



**Matematikos ir
informatikos
fakultetas**

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
FINANSŲ IR DRAUDIMO MATEMATIKOS
MAGISTRANTŪROS STUDIJŲ PROGRAMA

Akcijų indeksų pozicijų analizė, remiantis kainų dinamikos režimais

**Stock Index Position Timing Based on Regimes of
Price Dynamics**

Baigiamasis magistro darbas

Atliko: Evelina Veikutytė

VU el. p.: evelina.veikutyte@mif.stud.vu.lt

Vadovas: asist. dr. Gediminas Bagdonas

Darbo konsultantas: Romualdas Zovė, Lietuvos bankas

Vilnius

2023

Akcijų indeksų pozicijų analizė, remiantis kainų dinamikos režimais

Evelina Veikutytė

Santrauka

Šio darbo tikslas – sukonstruoti skirtingus režimus akcijų rinkose atpažįstančius mašininio mokymosi modelius ir palyginti jų veikimą su Lietuvos banko naudojamu autoregresiniu Markovo režimų pasikeitimo modeliu. Mašininio mokymosi modelių apmokymui naudojamas makroekonominių rodiklių rinkinys skelbiamas FRED puslapyje, gautų prognozių tikslumui įvertinti naudojami tikslumo (ACC), Brier (QPS), ploto po ROC kreive (AUC) bei Matthew koreliacijos koeficiento rodikliai. Režimų pasikeitimo modeliui naudojami makroekonominiai ir rinkos rodikliai gauti iš Bloomberg. Gautos modelių režimų prognozės naudojamos investavimo į S&P 500 akcijų indeksą strategijoms, kurių veikimas įvertinamas apskaičiavus gražos ir rizikos rodiklius. Tyrimo rezultatai parodė, kad mašininio mokymosi modeliai, jų apmokymui naudojant makroekonominių rodiklių rinkinį, nedavė reikšmingai geresnių investavimo rezultatų nei naudojant režimų pasikeitimo modelį.

Raktiniai žodžiai: mašininis mokymasis, režimų pasikeitimo modelis, akcijų indeksai, investavimo strategija

Stock Index Position Timing Based on Regimes of Price Dynamics

Evelina Veikutytė

Abstract

The work aims to construct machine learning models that recognize different regimes in stock markets and compare their results with the autoregressive Markov regime-switching model used by the Bank of Lithuania. The set of macroeconomic indicators used for training machine learning models is available on the FRED page. Metrics used to assess the accuracy of the obtained forecasts are accuracy (ACC), Brier (QPS), the area under the ROC curve (AUC) and Matthew's correlation coefficient. The regime-switching model uses macroeconomic and market indicators obtained from Bloomberg. Market regime forecasts are employed for creating investment strategies in the S&P 500 stock index, the performance of those strategies is evaluated by calculating return and risk indicators. The findings of the study show that machine learning models trained using a set of macroeconomic indicators do not produce significantly better investment results than a regime-switching model.

Key words: machine learning, regime switching model, stock market indices, investing strategy

Turinys

Įvadas	4
1 Teorinė dalis	8
1.1 Režimų nustatymas	8
1.2 Režimų nustatymo modeliai	9
1.2.1 Režimų pasikeitimo modelis	9
1.2.2 Sprendimų medžiai	10
1.2.3 Atsitiktinis miškas	12
1.2.4 Gradientinis stiprinimas	13
1.2.5 Logistinė regresija	14
1.2.6 Atraminių vektorių klasifikatorius	15
1.2.7 Dirbtinis neuroninis tinklas	17
1.3 Klasifikavimo modelių tikslumo vertinimo rodikliai	19
1.4 Rodikliai strategijos rezultatų palyginimui	21
2 Tiriamoji dalis	25
2.1 Duomenys	25
2.2 Režimų nustatymas	26
2.3 Duomenų paruošimas	28
2.4 Mašininio mokymosi modelių konstravimas	29
2.5 Modelių tikslumas ir investavimo rezultatai	34
2.6 Palyginimas su režimų pasikeitimo modeliu	40
Išvados	42
Literatūra	43
Priedai	47

Įvadas

Yra nemažai „amžinų“ klausimų pvz.: „Kas tai yra laimė?“, „Kas mes esam?“. D. Hilbertas paklaustas kokį klausimą jis užduotų Dievui atsakė, kad paklaustų „Kur yra Riemann‘o dzeta funkcijos nuliai?“. Ne paskutinis tarp amžinųjų klausimų yra pirkti ar ne kuri nors turta/aktyvą šiandien. Viena populiariausių aktyvų klasių, į kurią vienu ar kitu mastu investuoja faktiškai visi investuotojai yra akcijos, tai reiškia, kad šis klausimas akcijų indeksams, kurie atspindi visą akcijų rinką, yra vienas aktualiausių. Akcijos, kaip aktyvų klasė, investuotojui sukuria savo vertę, jei įmonės laukiami ateities pinigų srautai viršija laukiamus mokėjimus už įsipareigojimus. Įprastai įsipareigojimai turi iš anksto žinomą struktūrą (prisiimtos paskolos, išleistos obligacijos ir pan.), o laukiamų pajamų prognozės labai smarkiai priklauso nuo ekonomikos raidos scenarijaus. Investuotojai, nagrinėdami gražų ir jų kintamumo elgesį, pastebėjo, kad kintamumo padidėjimas akcijų rinkoje dažnai atspindi padidėjusį neapibrėžtumą ekonomikoje. Pavyzdžiui, Andrew Ang ir Geert Bekaert nustatė, kad akcijų gražos gali būti apibūdinamos dviem režimais: normaliu režimu ir režimu, kuriame gražos vidutiniškai mažesnės ir pasižymi didesniu kintamumu nei normaliaame režime [1] bei plėtodami šią idėją tyrė, kokį poveikį investuojant turi skirtingi režimai [2]. Anders Nielsen Goldman Sachs išleistame tyrime [25] nagrinėjamos galimybės investuoti į JAV akcijas ir Vyriausybės obligacijas atsižvelgiant į finansų rinkos būsenas ir yra išskiriami trys rinkos režimai: krizinis, atsigavimo ir stabilus. Krizinis režimas pasižymi neigiamomis akcijų gražomis ir dideliu kintamumu, atsigavimo režimas pasižymi geriausiomis akcijų ir obligacijų gražomis, o stabilaus režimo atveju pastebimas žemas kintamumo lygis ir teigiamos gražos tiek akcijoms, tiek obligacijoms. O štai Joseph Simonian ir Chenwei Wu nustatė skirtingas ekonomikos būsenas naudodami augimo, infliacijos ir finansinio svorto rodiklius. Toks pasirinkimas priimtas vadovaujantis Minsky (1986, 1992) teorija, kad ekonomikos būsenos stipriai priklauso nuo kredito būsenos ir finansinio svorto. Jų išskirtos rinkos būsenos gali būti prasmingai panaudotos priimant investavimo sprendimus [31].

Rinka visada yra kuriame nors režime ir režimo pasikeitimai yra atsitiktiniai. Pasikeitimai vyksta gana retai, todėl tuo galima pasinaudoti aktyviam valdymui, pavyzdžiui, jei šiandien rinka yra stabiliaame režime, tai didelė tikimybė, kad rytoj rinka bus tame pačiame režime. Kadangi atsigavimo režimas netrunka ilgai, tai pa-

prastumo dėlei šiame darbe toliau bus laikoma, kad akcijų rinka gali būti kriziniame arba stabiliam režime, o investavimu į akcijas bus laikomas investavimas į S&P 500 akcijų indeksą. Režimų atpažinimui yra naudojami įvairūs modeliai. Vienas iš jų – paslėptasis Markovo modelis (angl. *Hidden Markov Model*), kurio pritaikymą finansų rinkoje pristatė James D. Hamilton (1989) [13] išvystydamas režimų pasikeitimo modelį (angl. *regime-switching model*). Paslėptasis Markovo modelis naudojamas ganėtinai dažnai norint atpažinti režimus ir sukurti tuo paremtas investavimo strategijas (pavyzdžiui, Luca De Angelis ir Leonard J. Paas (2013)[8], Lorenzo Reus ir John M. Mulvey (2016)[29], Nguyet Nguyen (2018)[26]).

Tokį režimų pasikeitimo modelį naudoja Lietuvos bankas, siekdamas atpažinti kokiam režime yra akcijų rinka ir priimti atitinkamus investavimo sprendimus. Šiuo metu taikant modelio skaičiavimus praktiškai susiduriama su tokiais jo trūkumais: modelis yra mėnesinis; negalima pasakyti kokia šiandien yra akcijų rinkos būseną paėmus paskutinę laiko eilutės reikšmę (reikšmių vektorių), kadangi kiekvieną kartą reikia perskaičiuoti Markovo grandinę; modelio skaičiavimas yra laikui imlus procesas.

Kita vis plačiau taikoma modelių kategorija yra mašininio mokymosi modeliai. Režimų prognozavimui gali būti naudojami įvairūs klasifikavimo modeliai, pavyzdžiui, logistinė regresija, atsitiktiniai miškai, gradientinio stiprinimo modeliai. Šie modeliai patrauklūs gebėjimu atpažinti sudėtingas duomenų sąryšių struktūras. Be to, apmokius šiuos modelius režimo prognozavimui užtenka paskutinės naujausių duomenų reikšmės, priešingai nei paslėptąjo Markovo modelio atveju. Minėtus mašininio mokymosi modelius režimų nustatymui naudojo Junhan Gu ir John M. Mulvey [12] bei A. Sinem Uysal and John M. Mulvey [34]. Alexander James, Yaser S. Abu-Mostafa and Xiao Qiao recesijos prognozavimui ir atitinkamų investavimo sprendimo priėmimui naudojo atraminių vektorių modelį [17]. Kartu su prieš tai išvardintais modeliais papildomai atraminių vektorių modelį ir neuroninio tinklo modelį recesijos prognozavimui naudojo Alireza Yazdani [36].

Šiame darbe bus konstruojamas Lietuvos banko naudojamas režimų pasikeitimo modelis, logistinės regresijos modeliai, atsitiktinis miškas, gradientinio stiprinimo modeliai, atraminių vektorių klasifikatorius bei neuroninio tinklo modelis. Minėti modeliai bus naudojami stabilaus ir krizinio režimo atpažinimui akcijų rinkoje. Mašininio mokymosi modelių apmokymui bus naudojamas Michael W. McCracken ir Serena Ng sudarytas 127 makroekonominių rodiklių [rinkinys](#), skelbiamas FRED

puslapyje, o paslėptajame Markovo modelyje naudojami keli makroekonominiai ir rinkos rodikliai. Kadangi naudojami mašininio mokymosi modeliai yra prižiūrimo mokymosi tipo, tai iš anksto nustatytos režimų žymės modelių apmokymui bus gaunamos naudojant ℓ_1 trendo filtrą. Šio darbo tikslas yra palyginti ekonometrinio modelio veikimą su mašininio mokymosi modeliais. Pritaikius minėtus mašininio mokymosi modelius siekiama pagerinti investavimo į akcijas rezultatus ir atrasti geresnius duomenų tarpusavio sąryšius, padedančius atpažinti skirtingus režimus. Taip pat, bus palygintas mašininio mokymosi modelių veikimas darant prognozes mėnesiniu ir savaitiniu horizontu bei išanalizuota kaip skiriasi rezultatai naudojant visus makroekonominis rodiklius esančius rinkinyje ir padarius svarbiausių rodiklių atranką.

Darbo teorinėje dalyje pristatomas ℓ_1 trendo filtras, kuris naudojamas režimų žymių akcijų rinkoje nustatymui (šios žymės naudojamos prižiūrimo mokymosi modelių apmokymui), toliau pristatomas režimų pasikeitimo modelis, mašininio mokymosi modeliai ir jų tikslumo įvertinimo rodikliai, taip pat pristatomi rodikliai leidžiantys įvertinti ir palyginti investavimo į akcijas rezultatus, gautus remiantis skirtingais režimų akcijų rinkoje nustatymo modeliais. Praktinėje dalyje supažindinama su naudojamu makroekonominių rodiklių rinkiniu, aptariamas jų paruošimas tyrimui, aprašomas ℓ_1 trendo filtro pritaikymas, mašininio mokymosi modelių hiperparametrų parinkimas. Toliau aprašomas tų modelių skaičiavimo įgyvendinimas savaitiniu ir mėnesiniu horizontu, svarbesnių makroekonominių rodiklių atrinkimas, pateikiami mašininio mokymosi modelių tikslumo rodikliai, investavimo į S&P 500 akcijų indeksą strategijos aprašymas, rodiklių skaičiavimas ir rezultatai. Gauti rezultatai palyginami su investavimo rezultatais naudojant ekonometrinį režimų pasikeitimo modelį. Mašininio mokymosi modelių tikslumo rodiklių bei investavimo rezultatų rodiklių lentelės pateikiamos priedų skyriuje. Tyrimo išvados ir rekomendacijos - išvadų skyriuje.

