, kurio tikslumas siekė net 98 %. Katz, Bommarito ir Blackman (2016) sukūrė modelį, skirtą Jungtinių Amerikos Valstijų Aukščiausiojo Teismo elgesiui prognozuoti, pasiekdami 71,9 % tikslumą.

Nepaisant pasiektų rezultatų, prognozavimo modeliai gali būti tobulinami, kad būtų pasiekti dar aukštesni tikslumo rodikliai. Tokių modelių tobulinimas gali sudominti teismų stebėtojus, proceso dalyvius ir rinkas, nes net nedideli prognozavimo pasiekimai gali atnešti didelę finansinę naudą.

Darbo problema: kaip pagerinti finansinių paslaugų ginčų baigties prognozavimo procesą pasitelkiant mašininį mokymą?

Darbo objektas – ginčų dėl finansinių paslaugų baigties prognozavimas.

Darbo tikslas – pasiūlyti ginčų dėl finansinių paslaugų baigties prognozavimo modelį.

Darbo uždaviniai:

1. Apžvelgti teisinių ginčų dėl finansinių paslaugų baigties sampratą ir prognozavimą, šių ginčų priežastis ir technologinius jų baigties prognozavimui galimus sprendinius.
2. Paruošti duomenis apie ginčų dėl finansinių paslaugų baigtį mašininio mokymosi algoritmams bei nustatyti tam tinkamus parametrus.
3. Pasiūlyti ginčų dėl finansinių paslaugų baigties prognozavimo modelį.
4. Patikrinti pasiūlytą modelį, nustatyti jo efektyvumą bei pateikti pastebėjimus ar pasiūlymus tolimesniam modelio tobulinimui.

1. MOKSLINĖS LITERATŪROS APŽVALGA

Sprendžiant pirmąjį šio darbo uždavinį, keturiuose šios dalies skyriuose apžvelgiami teisinių ginčų dėl finansinių paslaugų baigties samprata bei prognozavimas (1.1.) šių ginčų priežastys (1.2.) ir technologiniai jų baigties prognozavimui sprendiniai (1.3.). Tokiu būdu yra nustatomas bendras šio darbo kontekstas ir teoriniai pagrindai.

1.1. Teisinių ginčų dėl finansinių paslaugų baigties samprata bei prognozavimas

Teisiniai ginčai finansų paslaugų srityje sudaro sudėtingą mazgą, kurį sudaro įvairūs teisiniai, finansiniai ir elgsenos elementai. Šie ginčai dažnai kyla dėl įvairių priežasčių, įskaitant nesutarimus, nepakankamą informacijos pateikimą, netinkamą ar klaidingą elgesį ir t.t. Siekiant efektyviai prognozuoti tokių ginčų baigtį, būtina turėti supratimą apie finansinių paslaugų teisinę bazę, įskaitant reglamentus ir precedento teisę. Taip pat svarbu suprasti, kaip galima įtraukti duomenų analizės ir dirbtinio intelekto technologijas, siekiant išgryninti paslėptus ryšius ir tendencijas, kurios padėtų suprasti ir numatyti ginčų išsprendimo rezultatus.

Žinoma, kad teisės mokslininkai jau ilgą laiko tarpą bando nuspėti galimus teismo procesų rezultatus, o tam, ypač pastaraisiais metais, mokslininkai pradėjo naudoti mašininio mokymosi pasiekimus šioje srityje. Šie pasiekimai naudojami tam, kad būtų galima prognozuoti tiek gamtos tiek socialinių procesų elgseną ateityje. Pavyzdžiui Brazilijos teismai kasmet susiduria su nemenku iššūkiu – sparčiai augančių bylų skaičiumi, tad natūralu, jog kyla poreikis pagerinti teisinės sistemos procesus. Šį iššūkį, pritaikius tris giliojo mokymosi architektūras ir apmokius 612 tūkstančių Brazilijos 5-ojo apygardos teismo apeliacinių skundų, priėmė Brazilijos mokslininkai. Savo straipsnyje pastarieji gautus modelio rezultatus lygino su 22 aukštos kvalifikacijos ekspertų prognozėmis ir gavo kiek geresnius rezultatus. Nustatyta jog, žmonių ekspertų koreliacijos koeficientas yra 0,1253, o geriausio modelio koreliacijos koeficientas siekė beveik 3 kartus didesnį skaičių – 0,3688. Taip buvo pagrįsta, jog natūralios kalbos apdorojimo ir mašininio mokymosi metodai yra perspektyvus metodas teisiniams rezultatams prognozuoti (Menezes-Neto ir Clementino, 2022).

Tuo tarpu Europoje, kai teismai pradėjo viešai skelbti sprendimus, taip pat tapo įmanoma atlikti duomenų analizę teisės srityje. Vienas iš tokių atvejų – automatinis teismo sprendimo prognozavimas remiantis Europos žmogaus teisių teismo (EŽTT) duomenimis. Dar 2020 metais mokslininkams pavyko pasinaudoti šiais duomenimis ir sukurti modelį, paremtą mašininio mokymosi metodais, kurio vidutinis tikslumas siekė 75%. Šiam rezultatui pasiekti mokslininkai nagrinėjo kalbos analizės bei automatinio informacijos išskyrimo panaudojimo galimybes, o konkrečiau – natūralios kalbos apdorojimo (angl. Natural language processing, NLP) metodus. Naudojant mašininį mokymąsi buvo atlikta kiekybinė analizė pagal byloje pavartotus žodžius ir frazes ir taip „apmokomas“ kompiuteris, kad galėtų prognozuoti teismo sprendimą. Autoriai ištyrė, kaip mašininis mokymasis gali būti naudojamas teisinių tekstų klasifikacijai, naudodami atraminių vektorių mašinas (AVM). Jie taip pat parodė, kad galima naudoti ir tradicinius NLP metodus, kad būtų galima atpažinti svarbiausius veiksnius, lemiančius teismo sprendimus (Medvedeva, Vols ir Wieling, 2019). Toks tyrimas demonstruoja, kad NLP yra labai naudinga teisinėje srityje ir yra puiki perspektyva tolimesniems metodų tobulinimams.

Teisminių bylų rezultatų prognozavimas gali padėti suprasti teisminių sprendimų priėmimo procesą. Bylų baigtis galima prognozuoti remiantis konkrečiai bylai būdingais teisiniais veiksniais, pavyzdžiui, įrodymų rūšimi ar neteisiniais veiksniais, pavyzdžiui, teismo ideologine kryptimi. Išsamią informaciją apie konkrečiai bylai būdingus teisinius veiksnius galima gauti iš teismų sprendimų. Tačiau šių veiksnių išskyrimas iš teisinių tekstų yra varginantis ir daug laiko reikalaujantis procesas (Shaikh, Sahu, ir Anand, 2020). Nepaisant to, autoriai savo darbe nustatė svarbius veiksnius, darančius įtaką su nužudymu susijusių bylų (paimtų iš Delio apygardos teismo) baigtims, ir parengė 86 bylų duomenų bazę, siekiant šiuos veiksnius naudoti kaip deskriptorius baigtims prognozuoti. Jie pritaikė įprastus mašininio mokymosi klasifikavimo algoritmus, o rezultatams gauti panaudojo kryžminį patvirtinimą „Leave – one - out“. Taip jiems pavyko nustatyti svarbius atitinkamų bylų veiksnius ir savo ruožtu numatyti jų baigtį.

Baigties numatymas svarbus ir, pavyzdžiui, dėl finansinės naštos bei papildomo laiko, kurio reikalauja teisminiai procesai, sumažinimo. Vienas iš tokių pavyzdžių – statybų pramonė. Statybų pramonė yra vienas iš pagrindinių ekonomikos sektorių, darančių didelę įtaką šalies augimui ir klestėjimui. Tačiau statybų pramonės indėlį į šalies ekonomiką stabdo vis daugiau ginčų, kurie kyla ir dažnai paaštrėja vykdant projektus. Ties tuo dirba dirbtinio intelekto srities

tyrėjai, kurie sukūrė priemones bei metodikas, skirtas modeliuoti teismų argumentaciją ir prognozuoti teisinių ginčų bylų baigtis.

Pasitelkiant mašininio mokymosi modelius, dar 2011 metais, sukurtas automatizuotas teisminių ginčų rezultatų prognozavimo metodas. Siekiant sukurti siūlomą metodą, Mahfouz, ir Kandil (2011) lygino trijų mašininio mokymosi metodų, t. y. atraminių vektorių mašinų (AVM), naivaus Bajeso ir taisyklių indukcijos bei neuroninių tinklų klasifikatorių (sprendimų medžių, sustiprintų sprendimų medžių ir projekcinės adaptyviosios rezonanso teorijos), rezultatai. Modeliai buvo apmokyti ir išbandyti naudojant 400 ginčų atvejų, užpildytų 1912-2007 m. laikotarpiu. Modelių prognozės pagrįstos reikšmingais teisiniais veiksniais, kuriais grindžiami nuosprendžiai ginčuose statybų sektoriuje. Iš devynių sukurtų MM modelių geriausiai pasirodė trečiojo laipsnio AVM polinominis modelis, kurio prognozavimo tikslumas siekė 98 %.

Dar vienas pavyzdys – remdamiesi mašininiu mokymusi bei ankstesniais darbais teisinio prognozavimo srityje buvo sukurtas modelis, skirtas Jungtinių Amerikos Valstijų Aukščiausiojo Teismo elgesiui prognozuoti. Šį modelį naudojant atsitiktinio miško metodą bei unikalia požymių inžinerija paruošė dar 2016 metais. Remiantis beveik dviejų šimtmečių istoriniais sprendimais nuo 1816 iki 2015 m. buvo pasiektas 71,9 % tikslumas. Šie rezultatai pagerina ankstesniuose darbuose pasiektą prognozavimo lygį ir jį pagerina, o pats modelis išsiskiria tuo, jog jį galima taikyti ne imties būdu visai teismo praeičiai bei ateičiai. Pasak autorių šie rezultatai bei modelis yra kiekybinio teisinio prognozavimo mokslo pažanga ir leidžia numatyti daugybę kitų galimų taikymo sričių (Katz, Bommarito, ir Blackman, 2016).

Nepaisant to, jog jau pasiekiami neblogi prognozavimo tikslumo rezultatai, autoriai drąsiai teigia jog procesą galima tobulinti ir pasiekti dar aukštesnių rezultatų. Pasak Katz, Bommarito, ir Blackman (2016) tokių modelių atsiradimas bei tobulinimas turėtų sudominti teismų stebėtojus, proceso dalyvius, piliečius ir rinkas. Atsižvelgiant į tai, kad teismų sprendimai gali turėti įtakos, pavyzdžiui, viešai prekiaujančioms bendrovėms, o net ir nedideli prognozavimo pasiekimai gali atnešti didelę finansinę naudą.

1.2. Galimos ginčų priežastys bei sprendimai Lietuvoje

Ginčų priežastys bei jų sprendimo būdai Lietuvoje gali labai skirtis priklausomai nuo ginčo pobūdžio ir įvairių teisinių bei institucinių kontekstų. Dažniausių ginčų priežastys yra šios:

1) Civiliniai ginčai;
2) Darbo ginčai;
3) Administraciniai ginčai;
4) Kriminaliniai ginčai;
5) Šeimos ginčai;
6) Komerciniai ginčai.

Ginčų sprendimo būdai priklauso nuo ginčo pobūdžio, šalių pageidavimų ir teisinio konteksto. Teisinės procedūros Lietuvoje yra reguliuojamos civilinio proceso kodeksu, baudžiamuoju procesu kodeksu, administracinių bylų teisenos kodeksu ir kitais teisės aktais, kurie nustato, kaip turi būti nagrinėjami ir sprendžiami įvairūs ginčai. Tuo tarpu finansinių paslaugų ginčų baigties prognozavimas reikalauja įvairialypio tyrimo, kuriame susipina teisiniai, finansiniai ir kiti elgsenos aspektai. Šiam tikslui reikia suprasti finansines paslaugas reglamentuojančius teisės aktus ir teisines sistemas. Be to, reikia įvertinti atskirų ginčų niuansus, įskaitant susijusių finansinių produktų ar paslaugų specifiką, aplinkybes, dėl kurių kilo ginčas, ir šalių elgesį bei motyvus.

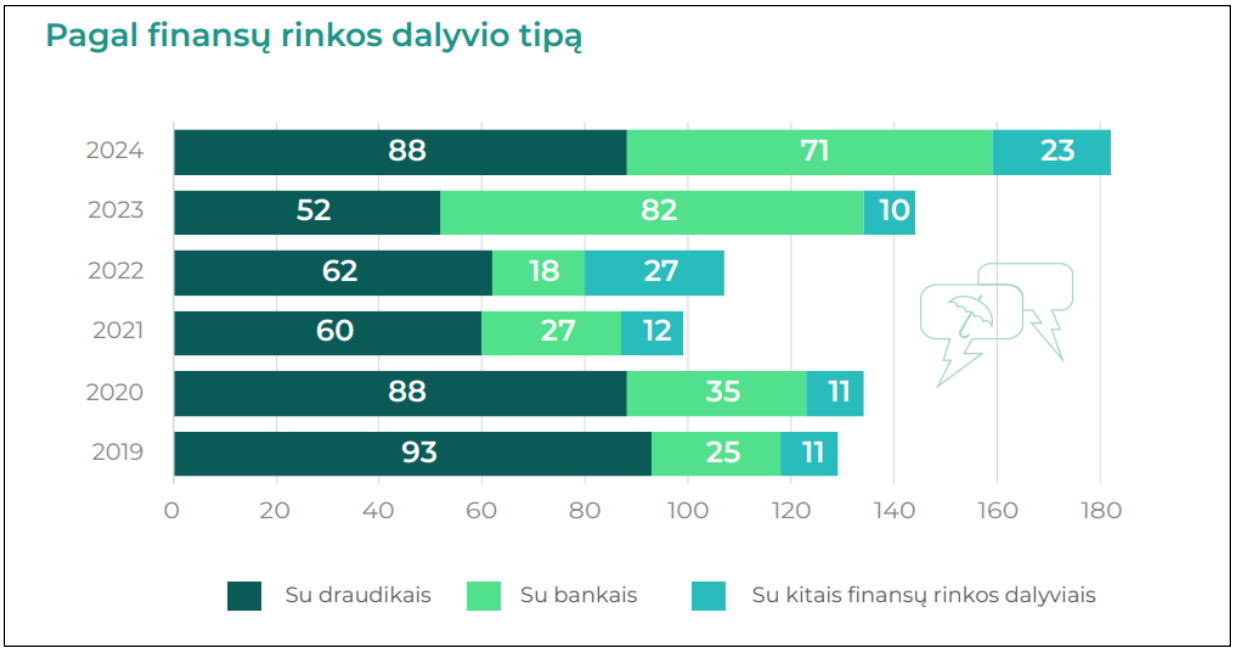
Išnagrinėjus Lietuvos banko svetainėje pateiktus duomenimis, išskiriamos šios galimos finansinių ginčų priežastys (Vartojimo ginčų neteisminio sprendimo ataskaita, 2023):

- Netinkamas paslaugos administravimas (pavyzdžiui, po pardavimo) – ginčai gali kilti dėl netinkamo ar nekokybiško paslaugų teikimo ar administravimo po to, kai klientas jau įsigijo finansinę paslaugą ar produktą;
- Pardavimo procesas – nesutarimai gali kilti dėl paties pardavimo proceso, įskaitant galimus klaidingus ar neskaidrius duomenis, teikiamus klientams arba sudarant sutartis;
- Mokesčių taikymas – ginčai gali atsirasti dėl mokesčių taikymo finansiniuose sandoriuose ar paslaugų teikime, kai klientas ir teikėjas turi skirtingas nuostatas ar supratimus;
- Netinkamas informacijos atskleidimas, reklama – nesutarimai gali kilti dėl netinkamo ar klaidinančio informacijos atskleidimo apie teikiamas finansines paslaugas ar produktus, įskaitant reklamą arba informaciją, pateiktą klientams;

- Nelicencijuota veikla – ginčai gali kilti, jei finansinių paslaugų teikėjas veikia be reikiamos licencijos ar leidimo, nesilaikant teisinių reikalavimų ir reglamentų;
- Kita – ši kategorija apima visas kitas galimas ginčų priežastis, kurios negali būti tiesiogiai priskirtos prie aukščiau minėtų kategorijų.