Darbe nagrinėjami duomenys nuo 1959 iki 2022 metų, ši imtis padalijama į tris dalis. Pirmoji dalis apima periodą nuo 1959 iki 1978 metų ir yra naudojama modelių hiperparametrų nustatymui. Antroji - nuo 1978 iki 2014 metų galo, ji naudojama modelių veikimo testavimui. Trečioji - nuo 2015 metų iki 2022 metų taip pat naudojama modelių veikimo testavimui, tačiau šiai imčiai modelių hiperparametrai perskaičiuojami iš naujo ir režimų žymės modelių apmokymui gaunamos nenaudojant duomenų iš ateities. Tiriami pradiniai duomenys yra mėnesiniai, tačiau jie

pertvarkomi taip, kad modelius būtų galima skaičiuoti savaitiniu horizontu. Pagrindiniai skaičiavimai mašininio mokymosi modelių konstravimui atlikti naudojantis Python programavimo kalba, režimų pasikeitimo modelio skaičiavimui naudojamas MATLAB Perlin-Ding paketas [28].

1 Teorinė dalis

1.1 Režimų nustatymas

Šiame skyrelyje pristatomas prižiūrimo mokymosi modeliams reikalingų režimų žymių (angl. *labels*), kurios toliau vadinamos režimais, nustatymo algoritmas. Režimų žymės gaunamos pasinaudojus ℓ_1 trendo filtravimo algoritmu, kurį 2009 m. pristatė Seung-Jean Kim [21], o 2014 m. apibendrino Ryan J. Tibshirani [32]. Toliau pateikiamas algoritmo apibrėžimas.

Tarkime, kad gražų eilutė r_t , $t = 1, \dots, n$ yra sudaryta iš lėtai kintančio trendo x_t ir sparčiau kintančios atsitiktinės komponentės z_t . Režimų žymės gaunamos apskaičiavus trendo komponentę x_t (kartais x_t apskaičiavimas vadinamas glodiniu arba filtravimu) arba atsitiktinę komponentę $z_t = r_t - x_t$ (liekanas). Norima, kad x_t eilutė būtų glodi (angl. *smooth*), o atsitiktinė komponentė z_t maža. Nauja laiko eilutė x_t gaunama išsprendus optimizavimo problemą

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \sum_{t=1}^n (r_t - x_t)^2 + \lambda \sum_{t=2}^n |x_{t-1} - x_t|, \quad (1)$$

kur $\lambda \geq 0$ yra parametras, nuo kurio priklauso x_t glodumas ir liekanų $r_t - x_t$ dydis. Minimizuojama funkcija neturi išvestinės, todėl problemos sprendimui taikomas Ryan J. Tibshirani ir Jonathan Taylor straipsnyje [33] pasiūlytas specifinis algoritmas.

Iš (1) formulės matoma, jog vektorius $x = (x_1, \dots, x_n)$, kurį norima rasti, minimizuoja dydį sudarytą iš dviejų dalių. Pirmoji dalis yra sudaryta iš liekanų kvadratų sumos, o antroji dalis yra baudos narys (angl. *penalty term*), kuris baudžia, jei iš eilės einančių gražų dyžiai ženkliai skiriasi. Antroji dalis skatina algoritmą nepaisyti staigių pasikeitimų laiko eilutėje, kurie gali būti sukelti triukšmo. Algoritmo rezultatas yra laužtinė funkcija (trendo laiko eilutė). Režimus galima gauti nustačius tam tikrą slenkstį, pavyzdžiui, neigiamos reikšmės gali reikšti krizinį režimą, o likusios reiškia stabilų režimą. Šiuos režimus galima pažymėti 1 ir 0 atitinkamai. Algoritmo pagrindinė idėja yra triukšmo išfiltravimas ir trendo radimas iš gražų laiko eilutės.

1.2 Režimų nustatymo modeliai

Toliau pristatomas režimų pasikeitimo modelis ir prižiūravimo mokymosi klasifikavimo modeliai. Neprižiūravimo mokymosi atveju modeliui pateikiami duomenys be pradinių klasių žymių. Šio tipo modeliai siekia atrasti nesuklasifikuotų duomenų struktūrą, pasinaudojant koku nors matematiniu kriterijumi. Pavyzdžiui, neprižiūravimo mokymosi modelis yra k -vidurkių klasterizavimas (angl. *k-means clustering*) padalijantis N stebėjimų aibę į k klasterių. Prižiūravimo mokymosi atveju modeliams pateikiami duomenys turintys žymes, tai yra kiekvienas požymių rinkinys turi jiems priskirtą žymę (pavyzdžiui, klasę 1 arba 0). Šių modelių algoritmų tikslas yra atrasti funkciją, kuri susietų požymių vektorius su jų žymėmis.

1.2.1 Režimų pasikeitimo modelis

Šiame darbe yra naudojamas autoregresinis Markovo režimų pasikeitimo modelis toliau vadinamas režimų pasikeitimo modeliu. Modelyje režimų pasikeitimas yra stochastinis, taigi iš anksto nėra žinoma ar įvyks režimo pasikeitimas, tačiau yra žinoma režimų pasikeitimo proceso dinamika, kuri apibrėžiama laike kintančia perėjimų matrica (tikimybės šioje matricoje taip pat gali būti fiksuotos). Toliau bus pateiktas modelio apibrėžimas, kai yra du režimai. Modelio apibendrinimą, kai režimų yra daugiau nei du, galima rasti James D. Hamilton knygoje [14].

Tarkime, kad $X_t = [x_{1,t}, \dots, x_{n,t}]'$ yra vektorius sudarytas iš makroekonominių ir rinkos rodiklių duomenų laiku t . Tada režimų pasikeitimo modelio autoregresinė VAR(1) dalis užrašoma:

$$X_t = \beta_{0,S_t} + \beta'_{S_t} X_{t-1} + \epsilon_t, \epsilon_t \sim \mathcal{N}(0, \Sigma_{S_t}),$$

čia

$$\Sigma_{S_t} = \begin{pmatrix} \sigma_{1,S_t}^2 & \cdots & \sigma_{1,n,S_t} \\ \vdots & \ddots & \vdots \\ \sigma_{1,n,S_t} & \cdots & \sigma_{n,S_t}^2 \end{pmatrix}.$$

Regresijos koeficientai ir kovariacijų matricos reikšmės yra priklausomos nuo režimo S_t . Nagrinėkime modelį su dvejais režimais $S_t = 1$ ir $S_t = 2$. Tada modelio koeficientai ir kovariacijų matrica yra $(\beta_{0,1}, \beta'_{1}, \Sigma_1)$ arba $(\beta_{0,2}, \beta'_{2}, \Sigma_2)$. Šiame modelyje tariama, kad režimų perėjimo tikimybės priklauso tik nuo prieš tai buvusio režimo.

Tikimybės aprašomos taip:

$$\begin{aligned}
 p_t &= P(S_t = 1 | S_{t-1} = 1, \psi_{t-1}) = p(\psi_{t-1}) \\
 1 - p_t &= P(S_t = 2 | S_{t-1} = 1, \psi_{t-1}) = 1 - p(\psi_{t-1}) \\
 q_t &= P(S_t = 2 | S_{t-1} = 2, \psi_{t-1}) = q(\psi_{t-1}) \\
 1 - q_t &= P(S_t = 1 | S_{t-1} = 2, \psi_{t-1}) = 1 - q(\psi_{t-1}),
 \end{aligned}$$

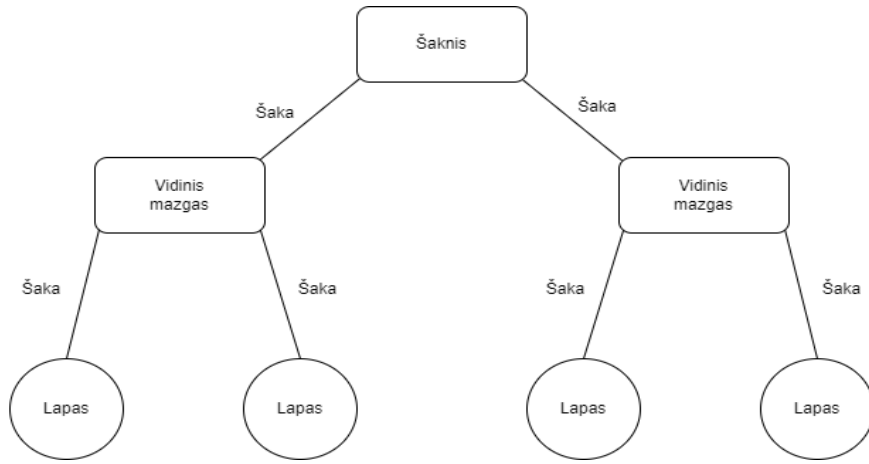
čia ψ_{t-1} yra informacija žinoma iki laiko $t - 1$, ji daro įtaką režimų pasikeitimo tikimybėms tarp $t - 1$ ir t periodų. Detalesnis koeficientų ir tikimybių apskaičiavimo algoritmas pateikiamas Perez ir Timmermann straipsnyje [27] arba Marcelo Perlin MATLAB paketo *MS_Regress* apraše [28].

1.2.2 Sprendimų medžiai

Toliau pristatomi mašininio mokymosi modeliai. Pirma apibrėžkime sprendimų medžio modelį, kuris yra kelių kitų mašininio mokymosi modelių pagrindas.

Sprendimų medis yra prižiūrimo mokymosi metodas, kuris gali būti naudojamas tiek regresijos, tiek klasifikacijos užduotims spręsti. Šis metodas padalija požymių erdvę į stačiakampių aibę ir kiekviename iš jų pritaiko paprastą funkciją, pavyzdžiui, konstantą, tai pavaizdavus grafiškai gaunama schema primenanti medį, o metodo terminai yra susiję su medžio dalimis. Pirmas medžio mazgas yra vadinamas šaknimi arba šaknies mazgu, iš mazgų išeina šakos, kurios baigiasi lapu arba veda į sekančius mazgus. Mazgai reiškia testą, šakos reprezentuoja testo rezultatą, o lapai klasifikavimo atveju parodo klasę, kuriai buvo priskirtas objektas.

Šiame darbe pateikiamas tikslesnis apibrėžimas remiantis Wu et al. straipsniu [35]. Toliau i -tojo objekto požymių vektorius bus žymimas $X_i = (x_{i_1}, \dots, x_{i_n})$. Sprendimų medis apibrėžiamas kaip $T : \mathbb{R}^n \rightarrow \mathbb{R}$. Vidiniai mazgai v_k yra susieti su Bulio funkcija $f_k(X_i) = \mathbb{1}_{\{x_{i_k} < t_k\}}$, kur t_k yra sprendimo slenkstis. Kiekvienas lapas l yra susietas su reikšme z_l . Turint i -tojo objekto požymių vektorių X_i , sprendimų medžio vertinimas pradedamas nuo šaknies mazgo ir kiekviename vidiniame mazge v_k yra įvertinama funkcija $f_k(X_i)$. Priklausomai nuo $f_k(X_i)$ reikšmės (1 arba 0) yra pasirenkama kairė arba dešinė medžio šaka. Procesas kartojamas, kol pasiekiamas lapas l , tada $T(X_i)$ gražina lapo reikšmę z_l .



1.1 pav.: Sprendimų medžio struktūra

Sprendimo medžio konstravimo metu treniravimo duomenys yra padalijami į poaibius. Iš pradžių turime visus treniravimo duomenis ir susijusias žymes. Tam, kad duomenys geriausiu įmanomu būdu būtų padalinti į atitinkamas klases, naudojami požymių atrankos metodai, kuriais pamatuojama kaip „švariai“ (angl. *pure*) buvo atskirtos klasės. Egzistuoja keli skirtingi sprendimų medžių sudarymo algoritmai, kiekviename iš jų naudojami skirtingi požymių atrankos metodai. **ID3** algoritmo atveju naudojamas informacijos prieaugis (angl. *information gain*), **C4.5** algoritmo atveju naudojamas informacijos prieaugio santykis (angl. *information gain ratio*), o **CART** algoritme naudojamas Gini indeksas (angl. *Gini index*). Šiame darbe buvo naudojamas atsitiktinio miško metodas, kuriame sprendimo medžiai sudaromi pagal **CART** algoritmą, todėl plačiau bus pristatomas tik Gini indeksas. Daugiau informacijos apie kitus rodiklius Jiawei Han knygoje [15].