Lietuvos bankas, remdamasis Lietuvos banko įstatymo 47 straipsniu, sprendžia fizinių asmenų, t. y. vartotojų, ginčus su tam tikrais finansų rinkos dalyviais bei kitais juridiniais asmenimis. Šie ginčai dažniausiai kyla dėl finansinių ir papildomų investicinių paslaugų teikimo arba dėl draudimo paslaugų sutarčių, kai joms taikoma Lietuvos Respublikos teisė. Be to, Lietuvos bankas nagrinėja vartotojų ginčus ir su kitais subjektais, kai jie susiję su Lietuvos Respublikos mokėjimų įstatyme numatytų reikalavimų laikymusi, jei ginčo nagrinėjimą numato finansų rinką reglamentuojantys teisės aktai.

1 pav. Išnagrinėtų ginčų skaičius



Šaltinis: https://www.lb.lt/lt/statistikos-duomenu-iliustracijos

1 pav. pateiktoje diagramoje matyti, kad išnagrinėtų finansinių ginčų skaičius nuolat kinta, o šiais metais jis pasiekė 182. Lyginant su praėjusiais metais, stebimas tam tikras skaičiaus

padidėjimas, kurį galėjo lemti įvairios priežastys. Pirmąjį šių metų ketvirtį išaugo nesutarimų su draudimo įmonėmis skaičius: palyginti su 2023 m. I ketvirčiu, gautų kreipimųsi padaugėjo 41 %, o išnagrinėtų ginčų – 48 %. Tuo tarpu nesutarimų su bankais skaičius sumažėjo – gauta 131 kreipimasis ir išnagrinėta 71 byla, kas sudaro atitinkamai 25 % ir 13 % mažėjimą. Visi ginčai buvo išspręsti.

Tačiau, kas gi lemia ginčo sprendimą ir kaip jis vyksta? Pavyzdžiui Lietuvos banko internetinėje svetainėje skelbiama, jog bankas, priėmęs nagrinėti ginčą, per 3 darbo dienas išsiunčia vartotojui pranešimą, kad Lietuvos banke buvo pradėta vartojimo ginčo neteisminio sprendimo procedūra. Atitinkamai pranešime vartotojas informuojamas dėl jo teisių bei pareigų.

Vartojimo ginčo neteisminio sprendimo procedūros metu ginčo šalys turi teisę teikti Lietuvos bankui įrodymus, dalyvauti jų tyrimo procese, pateikti paaiškinimus bei argumentus, prieštarauti kitos šalies teiginiams, užduoti klausimus ir teikti prašymus. Svarbu pabrėžti, kad šalys privalo sąžiningai naudotis savo teisėmis ir nesiekdamos vilkinti ginčo nagrinėjimo. Lietuvos bankas imasi priemonių užtikrinti, kad nebūtų piktnaudžiaujama procedūromis ir ginčo nagrinėjimas vyktų sklandžiai, siekdamas kuo greitesnio ir tinkamesnio sprendimo. Paprastai ginčai nagrinėjami rašytinės procedūros tvarka, remiantis ginčo šalių pateiktais įrodymais. Kaip skelbia finansų rinkos dalyvių priežiūros - ginčų nagrinėjimas vyksta laikantis pagrindinių principų: rungimosi, operatyvumo, koncentracijos, ekonomiškumo ir bendradarbiavimo (Finansų rinkos dalyvių priežiūra, 2019).

Galiausiai, išnagrinėjus ginčą, priimamas vienas iš šių sprendimų:

- patenkinti vartotojo reikalavimus;
- iš dalies patenkinti vartotojo reikalavimus;
- atmesti vartotojo reikalavimus.

Atsižvelgus į ginčo nagrinėjimo metu nustatytas aplinkybes ir jas pagrindžiančius įrodymus bei vadovaudamasis Lietuvos Respublikos teisės aktais ir teisės principais Lietuvos bankas priima sprendimą. Be to, nuo šių 2024 m. lapkričio mėn. įsigaliojo ir teisės aktų pakeitimai, suteikiantys galimybę Lietuvos bankui ir finansų rinkos dalyviams efektyviau spręsti su nustatytais veiklos trūkumais susijusius klausimus – dabar leidžiama sudaryti taikius susitarimus. „Iki šiol galiojęs teisinis reguliavimas suteikė galimybę Lietuvos bankui sudaryti taikos sutartis su finansų rinkos dalyviais tik teismo procese. Naujoji taikių susitarimų tvarka leis greičiau ir paprasčiau rasti

sprendimus, išvengiant teismo. Tai naudinga abiem pusėms – taupomas tiek laikas, tiek ištekliai, o pats procesas tampa sklandesnis ir konstruktyvesnis“, – teigė Arūnas Raišutis, Lietuvos banko Teisės ir licencijavimo departamento direktorius.

Panašus susitarimų institutas veikia ir kitose Europos šalyse, tokiose kaip Austrija, Airija, Belgija, Islandija, Kipras, Latvija, Lenkija bei daugelis kitų. Šių šalių finansų rinkos priežiūros institucijos, užbaigdamos priemonių taikymo bylas taikiu susitarimu, pasiekia rezultatą nuo 50 % iki 75 % atvejų (Lietuvos bankas, 2024).

1.3. Technologiniai teisinių ginčų dėl finansinių paslaugų baigties prognozavimo sprendimai

Legaltech - tai naujų technologijų taikymas teisės pasaulyje, siekiant atlikti užduotis, kurias iki šiol atlikdavo teisininkai ar kiti advokatų kontorose dirbantys darbuotojai (Salmerón-Manzano, 2021). Šių technologijų taikymas finansų paslaugų sektoriuje gali turėti įvairių pritaikymo sričių ir naudos – jis palengvina įvairius teisinius procesus. Kaip rašoma autorės straipsnyje LegalTech būtų galima apibrėžti kaip technologijų naudojimą teisinėms paslaugoms kurti:

- internetines paslaugas, kurios sumažina arba panaikina poreikį kreiptis į teisinį sektorių pačia tradiciškiausia forma;
- internetines paslaugas, kurios pagreitina procedūras ir pačių teisininkų užduočių valdymą, sumažindamos išlaidas ir laiką, kurį specialistas turi skirti daugeliui savo užduočių atlikti;
- internetinės paslaugos, kurios supaprastina ir pakeičia teisininkų ir potencialių klientų bendravimo formą.

Be to jau dabar prieinama ir naudojama nemažai priemonių ar procesų, kurie „palengvina“ teisininko darbą: duomenų automatizavimas, pažangūs pokalbių robotai ir praktikos valdymo įrankiai, prognozuojamasis dirbtinis intelektas, išmanios teisinės sutartys ar žinių valdymo bei tyrimo sistemos. Nenuostabu, jog tam naudojamas mašininis mokymasis ar didelių duomenų analizės metodai (Dubois, 2021).

Tiriant dirbtinio intelekto naudojimo patirtį advokatų kontorų veikloje visame pasaulyje Evgenia E. Frolova, ir Elena P. Ermakova pabrėžia, jog šiuo metu pagrindinės LegalTech naudojamos technologijos yra teismų sprendimų prognozavimas arba prognozavimo technologija

ir prognozavimo kodavimas. Savo straipsnyje autorės teigia, jog ateityje dirbtinio intelekto technologijos suteiks praktikuojantiems teisininkams konkurencinį pranašumą bylinėjantis ir leis jiems geriau aptarnauti savo klientus. Advokatų kontoros, naudojančios dirbtinį intelektą, bus paklausesnės, o kontoros, nesugebančios automatizuoti savo veiklos, gali prarasti klientus dėl didesnių tų pačių paslaugų kainų (Frolova, ir Ermakova, 2021).

Tuo tarpu Rodionov (2023) savo moksliniame straipsnyje pateikia vertingas įžvalgas apie galimus ateities pokyčius ir iššūkius, kurie gali kilti toliau vystantis LegalTech. Anot jo, tokių technologijų integracija taip pat kelia keletą etinių ir reguliavimo problemų, įskaitant duomenų privatumą ir saugumą, sąžiningumą bei skaidrumą, atsakomybę ir atskaitomybę. Remiantis savo tyrimo išvadomis autorius pateikia tokias rekomendacijas:

- Išsamus dirbtinio intelekto priemonių supratimas ir jų taikymas teisės srityje;
- Investavimas į dirbtinį intelektą;
- Pirmenybės teikimas nuolatiniam profesiniam tobulėjimui bei mokymams;
- Bendradarbiavimas su suinteresuotomis šalimis iš visos teisinės ekosistemos;
- Teisinių ir reguliavimo sistemų paisymas.

Akivaizdu, jog LegalTech tobulėjimas ir vystymas turi nemažai akivaizdžių privalumų. Be to, kad sumažės teisinių tyrimų veiklos išlaidos, bus sutaupoma nemažai laiko, todėl galima daugiau laiko skirti asmeniniam bendravimui su klientu. Visa tai gali reikšti, kad teisinės konsultacijos bus teikiamos už priimtinesnę kainą, atveriant prieigą socialiniams sluoksniams, kurie iki šiol negalėjo gauti teisinės pagalbos. Taip pat skaitmeninė pagalba pagelbėtų teisininkams pagerinti jų darbo kokybę bei efektyvumą. Tačiau svarbu pabrėžti ir galimus trūkumus. Būtų problematiška, jei teisininkas nevisiškai suprastų dirbtinio intelekto paruoštą rekomendaciją ar internetines paslaugas. Taip pat kai kurios teisinės problemos gali būti per daug kompleksiškos ar sudėtingos, kad būtų pilnai išspręstos naudojant tik technologijas. Taikant LegalTech platformas gali kilti ir saugumo rizikų susijusių su duomenų saugumu ir privatumu. Be to LegalTech įrankiai gali sukelti naujų juridinių iššūkių, ypač susijusių su duomenų privatumu, intelektinės nuosavybės teisėmis ar netgi etikos klausimais, susijusiais su dirbtiniu intelektu (Fries, 2016).

Prognozavimo sprendimai teisės srityje yra kuriami, kad padėtų teisininkams, teisėjams ir kitiems suinteresuotiems asmenims priimti sprendimus. Šie sprendimai - pagrįsti didelių duomenų

analize ar mašininio mokymosi algoritmais gali apimti įvairiausius metodus ar įrankius, kurie naudojami tiek akademiniuose tyrimuose tiek praktinių teismų procesuose.

Taigi apibendrinant aukščiau pateiktus faktus, galima teigti jog teisinių technologijų, kitaip LegalTech, taikymas yra neišvengiamas modernioje teisinėje praktikoje. LegalTech įrankiai keičia teisės sritį ir įmonės, kurios juos taikys taps patrauklesnės, nei tos, kurios savų procesų automatizuoti negebės.

1.3.1. Natūralios kalbos apdorojimo metodai

Nagrinėjant baigties prognozavimo sprendimus ir skaitant mokslinius straipsnius gan aiški tendencija, ypač teisės srityje, taikyti natūralios kalbos apdorojimu grindžiamus metodus. NLP yra sritis, kuri taiko kompiuterių kalbos supratimo ir generavimo technologijas, o teisinėje srityje NLP metodai yra naudojami siekiant efektyviau analizuoti ir interpretuoti didelius teksto duomenų kiekius, susijusius su teisės aktais ar teismo sprendimais.

Pagrindinis natūralios kalbos uždavinių sprendimo elementas yra žodžių įterpiniai (angl. word embeddings), nes jie suteikia galimybę tekstinius duomenis atvaizduoti mažos dimensijos erdvėje. Turint sakinį, sudarytą iš žodžių, posakių ar simbolių, žodžių įterpiniai diskretinį atvaizdavimą paverčia realių reikšmių vektoriumi. Pastaraisiais metais šie žodžių įterpiniai jau kelis kartus patobulinti, kad galėtų aprėpti dar didesnę imtį, pavyzdžiui, žodžius pakeičiant daliniais žodžiais ir juos sujungiant ar naudojant *BERT* (Condevaux, 2020). *BERT* yra NLP modelis, sukurtas „Google AI Language“ dar 2018 metais, skirtas dvipusiai suprasti kalbą ir analizuoti tekstą tiek iš kairės į dešinę, tiek iš dešinės į kairę. Pastarasis naudojamas tiek teksto klasifikacijai, klausimų atsakymui, teksto sąryšio analizei ir pan.

Apart to, nemažai jau atliktų tyrimų (teisės srityje) yra gan riboti, nes juose paprastai nagrinėjamas vienas mašininio mokymosi metodas ir vienas teismo tipas, o tai iš dalies apsunkina jau atliktų tyrimų palyginimą (Mumcuoglu, Ozturk, Ozaktas ir Koc, 2021). Emre Mumcuoglu, Ceyhun E. Ozturk, Haldun M. Ozaktas ir Aykut Kocdarbe darbo siūloma metodika apima NLP taikymą teisinėje srityje, o jie siekia numatyti bylos baigtį atsižvelgiant tik į teismo surašytą faktų aprašymą. Šiam prognozavimo uždaviniui įveikti nagrinėti apygardų apeliaciniai teismai ir konstitucinis teismas, o naudojami ir keli literatūroje gerai žinomi klasifikavimo metodai. Pastarieji apima sprendimų medžius, atsitiktinius miškus, atraminių vektorių mašinas ir

naujausiais gilaus mokymosi metodus. Palyginus ir išanalizavus tam tikrų metodų našumą buvo įgyvendinta proceso baigties prognozavimo sistema, kuri pasiekė puikius rezultatus. Taikant ir lyginant įvarius procesus pavyko pasiekti net 97% tikslumą.

Tad šiuolaikiniai NLP algoritmai, tokie kaip *BERT*, *GPT*, *LSTM* (angl. Long Short-Term Memory) ir kitos gilios mokymosi technologijos, leidžia giliau suprasti ir interpretuoti tekstą, įskaitant teisines temas.

1.3.2. BERT taikymas teisinių dokumentų analizei ir prognozėms

BERT yra dirbtinio intelekto modelis, kuris padeda „perprasti“ tekstus – tai tarsi smegenys, galinčios išanalizuoti daugybę sudėtingų žodžių ryšių. Kai jis pritaikomas teisei, atsiranda galimybė analizuoti didžiulius teisinių dokumentų kiekius ir net numatyti teismo sprendimus. Šio modelio naudojimas vis populiarėja ir galima surasti vis daugiau įvairių mokslinių straipsnių. Nors Lietuvoje šie metodai dar nėra plačiai įdiegti, jų taikymas galėtų reikšmingai pagerinti ginčų sprendimo efektyvumą, ypač finansinių paslaugų srityje.

Puikus *BERT* taikymo pavyzdys teisėje: PILOT sistema teisinių bylų prognozavimui. Šiam tyrimui atlikti buvo pasirinkti tiek JAV tiek Europoje naudojami teismų praktikos duomenys. Ši sistema naudoja *BERT* ir jo teisės srityje pritaikytą variantą *Legal-BERT*. Modelis taikomas sprendžiant teisinių precedentų analizę ir sprendimų prognozavimą. Vienas svarbiausių sistemos elementų – gebėjimas rasti panašius teisinius atvejus, įtraukti kontekstą ir jį išanalizuoti. Šio tyrimo metu prognozių tikslumas siekė ~ 75 – 80 %, priklausomai nuo bylos sudėtingumo ir įtraukiamos informacijos detalumo (Cao, Wang, Xiao ir Sun, 2024).

Panašiai, naudojant kelias *BERT* modelio variacijas: *BERT, DistilBERT, roBERTa* ir kt., šiais metais JAV mokslininkai bandė ištirti, kaip tekstų santraukos (angl. summary) veikia prognozavimo tikslumą. Nenuostabu jog modeliai, mokyti su pilnu tekstu ir santraukų versijomis, rodė skirtingus rezultatus. Pastebėta, kad *Legal-BERT* modelis pasiekė aukščiausią ~ 85 % tikslumą, kai buvo naudojami pilni tekstai, o santraukos rezultatai sumažino tikslumą iki ~ 78 % (Varshini, Varma, Prabhakar ir Pati, 2024).

Tuo tarpu Brazilijoje sukurtas modelis, kuris geba analizuoti, kokie bylos faktai dažniausiai lemia tam tikrą teismo sprendimą. *BERT* variantas, pritaikytas Portugalų kalbai, buvo taikomas analizuojant teismo sprendimus, ypač atsižvelgiant į bylos faktus (anlg. fact embeddings).

Pritaikius modelį, rezultatai parodė, kad teisinių dokumentų klasifikacijos tikslumas siekia 80 – 82 %, o sudėtingesnių prognozių atveju (pvz., nustatant nuobaudos tipą) – ~ 76 % (Lage-Freitas, et al., 2019).

Modelis, skirtas teismų sprendimų prognozavimui Kinijoje taip pat pademonstravo panašius rezultatus. Kinijos teismų sistemose *BERT* buvo taikomas didelio masto teisinių dokumentų analizei ir teismo sprendimų prognozėms. Šis modelis analizavo bylos faktus, teisės normas ir ankstesnių bylų ryšius taip siekdamas sukurti prognozę. Tyrimas parodė ~ 87 % tikslumą, kai buvo naudojami specializuoti *BERT* modeliai, optimizuoti teisės tekstams (Wang, Gao ir Chen, 2020).

Remiantis įvairiais moksliniais tyrimais, galima teigti, kad vidutinis BERT modelio pasiekiamas tikslumas teisinių dokumentų analizei ir prognozėms dažnai siekia apie 80 %. Taigi šių technologijų – *BERT* modelio ir jo variacijų integracija Lietuvoje taip pat galėtų palengvinti teisinių dokumentų apdorojimą, ginčų sprendimą ir prognozių generavimą, ypač finansinių paslaugų sektoriuje. Specialiai lietuvių kalbai pritaikytas *BERT* modelis padėtų greitai ir tiksliai analizuoti teisės dokumentus, užtikrindamas efektyvumą ir kokybę.

1.3.3. Neuroninių tinklų taikymas

Nemažai mokslinių darbų atlikta ir naudojant neuroninius tinklus. Nustatyta, kad pastarieji konkrečiose situacijose gali pasiekti didelį tikslumą ir taip pranokti standartinius bei plačiai naudojamus algoritmus, kaip, pavyzdžiui, logistinė regresija, medžių modeliai ar AVM. Pasak Condevaux (2020), nors naudojant neuroninius tinklus labai pagerėja modelio našumas tačiau taip pat gerokai padidėja ir skaičiavimų sąnaudos. Norint lengviau ir greičiau apmokyti modelį bei pasiekti kompromisą tam tikrais atvejais taikoma redukcija – o taip, autoriaus atliktame tyrime, prarandama vidutiniškai 1,5 taško tikslumo. Nepaisant to, neuroniniai tinklai iš esmės lenkia paprastus tiesinius modelius ir geba teisingai apdoroti įvairias sekas ar tvarkyti retus žodžius, naudojant iš anksto apmokytą žodžių įterpimą.

Neuroniniai tinklai taikomi ir Kaur bei Bozic (2019) straipsnyje, kuriame mokslininkai bando spręsti EŽTT sprendimų prognozavimo problemą. Mokslininkai taip pat palygina AVM apmokytų modelių rezultatus su konvulsinių neuroninių tinklų apmokytais modelių rezultatais ir pateikia pakankamai statistinių įrodymų, kurie leidžia nustatyti, kuris algoritmas geriau sprendžia

šią problemą. Nenuostabu, jog neuroniniai tinklai lenkia AVM metodus ir pateikia 82% procentų tikslumą, o šis yra 8% didesnis nei AVM gautų rezultatų. Ruošiant duomenis buvo naudojamos ir įvairios žodžių įterpinių sistemos, kaip, pavyzdžiui, FastText, GloVe ir Word2vec. Kadangi šios įterpinių sistemos jau turi iš anksto apibrėžtą žodyną, apibrėžtą naudojant skirtingų sričių duomenis, taip pat buvo apmokytas ir pritaikytas naujas žodžių įterpinys, kad būtų galima sukurti teisinės srities žodyną. Taip buvo atsižvelgta į tai, kad iš anksto apmokytuose įterpiniuose gali nebūti teisinės srities žodžių, todėl jie gali netinkamai atvaizduoti EŽTT duomenis. Visi šie atvaizdai buvo įdiegti iš anksto paruoštuose tekstiniuose failuose. Vėlesniuose etapuose buvo kuriami skirtingi neuroninių tinklų modeliai su kiekvienu iš šių žodžių įterpinių ir buvo lyginami, siekiant nustatyti jų veiksmingumą (Kaur ir Bozic, 2019).

Taigi neuroninių tinklų taikymas teisės procesų baigties prognozavimui yra akivaizdi pažanga, leidžianti išnaudoti gilųjį mokymąsi ir kitas sudėtingas struktūras teisės srityje. Nemažai mokslinių darbų yra skirta tyrimams, kuriuose neuroniniai tinklai rodo didelį tikslumą, pranokdami tradicinius algoritmus.

1.3.4. Duomenų integracija ir šaltiniai

Prognozavimo sistemoms integruoti naudojami duomenys iš skirtingų teismų, apimančių civilinius, baudžiamuosius, administracinius ir kitus teisinius procesus. Šie duomenys gali būti gauti iš teismo sprendimų, kurie yra viešai prieinami, taip pat iš teisinių dokumentų, pvz., bylų medžiagos, teismo sprendimų motyvų ir precedento atvejų.

Pavyzdžiui Kaur bei Bozic (2019) straipsnyje buvo naudojamas duomenų rinkinys iš ankstesnio tiriamojo darbo, kuris atliktas prognozuojant EŽTT teisinių bylų sprendimus. Pastarieji gauti iš 2017 metų duomenų bazės patalpintos HUDOC internetiniame puslapyje. Visi duomenų rinkinyje esantys tekstiniai failai buvo struktūrizuoto formato, kai kiekvienas bet kurio straipsnio failas turėjo tuos pačius poskyrius. Tuo tarpu Condevaux (2020) naudojo pirmosios instancijos teismų sprendimų tekstus, bei turėjo surinkti tikslinį kintamąjį, t. y. apeliacinių skundų rezultatus. Tam pasiekti prireikė bylos metaduomenų, kurie ne visada prieinami. Taip pat kilo iššūkiai dėl bylų pavadinimų, kurie ne visuomet buvo tikslūs. Norint įsitikinti duomenų kokybe keletą kartų buvo patikrinti bylų pavadinimai, kitaip, etiketės. Taip pat buvo naudojamas klasifikatorius, kuriame bent du iš trijų etikečių šaltinių turėjo sutapti, kad ta etiketė būtų laikoma galiojančia. Šie

iššūkiai pasitaiko, kai teismo atstovai elektroninėje sistemoje neteisingai tvarko bylas arba kai nepakeičia į sistemą įkeltos bylos pavadinimo.

Be to, siekiant užtikrinti duomenų kokybę ir tinkamą jų analizę, duomenys iš skirtingų teismų ir teisinių praktikos šaltinių gali būti standartizuojami. Tai apima teisinių terminų vienodinimą, bylų numerių formatavimą, datos ir kitų svarbių duomenų elementų standartizavimą.

Dar svarbu atkreipti dėmesį ir į apmokymų, patvirtinimo ir testavimo duomenų rinkinius. Paprastai tariant, mokymo duomenų rinkinys naudojamas modeliui mokyti, patvirtinimo duomenų rinkinys - parametrams ir (arba) architektūroms koreguoti, o testavimo duomenų rinkinys - galutiniams rezultatams pateikti. Labiausiai paplitęs duomenų skaidymo būdas - sulaikymo metodas, kai tyrėjai vieną atsitiktinai pasirinktą duomenų rinkinio dalį (paprastai 20%) laiko tvirtinimo ir testavimo duomenimis (Menezes-Neto ir Clementino, 2022).

Taigi norint užtikrinti, kad prognozavimo sistemos ar sprendimai būtų tinkamai paruošti ir pasirengę numatyti bylų baigtis, reikia remtis platesniu teisinių duomenų spektru bei supratimu. Duomenims paruošti gali prireikti bent kelių žingsnių. Šis procesas gali apimti automatizuotą teksto analizę, kad būtų išskirti svarbūs teisiniai terminai, faktai ir jų sąryšiai.

Šis skyrius akcentuoja pažangias technologijas, kurios revoliucionizuoja teisės procesų valdymą ir sprendimų priėmimą, siekiant efektyviai prognozuoti teisinius ginčus ir proceso baigtį.

2. GINČŲ DĖL FINANSINIŲ PASLAUGŲ BAIGTIES PROGNOZAVIMO MODELIO METODIKA

Sprendžiant antrąjį šio darbo uždavinį šiame skyriuje bus aptariami pagrindiniai aspektai, susiję su duomenų surinkimu (2.1.) ir duomenų paruošimu (2.3.) Taip pat bus aptariama kokia programinė įranga (2.2.) ir įvairus mašininio mokymosi metodai pasirinkti tolimesniam modelio realizavimui (2.4.). Tokiu būdu bus nustatomi tinkami metodai ir parametrai šio darbo uždavinio įvykdymui.

2.1. Duomenų rinkinys

Prognozavimo modelio kūrimas pradėtas nuo duomenų surinkimo, kurie apima informaciją apie finansinių paslaugų ginčų sprendimus. Šie duomenys buvo gauti iš viešai prieinamos Lietuvos Banko (toliau LB) skelbiamų rekomendacinių sprendimų duomenų bazės, kur pateikiami ginčai dėl įvairių finansinių paslaugų. Kiekvienas pateiktas ginčo sprendimas turi atitinkamus požymius, tokius, kaip ginčo priežastis, ieškovo reikalavimas, atsakovų argumentai ir galutinis sprendimas – teigiamas, dalinai teigiamas arba neigiamas ieškovo naudai.

Remiantis tam tikromis taisyklėmis LB nemokamai nagrinėja vartotojų ir finansų rinkos dalyvių ginčus ir pateikia rekomendacinio pobūdžio sprendimą dėl ginčo esmės, kuris nėra skundžiamas teismui. Iš LB internetinės svetainės buvo ištraukta 1654 PDF dokumentai, kuriuose nagrinėjami skirtingi ginčai dėl įvairių finansinių paslaugų. Duomenys surinkti nuo 2016 metų. Žemiau, 1-oje lentelėje matyti, koks yra šių duomenų rinkinio sprendimų pasiskirstymas.

1 lentelė

Nagrinėjamų duomenų sprendimų pasiskirstymas

LB rekomenduojamas sprendimas	Kiekis
Atmesti	1337
Dalinai patenkinti	146
Patenkinti	157

Šaltinis: sudaryta autorės

2.2. Programinė įranga

Duomenų apdorojimui, modeliavimo ir vertinimo procesams buvo naudojama *Python* programinė įranga. *Python* - pagrindinė programavimo kalba modelio kūrimui. Ši kalba žinoma kaip lengvai suprantama, paprasta bei universali. Tai bendrosios paskirties programavimo kalba, kuri populiari tarp įvairių sričių mokslinių tyrimų. Pastaruoju metu pastebimos ir ryškios tendencijos, kai *Python* programavimo kalba vis dažniausiai naudojama mašininiam mokymuisi ir dirbtiniam intelektui. Dėl savo universalumo ir paprastumo *Python* puikiai tinka tokioms užduotims atlikti, nes ją lengva perprasti bei naudoti, be to, yra daug mašininiam mokymuisi ir dirbtiniam intelektui pritaikytų bibliotekų (Rayhan ir Gross, 2023).

Taip pat naudojama *SpaCy* biblioteka, kuri yra populiari NLP biblioteka, skirta teksto analizei bei apdorojimui. Pastaroji suteikia įvairias galimybes apdoroti ir dirbti su tekstu, kaip, pavyzdžiui, teksto išskyrimas, žodžių kamienavimas (angl. lemmatization) ir nutraukimo žodžių (angl. stop-words) filtravimas. Atliktame tyrime *spaCy* biblioteka buvo naudojama būtent frazių tokenizavimui (angl. tokenization), nutraukimo žodžių pašalinimui ir normalizavimui, kuriam atlikti keli tam tikri žingsniai. Teksto valymui buvo naudojamas *spaCy* modelis *lt_core_news_lg*, kuris skirtas būtent lietuvių kalbai. Šis modelis turi integruotas kalbos taisykles, lematizaciją, POS (kalbos dalių) žymėjimą, morfologiją, vardų atpažinimą (NER) ir kitus NLP įrankius, pritaikytus būtent lietuvių kalbos ypatybėms. Taip pat įdiegtos ir kitos reikalingos bibliotekos, kaip, pavyzdžiui, *pdfplumber* naudojama PDF failų turinio skaitymui, *torch*, *numpy*, *os*, *transformers*.

2.3. Duomenų paruošimas

Duomenys iš duomenų bazės ištraukti ir pateikti PDF formatu, todėl pirmas žingsnis buvo jų apdorojimas ir struktūrizavimas. Naudojant *pdfplumber* iš PDF dokumentų išgautas tekstas, kuris bus naudojamas tolesniam apdorojimui. Toliau, naudojantis įvairiomis teksto analizės technikomis buvo išskirti svarbūs atributai iš analizuojamų finansinių ginčų sprendimų dokumentų: tai apėmė duomenų valymą, nutraukimo žodžių pašalinimą ir specialių juridinių terminų pritaikymą. Tekstas buvo valomas naudojant *spaCy* modelį - pašalinti nepageidaujami žodžiai ir skyrybos ženklai.

Daugelyje NPL sričių išskiriami du skirtingi teksto normalizavimo būdai - kamienavimas (angl. stemming) ir lematizavimas (angl. lemmatization). Lematizavimas pakeičia žodžio priesagą arba ją visiškai pašalina, kad būtų gauta tik pagrindinė žodžio forma. Kaip ir minėta anksčiau, naudojant *spaCy* - taikomas ir lematizavimas, kuris yra labiau pritaikytas apdoroti „teisinį“ tekstą, nes jis išlaiko daugiau konteksto nei kamienavimas. Tuo tarpu taikant kamienavimą gaunama tik apytikslė žodžio forma, vadinama kamienine arba normalizuota forma. Be to, lematizavimas yra sudėtingesnis procesas, tačiau jis naudingas daugelyje taikymo sričių (Toman, Tesar ir Karel, 2006).

Kadangi pateikti teisiniai tekstai gan aiškus ir struktūrizuoti, buvo sukurtos funkcijos, kurios leidžia atpažinti ir lengvai išgauti pastraipas, prasidedančias tam tikrais raktiniais žodžiais, pavyzdžiui, „n u s p r e n d ž i u“, „n u s t a t y t a“ ar „k o n s t a t u o j u“. Po tokių raktinių žodžių sekė LB rekomenduojamas sprendimas, ieškovo skundas ar įstatymai ir pataisos skirtos LB rekomenduojamiems sprendimams priimti. Tokiems sprendimas žymėti sukurta atskira funkcija, kuri apima sprendimų klasifikavimą pagal iš anksto apibrėžtas klases: neigiama, dalinai teigiama, teigiama, ir taip leidžia analizuojamą tekstą suporuoti su atitinkama klase. Kiekvienas ginčo sprendimas įvestas CSV faile turėjo savo atskirą žymą.

Taip pat, kad užtikrinti, jog duomenys ir toliau būtų lengvai prieinami bei apdorojami tolimesnėje analizėje, išgauti ginčų sprendimai ir priežastys buvo nuskaitomi ir įrašomi į atskirą CSV formato failą. Tam panaudota *pandas* biblioteka. Šis sprendimas leidžia optimaliau kaupti informaciją, kurią galima efektyviai panaudoti tolimesniuose modelio treniravimo ir testavimo etapuose. CSV failas yra patogus pasirinkimas, nes jame duomenys pateikiami struktūrizuota,

lengvai pasiekiama forma, kuri užtikrina, kad informacija būtų paruošta tolesniam naudojimui tiek kryžminės validacijos, tiek kitų vertinimo metodų taikymui.