Tegul D yra modelio apmokymo duomenų aibė, sudaryta iš objektų požymių vektorių ir atitinkamų žymų. Sakykime, kad yra m skirtingų žymų apibrėžiančių m skirtingų klasių $C_i, i = 1, \dots, m$. Tegul $C_{i,D}$ yra apmokymo duomenų dalis esanti D aibėje, kuri priklauso klasei C_i . Pažymėkime $|D|$ ir $|C_{i,D}|$ kaip skaičių objektų priklausančių D ir $C_{i,D}$ atitinkamai. Gini indeksas matuojantis D aibės priemaišą apibrėžiamas taip:

$$Gini(D) = 1 - \sum_{i=1}^m p_i^2,$$

kur $p_i = \frac{|C_{i,D}|}{|D|}$ yra tikimybė, kad objektas esantis D aibėje priklauso klasei C_i . Skaičiuojant Gini indeksą yra patikrinami visi įmanomi binariniai padalijimai su kiekvienu turimu požymiu. Pavyzdžiui, tegul A yra požymis, kuris įgyja k skirtingų

diskrečių reikšmių (a_1, \dots, a_k) esančių apmokymo duomenyse D . Tam, kad būtų nustatytas geriausias binarinis padalijimas pagal A , ištiriami visi galimi poabiai naudojant žinomas A reikšmes, išskyrus tuščią poabį ir visų aibės reikšmių poabį (a_1, \dots, a_k) , nes jie nereprezentuoja padalijimo. Taigi, patikrinami $2^k - 2$ galimi padalijimai. Jei požymis A padalija D aibę į D_1 ir D_2 , tai D aibės Gini indeksas yra:

$$Gini_A(D) = \frac{|D_1|}{|D|}Gini(D_1) + \frac{|D_2|}{|D|}Gini(D_2).$$

Su kiekvienu požymiu yra patikrinami visi galimi binariniai padalijimai. Kai požymis įgyja tik diskrečias reikšmes, tai poabius, su kuriuo gaunama mažiausia Gini indekso reikšmė, laikomas padalijimo poabiu. Tiems požymiams, kurie įgyja tolydžias reikšmes, turi būti patikrinti visi galimi padalijimo taškai. Galimi padalijimo taškai padalija aibę D į D_1 , kur $A \leq \textit{padalijimo_taškas}$ ir D_2 , kur $A > \textit{padalijimo_taškas}$. Tolydaus požymio padalijimo tašku laikomas tas taškas, su kuriuo gaunama mažiausia Gini indekso reikšmė.

1.2.3 Atsitiktinis miškas

Apibrėžus sprendimų medžius, galime pristatyti šiame darbe naudojamą atsitiktinio miško metodą (L. Breiman [5]). Šis modelis priskiriamas kolektyvinio mokymosi metodams kadangi jis yra sudarytas iš klasifikatorių kolektyvo, kur kiekvienas klasifikatorius yra sprendimo medis. Atsitiktiniai miškai konstruojami naudojant savirankos metodą (angl. *bootstrapping*) kartu su atsitiktiniu požymių parinkimu.

Tariant, kad D yra apmokymo duomenų aibė turinti d objektų, k sprendimo medžių (T_1, \dots, T_k) atsitiktiniame miške yra konstruojami taip: kiekvienai iteracijai $i = 1, \dots, k$ iš D yra parenkama D_i apmokymo duomenų aibė, parinkti objektai grąžinami į D , todėl aibėje D_i gali būti pasikartojančių objektų. Tegul F yra požymių skaičius naudojamas kiekviename mazge, F yra mažesnis nei visų turimų požymių skaičius. Konstruojant sprendimo medį T_i , kiekviename mazge atsitiktinai parenkama F požymių, kurie gali būti kandidatais padalijimui. Medžiai auginami iki maksimalaus galimo dydžio. Sukonstravus k medžių gaunamas medžių kolektyvas $\{T_i\}_1^k$. Turint naują objektą atsitiktinio miško klasifikacijos rezultatas yra ta klasė, už kurią nubalsavo didesnė dalis medžių.

Atsitiktinio miško modelis nėra jautrus išskirtims (angl. *outliers*) duomenyse, nebūtinai apmokymo duomenų standartizavimas, mažesnė persimokymo (angl. *over-*

fitting) rizika lyginant su sprendimų medžio algoritmu. Tačiau modelis turi ir trūkumų, atsitiktinio miško rezultatus sunkiau interpretuoti, o turint didelį duomenų rinkinį skaičiavimai gali trukti ganėtinai ilgai.

1.2.4 Gradientinis stiprinimas

Sekantis darbe naudojamas modelis yra gradientinis stiprinimas. Gradientinio stiprinimo metodas taip pat priklauso kolektyvinio mokymosi metodų klasei. Metodas turi tokį pavadinimą, nes nuostolio funkcijos optimizavimui yra naudojamas gradientinio nuolydžio metodas.

Stiprinimo metode modeliai konstruojami vienas po kito, kiekvienas mokosi iš prieš tai buvusio modelio klaidų. Pavyzdžiui, kiekvienam objektui iš apmokymo duomenų aibės gali būti priskiriamas svoris (pirmame žingsnyje svoriai vienodi) ir iteruojant sukuriami k klasifikatorių. Kai sukonstruojamas T_i klasifikatorius, svoriai atnaujinami taip, kad sekantis klasifikatorius T_{i+1} daugiau dėmesio skirtų objektams, kuriuos T_i suklasifikavo neteisingai. Galutinis klasifikatorius T^* apjungtų visų individualių klasifikatorių balsus, kiekvieno klasifikatoriaus balso svoris priklausytų nuo jo tikslumo funkcijos. Gradientinio stiprinimo metode nėra priskiriamų svorių, vietoj jų apskaičiuojami modeliai, kurie atsižvelgia į paklaidas. Šio metodo algoritmas aprašomas taip [16]:

1. Inicijuojama $f_0(X) = \arg \min_{\gamma} \sum_{i=1}^N \mathcal{L}(y_i, \gamma)$,
kur γ yra prognozė, y_i tikra objekto reikšmė, \mathcal{L} diferencijuojama iškili nuostolio funkcija, matuojanti atstumą tarp γ ir y_i .
2. Kiekvienam $m = 1, \dots, M$, kur M - medžių skaičius:
 - (a) apskaičiuojamas gradientas

$$r_{im} = - \left[\frac{\partial \mathcal{L}(y_i, f(X_i))}{\partial f(X_i)} \right] \Bigg|_{f=f_{m-1}}, i = 1, \dots, N.$$

N - objektų skaičius.

- (b) Gautoms paklaidoms konstruojamas sprendimų medis turintis lapus R_{jm} ,
 $j = 1, \dots, J_m$.
- (c) Randama γ_{jm} reikšmė, kuri minimizuoja nuostolio funkciją kiekviename lape J_m :

$$\gamma_{jm} = \arg \min_{\gamma} \sum_{x_i \in R_{jm}} \mathcal{L}(y_i, f_{m-1}(X_i) + \gamma).$$

(d) Atnaujinama $f_m(X) = f_{m-1}(X) + \sum_{j=1}^{J_m} \gamma_{jm} \mathbb{1}_{\{x \in R_{jm}\}}$

3. Galutinis modelis yra $\hat{f}(X) = f_M(X)$.

Modelio mokymosi procedūrą galima kontroliuoti naudojant mokymosi greitį $0 < \nu < 1$.

Tada (d) dalyje esanti formulė atrodo taip:

$$f_m(X) = f_{m-1}(X) + \nu \sum_{j=1}^{J_m} \gamma_{jm} \mathbb{1}_{\{x \in R_{jm}\}}.$$

Darbe naudojamas ir XGBoost (angl. *Extreme gradient boosting*) modelis, jo pagrindas taip pat yra gradientinis stiprinimas. XGBoost skiriasi nuo įprasto gradientinio stiprinimo modelio savo realizacija, dėl kurios skaičiavimai vyksta greičiau ir sumažinama persimokymo galimybė. Plačiau apie modelio realizaciją galima rasti čia [6].

Gradientinio stiprinimo modelio prognozių tikslumas dažnai geresnis nei atsitiktinio miško modelio. Šiam modeliui nebūtinai duomenų standartizavimas, duomenyse gali būti trūkstamų reikšmių – jų užpildyti nebūtina. Kaip ir atsitiktinio miško atveju, gradientinio stiprinimo modelio rezultatus sunku interpretuoti, o skaičiavimai trunka ilgai. Be to, šis modelis turi didesnę persimokymo riziką.

1.2.5 Logistinė regresija

Kitas šiame darbe naudojamas modelis yra logistinė regresija. Tai prižiūrimo mokymosi modelis, kuris gali būti naudojamas klasifikavimo užduotims spręsti.

Dvinarės logistinės regresijos modelis yra sudaromas tikimybių santykio logaritmui, kuris dar vadinamas logit funkcija:

$$\ln \left(\frac{P(y = 1)}{P(y = 0)} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n,$$

čia y yra priklausomas kintamasis įgyjantis reikšmes 0 arba 1, $\beta_0, \beta_1, \dots, \beta_n$ nežinomi koeficientai. $P(y = 0) = 1 - P(y = 1)$. Pažymėjus

$$Z = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n,$$

gaunama

$$\ln \left(\frac{P(y = 1)}{1 - P(y = 1)} \right) = Z$$
$$\frac{P(y = 1)}{1 - P(y = 1)} = e^Z.$$

Be to,

$$P(y = 1) = \frac{1}{1 + e^{-Z}}. \quad (2)$$

(2) formulė dar vadinama sigmoidės funkcija, kuri bet kurį realų skaičių atvaizduoja $[0, 1]$ intervale. Toliau $P(y = 1)$ bus žymima $p(X)$.

Logistinės regresijos koeficientų radimui yra skaičiuojama didžiausio tikėtinumo funkcija ℓ :

$$\ell(\beta) = \sum_{i=1}^N [y_i \ln(p(X_i)) + (1 - y_i) \ln(1 - p(X_i))] \quad (3)$$

Plačiau apie (3) funkcijos maksimizavimą skaitykite čia [16].

Šiame darbe bus naudojama logistinė regresija su elastinio tinklo (angl. *elastic net*), L1 bei L2 regularizacija. Naudojant elastinio tinklo regularizaciją, minimizuojama funkcija yra

$$\frac{\ell(\beta)}{2N} + \lambda((1 - \alpha) \frac{1}{2} \|\beta\|_2^2 + \alpha \|\beta\|_1),$$

kai $\alpha = 1$, turime L1 (Lasso) regularizaciją, kai $\alpha = 0$, turime L2 (Ridge) regularizaciją. Naudojant L1 regularizaciją, mažiau svarbių požymių svoriai tampa 0, o naudojant L2 regularizaciją svorių dydžiai sumažinami, bet nebūtinai tampa 0. Naudojant elastinį tinklą, kuris yra L1 ir L2 regularizacijų kombinacija, vieni mažesnę svarbą turinčių požymių svoriai paverčiami 0, kiti svoriai sumažinami link 0.

Logistinės regresijos modelio privalumas yra ganėtinai greitas apmokymas ir aiškus rezultatų interpretavimas. Šis modelis duoda gerus prognozavimo rezultatus, kai duomenų rinkinio stebėjimai gali būti atskirti tiesiškai. Naudojant logistinės regresijos modelį rekomenduojama duomenis standartizuoti.

1.2.6 Atraminių vektorių klasifikatorius

Dar vienas darbe naudojamas modelis yra atraminių vektorių metodas (angl. *Support Vector Machines*), tai vienas iš populiarių klasifikavimo metodų, kurį galima pritaikyti tiek tiesinėms, tiek netiesinėms duomenų struktūroms.