2.4. Modelyje naudojamų metodų pristatymas

Modelio tikslas – prognozuoti finansinių ginčų baigtį, remiantis pateiktais duomenimis. Tam pritaikyti ir panaudoti tam tikri metodai bei funkcijos, kurios aprašomos tolimesniuose skyreliuose.

2.4.1. Natūralios kalbos apdorojimas

Natūralios kalbos apdorojimas, toliau NLP, tai tam tikri metodai, kurie padeda analizuoti ir apdoroti natūralios kalbos tekstus pateiktus skaitmeniniu formatu. Paprastai tariant, šių metodų tikslas yra suprasti analizuojamų dokumentų turinį. Pastaruoju metu natūralios kalbos apdorojimo metodikos dar labiau suklestėjo ir jau paskelbta daugybė tiriamųjų darbų, kuriuose sprendžiami įvairūs šios srities uždaviniai, kaip, pavyzdžiui, teksto klasifikavimas, įvairių esybių atpažinimas arba tiesiog apibendrinimas (Gonzalez - Carvajal ir Garrido - Merchan, 2021).

Pasak Gonzalez - Carvajal ir Garrido - Merchan galime išskirti skirtingus dviejų tipų požiūrius į NLP problemas. Pirmas - NLP plačiai naudojami lingvistiniai metodai, kurie naudoja srities ekspertų svarbiomis nustatytas teksto savybes. Tai gali būti tam tikri žodžių junginiai ar n-gramos (angl. n-grams), gramatinės savybės, tam tikros žodžių reikšmės ar pozicionavimas ir panašiai. Tokie požymiai gali būti sukurti rankiniu būdų konkrečiai problemai spręsti. Antrasis požiūris – mašininiu ar giliuoju mokymusi grindžiami metodai, kurie klasikiniu būdų analizuoja tekstą ir automatiškai išveda teksto savybes, kurios svarbios klasifikavimui. Nors abu metodai turi savo trūkumų ar privalumų, pastaruoju metu dėl didesnių skaičiavimų galimybių dominuoja mašininiu mokymusi paremti metodai.

Taip pat NLP vis dažniau taikomas finansų pramonėje siekiant automatizuoti įvairias užduotis, pavyzdžiui, teksto klasifikavimą, nuotaikų analizę ir natūralios kalbos generavimą. Vienas iš pagrindinių NLP taikymų finansų srityje yra teksto klasifikavimas, kuris naudojamas finansiniams dokumentams, pavyzdžiui, naujienų straipsniams ir pelno ataskaitoms, automatiškai suskirstyti į iš anksto nustatytas kategorijas ().

2.4.2. Dvikryptis transformacinio kodavimo būdas – *BERT*

Pastaruoju metu išpopuliarėjo išankstinis NLP modelių mokymas siekiant spręsti įvairias kalbos modeliavimo užduotis. Tokie iš anksto apmokyti modeliai kaip *BERT* (Devlin et al., 2019) ir *ELMo* (Peters et al., 2018a) pažengė į priekį sprendžiant įvairiausias užduotis, o tai rodo, kad šie modeliai išankstinio apmokymo proceso metu įgyja vertingų kalbinių ir semantinių kompetencijų. Nors ir nustatyta, kad kalbos modelių išankstinio mokymo nauda yra didelė, dar reikia suprasti, ko apie kalbą šie modeliai išmoksta išankstinio mokymosi metu ir kaip tai veikia (Ettinger, 2020).

Būtent dvikryptis transformacinio kodavimo būdas – *BERT* (angl. Bidirectional Encoder Representations from Transformers) – tai 2018 m. kompanijos „Google AI“ tyrėjų Jacob Devlin, Ming-Wei Chang, Kenton Lee ir Kristina Toutanova sukurtas modelis, skirtas pasiekti aukščiausio lygio tikslumą sprendžiant įvairias natūraliosios kalbos apdorojimo (NLP) užduotis. Šis modelis leidžia tiksliai nustatyti duomenų reprezentacijas, atsižvelgiant į konkretų uždavinį ar duomenų rinkinį, o tai pašalina būtinybę kurti ir mokyti naujus, labai specifinius modelius.

Dėl savo universalumo ir efektyvumo *BERT* atvėrė galimybes kurti modernius NLP sprendimus, ir ši technologija greitai tapo standartu tiek akademinėje, tiek pramonės srityje (Gonzalez-Carvajal ir Garrido-Merchan, 2021). Be to, lentelėje apačioje (žr. 2 lentelę) pateiktas skirtingų metodikų tikslumo palyginimas rodo, kad būtent *BERT* pasiekė aukščiausią tikslumą – 0,9387 lyginant su kitomis gerai žinomomis technikomis. Autoriai taip pat pabrėžia, kad šių rezultatų pasiekimas taikant tradicinius metodus buvo kur kas sudėtingesnis nei naudojant *BERT* modelį.

2 lentelė

Tikslumas, gautas lyginant skirtingas metodikas

Modelis	Tikslumas
BERT	0,9387
Balsavimo klasifikatorius (angl. Voting Classifier)	0,9007
Logistinė regresija (angl. Logistic Regression)	0,8949
Linijinis atraminių vektroių klasifikatorius (angl. Linear SVC)	0,8989
Daugianominis Naivusis Bajesas (angl. Multinomial NB)	0,8771
Keterinis klasifikatorius (angl. Ridge Classifier)	0,8990
Pasyvus-agresyvus klasifikatorius (angl. Passive Aggresive Classifier)	0,8930

Šaltinis: Gonzalez-Carvajal, Garrido-Merchan (2023). Comparing BERT against traditional machine learning text classification, p. 6.

BERT naudoja transformerį - dėmesio mechanizmą (angl. attention mechanism), kuris išmoksta kontekstinius ryšius tarp teksto žodžių. Transformerį sudaro du atskiri mechanizmai – koduotojas (angl. encoder), kuris skaito teksto įvestį, ir dekoderis (angl. decoder), kuris sukuria užduoties prognozę. Kadangi *BERT* tikslas sukurti kalbos modelį, reikalingas tik kodavimo mechanizmas. Skirtingai nuo kryptinių modelių, kurie teksto įvestį skaito nuosekliai iš kairės į dešinę arba iš dešinės į kairę, transformeriais remtas metodas skaito visą žodžių seką iš karto. Ši savybė leidžia modeliui mokytis žodžio konteksto iš visos jo aplinkos (Horev, 2018).

Pasak R. Horev *BERT* gali būti naudojamas įvairioms kalbos užduotims atlikti:

- Klasifikavimo užduotys, pavyzdžiui, nuotaikų ar sentimentų analizė. Šios užduotys atliekamos panašiai kaip ir kito sakinio klasifikavimas, pridedant klasifikavimo sluoksnį prie transformerio išvesties.
- Klausimų atsakymo užduotys. Gavus klausimą apie žodžių ar teksto seką reikia nurodyti sekantį sekos atsakymą. Naudojant *BERT*, klausimų ir atsakymų modelį galima apmokyti išmokstant vos du papildomus vektorius, žyminčius atsakymo pradžią ir pabaigą.
- Įvardytų subjektų atpažinimo (angl. named entity recognition, *NER*) užduotys. Tai užduotys kai, gaunama teksto seka ir joje reikia pažymėti įvairių tipų subjektus, kaip, pavyzdžiui, asmuo, organizacija, data ar pan., esančius analizuojame tekste. Naudojant *BERT*, *NER* modelį galima apmokyti pateikiant kiekvieno simbolio išvesties vektorių į klasifikavimo sluoksnį, kuriame numatoma *NER* etiketė.

Dar dvi labai svarbios *BERT* modelio dalys: išankstinis mokymas (angl. prie-training) ir derinimas (angl. fine-tuning). Pirmoji turėtų suteikti nuo užduoties nepriklausomų žinių, o antroji - išmokyti modelį labiau pasikliauti frazėmis, naudingomis atliekant užduotį. Tyrimai siūlo įvairias metodikas, kad pagerintų *BERT* derinimo procesą. Pavyzdžiui, kai kurie mokslininkai siūlo naudoti daugiau sluoksnių ir derinti informaciją iš gilesnių sluoksnių, o kiti rekomenduoja dviejų etapų tikslinimą, įterpiant papildomą prižiūrimą mokymą tarp išankstinio mokymo ir derinimo etapų (Rogers, Kovaleva, Rumshisky, 2021).

Taip pat išskiriamas ir *mBERT* (angl. multilingual *BERT*), ankščiau minėtų „Google AI“ tyrėjų pristatytas daugiakalbis *BERT* modelis. Pastarasis buvo apmokytas su 104-ių pasaulinių kalbų duomenimis ir gali būti taikomas kaip universalus kalbos modelis, leidžiantis apdoroti NLP

užduotis įvairiomis kalbomis be papildomo mokymo (Libovicky, Rosa ir Fraser, 2019). Šis modelis pasirinktas dėl to, nes yra apmokytas ir lietuvių kalbos duomenimis.

Tuo tarpu *DistilBERT*, tai modelis, kuris iš esmės turi tą pačią bendrą struktūrą kaip ir *BERT*, tačiau yra dar labiau optimizuotas. 2020 m. mokslininkai Victor Sanh, Lysandre Debut, Julien Chaumond ir Thomas Wolf savo tyrime parodė, kad *DistilBERT* gali sumažinti originalaus BERT modelio dydį apie 40 %, tuo pačiu išlaikydamas apie 97 % *BERT* kalbos supratimo gebėjimų. Dėl šios optimizacijos *DistilBERT* ne tik užima mažiau vietos, bet ir veikia apie 60 % greičiau, todėl yra labai tinkamas taikymams, kuriems reikia didelio efektyvumo su nedideliu skaičiavimo resursų naudojimu (V. Sanh et al., 2020).

DistilBERT treniruojamas taikant žinių distiliavimo metodą, kai „studento“ modelis mokosi atkartoti didesnio modelio, vadinamo „mokytoju“, elgesį. Tai atliekama su distiliavimo praradimo funkcija, kuri naudojama lyginant tikimybes tarp „mokytojo“ ir „studento“. Palyginus modelių dydį ir greitį, *DistilBERT* yra žymiai mažesnis ir greitesnis. Jis turi 40 % mažiau parametrų nei *BERT* ir yra 60 % greitesnis. Iniciacijos laiko matavimai rodo, kad DistilBERT yra 71 % greitesnis nei *BERT*, o visas modelis sveria tik 207 MB (Tsang, 2022). Nors *DistilBERT* ir yra lengvesnė ir greitesnė *BERT* versija, atrodo, jog ji praranda dar didesnę dalį jau ribotų *BERT* gebėjimų atspindėti kalbos formalumus, be to, šis modelis rodo stipresnius išankstinius nusistatymus (Staliūnaitė ir Iacobacci, 2020).

Modeliui realizuoti naudojami ir kiti metodai, pavyzdžiui:

1. Slankiojančio lango (angl. sliding window) metodas yra plačiai naudojama technika, siekiant apdoroti įvairių sekų duomenis, kaip, pavyzdžiui, tekstas, vaizdeliai ar skirtingi signalai. Šis metodas leidžia analizuoti ir apdoroti ilgus duomenų segmentus, kurie negali būti tvarkomi kaip vientisas blokas dėl įvairių priežasčių, pavyzdžiui, atminties ar apdorojimo ribojimų. Tad šis metodas pasirinktas tam, kad padėtų išvengti *BERT* 512 tokenų (angl. tokens) ribojimo. Metodui įgyvendinti sukurta funkcija, kuri suskaido ilgus teisinius tekstus į 512 tokenų ilgio dalis ir taiko tam tikrą persidengimą tarp šių dalių, siekiant gauti optimalesnius rezultatus. Toliau, padalinti segmentai klasifikuojami naudojant *BERT* ir gautos prognozės apibendrinamos.
2. TF-IDF (angl. term frequency inverse document frequency) yra skaitinė statistika, kuri parodo raktinių žodžių svarbą tam tikriems dokumentams arba, nurodo tik tuos

raktinius žodžius, pagal kuriuos galima identifikuoti arba suskirstyti tam tikrus dokumentus. Taigi TF-IDF yra dviejų skirtingų žodžių derinys: terminų dažnumo ir atvirkštinio dokumentų dažnumo. TF arba terminų dažnis naudojamas matuoti, kiek kartų ieškomas terminas yra dokumente, o atvirkštinis dokumento dažnis dažniems žodžiams priskiria mažesnį, o ne dažniems žodžiams priskiria didesnį svorį. Visgi pagrindinis TF-IDF apribojimas yra tas, kad algoritmas negali identifikuoti žodžių vos šiek tiek pakitus jų laikui, pavyzdžiui, algoritmas traktuos „eiti“, „ėjo“ ir „eina“ kaip tris skirtingus nepriklausomus žodžius, todėl galiausiai galima gauti netikėtus rezultatus (Qaiser ir Ali, 2018).

3. *SMOTE* arba sintetinės mažumos perteklinės atrankos metodas (angl. synthetic minority over-sampling technique) yra vienas iš pagrindinių įrankių, naudojamų nesubalansuotiems duomenims apdoroti. Šiuo metodu sukuriami nauji sintetinių duomenų pavyzdžiai, kurie gaunami atlikus tiesinę interpoliaciją (angl. interpolation) tarp mažumos klasės duomenų ir jų K artimiausių kaimynų. Visgi svarbu atkreipti dėmesį, kad *SMOTE* sugeneruoti pavyzdžiai ne visada tiksliai atspindi pradinį mažumos klasės pasiskirstymą (Elreedy, Atiya ir Kamalov, 2023).

2.5. Modelio realizavimo ir testavimo žingsniai

Siekiant įgyvendinti šio darbo tikslą ir pasiūlyti finansinių ginčų baigties prognozavimo modelį, jo paruošimas ir realizavimas buvo suskirstytas į šiuos, toliau vardinamus žingsnius:

1. Duomenų paruošimas: Duomenys buvo surinkti iš LB teikiamų duomenų - PDF dokumentai, iš kurių naudojant *pdfplumber* biblioteką buvo išgautas tekstas. Po to buvo atliktas teksto valymas: pašalintos nereikšmingos žodžių formos, tokios kaip nutraukimo žodžiai naudojant *spaCy*. Tam buvo naudojamas būtent lietuvių kalbos modelis *lt_core_news_lg*, kuris lematizavo ir padėjo pašalinti nereikšmingus žodžius. Be to buvo sukurta speciali funkcija, kad būtų galima lengvai identifikuoti ir iš analizuojamų tekstų ištraukti ginčų sprendimų dalis. Tai atlikta remiantis tam tikromis nustatytomis frazėmis, kurios rodo sprendimo dalių pradžią ir pabaigą.

2. Teksto segmentavimas: siekiant tekstą išlaikyti tinkamą BERT modeliui, tekstas išgautas iš PDF dokumentų buvo padalintas į mažesnes dalis naudojant slankiojančio lango metodą (angl. sliding window). Tai užtikrino, kad kiekvienas tekstas - segmentas atitiktų 512 tokenų (angl. tokens) ilgį, kas leido išlaikyti kontekstą tarp padalintų gretimų segmentų.
3. Modelio treniravimas: modelio treniravimui naudojamas iš anksto ištreniruotas *BERT* (vėliau ir kiti modeliai, kaip *DistilBERT*) modelis *bert-base-multilingual-cased*. Treniravimo procesas apėmė kelis žingsnius: *BERT* modelio pritaikymas, klasifikacija, treniravimo parametrų nustatymas. Atsižvelgiant į duomenų kiekį buvo nustatyti tam tikri parametrai, kaip, pavyzdžiui, mokymosi sparta (angl. learning rate), kuri pasirinkta naudoti mažesnė nei įprastai, kad modelis per daug neprisitaikytų. Taip pat pridėtas ir svorio nurašymas (angl. weight decay), siekiant sumažinti perpratimo riziką, reguliuojant svorius. Įtrauktas ir ankstyvas sustabdymas (angl. early stopping), kad treniravimas būtų sustabdytas, jei per kelias iteracijas vertinimo metrikos nepagerėtų.
4. Svorių peržiūra: atsižvelgiant į didelį klasių disbalansą: 157 teigiamų, 146 dalinių teigiamų ir 1337 neigiamų sprendimų, apsvarstytas klasės svorių naudojimas modelio praradimo funkcijoje (angl. loss function). Taip modelis turėtų geriau prisitaikyti prie nevienodo duomenų pasiskirstymo ir išvengti pernelyg didelio prisitaikymo dominuojančiai klasei. Klasės svoriai gali būti nustatyti tiek automatiškai tiek rankiniu būdu pagal atvirkštinį klasių pasiskirstymą. Modelis *DistilBERT* ir dauguma kitų *transformers* modelių leidžia naudoti funkciją *class_weights*, juos nustatant praradimo funkcijoje. Svoriai apskaičiuoti pagal paprastą formulę:

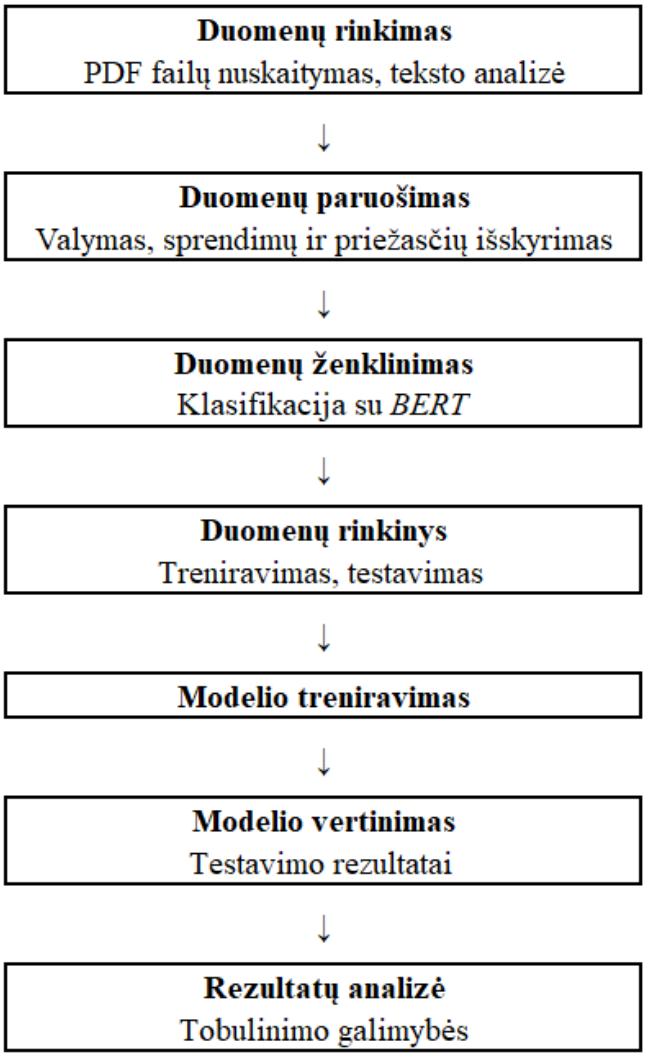
$$svoris_i = \frac{bendras\ pavyzdžių\ skaičius}{klasės\ i\ pavyzdžių\ skaičius} \quad (1)$$

5. Modelio vertinimas: kad įvertinti modelio veikimo efektyvumą ir tikslumą apskaičiuoti svarbiausi rodikliai: tikslumas (angl. accuracy), rodantis, kiek teisingų prognozių modelis atliko iš visų duomenų rinkinio. Svertinis F1 balas (angl. weighted F1), kuris apjungia tikslumą ir atkūrimą (angl. recall).

Žemiau, 2 paveiksle, pateikiama ir vaizdinė modelio įgyvendinimo diagrama, kuri vizualizuoja modelio realizavimo ir testavimo procesą. Šis procesas apima septynis pagrindinius žingsnius. Pirmasis žingsnis — duomenų rinkimas. Jame atliekamas PDF failų nuskaitymas ir teksto analizė, siekiant išgauti reikalingą informaciją. Antrajame žingsnyje vyksta duomenų paruošimas, kuomet surinkti duomenys yra valomi, o sprendimai ir priežastys išskiriamos iš teisinių tekstų. Trečiame žingsnyje, kuris apibrėžiamas kaip duomenų ženklinimas, paruošti duomenys yra klasifikuojami naudojant *BERT* modelį.

Kiekvienam sprendimui priskiriama atitinkama klasė (angl. label), pavyzdžiui, teigiama, neigiama ar dalinai teigiama. Ketvirtame žingsnyje paruošti duomenys padalijami į treniravimo, validavimo ir testavimo rinkinius, kurie bus naudojami tolesniame modelio treniravimo procese. Penktame etape modelis mokomas atpažinti skirtingus sprendimų tipus, pasitelkiant paruoštus duomenis ir *Hugging Face Trainer*. Po treniravimo šeštame žingsnyje modelis vertinamas testavimo rinkinyje, leidžiančiame įvertinti jo našumą, skaičiuojant tikslumą, F1 balą ir kitas metrikas. Galiausiai, septintame etape atliekama rezultatų analizė, kurioje nagrinėjami modelio veiklos rezultatai ir nustatomos galimos tobulinimo kryptys.

2 pav. Modelio realizavimo diagrama



Šaltinis: sudaryta autorės

Taigi, kiekvienas šis žingsnis yra labai svarbus ir prisideda prie galutinio modelio efektyvumo, todėl juos reikia kruopščiai suplanuoti ir įgyvendinti. Tik tuomet galime pasiekti norimus rezultatus ir užtikrinti, kad modelis veiktų taip, kaip numatyta.

3. FINANSINIŲ PASLAUGŲ BAIGTIES PROGNOZAVIMO MODELIO REALIZAVIMAS

Šioje darbo dalyje aprašomas trečiojo uždavinio sprendimas. Skyriuje (3.1.) aptariami siūlomo modelio pradiniai įgyvendinimo etapai, o tolimesnėje dalyje (3.2.) pristatomas ir siūlomas modelis. Taip pat šioje dalyje (3.3.) aptariami ir palyginami gauti finansinių ginčų baigties prognozavimo modelio rezultatai.

3.1. Duomenų užkrovimas, analizė

Kaip ir aprašyta 2.1 poskyryje duomenys gauti iš LB viešai pateiktų finansų ginčų sprendimų. Šį duomenų rinkinį sudarė 1640 PDF failai, kuriuos pavyko nuskaityti, o iš jų rekomenduojami sprendimai pasiskirstę taip, kaip buvo nurodyta 1-oje lentelėje: 1337 siūlymų atmesti, 146 siūlymai dalinai patenkinti ir 157 siūlymai patenkinti ieškovo prašymą. PDF failai sukelti į „Google drive“ ir atlikus tam tikrus žingsnius išsaugoti CSV formatu, kad būtų patogu juos naudoti tolimesniuose žingsniuose.

Sėkmingai teksto analizei atlikti tekstas taip pat išvalytas: naudojant reguliariųjų išraiškų - *regex* metodus teksto valymui buvo panaikinti papildomi tarpai, specialieji simboliai, nereikšmingos frazės ar žodžiai , tokie kaip „ir“, „dėl“, „kai“, „kaip“ ir kt.. Tekstas taip pat buvo konvertuotas tik į mažąsias raides. Tai leido sutelkti dėmesį į esminius žodžius ir frazes, galinčias turėti įtakos modelio mokymui.

3.1.1. Sentimentų analizė

Naudojant *BERT* modelį sentimentų analizei, buvo siekiama kiekvieną sprendimo fragmentą priskirti vienai iš trijų kategorijų: „atmesti“, „dalinai patenkinti“ arba „patenkinti“. Tam naudotas iš anksto apmokytas *BERT* modelis, specialiai pritaikytas daugiakalbei sentimentų analizei. Šis modelis padėjo ne tik įvertinti dokumentų sentimentus (žr. 3 lentelę), bet ir nustatyti kiekvieno sprendimo polinkį, atsižvelgiant į ieškovo prašymą.

3 lentelė

Sentimentų analizės rezultatai

Sentimentas	Kiekis
Neutralus	1050
Teigiamas	325
Labai teigiamas	158
Neigiamas	74
Labai neigiamas	33

Šaltinis: sudaryta autorės

Atlikus sentimentų analizę su *BERT*, gauti sprendimų sentimentų duomenys rodo, kaip dažnai įvairiuose ginčuose pasitaiko teigiami, neigiami ar neutralūs sprendimai. Daugiausiai sprendimų, net 1050, pasižymi neutraliu sentimentu, tai reiškia, kad šie sprendimai nebuvo itin palankūs ieškovui ar priešingai – neigiami atsakovui. Neutralus vertinimas gali rodyti objektyvų požiūrį į abi ginčo šalis ir pagrindinių šalių interesų subalansavimą. Tai taip pat gali reikšti, kad daugelis atvejų sprendžiami laikantis procedūrinio teisingumo principų, nedarant reikšmingo poveikio nė vienai šaliai. Teigiamas ir labai teigiamas sentimentas taip pat pasitaiko gana dažnai (atitinkamai 325 ir 158 kartai), kas rodo palankesnį sprendimą ieškovui ar bent jau dalinį reikalavimo patenkinimą. Šie skaičiai gali atspindėti atvejus, kuriuose ginčas išspręstas taip, kad ieškovas gauna tam tikrą naudingą rezultatą, net jei ir ne visapusišką. Neigiamas ir labai neigiamas sentimentas pasitaiko rečiausiai (atitinkamai 74 ir 33 kartai), kas reiškia, jog šie ginčai dažniausiai nesprendžiami ieškovo naudai. Tokio tipo sprendimai gali rodyti atvejus, kai ieškovo reikalavimai neatitinka teisinės ar faktinės bylos aplinkybių, todėl ginčas baigiasi atmetimu.

Ši sentimentų pasiskirstymo analizė suteikia svarbių įžvalgų, kaip Lietuvos banko sprendimuose atsižvelgiama į ieškovų ir atsakovų interesus. Daugumos sprendimų neutralus pobūdis atspindi subalansuotą požiūrį į ginčo šalis, o santykinai nedidelis neigiamų sprendimų skaičius leidžia daryti prielaidą, kad dažniau stengiamasi ieškovams suteikti tam tikras teigiamas sprendimo baigtis.

Deja, finansų ginčų atveju, sentimentų analizė turi ribotas galimybes prognozuoti faktinę baigtį, nes finansų teisė dažnai grindžiama objektyviais kriterijais, o ne subjektyviais ar emociniais

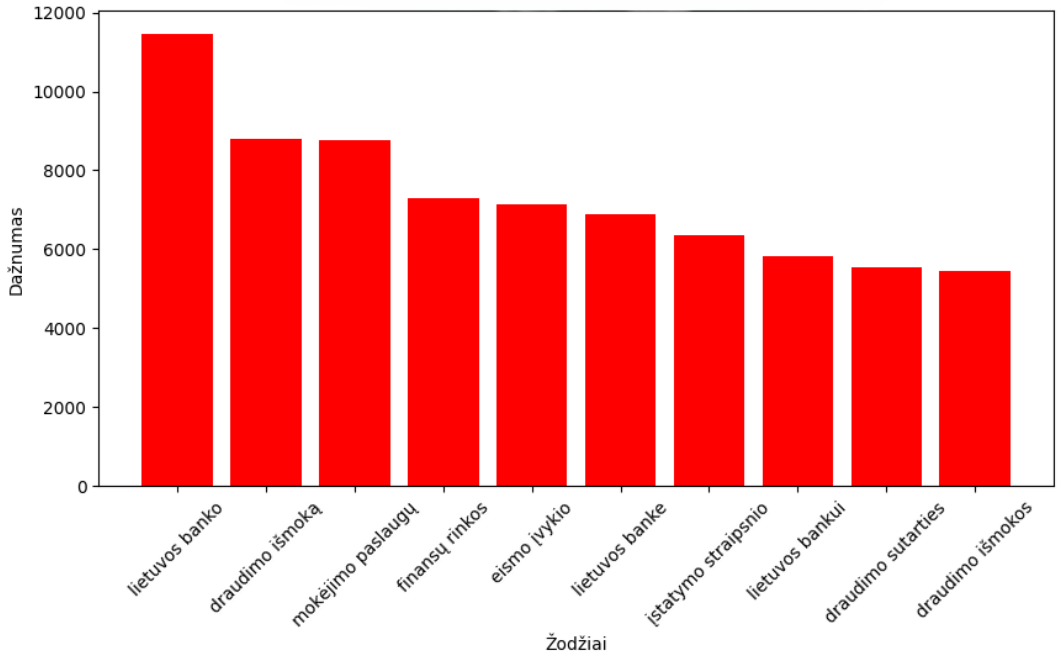
tekstiniais požymiais. Todėl sentimentų analizę galima vertinti kaip pagalbinį įrankį, suteikiantį papildomų įžvalgų apie sprendimų toną, bet ne apie tiesiogines finansines pasekmes.

3.1.2. Įvairių dydžių frazių analizė

Gilesnei tekstų analizei pasitelkta n-gramų analizė, atlikta siekiant identifikuoti dažniausiai pasitaikančias frazes Lietuvos Banko ginčų sprendimuose. Šio proceso metu buvo išskirtos dažniausiai pasikartojančios žodžių sekos, siekiant labiau suprasti dominuojančias temas ir kalbos struktūrą nagrinėtuose tekstuose.

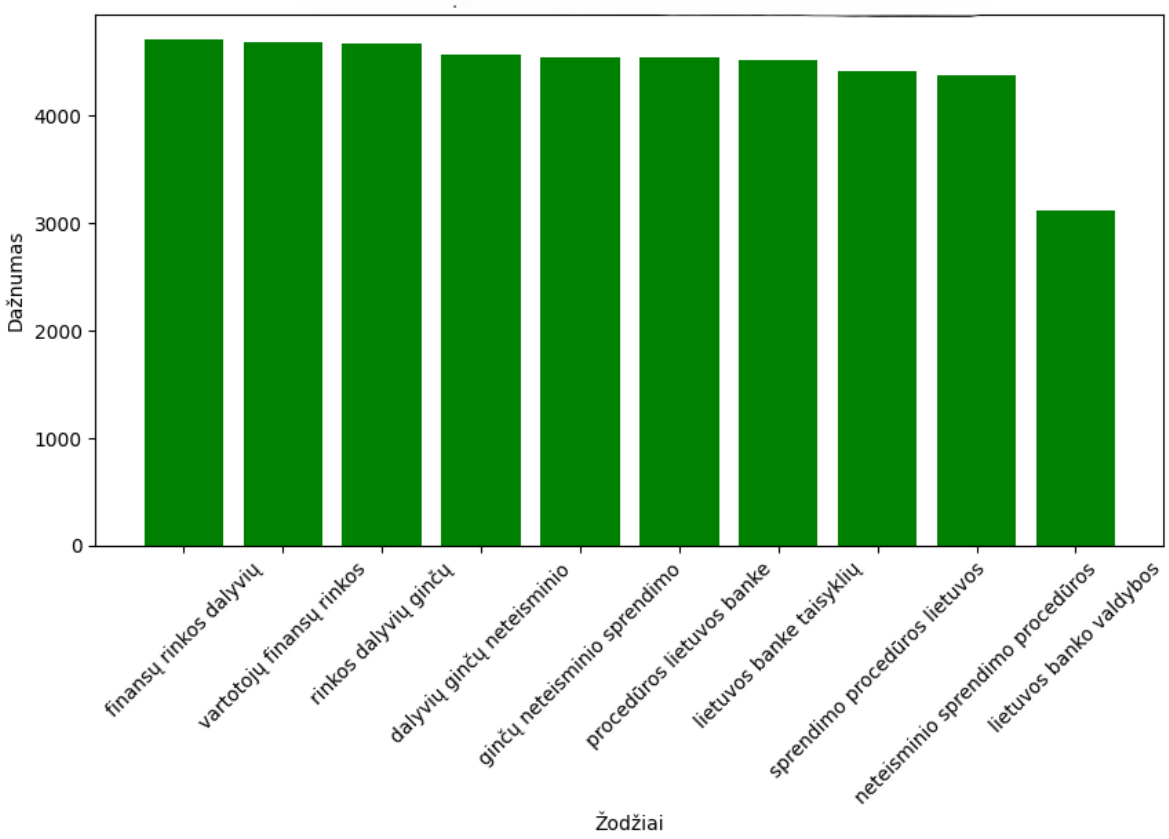
N-gramų analizėje buvo sukurtos dviejų, trijų keturių žodžių junginių sekos, siekiant atskleisti dažniausiai vartojamus žodžių junginius ginčų sprendimų kontekste. N-gramų kūrimui naudotas *CountVectorizer* modulis, kuris sukūrė žodžių sekų sąrašus iš paruošto teksto. Po to, kiekvienai n-gramų sekai, buvo apskaičiuotas pasikartojimų dažnis. Naudojant *matplotlib* biblioteką, vizualizuotos dažniausiai pasikartojančios frazės kiekvienai n-gramai, o jų dažniai išdėstyti mažėjančia tvarka. Tokia vizualizacija leido lengvai pastebėti dažniausiai pasitaikančias frazes, kurios galėtų būti naudingos identifikuojant ginčų sprendimams būdingus kalbinius modelius. Atitinkamai, šios vizualizacijos pateikiamos 3-iame, 4-ame ir 5-ame paveiksluose.