Naudojant šį metodą, duomenys, kuriuos reikia suklasifikuoti, yra atskiriami hiperplokštuma, kuri randama naudojant atraminius vektorius (angl. *support vectors*) ir paraštes (angl. *margins*), kurios priklauso nuo atraminių vektorių. Atstumas nuo atraminių vektorių iki hiperplokštumos turi būti didžiausias, kurį galima pasiekti. Be to, tikimasi, kad esant didesniai atstumui tarp hiperplokštumos paraščių, nematyti duomenys bus suklasifikuojami geriau [15], todėl apmokymo metu ieškoma

hiperplokštumos su didžiausiu įmanomu atstumu tarp paraščių. Tuo atveju, kai apmokymo duomenys D turi dvi klases $y_i \in \{1, -1\}$, kurios gali būti atskiriamos tiesiškai, jas atskirianti hiperplokštuma gali būti užrašyta kaip:

$$\mathbf{W} \cdot \mathbf{X} + b = 0,$$

kur $\mathbf{W} = \{w_1, \dots, w_n\}$ yra svorių vektorius, n - požymių skaičius, \mathbf{X} apmokymo objektai, b - skaliaras. Klasifikuojant siekiama, kad

$$\mathbf{W} \cdot \mathbf{X}_i + b > 0, \text{ kai } y_i = 1,$$

$$\mathbf{W} \cdot \mathbf{X}_i + b < 0, \text{ kai } y_i = -1.$$

Svoriai gali būti koreguojami taip, kad hiperplokštumos apibrėžiančios paraščių kraštines būtų užrašytos taip:

$$H_1 : \mathbf{W} \cdot \mathbf{X}_i + b \geq 1, \text{ kai } y_i = 1, \tag{4}$$

$$H_2 : \mathbf{W} \cdot \mathbf{X}_i + b \leq -1, \text{ kai } y_i = -1. \tag{5}$$

Taip pat (4) ir (5) formules galime užrašyti

$$y_i(\mathbf{W} \cdot \mathbf{X}_i + b) \geq 1, \forall i.$$

Gaunama, kad objektai esantys ant arba virš H_1 , priklauso 1 klasei, o objektai esantys ant arba po H_2 priklauso -1 klasei. Apmokymo objektai esantys ant H_1 ir H_2 vadinami atraminiais vektoriais. Atstumas skiriantis hiperplokštumą nuo bet kurio taško ant H_1 yra $\frac{1}{\|\mathbf{W}\|}$, $\|\mathbf{W}\| = \sqrt{\mathbf{W} \cdot \mathbf{W}} = \sqrt{w_1^2 + w_2^2 + \dots + w_n^2}$ - euklidinė norma. Pagal apibrėžimą šis atstumas lygus atstumui nuo bet kurio taško ant H_2 iki skiriančios hiperplokštumos, taigi atstumas tarp H_1 ir H_2 yra $\frac{2}{\|\mathbf{W}\|}$. Šio atstumo maksimizavimui sprendžiamas kvadratinio optimizavimo uždavinys, plačiau apie tai šioje knygoje [16].

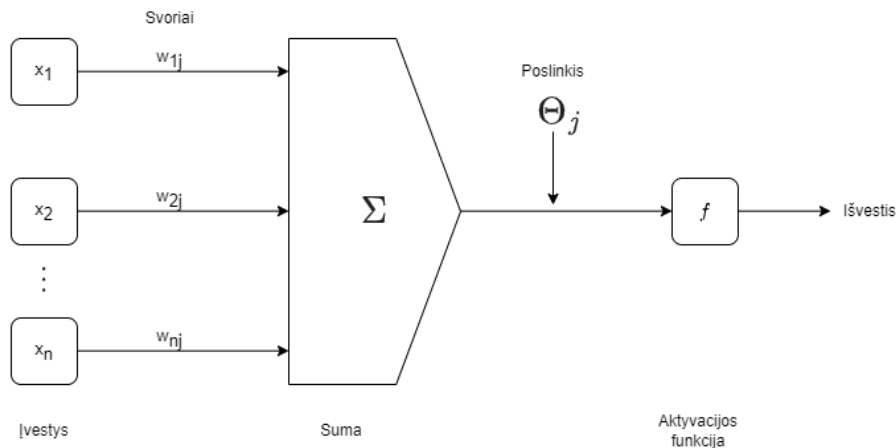
Jei duomenų klasių negalima atskirti naudojant tiesinę hiperplokštumą, tai duomenys gali būti transformuojami į aukštesnės dimensijos erdvę ir tiesiškai atskirianti hiperplokštuma randama naujoje erdvėje, tačiau čia galimi keblumai pasirenkant tinkamą transformaciją ir tokių skaičiavimų atlikimas gali ilgai užtrukti [15]. Šios problemos sprendžiamos naudojant branduolio metodą, detalesnė informacija pateikiama [20].

Atraminų vektorių klasifikatorius veikia ganėtinai gerai, kai egzistuoja aiški klasių atskyrimo riba (angl. *margin*), šis modelis pasižymi greitu prognozavimu. Deja,

atraminių vektorių klasifikatorius netinka didelės apimties duomenų rinkiniams ir duomenims turintiems daug triukšmo. Be to, modeliui reikalingas duomenų standartizavimas, o modelio apmokymas (priešingai nei prognozavimas) yra laikui imlus procesas.

1.2.7 Dirbtinis neuroninis tinklas

Dirbtinis neuroninis tinklas yra matematinis modelis, kuris bando imituoti biologinių neuroninių tinklų struktūrą ir funkcijas. Su šiuo modeliu gali būti sprendžiamos tiek regresijos, tiek klasifikavimo užduotys. Kiekvieno dirbtinio neuroninio tinklo pagrindinė sudedamoji dalis yra dirbtinis neuronas - paprastas matematinis modelis (funkcija). Neuroną sudaro daugybės, sudėties ir aktyvacijos funkcijos komponentės [15].



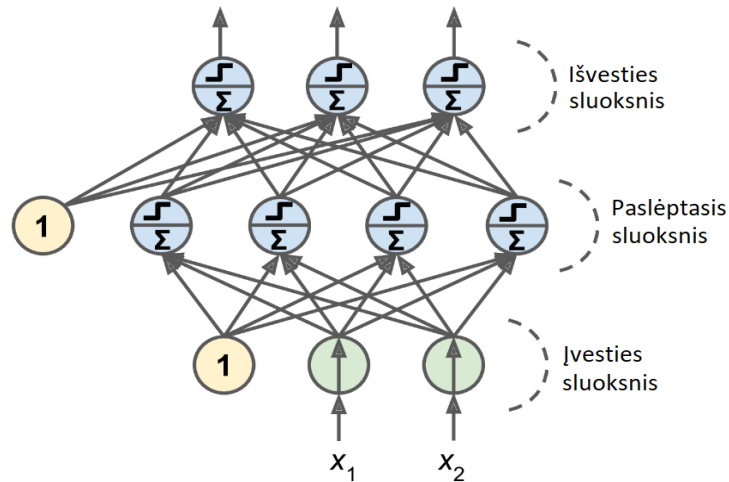
1.2 pav.: Paprastas dirbtinis neuronas

Dirbtinis neuronas gauna įvesties duomenų, kurie padauginti iš jiems priskirtų svorių, sumą su pridėtu poslinkiu (angl. *bias*). Sumos rezultatas transformuojamas naudojant aktyvacijos funkciją, gauta reikšmė yra neurono išvestis. Matematinis užrašas atrodo taip [22]:

$$y(k) = F \left(\sum_{i=0}^m w_i(k) \cdot x_i(k) + b \right),$$

kur $x_i(k)$ yra i -toji įvesties reikšmė laiku k , $w_i(k)$ yra i -tojo svorio reikšmė laiku k , b - poslinkis, F - aktyvacijos funkcija, $y(k)$ - išvesties reikšmė laiku k . Aktyvacijos funkcijos pasirinkimas priklauso nuo sprendžiamos problemos, dažniausiai renkamasi iš šių aktyvacijos funkcijų: laiptinė funkcija (angl. *step function*), tiesinė funkcija (angl. *linear function*), sigmoidė (angl. *sigmoid*).

Neuroninį tinklą sudaro tarpusavyje sujungti neuronai su jiems priskirtais svoriais, kuriuos modelis apmokymo metu koreguoja taip, kad galėtų teisingai prognozuoti gauto objekto klasę. Šiame darbe naudojamas daugiasluoksnis tiesioginio sklidimo (angl. *feed-forward*) neuroninis tinklas, kuris sudarytas iš įvesties sluoksnio, vieno ar daugiau paslėptųjų sluoksnių ir išvesties sluoksnio. Dirbtinis neuroninis tinklas turintis daugiau nei vieną paslėptąjį sluoksnį vadinamas giliuoju neuroniniu tinklu.



1.3 pav.: Daugiasluoksnis neuroninis tinklas [11]

Į pirmą (įvesties) sluoksnį patenka objekto, priklausančio apmokymo duomenų aibeii, požymiai. Šie įvesties elementai toliau yra padauginami iš svorių ir patenka į antrąjį (paslėptąjį) sluoksnį, kuriame yra prieš tai aptarti neuronai. Išvestis iš paslėptojo sluoksnio neuronų gali keliauti į sekantį paslėptąjį sluoksnį arba į išvesties sluoksnį, kuris grąžina modelio prognozę. Aprašytas modelis vadinamas tiesioginio sklidimo modeliu, nes svoriai negrįžta į prieš tai buvusių sluoksnius. Taip pat galima paminėti, jog šiame modelyje sluoksniai vadinami pilnai sujungtais (angl. *fully-connected*), nes vieno sluoksnio neuronai yra sujungti su šalimais esančių sluoksnių neuronais.

Neuroninis tinklas yra apmokomas naudojant atgalinės sklaidos metodą (angl. *backpropagation*). Kiekvienam modelio apmokymo objektui neuroninis tinklas pateikia prognozę, tada apskaičiuojama pasirinkta paklaidos (angl. *error*) funkcija (*MSE* arba entropija), kuri apskaičiuoja neatitikimą tarp neuroninio tinklo prognozių ir tikrų reikšmių. Toliau apskaičiuojama kiek kiekvienas neuronas paskutiniame paslėptajame sluoksnyje prisidėjo prie išvesties neuronų padarytų klaidų, tai įgyvendinama apskaičiuojant paklaidos funkcijos gradientus kiekvieno svorio atžvilgiu,

šis žingsnis kartojamas grįžtant per visus sluoksnius atgal. Galiausiai pritaikomas gradientinio nusileidimo metodas siekiant rasti svorius minimizuojančius paklaidos funkcijos reikšmę. Detalesnis metodo paaiškinimas pateikiamas čia [16].

Neuroniniai tinklai yra lankstus modelis, tinkamas tiesiškai neatskiriamiems duomenims ir darbui su didelės apimties duomenų rinkiniais. Tačiau šio modelio apmokymas gali trukti ganėtinai ilgai, geresniam veikimui reikalingi dideli duomenų kiekiai, o gautą rezultatą labai sunku interpretuoti. Naudojant šį modelį duomenis būtina standartizuoti.

1.3 Klasifikavimo modelių tikslumo vertinimo rodikliai

Toliau bus pateikiami keli rodikliai skirti mašininio mokymosi klasifikavimo modelių tikslumui įvertinti. Rodikliai parinkti remiantis A. Sinem Uysal ir John M. Mulvey straipsniu[34].

Pirmasis rodiklis yra **tikslumas** (angl. *accuracy*), toliau bus žymimas *ACC*. Turint režimo prognozės tikimybę \hat{y}_t (šiuo atveju bus žiūrima į krizinio režimo prognozės tikimybes) ir tikrą režimo žymę $y_t \in \{1, 0\}$, rodiklis apskaičiuojamas taip:

$$ACC = \frac{1}{N} \sum_{t=1}^N (\mathbb{1}_{\{\hat{y}_t \geq c\}} y_t + (1 - \mathbb{1}_{\{\hat{y}_t \geq c\}})(1 - y_t)), \quad (6)$$

čia N yra visų stebėjimų skaičius, $c = 0.5$ yra slenkstis atskiriantis režimus, jei $\hat{y}_t \geq c$ bus laikoma, kad modelis prognozuoja krizinį režimą, kuris žymimas 1, kitu atveju laikoma, kad modelio prognozė - stabilus režimas, žymimas 0. Kaip matoma iš (6) formulės, tikslumo rodiklis apskaičiuojamas sudedant visų teisingai suklasifikuotų stebėjimų skaičių ir padalijant jį iš visų stebėjimų skaičiaus.

Antras naudojamas rodiklis yra **Brier rodiklis**, angliškai dar vadinamas *Quadratic Probabilistic Score*, todėl bus žymimas *QPS*. Šis rodiklis pamatuoja prognozuojamų tikimybių tikslumą.

$$QPS = \frac{1}{N} \sum_{t=1}^N (\hat{y}_t - y_t)^2,$$

QPS įgyja reikšmes tarp 0 ir 1. Kuo mažesnė rodiklio reikšmė, tuo geresnės tikimybių prognozės. Svarbu atkreipti dėmesį, kad modelių, turinčių tokią pačią *ACC* reikšmę, *QPS* rodiklių reikšmės gali nesutapti. Platesnę informaciją apie Brier rodiklių naudojimą galima rasti čia [4].

Geresniam klasifikavimo modelių daromų klaidų įvertinimui buvo skaičiuojamos **maišaties matricos** (angl. *confusion matrix*):

		Prognozė	
Tikra reikšmė	0	TN	FP
	1	FN	TP
		0	1

1.4 pav.: Maišaties matrica

TN - true negative, teisingai suklasifikuotų 0 klasei priklausančių objektų skaičius.

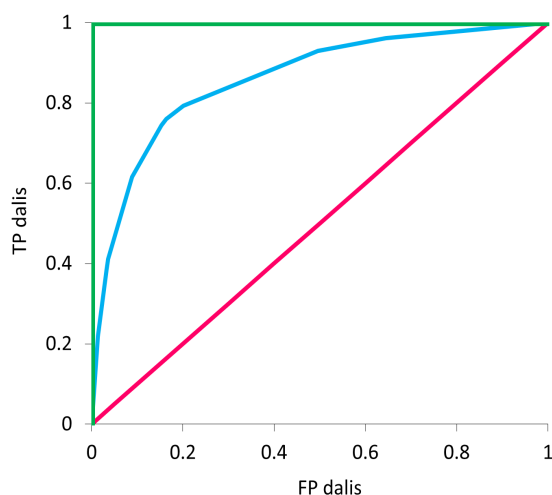
TP - true positive, teisingai suklasifikuotų 1 klasei priklausančių objektų skaičius.

FN - false negative, neteisingai 0 klasei priskirtų objektų skaičius.

FP - false positive, neteisingai 1 klasei priskirtų objektų skaičius.

TP dalis = $TP / (TP + FN)$, FP dalis = $FP / (FP + TN)$.

Remiantis šia matrica galima apskaičiuoti kitus rodiklius. Vienas jų yra **plotas po ROC** (angl. *Receiver operating characteristic*) **kreive**, žymimas *AUC* (angl. *area under the curve*). ROC kreivė yra grafikas atvaizduojantis TP dalį y ašyje ir FP dalį x ašyje esant skirtingiems c slenksčiams. Tobulo modelio atveju plotas po ROC kreive lygus 1, o kreivė yra apverstos L raidės formos. Jei *AUC* reikšmė 0.5, tai modelis prognozuoja klases atsitiktinai, o kreivės forma yra tiesė $y = x$.



1.5 pav.: ROC kreivės pavyzdžiai

ROC kreivės naudingos, kai duomenys yra nesubalansuoti (vienos klasės stebėjimų yra gerokai daugiau nei kitos), nes remiantis kreive galima pasirinkti tinkamiausią c slenkstį. Daugiau informacijos apie ROC kreives galima rasti Tom Fawcett straipsnyje [10].

Paskutinis modelių tikslumo vertinimo rodiklis yra **Matthew koreliacijos koeficientas** (angl. *Matthew's correlation coefficient*), žymimas MCC :

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}.$$

Didžiausia reikšmė, kurią gali įgyti rodiklis yra 1, mažiausia -1, 0 reiškia atsitiktines klasių prognozes. Šis rodiklis išvengia nesubalansuotų duomenų problemos ir įgyja aukštas reikšmes tik tada, kai klasifikavimo modelis geba teisingai suklasifikuoti daugumą objektų priklausančių 0 klasei ir daugumą objektų priklausančių 1 klasei. Išsamesnė informaciją apie MCC rodiklį pateikiama Davide Chicco ir Giuseppe Jurman straipsnyje [7]

1.4 Rodikliai strategijos rezultatų palyginimui

Tam, kad būtų galima įvertinti strategijos generuojamą grąžą ir riziką, bus konstruojami du kontroliniai portfeliai. Vieno kontrolinio portfelio pastovus svoris į S&P 500 indeksą apskaičiuojamas taip, kad standartinis nuokrypis virš nerizikingos investicijos sutaptų su strategijos standartiniu nuokrypiu virš nerizikingos investicijos. Kito kontrolinio portfelio pastovus svoris yra toks, kad portfelio β reikšmės sutaptų. Toliau pateikiami skaičiuoti rodikliai, kurie leidžia palyginti strategijų rezultatus tarpusavyje, įvertinti grąžos ir standartinio nuokrypio santykį bei riziką.

Pirma apibrėžiamas **vidutinės grąžos** skaičiavimas. Tarkime investicijos kaina laiku t yra P_t , tada vieno periodo paprastoji grąža r_t apskaičiuojama taip:

$$r_t = \frac{P_t}{P_{t-1}} - 1,$$

o vidutinė grąža pasirinktam horizontui h yra:

$$\left(\prod_{t=1}^n (1 + r_t) \right)^{h/n} - 1,$$

kur n - visa investavimo trukmė. Be to, vidutinė grąža virš nerizikingos investicijos apskaičiuojama taip:

$$E(r_p - r_f) = \left(\prod_{t=1}^n \frac{1 + r_{p_t}}{1 + r_{f_t}} \right)^{h/n} - 1,$$

čia r_p portfelio gražos, r_f nerizikingos investicijos gražos. Rodiklių lentelėse pateikiama vidutinė metinė graža virš nerizikingos investicijos ir žymima *Mean*.

Turint gražas galima apskaičiuoti **istorinį imties standartinį nuokrypį**:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (r_i - r)^2},$$

čia r_i yra i - toji laiko eilutės graža, r - aritmetinis imties vidurkis, n - imties dydis. Šis dydis skaičiuojamas metiniam horizontui ir rodiklių lentelėse žymimas *Stdev*.

Kitas naudojamas rodiklis yra **Šarpo santykis** (angl. *Sharpe ratio*), kuris buvo pristatytas 1966 metais [30]. Toliau bus žymimas *SharpeR*.

$$SharpeR = \frac{E(r_p) - r_f}{\sigma_p},$$

čia $E(r_p)$ ir r_f yra atitinkamai vidutinė portfelio bei nerizikingos investicijos graža, σ_p - portfelio standartinis nuokrypis. Apskaičiavus skirtumą $E(r_p) - r_f$ gaunama vidutinė graža virš nerizikingos investicijos. Lyginant investicijas ar strategijas pagal *SharpeR* rodiklį, patrauklesnė būtų ta, kurios *SharpeR* rodiklis didesnis.

Apskaičiavus Šarpo santykį, toliau skaičiuojamas **tikimybinis Šarpo santykis** (angl. *Probabilistic Sharpe Ratio*), jį 2012 metais pristatė Bailey ir López de Prado. Šis rodiklis yra tikimybė, kad *SharpeR* viršija kažkokį pasirinktą slenkstinį *SharpeR** rodiklį [3]. Rodiklis bus žymimas *ProbSR*. Tikimybinio Šarpo santykio skaičiavimas yra naudingas, nes gražų skirstinys dažniausiai nėra normalusis, o nuo to priklauso *SharpeR* įvertio tikslumas, nors pačio *SharpeR* rodiklio skirstinys yra normalusis. *ProbSR* rodiklis apskaičiuojamas taip: jei tikrasis *SharpeR* = *SharpeR**, tai statistika $\hat{z}[SharpeR^*]$ - asimptotiškai pasiskirsčiusi kaip standartinis normalusis skirstinys

$$\hat{z}[SharpeR^*] = \frac{(SharpeR - SharpeR^*)\sqrt{n-1}}{\sqrt{1 - \hat{\gamma}_3 SharpeR + \frac{\hat{\gamma}_4 - 1}{4} SharpeR^2}} \xrightarrow{a} Z,$$

čia n yra gražų eilutės ilgis, $SharpeR$ yra gražų eilutės *SharpeR* įvertis, $\hat{\gamma}_3$ - gražų asimetrijos koeficientas, $\hat{\gamma}_4$ - gražų eilutės kurtosis (ekscesas + 3). Rodiklių dažnis turi sutapti su gražų eilutės dažniu, tai yra turint mėnesines gražas, *SharpeR* rodiklio įvertis $SharpeR$ irgi mėnesinis. Toliau apskaičiuojamas $SharpeR$ įvertio standartinis nuokrypis [9]:

$$\hat{\sigma}_{SharpeR} = \sqrt{\frac{1 - \hat{\gamma}_3 SharpeR + \frac{\hat{\gamma}_4 - 1}{4} SharpeR^2}{n-1}}.$$

Tarus, kad $SharpeR^* = 0$, $SharpeR$ įvertinamas pagal hipotezės testą:

$$\begin{cases} H_0 : SharpeR = 0 \\ H_1 : SharpeR > 0. \end{cases}$$

Apskaičiuojama, tikimybė, jog $SharpeR$ viršija pasirinktą slenkstinį $SharpeR^*$ rodiklį, žymima $ProbSR$:

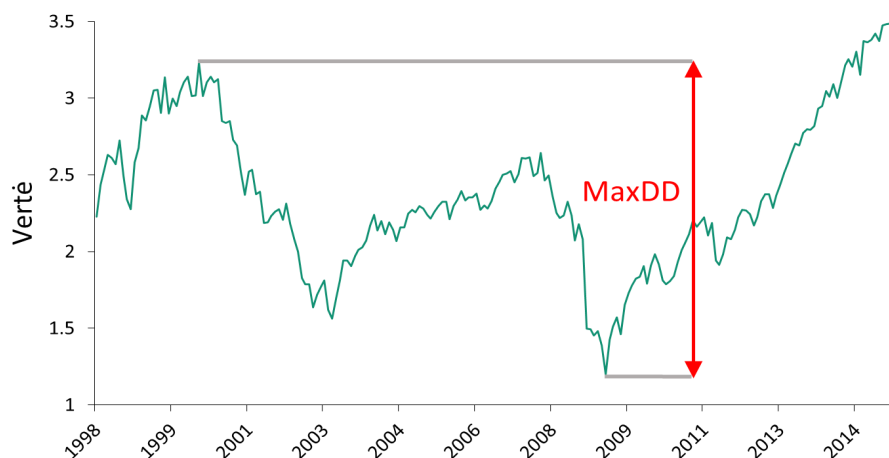
$$ProbSR(\hat{z}[0]) = Z \left[\frac{SharpeR\sqrt{n-1}}{\sqrt{1 - \hat{\gamma}_3 SharpeR + \frac{\hat{\gamma}_4 - 1}{4} SharpeR^2}} \right],$$

čia Z - standartinio normalaus skirstinio pasiskirstymo funkcija. Pasirinkus C pasikliovimo lygmenį, pavyzdžiui, 90 %, tikrinama ar $ProbSR(\hat{z}[0]) \geq C$. Jei $ProbSR(\hat{z}[0]) \geq C$, tai nulinė hipotezė atmetama.

Kitas šiame darbe naudojamas rodiklis yra **didžiausias vertės smukimas** (angl. *maximum drawdown*), toliau bus žymimas $MaxDD$. Šis rodiklis parodo koks buvo didžiausias investicijos ar portfelio vertės smukimas matuojant nuo didžiausios buvusios vertės iki mažiausios vertės, prieš naujos didžiausios vertės pasiekimą.

$$MaxDD = \frac{D - M}{D},$$

kur D - didžiausia vertė prieš didžiausią nuosmukį, M - mažiausia vertė iki naujos didžiausios vertės.



1.6 pav.: Didžiausio vertės nuosmukio vizualizacija

Lyginant dvi investicijas turinčias tokią pačią vidutinę grąžą, investicija, kurios $MaxDD$ mažesnis yra patrauklesnė.

Strategijų rizikos palyginimui taip pat bus skaičiuojama **vertės pokyčio rizika** (angl. *Value at Risk*). Šis rodiklis parodo kokio didžiausio nuostolio, kuris nebūtų viršijamas su tam tikru pasiklovimo lygmeniu, galima tikėtis per pasirinktą laikotarpį. Šiame darbe skaičiuojama metinio horizonto vertės pokyčio rizika su 95% pasiklovimo lygmeniu, apskaičiuotas rodiklis bus žymimas VaR_{95} . Toliau pateikiamas tikslesnis apibrėžimas. Pažymėkime nuostolį (angl. *loss*) raide L , tada VaR su pasiklovimo lygmeniu C yra toks kvantilis, jog

$$P(L \leq VaR) = 1 - C.$$

Lyginant strategijų, kurių vidutinės grąžos panašios ar vienodos, rezultatus tarpusavyje, patrauklesnės yra strategijos turinčios žemesnę vertės pokyčio riziką. Detalesnę informaciją apie vertės pokyčio rizikos skaičiavimą ir taikymą galima rasti Philippe Jorion knygoje[18].

Dar vienas naudojamas rodiklis yra β **koeficientas**, susijęs su finansų aktyvų įkainojimo modeliu (angl. *capital asset pricing model* - CAPM), kurį 1964 m. pristatė William Sharpe. Modelio formulė[19]:

$$E(r_i) = r_f + \beta_i(E(r_m) - r_f),$$

čia r_i investicijos i grąža, r_f nerizikingos investicijos grąža, r_m rinkos grąža (šiam darbe rinka bus laikomas S&P 500 indeksas), $E(r_i)$ - tikėtina investicijos i grąža, $E(r_m) - r_f$ - tikėtina rinkos rizikos premija, $\beta_i = \frac{cov(r_i, r_m)}{var(r_m)}$.

Koeficientas β yra sisteminės rizikos matas. Jei $\beta = 1$, tai investicijos rizika sutampa su rinkos rizika, jei $\beta < 1$ ($\beta > 1$), tai rizika mažesnė (didesnė) už rinkos riziką.

Paskutinis naudojamas rodiklis, kuris panašus į Šarpo santykį, yra **Treynor santykis** (angl. *Treynor ratio*), kuris žymimas $TreynorR$.

$$TreynorR = \frac{E(r_p) - r_f}{\beta_p},$$

čia β_p yra portfelio beta reikšmė. Šis rodiklis priklauso sisteminės rizikos β_p , kuo didesnė Treynor santykio reikšmė, tuo investicija yra patrauklesnė.