3 pav. Dažniausiai pasitaikančios 2-jų žodžių frazės (bigramos)



Šaltinis: sudaryta autorės

4 pav. Dažniausiai pasitaikančios 3-jų žodžių frazės (trigramos)

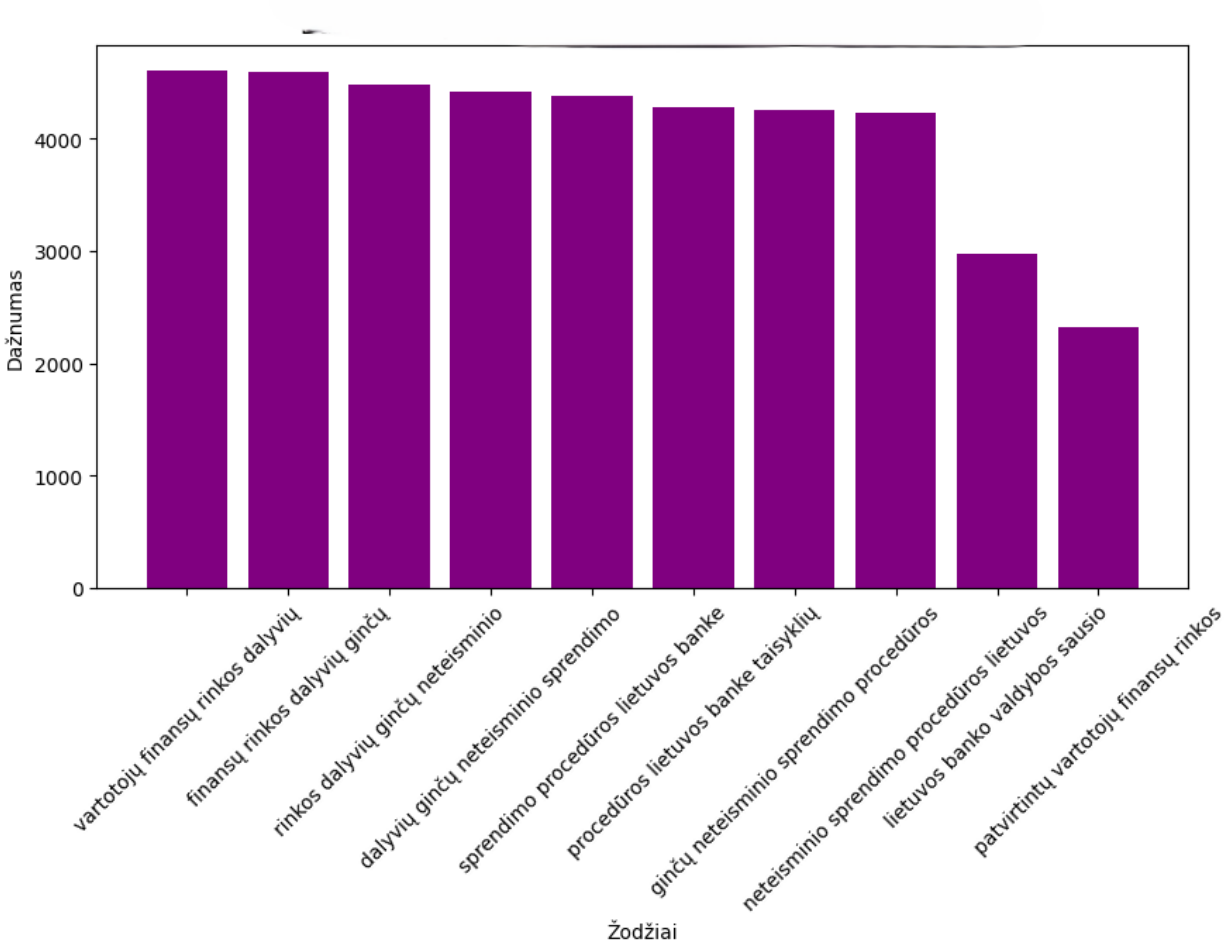


Šaltinis: sudaryta autorės

Analizuojant bigramas (žr. 3-ią paveikslą), išryškėjo frazės, tokios kaip „lietuvos banko“, „draudimo išmoką“ ir „mokėjimo paslaugų“, kurios dažniausiai kartojosi dokumentuose. Tai rodo, kad LB ir finansų sektorius yra svarbūs analizuojamose dokumentuose. Kai pažvelgiame į trigramas (žr. 4-ą paveikslą), pavyzdžiui, „finansų rinkos dalyvių“ ir „vartotojų finansų rinkos“, matome, kad daugiausia dėmesio skiriama dalyviams ir jų santykiams su finansų rinka, kas ypač aktualu nagrinėjant ginčų sprendimą.

Be to, keturgramos (žr. 5-ą paveikslą), tokios kaip „vartotojų finansų rinkos dalyvių“ ir „dalyvių ginčų neteisminio sprendimo“, atskleidžia sudėtingesnius žodžių junginius, kurie gali būti itin svarbūs analizuojant teisines ir reguliacines problemas finansų srityje.

5 pav. Dažniausiai pasitaikančios 4-ių žodžių frazės (keturgramos)



Šaltinis: sudaryta autorės

Apskritai šios n-gramos ne tik leidžia geriau suprasti, apie ką kalbama dokumentuose, bet ir padeda pamatyti, kaip skirtingos suinteresuotosios šalys yra tarpusavyje susijusios ir kokį vaidmenį jos atlieka finansų sistemoje. Detaliau, t. y. kokie žodžiai dažniau pasitaiko atitinkamose klasėse (esant skirtingiems sprendimams) galima matyti paveikslėlyje žemiau (žr. 6-ą paveikslą).

6 pav. Dažniausiai pasitaikantys žodžiai skirtingoms klasėms

Šaltinis: sudaryta autorės

6 paveiksle pateikti dažniausiai pasikartojantys žodžiai rodo, kokie aspektai bei temos dažniausiai aktualūs konkrečių sprendimų priėmimo procese. Kiekviena išskirta sprendimo klasė turi sau aktualesnius žodžius, kurie atspindi sprendimų esmę ar priežastis. Matyti, jog „Atmesti“ klasė, kur priimamas nepalankus sprendimas pareiškėjui, labiausiai susijusi su ginčais, bankais ir draudimo sąlygomis. Tuo tarpu klasėje „Patenkinti“ vyrauja žodžiai, susiję su teisingu draudimo sutarties vykdymu ir įvykiais, kurie buvo apmokėti pagal taisykles. Klasė „Dalinai patenkinti“ apima sprendimus, kuriuose dažnai pasikartoja žodžiai dalinė nauda ar patenkinimas ir dažniausiai kyla dėl nesutarimų, dėl apmokėjimo sąlygų ar draudimo sąlygų.

Tokie žodžių analizės duomenys yra ypač naudingi siekiant geriau suprasti, kokie aspektai turi didžiausią įtaką sprendimo priėmimui. Taip pat tam, kad būtų galima ir toliau kiekvienoje iš trijų klasių tobulinti sprendimų priėmimo procesus ar numatyti kitus, ateityje galimus ginčus.

3.2. Modelio gautų rezultatų apibendrinimas

Šiame poskyryje bus aptartas modelis ir jo variacijos, kurios skirtos finansų ginčo baigties klasifikavimo – prognozavimo užduočiai spręsti.

3.2.1. *BERT* ginčo baigties prognozavimo modelis

Pirmiausia kodo vykdymo metu buvo įdiegtos reikiamos, anksčiau minėtos bibliotekos – *transformers*, *datasets*, *pandas*, *scikit-learn*, ir kt., kurios užtikrina *BERT* modelio veikimą. Taip pat įdiegtas *transformers* paketas, kuris suteikia prieigą prie daugelio jau paruoštų modelių, skirtų natūraliosios kalbos apdorojimui, tarp jų ir *BERT*. Biblioteka *datasets* panaudota duomenų rinkinių importavimui ir valdymui, o *pandas* ir *scikit-learn* buvo reikalingos duomenų apdorojimui ir modelio metrikų skaičiavimui.

Duomenys buvo įkelti, kaip tekstai iš CSV failų su stulpeliais, žyminčiais sprendimų tekstus, priežastis ir etiketes. Šie tekstai, vėliau buvo panaudoti *BERT* modelio mokymui ir validavimui. Nurodytos klasifikavimo ar klasių etiketės (*Label* stulpelis) pažymi, ar vartotojo pateiktą reikalavimą reikėtų patenkinti, todėl šis uždavinys gali būti priskirtas binariniui klasifikavimui.

Siekiant sukurti gerą modelį buvo pasirinktos atitinkamos hiperparametrų reikšmės: mokymosi greitis (angl. learning rate) – 0,00001. Tai populiarus greičio dydis, naudojamas *BERT* modeliuose, užtikrinantis stabilų ir efektyvų mokymąsi. Atsižvelgiant į modelio stabilumo stebėjimą pasirinktos 3 epochos. Kad būtų išvengta per didelio apsimokymo (angl. overfitting) ir pagerėtų konvergencija pasirinktas optimizavimo algoritmas *AdamW*.

Toliau buvo įkeltas *BERT* modelis *bert-base-multilingual-cased* su paruoštu tokenizatoriumi (pastarasis paverčia tekstus į žodžių ID sekas, kurios toliau naudojamos modelyje). Treniruojant modelį, buvo naudojami *TrainingArguments*, kurie nurodo hiperparametrus ir treniravimo sąlygas:

- Epochos (angl. epochs) - nusako, kiek kartų modelis pereis visą treniravimo duomenų rinkinį;
- Vertinimo strategija (angl. evaluation strategy) – periodiškai vykdoma modelio validacija.

Kad būtų galima stebėti treniravimo eigą, buvo naudojamas *wandb* įrankis, leidžiantis realiu laiku analizuoti treniravimo metrikas, tokias kaip apmokymo nuostoliai (angl. training loss), validavimo nuostoliai (angl. validation loss), tikslumas (angl. accuracy) ir F1.

Taigi treniruojant modelį per 3-is epochas buvo gauti rezultatai, pateikti 4-oje lentelėje.

4 lentelė

***BERT* modelio treniravimo rezultatai**

Epocha	Tikslumas	F1	Validavimo nuostoliai
1	0,817073	0,734817	0,613268
2	0,817073	0,734817	0,643764
3	0,817073	0,734817	0,625804

Šaltinis: sudaryta autorės

Pirmoje epochoje tikslumas siekė 81,7 %, o F1 balas buvo 0,7348. Treniruojant modelį toliau, antroje ir trečioje epochoje validavimo tikslumas ir F1 balas išliko stabilūs. Tai gali rodyti, kad modelis pradėjo „užsistovėti“ ir nepagerėjo. Tuo tarpu patenkinamas F1 balas rodo, kad modelis gana gerai atpažįsta klasifikacijos klases, tačiau dar yra vietos tobulėti.

Pasibaigus *BERT* modelio treniravimui, modelio vertinimas buvo atliktas naudojant *eval_loss, eval_accuracy* ir *eval_f1* parametrus. Vertinimo rezultatai pateikti 5-oje lentelėje.

5 lentelė

***BERT* modelio rezultatai**

Eval_loss	Eval_accuracy	Eval_f1
0,61639	81,7 %	0,7348

Šaltinis: sudaryta autorės

Modelio vertinimas yra esminė proceso dalis, kuri leidžia suprasti, kaip gerai mokomas modelis atlieka savo užduotį. Tai ypač svarbu klasifikavimo uždaviniams spręsti, kai tikslumas ir klaidos įvertinimas gali turėti didelės reikšmės. *Eval_loss* arba validavimo nuostolių vertė nurodo modelio klaidų skaičių, remiantis jam pateiktais duomenimis. Gauta vertė - 0,616 yra vidutinė ir ji rodo, kad modelis dar turi erdvės tobulėti.

Tuo tarpu *Eval_accuracy* arba validavimo tikslumas nurodo, koks procentas modelio prognozių buvo teisingas. Gautas 81,7 % tikslumas rodo, kad modelis didžiąją dalį laiko sėkmingai prognozuoja reikiamas etiketes. Tai gana geras rezultatas, ypač natūralios kalbos apdorojimo užduotyse. Tačiau tuo pat metu tai rodo, kad 18,3 % atvejų modelis klysta. Tai gali būti reikšminga, ypač kai modelis taikomas teisinėje srityje, kur netinkamos prognozės gali turėti rimtų pasekmių.

Atsižvelgiant į tai kad duomenų klasės yra išbalansuotos, svarbu atsižvelgti ir į *Eval_f1* vertę. Šis balas suteikia geresnį bendrą vaizdą apie modelio efektyvumą, palyginti su paprastu tikslumu. F1 balas, kuris gautas 73,48 % rodo gan gerą *BERT* modelio balansą.

Siekiant dar giliau išanalizuoti *BERT* modelį (šio modelio pavyzdinis kodas pateiktas priede) buvo panaudota 3 lygių kartotinė kryžminė validacija. Atsižvelgiant į didelį disbalansą tarp klasių šis metodas turėtų suteikti tikslesnį vertinimą, mat modelis testuojamas su skirtingais duomenų pogrupiais. Duomenys skaidomi į 3-ias dalis ar lygius (angl. folds), kad leistų įvertinti modelio veikimą su įvairiais duomenų pogrupiais. Šiam bandymui taip pat nustatytos 3 epochos, mokymosi greitis išlieka 0,00001. Įvertinus tarpinius rezultatus (žr. 6 lentelę) apskaičiuojami ir vidutiniai rodiklių dydžiai (žr. 7 lentelę).

6 lentelė

BERT **modelio 3 lygių kryžminės validacijos rezultatai epochoms**

Lygis	Epocha	Tikslumas	F1	Validavimo nuostoliai
1	1	0,815850	0,731954	0,649135
1	2	0,815850	0,731954	0,618176
1	3	0,815850	0,731954	0,601690
2	1	0,816850	0,734506	0,633698
2	2	0,816850	0,734506	0,611354
2	3	0,816850	0,734506	0,593771
3	1	0,816514	0,734038	0,624371
3	2	0,816514	0,734038	0,631237
3	3	0,816514	0,734038	0,606729

Šaltinis: sudaryta autorės

6-oje lentelėje pateikti kiekvienam kryžminės validacijos lygiui ir epochai atitinkantys rezultatai. Tikslumo rodiklis visomis epochomis išlieka labai stabilus, svyruoja nuo 81,5 % iki 81,7 %. Tai rodo, kad modelis užtikrintai prognozuoja didžiąją dalį teisingų sprendimo klasių. F1 balas taip pat gana stabilus ir svyruoja nuo 0,731 iki 0,734. Tai rodo gerą pusiausvyrą tarp tikslumo ir atsikvėpimo (angl. recall), kas yra ypač svarbu turint išbalansuotas klases. Tuo tarpu nuostolių reikšmės kiekviename lygyje mažėja tolygiai. Tai rodo, kad modelis nuosekliai mokosi mažinti klaidų skaičių, nors reikšmių skirtumai ir yra nedideli.