2 Tiriamoji dalis

2.1 Duomenys

Šiame darbe režimų prognozavimui mašininio mokymosi klasifikavimo modeliais yra naudojamas Michael W. McCracken ir Serena Ng sukurtas makroekonominių rodiklių rinkinys¹, kurį sudaro 127 rodikliai. Duomenų rinkinyje esantys rodikliai suskirstyti į 8 grupes:

1. Produkcija ir pajamos (angl. *Output and Income*),
2. Darbo rinka (angl. *Labor Market*),
3. Vartojimas ir užsakymai (angl. *Consumption and Orders*),
4. Užsakymai ir atsargos (angl. *Orders and Inventories*),
5. Pinigai ir kreditas (angl. *Money and Credit*),
6. Palūkanų normos ir valiutų kursai (angl. *Interest rates and Exchange Rates*),
7. Kainų indeksai (angl. *Prices*),
8. Akcijų rinka (angl. *Stock Market*).

Pilnas rodiklių sąrašas pateikiamas priedų skylyje. Naudojami duomenys - vintažiniai (vintažiniai duomenys, tai laiko eilučių rinkinys, kur kiekvienai datai priskirti stebėjimai buvo naujaisi tuo laiku turėti duomenys ir vėliau nebuvo atnaujinti), laiko eilučių pradžia yra 1959 metai. Duomenų surinkimo metu atsižvelgta į tai, kad kai kurių rodiklių laiko eilučių duomenys nebeteikiami, todėl jos sujungtos su atitinkamais naujais rodikliais, kurių laiko eilutės trumpesnės. Pavyzdžiui, OILP-RICE rodiklio duomenų pradžia yra 1946 sausio mėnesį, bet nuo 2013 liepos mėnesio duomenys nebeteikiami, taigi vietoj šio rodiklio gali būti naudojamas WTI rodiklis, tačiau jo pradžia yra nuo 1986 sausio mėnesio, ši problema sprendžiama sujungiant šias dvi eilutes. Be to, naudojant šį duomenų rinkinį, norint laiko eilutes paversti stacionariomis, yra naudojamos tokios duomenų transformacijos pagal autorių rekomendacijas: 1. Duomenys nekeičiami, 2. Δx_t , 3. $\Delta^2 x_t$, 4. $\log(x_t)$, 5. $\Delta \log(x_t)$, 6. $\Delta^2 \log(x_t)$, 7. $\Delta(x_t/x_{t-1} - 1)$, čia x_t - rodiklio laiko eilutė. Kokios transformacijos pritaikomos atitinkamiems duomenims pažymėta priedų skyrelio 1 bei 2 lentelėse.

¹<https://research.stlouisfed.org/econ/mccracken/fred-databases/>

Plačiau apie šį duomenų rinkinį skaitykite Michael W. McCracken ir Serena Ng straipsnyje [23]. Naudojamų rodiklių duomenys yra mėnesiniai, nagrinėjama rodiklių imtis nuo 1959 sausio iki 2022 balandžio mėnesio. Režimų nustatymui su ℓ_1 filtru naudojamos savaitinės S&P 500 kainos indekso gražos. Nerizikinga investicija laikoma JAV Vyriausybės 3 mėn. obligacija.

Režimų pasikeitimo modelyje naudojami tiek makroekonominiai, tiek rinkos rodikliai gaunami iš Bloomberg. Modelyje naudojami rinkos rodikliai: JAV vyriausybės 3 mėnesių bei 2, 5 ir 10 metų obligacijos palūkanų norma, BAA ir AAA reitingą turinčių korporacijų obligacijų palūkanų norma, JAV vyriausybės 1+ metų obligacijų indekso gražos, S&P 500 indekso gražos ir JAV vyriausybės 3 mėnesių obligacijos gražos, šie duomenys atnaujinami kiekvieną darbo dieną. Režimų pasikeitimo modelyje naudojami makroekonominiai rodikliai: Čikagos Fed nacionalinio aktyvumo indekso 3 mėnesių slenkantis vidurkis (atnaujinamas apytiksliai 20 mėnesio dieną), Čikagos pramonės pasitikėjimo rodiklis (atnaujinamas pirmą mėnesio d.d.), JAV nedarbo lygis (atnaujinamas mėnesio pirmos savaitės penktadienį), pramonės produkcija (atnaujinama paskutinį mėnesio antradienį) bei JAV vartotojų pasitikėjimo rodiklis (atnaujinamas mėnesio 14-16 dienomis). Vintažinių duomenų laiko eilutės pradžia yra 1977 metais. Naudojamų rodiklių sąrašas taip pat pateikiamas priedų skyriuje esančioje 3 lentelėje.

2.2 Režimų nustatymas

Kadangi režimų prognozavimui bus naudojami teorinėje dalyje pristatyti prižiūrimo mokymosi klasifikavimo modeliai, tai sekantis žingsnis - klasių žymių nustatymas, šios žymės reikalingos modelių apmokymui. Toliau pristatomas klasių žymių (šiuo atveju klasės yra stabilus ir krizinis režimai) nustatymo būdas.

Režimo žymių nustatymui buvo naudojamas ℓ_1 trendo filtras pristatytas teorinėje darbo dalyje. Optimizavimo problemos (1) formulėje gražų eilutės r_t apskaičiavimui naudojamas S&P 500 kainos indeksas, o laiko eilutės r_t trendo nustatymui patikrinti du būdai pasirenkant λ reikšmes 0.1 ir 0.15. Jei λ parametras per mažas, tai skaičiuojant trendą įtraukiamas papildomas triukšmas ir neteisingai identifikuojami režimo pasikeitimai, jei parametras per didelis, tai neužfiksuojami svarbūs buvusių krizių periodai [24]. Išnagrinėjus gaunamas gražų laiko eilutės trendo reikšmes, nuspręsta naudoti $\lambda = 0.1$ reikšmę. Trendas buvo skaičiuojamas naudojant tiek

mėnesines, tiek savaitines gražas, tačiau pastebėta, jog trendo, o tuo pačiu ir režimo, pasikeitimą geriau identifikuoja pseudo savaitinės gražos (toliau vadinamos savaitinėmis gražomis). Savaitinės gražos yra skaičiuojamos 4, 11, 18 ir 25 mėnesio dienoms. Gražos pasirinktoms dienoms buvo apskaičiuotos iš dieninių duomenų išfiltravus 4, 11, 18 ir 25 dienų S&P 500 indekso kainų uždarymo reikšmes kiekviename mėnesyje. Jei kažkuri diena yra ne darbo, tai imama prieš tai buvusios darbo dienos uždarymo reikšmė.

Mašininio mokymosi modelių rezultatai bus vertinami dviejose imtyse, nuo 1959 metų iki 2014 metų galo bei nuo 2015 metų pradžios iki 2022 metų balandžio mėnesio, todėl trendo skaičiavimas taip pat padalijamas į dvi dalis. Pirmoje imtyje trendas skaičiuojamas paduodant visas turimas S&P 500 kainos indekso gražas nuo 1959 metų iki 2014 metų galo. Antroje imtyje trendas skaičiuojamas paduodant po vieną naują gražų reikšmę nuo 2015 metų pradžios, kiekvieną kartą gavus naują reikšmę, trendo eilutė perskaičiuojama ir paskutinė gauta reikšmė prijungiama prie prieš tai buvusių trendo eilutės reikšmių. Antra dalis bus naudojama modelių veikimo testavimui, kai režimai nustatomi neturint informacijos apie ateitį.

Apskaičiavus trendą, reikia nustatyti slenkstį, pagal kurį krizinis režimas atskiriamas nuo stabilaus režimo. Šiame darbe patikrinami keli skirtingi slenksčio nustatymo variantai. Pirmasis iš jų, kai slenkstis yra 0, tai yra, jei trendo reikšmė mažesnė už 0, tai laikoma, jog tai yra krizinis režimas, kuris žymimas 1, kitu atveju režimas stabilus ir žymimas 0. Gauti režimai matomi priedų skyriuje esančiame 2 pav.

Likę variantai yra tokie: iš trendo eilutės reikšmių apskaičiuojami 10-tas, 20-tas, 30-tas ir 40-tas procentiliai, kiekvieną iš jų laikant režimo pasikeitimo slenksčiu gaunami skirtingi galimi režimų atskyrimai. Gautų režimų grafikai pateikiami priedų skyriuje esančiuose 3, 4, 5 ir 6 pav.

Buvo nuspręsta, kad geriausiai režimai atskiriami slenksčiu pasirinkus 30-tą procentilį, nes su šiuo slenksčiu išskiriamas 1976 metais prasidėjęs akcijų kritimo laikotarpis, geriau matomas 1990-ųjų recesijos periodas, išskiriama 1998 – 1999 metais vykusios Rusijos finansinė krizė ir 2011 rugpjūtį prasidėjęs akcijų rinkos kritimas dėl Euro zonos skolų krizės plitimo baimės ir nerimo dėl lėtėjančių JAV ekonominio augimo rodiklių. Papildomai buvo patikrinti 25-tas ir 35-tas procentiliai, tačiau jie nedavė geresnių režimų atskyrimo rezultatų.

2.3 Duomenų paruošimas

Sekantis žingsnis po režimų žymių nustatymo yra makroekonominių duomenų paruošimas. Šiame darbe makroekonominių duomenų rinkinys yra požymiai, kurie pateikiami klasifikavimo modelių apmokymui ir prognozėms gauti.

Makroekonominiai rodikliai yra mėnesiniai, todėl prieš tai aptarto trendo filtro režimų žymės, kurios yra savaitinės, paverčiamos keturiais mėnesinių režimų rinkiniais, išskirstant žymes į 4, 11, 18 ir 25 mėnesių dienų rinkinius (tokios dienos parinktos remiantis Lietuvos banko režimų pasikeitimo modelio skaičiavimo procedūra). Makroekonominių rodiklių rinkinys atnaujinamas iki kiekvieno mėnesio 11 dienos. Rodikliai esantys rinkinyje vėluoja apytiksliai mėnesį, pavyzdžiui, balandžio 11 dieną turimas naujas rodiklių rinkinys, kuriame yra duomenys už vasario mėnesį. Remiantis makroekonominių rodiklių atnaujinimo datomis, yra sudaromi keturi duomenų rinkiniai. Tiek rinkinių sudaroma dėl to, kad norima sukurti modelius, kuriuos galima naudoti skaičiuojant tiek mėnesiniu, tiek savaitiniu dažniu. Duomenų rinkiniai sudaromi tokiu būdu, kad atitinkama mėnesio dieną rinkinyje būtų iki tos dienos išėję rodikliai. Pavyzdžiui, modelio prognozę skaičiuojant birželio 4 dieną, naujausi makroekonominiai duomenys, esantys FRED puslapio rinkinyje, yra už kovo mėnesį. Tada bus laikoma, kad naujausia mėnesinio rinkinio eilutės data yra kovo 4 d. (iš tikrųjų duomenys yra už visą balandžio mėnesį, bet toks datų žymėjimas taikomas dėl patogumo vėliau gautas prognozes sujungti į savaitinį dažnį). Be makroekonominių duomenų, rinkinyje yra ir rinkos rodikliai, peržiūrima, kad atitinkamomis dienomis laiko eilutės būtų sudėliotos taip, jog prognozių skaičiavimo metu duomenyse atsispindėtų naujausia tuo metu buvusi rinkos informacija. Skaičiuojant birželio mėnesio 11, 18 ir 25 dienomis naujasi makroekonominiai duomenys rodiklių rinkinyje būtų už balandžio mėnesį.

Makroekonominių rodiklių laiko eilutės transformuojamos taip, kad būtų stacionarios. Duomenų transformacijos parenkamos pagal autorių rekomendacijas, kurias galima rasti šio darbo priedų skiltyje.

Iš pilno rodiklių rinkinio dėl duomenų trūkumo buvo išmesti šie rodikliai: New Orders for Consumer Goods, New Orders for Nondefense Capital Goods, Consumer Sentiment Index, Trade Weighted U.S. Dollar Index, Crude Oil, spliced WTI and Cushing, VIX. Be to, išmetami pirmi laiko eilučių duomenų metai, nes nėra New Private Housing Permits (SAAR); New Private Housing Permits, Northeast (SAAR);

New Private Housing Permits, Midwest (SAAR); New Private Housing Permits, South (SAAR); New Private Housing Permits, West (SAAR) rodiklių duomenų.

Toliau makroekonominiai rodikliai apjungiami su turimomis režimų žymėmis. Sutvarkius duomenis, naujos laiko eilutės prasideda nuo 1960-02-29, laikant, kad rodikliai yra už mėnesio paskutinę dieną. Tariama, kad naujas makroekonominių rodiklių rinkinys gaunamas mėnesio 11 dieną ir nesikeičia iki sekančio mėnesio 4 dienos. Prie prieš tai aprašytų sudarytų rinkinių laiko eilučių (4, 11, 18 ir 25 dienai), priskiriamos režimų žymės, kurios reikalingos mašininio mokymosi modelių apmokymui. Prijungimas daromas taip: tarkime turime 4 mėnesio dienos rinkinį, jame duomenys yra mėnesiniai, iš apskaičiuotų savaitinių režimų žymių išfiltruojamos 4 mėnesių dienos ir tomis dienomis buvę režimai prijungiami prie viso rodiklių rinkinio atsižvelgiant į laikotarpius už kuriuos yra paskelbti rodikliai. Pavyzdžiui, prie rodiklių eilutės, kuri yra už kovo mėnesį prijungiama režimo žymė gauta kovo 4 dieną (laikoma, kad gautas režimas nesikeičia visą mėnesį). Ta pati procedūra atliekama 11, 18 ir 25 dienos rinkiniams.

Šiame darbe nagrinėti modeliai yra skaičiuojami tiek savaitiniu, tiek mėnesiniu dažniu. Skaičiuojant mėnesiniu dažniu, laikoma, kad visi mėnesiniai makroekonominių rodiklių duomenys gaunami 11 dieną, tai yra panaudojamas vienas iš turimų rodiklių duomenų rinkinių. Skaičiuojant savaitiniu dažniu, tie patys modeliai skaičiuojami 4 kartus naudojant 4 rodiklių rinkinius, kurie yra 4, 11, 18 ir 25 mėnesių dienoms. Skaičiuojant tokiu būdu duomenys rinkiniuose yra mėnesiniai, bet gautos prognozės apjungiamos ir gaunamas savaitinis dažnis.

2.4 Mašininio mokymosi modelių konstravimas

Paruošus makroekonominius duomenis ir režimų žymes, turi būti atskirtos apmokymo ir testavimo imtys. Toliau pateikiamas detalus aprašymas kaip buvo padarytas šių imčių atskyrimas ir kaip apmokomi klasifikavimo modeliai.

Konstruojant teorinėje dalyje aprašytus mašininio mokymosi modelius ir tikrinant jų tikslumą, duomenų laiko eilutės padalinamos į tris dalis. Duomenys nuo 1960 kovo mėnesio iki 1978 sausio mėnesio naudojami hiperparametrų parinkimui, kurie bus naudojami modeliuose antroje imtyje. Likę duomenys padalijami į dvi dalis, nuo 1978 vasario mėnesio iki 2014 gruodžio mėnesio ir nuo 2015 sausio mėnesio iki 2022 balandžio mėnesio. Abi dalys naudojamos modelių veikimo testavimui, tačiau

antroje dalyje hiperparametrai perskaičiuojami naudojant duomenis nuo 1960 kovo mėnesio iki 2014 gruodžio mėnesio, o režimų žymės yra gaunamos prie režimų nustatymo aprašytu būdu, kai trendas perskaičiuojamas kiekvieną kartą gavus naują reikšmę. Taip elgiamasi siekiant patikrinti kokie būtų buvę rezultatai, jei modeliai režimų prognozavimui būtų naudoti nuo 2015 metų pradžios, kai nėra duomenų apie ateitį, priešingai nei imtyje nuo 1978 vasario mėnesio iki 2014 gruodžio mėnesio, kurioje režimų nustatymui buvo panaudota visa imtis nuo 1959 metų iki 2014 metų galo.

Padalinus duomenis į aukščiau aprašytas imtis, pereinama prie optimalių hiperparametrų parinkimo, tam naudojama tinklelio paieška (angl. *grid search*). Modelių įvertinimui su įvairiais hiperparametrų rinkiniais naudojama stratifikuota k -dalių kryžminė patikra (angl. *stratified k-fold cross validation*)², kadangi duomenys yra nesubalansuoti, tai yra krizinių režimų dalis sudaro apie 30% duomenų. Dažnai naudojama k reikšmė yra tarp 5 ir 10, šiame darbe pasirinkta reikšmė yra 10. Atliekant kryžminę patikrą, modelio tinkamumo vertinimo rodikliu pasirinktas Brier rodiklis, nes naudojant šį rodiklį baudžiamas netikslus režimų tikimybių prognozavimas.

Kaip jau minėta, šiame darbe modeliai skaičiuojami tiek mėnesiniu, tiek savaitiniu horizontu, bet savaitinio horizonto modelių sudarymui naudojami 4 duomenų rinkiniai, kuriuose duomenys mėnesiniai. Duomenų rinkinių struktūra panaši, todėl nuspręsta optimalių hiperparametrų ieškoti pasirinkus vieną iš rinkinių, tiksliau 11 mėnesių dienos rinkinį, nes jis bus naudojamas ir mėnesinio horizonto modeliams.

Nustačius optimalius modelių hiperparametrus, pereinama prie modelių testavimo. Modeliai testuojami naudojant slenkantį 150 stebėjimų langą. Pavyzdžiui, imama 150 stebėjimų iki 1978 vasario mėnesio (neimtina), modelis apmokomas naudojant stebėjimus patenkančius į tą langą. Apmokius modelį, daroma prognozė iš 1978 vasario mėnesio duomenų. Tokiu būdu skaičiuojamos prognozės iki 2014 metų gruodžio mėnesio, po to optimalūs hiperparametrai parenkami iš naujo ir testavimas vykdomas kitoje imtyje, nuo 2015 metų pradžios iki 2022 metų balandžio mėnesio. Testavimui pasirinktas slenkantis langas, nes laikui bėgant gali keistis rodiklių tarpusavio sąryšiai, lango dydis yra 150 stebėjimų, nes į tokį langą patenka bent keli krizinio režimo stebėjimai, kitu atveju negalima tinkamai apmokyti modelių.

²Naudojant stratifikuotą k -dalių kryžminę patikrą kaip ir įprastos k -dalių kryžminės patikros atveju duomenys padalijami į k lygių dalių, bet šiose dalyse išlaikomas pradiniuose duomenyse esantis klasių santykis.

Modeliai taikomi naudojant visus makroekonominis rodiklius ir ekspertiniu vertinimu atrinktą dalį svarbiausių rodiklių. Taigi viso bus keturi skirtingi modelių taikymo variantai, modeliai skaičiuojami mėnesiniu horizontu su visais galimais rodikliais ir dalimi rodiklių bei modeliai skaičiuojami savaitiniu horizontu su visais galimais rodikliais ir dalimi rodiklių.

Turint sutvarkytus makroekonominis duomenis, režimų žymes ir atskirtas apmokymo bei testavimo imtis, sekantys žingsniai yra modelių hiperparametrų nustatymas, modelių apmokymas ir gautų prognozių rezultatų vertinimas.

Logistinės regresijos modeliui su elastinio tinklo regularizacija ieškomi hiperparametrai yra α ir λ reikšmės. Pasinaudojus tinklelio paieška (paieškoje α reikšmės kito nuo 0.1 iki 0.95, o λ nuo 0.001 iki 10) buvo nustatyta, jog optimalios hiperparametrų reikšmės imtyje iki 1978 metų yra $\alpha = 0.95$ ir $\lambda = 1$ (imtyje iki 2015 $\alpha = 0.95$ ir $\lambda = 10$). Kuo λ reikšmė mažesnė, tuo didesnė regularizacija. Modeliui su **L1** (LASSO) regularizacija, kai $\alpha = 1$, imtyje iki 1978, nustatyta optimali hiperparametro λ reikšmė yra 10 (imtyje iki 2015 $\alpha = 1$). Logistinės regresijos atveju su **L2** (Ridge) regularizacija, kai $\alpha = 0$, gauta, jog optimali hiperparametro λ reikšmė imtyje iki 1978 yra 100 (imtyje iki 2015 nepasikeitė). Modelių su L1 ir L2 regularizacija tinklelio paieškoje reikšmės kito nuo 0.0001 iki 100.

Atsitiktinio miško modeliui nustatyti optimalūs hiperparametrai:

1. didžiausias galimas sprendimo medžių gylis (*max_depth*): 50,
2. požymių skaičius, ieškant geriausio padalijimo, (*max_features*): \sqrt{n} [3], kur n visų požymių skaičius,
3. mažiausias galimas objektų skaičius mazge (*min_samples_leaf*): 1 [3],
4. mažiausias objektų, reikalingų padalijimui, (*min_samples_split*) skaičius: 6 [9],
5. sprendimo medžių skaičius modelyje (*n_estimators*): 600[800].

Lenktiniuose skliausteliuose nurodyti atitinkami parametrų pavadinimai naudojant Python *scikit-learn* bibliotekos paketo *model_selection* metodą *GridSearchCV*. Laužtiniuose skliausteliuose nurodytos reikšmės yra gautos nustatant optimalius hiperparametrus imtyje iki 2015 metų. Ieškant optimalių hiperparametrų tinklelyje buvo tokie parametrai:

1. *max_depth*: [10, 30, 50, 60, 70, 80, 90, None],
2. *max_features*: [\sqrt{n} , 3, 4, 5, 7],
3. *min_samples_leaf*: [1, 2, 3, 4],
4. *min_samples_split*: [3, 5, 6, 7, 9, 10],
5. *n_estimators*: [200, 400, 500, 600, 700, 800, 900, 1000].

Gradientinio stiprinimo modeliui nustatyti optimalūs hiperparametrai:

1. mokymosi greitis (*learning_rate*): 0.01,
2. *max_depth*: 1,
3. *max_features*: \sqrt{n} ,
4. *min_samples_leaf*: 1 [7],
5. *min_samples_split*: 2,
6. *n_estimators*: 500 [200],
7. apmokymo objektų dalis parenkama individualių medžių konstravimui (*subsample*): 0.8.

Ieškant optimalių hiperparametrų buvo patikrinti tokie variantai:

1. *learning_rate*: [0.001, 0.01, 0.02, 0.05, 0.1, 0.3],
2. *max_depth*: [1, 3, 5, 7, 9],
3. *max_features*: [\sqrt{n} , 5, 7],
4. *min_samples_leaf*: [1, 2, 3, 5, 7],
5. *min_samples_split*: [2, 3, 5, 7, 9, 11],
6. *n_estimators*: [100, 200, 400, 500, 600, 1000],
7. *subsample*: [0.8, 0.9, 1].

XGBoost modelio optimalūs hiperparametrai:

1. požymių dalis imama sprendimų medžių konstravimui (*colsample_bytree*): 0.3,
2. *learning_rate*: 0.05 [0.01],
3. *max_depth*: 6 [3],
4. *n_estimators*: 100,
5. *subsample* : 0.8.

Optimalių reikšmių ieškota tokiame tinklelyje:

1. *colsample_bytree*: [0.3, 0.5, 0.7, 1],
2. *learning_rate*: [0.01, 0.05, 0.1, 0.2, 0.3],
3. *max_depth*: [2, 3, 5, 6, 10],
4. *n_estimators*: [100, 200, 500, 1000],
5. *subsample*: [0.8, 0.9, 1].

Atraminių vektorių klasifikatoriaus modelio optimalūs hiperparametrai:

1. Paraštės reguliavimo parametras (C): 10,
2. Polinomo laipsnis, kai branduolio funkcija polinomas (*degree*): 2 [1],
3. Branduolio funkcija (*kernel*): 'poly'.

Galimų parametrų tinklelis:

1. C : [0.1, 1, 10, 100],
2. *degree*: [1, 2, 3],
3. *gamma*: [0.001, 0.01, 0.1, 1],
4. *kernel*: ['linear', 'poly', 'rbf', 'sigmoid'].

Tinklelyje esantis *gamma* parametras naudojamas 'rbf' branduolio atveju, jis kontroliuoja skirtingų klasių atskyrimo krašto (angl. *decision boundary*) kreivumą.

Paskutinis naudotas mašininio mokymosi modelis yra **dirbtinis neuroninis tinklas**. Pirmas žingsnis konstruojant šį modelį - struktūros nustatymas. Buvo patikrinti keli modeliai keičiant paslėptųjų sluoksnių skaičių nuo 1 iki 6, juose pritaikant dalies neuronų išmetimo techniką (angl. *dropout*). Galutinis modelio variantas turi 5 paslėptuosius sluoksnius be neuronų išmetimo, nes tai nepagerino rezultatų. Aktyvacijos funkcijos paslėptuosiuose sluoksniuose yra 'ReLU', o išėjimo sluoksnyje 'sigmoid'. Taip pat buvo ieškoma optimalaus apmokymo duomenų kiekio, kuris pateikiamas modelio treniravimui vienu kartu (angl. *batch-size*). Patikrinti keli variantai imant 16, 32, 64, 128 reikšmes, geriausi rezultatai gauti pasirinkus 32. Gradientinio nusileidimo metodo įgyvendinimui naudojamas 'Adam' optimizatorius, nustatytas mokymosi greitis yra 0.0003.