7 lentelė

BERT **modelio 3 lygių kryžminės validacijos rezultatai**

Vidutinis tikslumas	Vidutinis F1 balas
0,8161± 0,0008	0,7335± 0,0011

Šaltinis: sudaryta autorės

7 lentelėje pateikti vidutiniai rezultatų rodikliai per visus lygius, kartu su paklaidomis. Vidutinis tikslumas siekia 81,61 % su ± 0,0008 paklaida, o tai rodo, kad modelio tikslumas išlieka stabilus nepriklausomai nuo lygio. Vidutinis F1 balas yra 0,7335 su ± 0,0011 paklaida, o tai rodo panašų stabilumą tarp lygių ir patvirtina modelio gebėjimą atskirti skirtingas klases.

Lyginant modelio rezultatus naudojant kryžminę validaciją ir ne, galima pastebėti, kad pastarosios taikymas modelio rezultatams reikšmingos įtakos neturėjo. Kryžminės validacijos metodas leido užtikrinti stabilesnį modelio našumą, tačiau šiek tiek sumažino tikslumą (nuo 81,7 % iki 81,6 %) ir F1 balą (nuo 0,734 iki 0,731).

Nors *BERT* modelio pasiekti rezultatai gan geri ir priimtini, jie taip pat rodo, kad modelis dar gali būti tobulinamas. Pirmiausia pasirinktas būdas tolesniam eksperimentavimui: kito modelio pritaikymas. Tam pasirinktas kiek „lengvesnis“ *BERT* modelio variantas *DistilBERT*. Kaip minėta 2 skyriuje šis modelis veikia greičiau ir naudoja mažiau resursų.

3.2.2. *DistilBERT* ginčo baigties prognozavimo modelis

Priešingai nei 3.2.1. skyrelyje vietoje *bert-base-multilingual-cased BERT* modelio variacijos buvo panaudotas *distilbert-base-multilingual-cased*. Šio eksperimento tikslas buvo įvertinti, ar ši „lengvesnė“ *BERT* modelio versija gali pasiekti panašius rezultatus, kaip ir pilnas modelis.

Duomenų paruošimas bei skaidymas identiškas pirmajam eksperimentui, t. y. BERT ginčo baigties prognozavimo modeliui. Atitinkamai klasės, kaip ir pirmu atveju, sužymėtos skaičiais 0,1 ir 2. Svoriai bei parametrai taip pat išlaikyti tokie patys. Taigi mokymai, paremti tokiais pačiais parametrais pademonstravo rezultatus, pateiktus 8-oje lentelėje.

8 lentelė

***DistilBERT* modelio treniravimo rezultatai**

Epocha	Tikslumas	F1	Validavimo nuostoliai
1	0,811073	0,714517	0,598427
2	0,811073	0,714517	0,601676
3	0,811073	0,714517	0,604864

Šaltinis: sudaryta autorės

Tuo tarpu modelio rezultatai (žr. 9 lentelę) taip pat gan panašūs į pirmojo eksperimento rezultatus. *Eval_loss* arba validavimo nuostolių gauta vertė - 0,6030 taip pat vidutinė ir rodo, kad šis modelis dar gali būti tobulinamas. *Eval_accuracy* arba validavimo tikslumas nurodo gautas 81,1 % tik 0,6 % mažesnis nei naudojant *BERT* modelį (81,7 %). *Eval_f1* vertė lygi 71,45 % ir yra 2 % procentais mažesnė nei *BERT* atveju - 73,48 %. Akivaizdu, kad skirtumai yra labai nedideli, tad realiuose taikymuose šie skirtumai yra neesminiai, ypač jei svarbi greitesnė ir resursus taupanti modelio versija.

Modelio treniravimo metu galima pastebėti, kad skirtingose epochose modelio veikimas išlieka stabilus, o tai rodo, kad modelis taip pat gali patirti „užsistovėjimą“, kaip ir *BERT* modelis.

9 lentelė

***DistilBERT* modelio rezultatai**

Eval_loss	**Eval_accuracy**	**Eval_f1**
0,6030	81,1 %	0,7145

Šaltinis: sudaryta autorės

Nepaisant minimaliai prastesnių rezultatų *DistilBERT* pasižymi keliais privalumais:

- Greitesnis treniravimas: svarbus realaus laiko taikymams;
- Mažesnės resursų sąnaudos: naudoja mažiau kompiuterio išteklių tad lengviau pritaikomas;
- Išlaikytos *BERT* savybės: pateikia panašius rezultatus, nors ir yra optimizuotas.

Toliau, vėl pamėginta panaudoti kiek didesnę 5 lygių kartotinę kryžminę validaciją, kad dar tiksliau įvertinti *DistilBERT* modelį. Duomenys skaidomi per 5-ias skirtingas dalis, o kiekviena dalis apdorojama atskirai. Daugiau lygiu pasirinkta todėl, nes modelis veikia „greičiau“ ir tai sutaupė laiko resursų. Vėlgi nustatytos 3 epochos, mokymosi greitis išliko 0,00001. Galiausiai gauti rezultatai apibendrinami, kad gauti bendrą išvadą apie tai, kaip veikia *DistilBERT* modelis naudojant kitą vertinimo metodą.

10 lentelė

***DistilBERT* modelio 5 lygių kryžminės validacijos rezultatai epochoms**

Lygis	Epocha	Tikslumas	F1	Validavimo nuostoliai
1	1	0,810024	0,71057	0,631211
1	2	0,810024	0,71057	0,608742
1	3	0,810024	0,71057	0,606443
2	1	0,812073	0,714817	0,631773
2	2	0,812073	0,714817	0,604744
2	3	0,812073	0,714817	0,603211
3	1	0,811514	0,714038	0,606729
3	2	0,811514	0,714038	0,611234
3	3	0,811514	0,714038	0,608446
4	1	0,811514	0,714038	0,610485
4	2	0,811514	0,714038	0,601229
4	3	0,811514	0,714038	0,593866
5	1	0,811514	0,714038	0,609784
5	2	0,811514	0,714038	0,60591
5	3	0,811514	0,714038	0,605556

Šaltinis: sudaryta autorės

10 lentelėje pateikti išsamūs validavimo rezultatai skirtingoms epochoms ir lygiams. Lentelėje galima matyti, kad kiekvienoje iš trijų epochų ir kiekviename lygyje tikslumo rodikliai išliko nuoseklūs ir kito nedaug. Pavyzdžiui, pirmame lygyje modelis visose epochose pasiekė 0,764 tikslumą ir 0,6805 F1 balą. Atitinkamai 3, 4 ir 5 lygiuose nei tikslumas nei F1 balas visai nepakito, o kiek „judėjo“ nuostolių dydis. Validavimo nuostoliai per epochas mažėjo, kas galėtų reikšti, kad modelis kiekvienoje epochoje gerėjo.

Galutiniai kryžminės validacijos rezultatai pateikti 11-oje lentelėje, kurioje matyti, kad vidutinis modelio tikslumas buvo lygus 0,8111 ± 0,0011, o F1 balas — 0,7135 ± 0,0015. Maži dispersijos rodikliai rodo, kad modelis gan stabilus ir geba veikti skirtinguose validacijos lygiuose. Be to, rezultatai rodo, jog šiam bandymui pasirinkti modelio hiperparametrai (angl. hyperparameters) leidžia pasiekti stabilų ir patikimą rezultatą.

11 lentelė

***DistilBERT* modelio 5 lygių kryžminės validacijos rezultatai**

Vidutinis tikslumas	Vidutinis F1 balas
0,8111 ± 0,0011	0,7135 ± 0,0015

Šaltinis: sudaryta autorės

Taigi įvertinus gautus rezultatus galima pastebėti, kad *DistilBERT* modelis pasižymi beveik tokiais pačiais tikslumo ir F1 balo rezultatais kaip ir *BERT* modelis, tačiau jį galima greičiau treniruoti, o jo veikimui reikia mažiau kompiuterio resursų. Maži tikslumo ir F1 balo skirtumai tarp šių modelių reiškia, kad *DistilBERT* yra praktiškas pasirinkimas, jei svarbi spartesnė, mažiau resursų reikalaujanti alternatyva, ypač realaus laiko sistemose.

3.2.3. Ginčo baigties prognozavimo modelis naudojant logistinę regresiją

Įvertinti aukščiau pateiktų ir aprašytų metodų gebėjimams ir palyginti *BERT* bei *DistilBERT* taikymo galimybes atliktas „paprastesnis“ bandymas. Pastarasis yra supaprastina klasifikavimo užduotis, kur vietoje sudėtingų metodų, kaip, pavyzdžiui *BERT* transformatoriai, naudojamas paprastas požiūris su TF-IDF ir logistine regresija. Pagrindinis šio metodo tikslas - įvertinti paprastą tekstų apdorojimo būdą prieš pereinant prie pažangesnių metodų.

Pirmiausia naudojamas paketas *pdfplumber*, kuris leidžia nuskaityti ir apjungti tekstą iš analizuojamų PDF dokumentų. Toliau, teksto pavertimui skaitine forma naudojamas *TfidfVectorizer*, kuris kuria teksto reprezentaciją pagal žodžių svarbą (TF-IDF) rinkinyje ir taip leidžia išgauti svarbiausias tekstų savybes. Pagal šias savybes pirmiausia apmokomas paprastas logistinės regresijos modelis.

12 lentelė

Logistinės regresijos modelio rezultatai

	Tikslumas	Atkūrimas	F1 balas
Atmesti	0,84	0,79	0,82
Dalinai patenkinti	0,16	0,21	0,18
Patenkinti	0,16	0,19	0,17

Šaltinis: sudaryta autorės

Lentelėje aukščiau (žr. 12 lentelę) pateikti rezultatai rodo, kaip modeliui pavyko klasifikuoti skirtingas klases. Akivaizdu, kad modelis geriausiai klasifikuoja „Atmesti“ klasę - 84 % tikslumu, tačiau patiria sunkumų su „Dalinai patenkinti“ ir „Patenkinti“ klasėmis ir šias geba prognozuoti tik 16 % tikslumu. Panašias išvadas galima daryti ir iš kitų dydžių, o tai rodo, kad bendras modelio našumas nėra geras, nes modelis blogai atpažįsta mažesnes klases. Tai gali būti didelio klasių disbalanso pasekmė. Bendras logistinės regresijos modelio pasiektas rezultatas: 68,29 % tikslumas ir 69,85 % F1 balo reikšmė rodo vidutinio lygio rezultatus, tačiau atsižvelgiant į tai, kad duomenys nesubalansuoti tai gali reikšti, kad modelis klasifikuoja tik dominuojančią klasę ir taip pasiekia aukštus rezultatus.

13 lentelė

Logistinės regresijos modelio rezultatai naudojant *SMOTE*

	Tikslumas	**Atkūrimas**	**F1 balas**
Atmesti	0,83	0,87	0,85
Dalinai patenkinti	0,19	0,16	0,14
Patenkinti	0,21	0,19	0,17

Šaltinis: sudaryta autorės

Logistinės regresijos modelio, pritaikius *SMOTE* metodą, rezultatai: 73,17 % tikslumas ir 71,58 % F1 balo reikšmė. Nors bendras tikslumas padidėjo (73,17 %), F1 balo reikšmės mažesnėse klasėse vis tiek išlieka mažos (žr. 13 lentelę), ypač „Dalinai patenkinti“ ir „Patenkinti“ klasėse. Šios klasės vis tiek neturi pakankamai informacijos ir išlieka sunkiai prognozuojamos, net su dirbtiniu pavyzdžių generavimu. Trumpai tariant, *SMOTE* taikymas padeda šiek tiek pagerinti klasių prognozes, tačiau net ir po jo taikymo nedominuojančios klasės išlieka sunkiai atpažįstamos.

Taigi, paprastesnis teksto analizės metodas su TF-IDF ir logistine regresija nėra pakankamai stiprus, kad sėkmingai prognozuotų reikiamas išvadas. Sunku pasiekti gerus rezultatus su šiais metodais be papildomų pažangių priemonių, tokių kaip *BERT* ar *DistilBERT*, kurie geriau atpažįsta kontekstą ir gali efektyviau apdoroti natūralios kalbos ypatybes.

3.3. Rezultatų apžvalga ir rekomendacijos modelio tobulinimui

Atlikus keletą eksperimentų buvo įvertinti skirtingi modeliai ir iš to suformuluotos tam tikros išvados bei idėjos, kaip būtų galima toliau tobulinti ginčų baigties prognozavimo modelį. Žemiau, 14-oje lentelėje, pateikiama gautų modelių rezultatų apžvalga, siekiant išsiaiškinti, kurie modeliai pasiekė geriausius rezultatus ir kokios būtų galimos modelio tobulinimo kryptys.

14 lentelė

Modelių rezultatų palyginimas

Modelis	Tikslumas (%)	F1 balas
BERT modelio rezultatai	81,70%	0,7348
BERT modelio 3 lygių kryžminės validacijos rezultatai	81,61%	0,7335
DistilBERT modelio rezultatai	81,10%	0,7145
DistilBERT modelio 5 lygių kryžminės validacijos rezultatai	81,11%	0,7135
Logistinės regresijos modelio rezultatai	73,80%	0,7217
Logistinės regresijos modelio rezultatai naudojant SMOTE	75,50%	0,7256

Šaltinis: sudaryta autorės

Gauti rezultatai rodo, kad geriausius tikslumo rodiklius pasiekė *BERT* modelis, tačiau *DistilBERT,* kaip mažesnė ir greitesnė modelio versija, išlieka labai konkurencinga. Tuo tarpu logistinė regresija pasiekė žemiausius rezultatus, tačiau jos našumas buvo pagerintas naudojant *SMOTE* metodą. Tai rodo, kad logistinė regresija yra naudingas modelis nelygių duomenų atveju, tačiau jos galimybės yra ribotos lyginant su giluminiais modeliais, kurie geriau apdoroja sudėtingą tekstinę informaciją ir supranta semantiką.

Taigi, Remiantis gautais rezultatais ir modelių palyginimu, pateikiamos kelios rekomendacijos, kurios galėtų padėti toliau gerinti ginčų baigties prognozavimo modelio tikslumą:

1. Duomenų papildymas. Kadangi buvo pastebėtas klasių disbalansas, kuris galėjo lemti neprognozuojamus rezultatus, rekomenduojama papildyti duomenų rinkinį. Baigčių

klasifikacijų, kurios yra mažiau randamos analizuojamuose dokumentuose pavyzdžių pridėjimas pagerintų modelio gebėjimą tiksliau klasifikuoti visų klasių pavyzdžius.

2. Hiperparametrų derinimas. Naudojant sudėtingus modelius, kaip *BERT* ir *DistilBERT*, būtų naudinga atlikti hiperparametrų optimizavimą. Tai apima modelio parametrų, tokių kaip mokymosi greitis, partijų dydis ar epochų skaičius optimizavimą. Gerai sureguliavus šiuos parametrus, galima ne tik pasiekti greitesnį modelio mokymą, bet ir pagerinti jo tikslumą.
3. Modelių derinimas tarpusavyje. Norint pasiekti dar geresnių rezultatų būtų galima derinti kelis modelius. Tokie hibridiniai modeliai galėtų pagerinti rezultatus atsižvelgiant į tai, kad kiekvienas modelis turi savo stipriąsias puses, kurias papildytų vienas kitam.
4. Teksto apdorojimo tobulinimas. Norint pagerinti modelio tikslumą, svarbu gilintis į teksto apdorojimo metodus. Teisiniai tekstai turi specifinę semantiką, kurią svarbu gerai suprasti, kad modelis galėtų tiksliai apdoroti ir klasifikuoti ginčų baigtis.

IŠVADOS IR PASIŪLYMAI

1. Išnagrinėjus mokslinę literatūrą, pastebėta, kad dirbtinis intelektas, ypač giluminio mokymosi modeliai kaip *BERT*, vis dažniau naudojami teisinių ginčų analizei ir baigties prognozavimui. Šie modeliai ne tik padeda apdoroti didžiulius duomenų kiekius, bet ir leidžia geriau suprasti tekstų prasmę. Tokie metodai jau taikomi skirtingose šalyse, pavyzdžiui, JAV, Kinijoje ir Brazilijoje, kur jie padėjo pasiekti reikšmingų rezultatų.
2. Ginčų baigties prognozavimo modelio kūrimo metu buvo atliktas nuoseklus duomenų paruošimas ir klasifikavimo klasių disbalanso sprendimai, siekiant užtikrinti kuo didesnį modelio tikslumą. Naudoti giluminio mokymosi algoritmai leido detaliai analizuoti tekstus ir efektyviai atpažinti sudėtingus semantinius ryšius. Gautas geriausiai pasirodžiusio *BERT* modelio tikslumas – 81,7 % – rodo, kad šis algoritmas yra tinkamas finansinių ginčų baigties prognozavimui.
3. Nustatyta, kad atlikto tyrimo rezultatai ir *BERT* modelio, kurio tikslumas buvo pasiektas didžiausias, efektyvumas yra konkurencingas pasauliniu mastu. Pavyzdžiui Kinijoje *BERT* modelio tikslumas prognozuojant teismų sprendimus siekė 87 %, tačiau šie rezultatai buvo paremti didelės apimties ir itin struktūruotais duomenimis, kas gali pagerinti modelio veikimą. Brazilijoje teisinių tekstų klasifikacijos tikslumas naudojant *BERT* svyravo tarp 80 % ir 82 %, labai artimai atlikto tyrimo rezultatams. Šie palyginimai rodo, kad tyrimo metu pasiektas rezultatas (81,7 %) yra labai artimas tarptautinei praktikai, ypač lyginant su Brazilijoje ir JAV pasiektais rezultatais.
4. Atsižvelgiant į atlikto tyrimo rezultatus, rekomenduojama toliau optimizuoti modelį, įtraukiant įvairius duomenų šaltinius ir papildomus teisinius kintamuosius. Tai galėtų apimti duomenis apie teismų procesų dinamiką, šalių teisinius argumentus, sprendimų priėmimo laiką ir kitus faktorius, galinčius turėti įtakos ginčo baigčiai. Taip pat rekomenduojama „giliau“ spręsti klasių disbalanso problemą ir tai bandyti įveikti naudojant lygesnius klasių rinkinius.
5. Nors tyrimas orientuotas į finansinių ginčų prognozavimą, sukurtas modelis galėtų būti pritaikytas ir kitose teisės srityse, tokiose kaip darbo teisė ar šeimos teisė. Pateiktas modelis galėtų tapti universaliu įrankiu prognozuoti įvairių teisinių ginčų baigtis, todėl ateityje rekomenduojama išplėsti taikymo sritį ir pritaikyti jį įvairiuose teisiniuose

kontekstuose. Tyrimas parodė, kad technologijų integracija į teisinį sektorių gali pagerinti ginčų sprendimo efektyvumą. Rekomenduojama skatinti teisininkų ir technologijų specialistų bendradarbiavimą, siekiant sukurti pažangias, automatizuotas sistemas, kurios leistų greičiau ir tiksliau spręsti ginčus.

LITERATŪRA

1. Condevaux, C. (2020). Neural legal outcome prediction with partial least squares compression. Stats, 3(3), 396–411. https://doi.org/10.3390/stats3030025
2. Dubois, C. (2021). How do lawyers engineer and develop legaltech projects?: A story of opportunities, platforms, creative rationalities, and strategies. Law, Technology and Humans, 3(1), 68–81. https://doi.org/10.3316/informit.950113855587667
3. Finansų rinkos dalyvių priežiūra. (2019). Prieiga per: https://www.lb.lt/lt/apie-prieziuros-veikla
4. Frolova, E. E., & Ermakova, E. P. (2021). Utilizing Artificial Intelligence in Legal Practice. In Smart Technologies for the Digitisation of Industry: Entrepreneurial Environment (pp. 17–27). Springer. https://doi.org/10.1007/978-981-16-4621-8_2
5. Jacob de Menezes-Neto, E., & Clementino, M. B. M. (2022). Using deep learning to predict outcomes of legal appeals better than human experts: A study with data from Brazilian federal courts. PLoS ONE, 17(7), e0272287. https://doi.org/10.1371/journal.pone.0272287
6. Katz, D. M., Bommarito II, M. J., & Blackman, J. (2016). A general approach for predicting the behavior of the Supreme Court of the United States. PLOS ONE, 12(4). https://doi.org/10.1371/journal.pone.0174698
7. Kaur, A., & Božić, B. (2019). Convolutional neural network-based automatic prediction of judgments of the European Court of Human Rights. School of Computer Science, Technological University Dublin. Prieiga per: https://ceur-ws.org/Vol-2563/aics_42.pdf
8. Mahfouz, T., & Kandil, A. (2011). Litigation outcome prediction of differing site condition disputes through machine learning models. Journal of Computing in Civil Engineering, 26(3). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000148
9. Medvedeva, M., Vols, M., & Wieling, M. (2020). Using machine learning to predict decisions of the European Court of Human Rights. Artificial Intelligence and Law, 28(2), 237–266. https://doi.org/10.1007/s10506-019-09255-y
10. Mumcuoğlu, E., Öztürk, C. E., Ozaktas, H. M., & Koç, A. (2021). Natural language processing in law: Prediction of outcomes in the higher courts of Turkey. Information Processing & Management, 58(5), 102684. https://doi.org/10.1016/j.ipm.2021.102684

11. Rodionov, A. (2023). Harnessing the Power of Legal-Tech: AI-Driven Predictive Analytics in the Legal Domain. Journal Title, Volume(Issue), Page Range. Prieiga per: https://irshadjournals.com/index.php/ujldp/article/view/69
12. Salmerón-Manzano, E. (2021). Legaltech and Lawtech: Global Perspectives, Challenges, and Opportunities. Laws, 10(2), 24. https://doi.org/10.3390/laws10020024
13. Shaikh, R. A., Sahu, T. P., & Anand, V. (2020). Predicting Outcomes of Legal Cases based on Legal Factors using Classifiers. Procedia Computer Science, 167, 2393–2402. https://doi.org/10.1016/j.procs.2020.03.292
14. Vartojimo ginčų neteisminio sprendimo ataskaita. (2023). Finansų rinkos dalyvių veikla. Prieiga per: https://www.lb.lt/uploads/publications/docs/44790_045b41c44211bce194016c9c5d871e51.pd
15. Rayhan, Abu. (2023). The Rise of Python: A Survey of Recent Research. https://doi.org/10.13140/RG.2.2.27388.92809. Prieiga per: https://www.researchgate.net/publication/373633075_The_Rise_of_Python_A_Survey_of_Recent_Research
16. Shen, Y. (2024). An Effective Sentimental Analysis Model Based on spaCy. Highlights in Science, Engineering and Technology, 85, 1065–1072. https://doi.org/10.54097/91axqs95 Prieiga per: https://drpress.org/ojs/index.php/HSET/article/view/18558
17. Gonzalez-Carvajal, S., & Garrido-Merchan, E. C. (2023). Comparing BERT against traditional machine learning text classification. Journal of Computational and Cognitive Engineering. https://doi.org/10.47852/bonviewJCCE3202838 Prieiga per: https://ojs.bonviewpress.com/index.php/JCCE/article/view/838
18. Libovicky, J., Rosa, R., & Fraser, A. (2019). How language-neutral is multilingual BERT? arXiv. https://doi.org/10.48550/arXiv.1911.03310 Prieiga per: https://arxiv.org/abs/1911.03310
19. Toman, M., Tesar, R., & Jezek, K. (2006). Influence of Word Normalization on Text Classification. Prieiga per: https://www.researchgate.net/publication/250030718_Influence_of_Word_Normalization_on_Text_Classification

20. Wieczorek, J., Guerin, C., & McMahon, T. (2022). K-fold cross-validation for complex sample surveys. Statistics in Medicine, 41(4), 686–704. https://doi.org/10.1002/sta4.454
21. Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2020). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. 5th Workshop on Energy Efficient Machine Learning and Cognitive Computing - NeurIPS 2019. https://doi.org/10.48550/arXiv.1910.01108 Prieiga per: https://arxiv.org/abs/1910.0110
22. Qaiser, S., & Ali, R. (2018). Text mining: Use of TF-IDF to examine the relevance of words to documents. International Journal of Computer Applications, 181(1), 25–29. Prieiga per: https://www.researchgate.net/profile/Shahzad-Qaiser/publication/326425709_Text_Mining_Use_of_TF-IDF_to_Examine_the_Relevance_of_Words_to_Documents/links/5b4cd57fa6fdcc8dae245aa3/Text-Mining-Use-of-TF-IDF-to-Examine-the-Relevance-of-Words-to-Documents.pdf
23. Elreedy, D., Atiya, A. F., & Kamalov, F. (2023). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. Journal Name, Volume(Issue), page range. https://doi.org/DOI Prieiga per: https://link.springer.com/article/10.1007/s10994-022-06296-4
24. Lietuvos bankas. (2024, lapkričio 16 d.). Taikūs susitarimai – konstruktyvi ir efektyvi priemonė užtikrinant finansų sistemos patikimumą. Lietuvos bankas. Prieiga per: https://www.lb.lt/lt/naujienos/taikus-susitarimai-konstruktyvi-ir-efektyvi-priemone-uztikrinant-finansu-sistemos-patikimuma
25. Horev, R. (2018, lapkričio 10). BERT explained: State of the art language model for NLP. Towards Data Science. https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b627
26. Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A Primer in BERTology: What We Know About How BERT Works. Transactions of the Association for Computational Linguistics, 8, 842–866. https://doi.org/10.1162/tacl_a_00349
27. Ettinger, A. (2020). What BERT Is Not: Lessons from a New Suite of Psycholinguistic Diagnostics for Language Models. Transactions of the Association for Computational Linguistics, 8, 34–48. https://doi.org/10.1162/tacl_a_00298

28. Staliūnaitė, I., & Iacobacci, I. (2020). Compositional and lexical semantics in RoBERTa, BERT and DistilBERT: A case study on CoQA. arXiv. https://doi.org/10.48550/arXiv.2009.08257 Prieiga per: https://arxiv.org/abs/2009.08257
29. Tsang, S.-H. (2022, kovo 5). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. Medium. https://sh-tsang.medium.com/review-distilbert-a-distilled-version-of-bert-smaller-faster-cheaper-and-lighter-5b3fa180169e
30. Varshini, S. H., Varma, G. S., Prabhakar, P., & Pati, P. B. (2024). Comparative analysis of legal outcome prediction with detailed and summarized text. In 2024 IEEE 9th International Conference for Convergence in Technology (I2CT) (pp. 1–6). https://doi.org/10.1109/I2CT61223.2024.10544203
31. Lage-Freitas, G., et al. (2019). Predicting Brazilian court decisions: Litigation prediction of judgment. arXiv. https://arxiv.org/abs/1905.10348 Prieiga per: https://publications.cohubicol.com/typology/predicting-brazilian-court-decisions/
32. Cao, L., Wang, Z., Xiao, C., & Sun, J. (2024). PILOT: Legal case outcome prediction with case law. arXiv. https://doi.org/10.48550/arXiv.2401.15770 Prieiga per: https://ar5iv.labs.arxiv.org/html/2401.15770v3
33. Wang, Y., Gao, J., & Chen, J. (2020). Deep learning algorithm for judicial judgment prediction based on BERT. In 2020 5th International Conference on Computing, Communication and Security (ICCCS) (pp. 1–6). https://doi.org/10.1109/ICCCS49678.2020.9277068

PRIEDAS

BERT MODELIO PROGRAMINIS KODAS

```
# 1. Reikalingų paketų diegimas
!pip install transformers pandas scikit-learn datasets

import pandas as pd
import numpy as np
import torch
from sklearn.model_selection import train_test_split
from sklearn.utils.class_weight import compute_class_weight
from transformers import BertTokenizer, BertForSequenceClassification, Trainer,
TrainingArguments, EarlyStoppingCallback
from sklearn.metrics import accuracy_score, f1_score

# 2. Duomenų nuskaitymas iš Google Sheets
sheet_id = '1p7tMbX-L5sR7bArB70qFQy9Nu3gkPIhFRIl1T2nDvs4'
sheet_url                                                                    =
https://docs.google.com/spreadsheets/d/{sheet_id}/export?format=csv'

# Duomenys iš Google Sheets
try:
    data = pd.read_csv(sheet_url)
    print("Duomenys sėkmingai įkelti:")
    print(data.head())
except Exception as e:
    print("Įkeliant duomenis įvyko klaida:", e)

# 3. Duomenų paruošimas
data['Label'] = data['Label'].map({'atmesti': 0, 'dalinai patenkinti': 1,
'patenkinti': 2})
train_texts, temp_texts, train_labels, temp_labels = train_test_split(
    data['Reason'], data['Label'], test_size=0.2, stratify=data['Label'],
random_state=42
)
```

```
    # 10% iš temp_texts bus naudojami validavimui
    val_texts, test_texts, val_labels, test_labels = train_test_split(
        temp_texts,     temp_labels,     test_size=0.5,     stratify=temp_labels,
random_state=42
    )

    # Indeksų paruošimas
    train_labels.reset_index(drop=True, inplace=True)
    val_labels.reset_index(drop=True, inplace=True)
    test_labels.reset_index(drop=True, inplace=True)

    # Pavertimas į sąrašą
    if isinstance(train_texts, pd.Series):
        train_texts = train_texts.tolist()
    if isinstance(val_texts, pd.Series):
        val_texts = val_texts.tolist()
    if isinstance(test_texts, pd.Series):
        test_texts = test_texts.tolist()

    train_texts = [str(text) for text in train_texts]
    val_texts = [str(text) for text in val_texts]
    test_texts = [str(text) for text in test_texts]

    # 4. Modelio kūrimas
    tokenizer = BertTokenizer.from_pretrained('bert-base-multilingual-cased')
    train_encodings   =   tokenizer(train_texts,   truncation=True,   padding=True,
max_length=128)
    val_encodings    =    tokenizer(val_texts,    truncation=True,    padding=True,
max_length=128)
    test_encodings    =    tokenizer(test_texts,    truncation=True,    padding=True,
max_length=128)


    #class FinancialDisputeDataset(torch.utils.data.Dataset):
        def __init__(self, encodings, labels):
            self.encodings = encodings
            self.labels = labels
```

```
        def __getitem__(self, idx):
            item = {key: torch.tensor(val[idx]) for key, val in self.encodings.items()}
            item['labels'] = torch.tensor(self.labels[idx])
            return item

        def __len__(self):
            return len(self.labels)

    train_dataset = FinancialDisputeDataset(train_encodings, train_labels)
    val_dataset = FinancialDisputeDataset(val_encodings, val_labels)
    test_dataset = FinancialDisputeDataset(test_encodings, test_labels)

    # 5. Klasės svorių nustatymas
    class_weights = compute_class_weight('balanced', classes=np.array([0, 1, 2]), y=train_labels)
    class_weights = torch.tensor(class_weights, dtype=torch.float).to('cuda' if torch.cuda.is_available() else 'cpu')

    # 6. Mokymas
    model = BertForSequenceClassification.from_pretrained('bert-base-multilingual-cased', num_labels=3)
    model.to('cuda' if torch.cuda.is_available() else 'cpu')

    training_args = TrainingArguments(
        output_dir='./results',
        num_train_epochs=3,
        per_device_train_batch_size=8,
        per_device_eval_batch_size=8,
        warmup_steps=500,
        weight_decay=0.1,
        learning_rate=2e-5,
        logging_dir='./logs',
        evaluation_strategy='epoch',
        save_strategy='epoch',
        load_best_model_at_end=True,
```

```
)

# Funkcija metrikoms apskaičiuoti
def compute_metrics(p):
    preds = np.argmax(p.predictions, axis=1)
    accuracy = accuracy_score(p.label_ids, preds)
    f1 = f1_score(p.label_ids, preds, average='weighted')
    return {
        'accuracy': accuracy,
        'f1': f1,
    }

# Pridedame ankstyvojo stabdymo funkciją
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
    eval_dataset=val_dataset,
    compute_metrics=compute_metrics,
    callbacks=[EarlyStoppingCallback(early_stopping_patience=2)]
)

# Apmokome modelį
trainer.train()

# 7. Vertinimas
results = trainer.evaluate(eval_dataset=test_dataset)
print("Modelio vertinimo rezultatai:", results)
```