Sukonstravus aukščiau aprašytus modelius, jų apmokymui buvo naudojama visa makroekonominių rodiklių aibė. Toliau keliose skirtingose imtyse buvo patikrinta, kokius rodiklius kaip svarbiausius išskiria mašininio mokymosi modeliai. Apžvelgus gautus rezultatus ir remiantis ilgamete Lietuvos banko tarnautojų patirtimi, buvo nuspręsta iš visos turimų makroekonominių rodiklių aibės modeliui apmokymui paimti makroekonominių rodiklių poaibį. Atrinktų makroekonominių rodiklių dalis: Real Personal Income; IP Index; Civilian Unemployment Rate; All Employees: Total nonfarm; Avg Hourly Earnings: Goods-Producing; Avg Hourly Earnings: Construction; Avg Hourly Earnings: Manufacturing; Housing Starts: Total New Privately Owned; Real personal consumption expenditures; New Orders for Durable Goods; Total Business Inventories; Real M2 Money Stock; Total Nonrevolving Credit; Effective Federal Funds Rate; Moody's Seasoned Baa Corporate Bond Yield Minus 10-Year Treasury Rate; 3-Month Commercial Paper Minus FEDFUNDS; 10-Year Treasury Rate Minus 1-Year Treasury Rate; Moody's Baa Corporate Bond Minus FEDFUNDS; CPI: All Items Less Food; CPI: All items less shelter; S&P's Common Stock Price Index: Composite.

Kadangi skaičiuojant mašininio mokymosi modelius gaunama ne tik režimo prognozė, bet ir prognozuojamo režimo tikimybė, tai nuspręsta tai pritaikyti skaičiuojant modelių tikimybių prognozių medianą ir gautą rezultatą taip pat naudoti strategijų formavimui. Šiuo atveju buvo naudojama mediana, o ne vidurkis, nes šiam rodikliui išsiskiriančios reikšmės (angl. *outliers*) daro mažesnę įtaką.

2.5 Modelių tikslumas ir investavimo rezultatai

Toliau aptariami mašininio mokymosi modelių tikslumo rodikliai, kai modeliai apmokomi naudojant visus galimus požymius savaitiniu ir mėnesiniu horizontu. Visos lentelės su modelių tikslumo rodikliais pateikiamos prieduose. Lentelėse naudojamos modelių pavadinimų santrumpos:

LR-elastic net - logistinė regresija su L1 ir L2 reguliarizacija,

LR-l1 - logistinė regresija su L1 reguliarizacija,

LR-l2 - logistinė regresija su L2 reguliarizacija,

RF - atsitiktinis miškas (angl. *random forest*),

GBM - gradientinis stiprinimas (angl. *gradient boosting*),

XGBM - kitaip implementuota gradientinio stiprinimo versija angliškai modelis

vadinamas *Extreme gradient boosting*,

SVC - atraminių vektorių klasifikatorius (angl. *support vector classifier*),

DNN - gilusis neuroninis tinklas (angl. *deep neural network*),

MEDIAN - aukščiau išvardintų modelių rezultatų mediana.

Žiūrint į modelių, kurie buvo konstruoti prognozavimui mėnesiniu horizontu naudojant visus makroekonominius rodiklius, rezultatus (žr. 4 lentelę) imtyje nuo 1978 vasario mėnesio iki 2014 metų galo, matoma, kad geriausias *ACC*, *MCC* ir *AUC* rodiklius (0.828, 0.559 ir 0.845) turėjo **LR-11** modelis, o geriausią *QPS* rodiklį (0.137) turėjo modelių mediana. Prasčiausiai lentelėje atrodo **DNN** ir **LR-elastic net** modeliai. Žiūrint į maišaties matricas (žr. 5 lentelę), matoma, jog daugiausiai kartų krizinį režimą teisingai prognozavo **SVC**, **LR-11** ir **LR-12** modeliai, tačiau **SVC** modelis daugiau klaidingai prognozavo krizinį režimą esant stabiliam režimui. Taip pat pastebima, kad krizinį režimą prastai atpažino **LR-elastic net** modelis.

Panaši situacija matoma ir analogiškiems modeliams, kurie prognozuoja savaitiniu horizontu (žr. 6 lentelę). Geriausias tikslumo rezultatus turi **LR-11** modelis, žemiausias *QPS* rodiklis gaunamas paėmus modelių medianą, o prasčiausi tikslumo rodikliai yra **LR-elastic net** ir **DNN** modelių atveju. Nagrinėjant maišaties matricas (žr. 7 lentelę), matoma, kad daugiausiai kartų teisingai krizinį režimą prognozavo **LR-11**, **LR-12** ir **SVC** modeliai, bet **SVC** modelis daugiau kartų neteisingai klasifikavo stabilų režimą.

Toliau nagrinėjami modelių tikslumo rezultatai, kai modelių apmokymui buvo pasirinkta dalis makroekonominių rodiklių. Nagrinėjant mėnesinio horizonto modelių tikslumo rodiklius (žr. 8 lentelę), geriausiai atrodo **RF**, **XGBM** ir modelių mediana. Žemiausias *QPS* rodiklis (0.145) pasiektas **RF** ir **XGBM** modelių atvejais, tačiau **RF** modelis turi šiek tiek geresnį *MCC* rodiklį (**RF** rodiklis 0.508, o **MEDIAN** 0.463). Prasčiausi tikslumo rezultatai yra su **SVC** ir **DNN** modeliais. Žiūrint į maišaties matricas (žr. 9 lentelę) matoma, kad dažniausiai krizinį režimą teisingai prognozavo **RF** ir **XGBM** modeliai.

Savaitinio horizonto modelių atveju (žr. 10 lentelę), tikslumo rodikliai geriausiai atrodo **LR-12** ir modelių medianos atveju, prasčiausiai atrodo **DNN** modelio rodikliai. Žemiausias *QPS* rodiklis pasiektas su modelių mediana, geriausias *MCC* rodiklis yra su **LR-11** modeliu, o aukščiausias *AUC* rodiklis yra **RF** modelio atveju. Nagrinėjant maišaties matricas (žr. 11 lentelę), pastebima, kad dažniausiai krizi-

nį režimą teisingai atpažino **LR-11** ir **SVC** modeliai, tačiau **SVC** modelis žymiai daugiau kartų neteisingai klasifikavo stabilų režimą.

Kitas žingsnis yra gautų prognozių panaudojimas investavimo sprendimų į S&P 500 indeksą priėmimui. Bus konstruojamos aktyvaus valdymo strategijos ir atitinkami kontroliniai portfeliai (lyginamieji indeksai), kurių svoris investuojant į S&P 500 indeksą yra pastovus. Norint įvertinti gautų prognozių tikslumą ir naudingumą investuojant bei patikrinti makroekonominių duomenų rinkinio skelbimo vėlavimo įtaką investavimo rezultatams, bus tikrinami du investavimo būdai.

Pirmas būdas yra investuoti pagal atitinkamas datas, kada iš tikrųjų buvo paskelbti duomenys, iš jų gauta režimo prognozė ir priimtas atitinkamas sprendimas. Mėnesinių modelių atveju sprendimas priimamas taip, jei paskutinė prognozė buvo, kad akcijų rinka yra kriziniame režime, tai svoris į S&P 500 indeksą sekančiam mėnesiui bus 0, ir 1 priešingu atveju. Svarbu atkreipti dėmesį į duomenų paskelbimo datas ir už kokį mėnesį yra tie duomenys, pavyzdžiui, kaip buvo minėta 2.3 skyrelyje, balandžio 11 dieną yra paskelbtas naujas makroekonominių rodiklių rinkinys, tačiau tame rinkinyje paskutinė esanti duomenų eilutė atitinka vasario mėnesio rodmenis. Taigi, strategija bus konstruojama taip: balandžio 11 dieną gaunama būsimos režimo prognozė naudojant naujausius makroekonominius rodiklius, kurie yra už vasario mėnesį, jei modelis rodo, kad akcijų rinkos režimas krizinis, tai pozicija į S&P 500 indeksą bus 0 ir ji laikoma iki gegužės 11 dienos, kai vėl bus išleisti naujausi duomenys ir atlikta nauja režimo prognozė.

Kitas investavimo būdas yra neįgyvendinamas realiame gyvenime, tačiau juo norima patikrinti kokią įtaką investavimo rezultatams daro makroekonominių duomenų atsilikimas nuo realios rinkos situacijos. Investuojant šiuo būdu yra „pažiūrima į ateitį“, tai yra balandžio 11 dieną sprendimai priimami naudojant režimo prognozę, kuri buvo gauta iš kovo mėnesio duomenų rinkinio, o šis rinkinys iš tikrųjų būtų pasiekiamas tik gegužės 11 dieną.

Investavimas turint savaitinius duomenis šiek tiek skiriasi. Kadangi duomenų dažnis didesnis, norint išvengti dažno akcijų pozicijų keitimo ir iš to atsirandančių prekybos kaštų tokiose situacijose, kai vieną savaitę modelis prognozuoja stabilų režimą, sekančią savaitę prognozuoja krizinį režimą, o vėliau vėl stabilų, įvedamos prekiavimo taisyklės. Akcijos parduodamos, jei 2 iš 3 paskutinių savaitių prognozių akcijų rinkoje buvo krizinis režimas. Akcijos perkamos, jei 2 iš eilės einančių savaitių prognozės buvo stabilus režimas akcijų rinkoje.

Sukonstravus aprašytas strategijas, sudaromi du lyginamieji indeksai. Pirmas indeksas vadinamas rizika svertu indeksu, jo rizika virš nerizikingos investicijos sutampa su strategijos rizika virš nerizikingos investicijos. Pastovus indekso svoris apskaičiuojamas taip:

$$w = \frac{\sigma(r_p - r_f)}{\sigma(r_{SPX} - r_f)},$$

čia $\sigma()$ reiškia eilutės esančios skliaustelių viduje standartinį nuokrypį, $r_p - r_f$ - strategijos grąža virš nerizikingos investicijos, $r_{SPX} - r_f$ yra S&P 500 indekso grąža virš nerizikingos investicijos. Kito lyginamojo indekso svoris yra lygus strategijos β reikšmei, jos aprašymas pateikiamas 1.4 skyrelyje. Šis indeksas vadinamas laiku svertu indeksu.

Strategijų ir lyginamųjų indeksų investavimo rezultatai pateikiami priedų skyrelyje 2.6. Strategija, kai investuojama atsižvelgiant į duomenų išleidimo datas bus vadinama *Lag*, kita strategija, kai „pažiūrima į ateitį“ bus vadinama *Foresight*. Priedų lentelėse pateikiami strategijų pavadinimų trumpiniai:

Regime_full_sample - strategija, kai naudojama modelių apmokymo žymės imtyje nuo 1978 iki 2014 metų,

Regime_bt_sample - strategija, kai naudojama modelių apmokymo žymės imtyje nuo 2015 iki 2022 metų,

Regime_model - strategija, kai naudojama modelių apmokymo žymės imtyje nuo 1986 iki 2022 metų (palyginimui su režimų pasikeitimo modelio rezultatais),

Regime_backtest - strategija, kai naudojamos režimų žymės apskaičiuotos su ℓ_1 filtru nesinaudojant informacija iš ateities (kaip imtyje nuo 2015 metų),

LR_enet, **LR_L1**, **LR_L2**, **RF**, **GBM**, **XGBM**, **SVC**, **DNN**, **Median** yra strategijos gautos naudojant atitinkamų modelių prognozes (LR_enet yra LR_elastic net).

Papildomi užrašai **TWBmk** ir **RiskBmk** reiškia strategijai paskaičiuotą laiku svertą (TWBmk) ir rizika svertą (RiskBmk) lyginamąjį indeksą. **S&P 500** yra strategija, kai svoris į S&P 500 indeksą yra 1 visame investavimo laikotarpyje. Lentelėse pateikiami teorinėje dalyje aprašyti rodikliai *Mean*, *Stdev*, *SharpeR*, *VaR95*, *TreynorR* - metiniai.

Pirma apžvelgiami strategijų rezultatai su mėnesinio horizonto modelių prognozėmis, kai prognozėms naudojami visi makroekonominiai duomenys. Analizuojamas periodas yra nuo 1978 vasario iki 2014 metų gruodžio mėnesio. Lentelėse pateikia-

ma strategija *Regime_full_sample* yra skaičiuota tik režimų žymių, skirtų modelių apmokymui, tinkamumo patikrinimui, todėl jos rezultatai nebus lyginami su kitomis strategijomis. Žiūrint į *Lag* strategijos rezultatus (žr. 12 lentelę) matoma, kad yra strategijų, kurių vidutinė metinė grąža šiek tiek didesnė (DNN, RF) arba standartinis nuokrypis mažesnis nei S&P 500 indekso. Didesnės dalies strategijų *SharpeR* rodiklis aukštesnis nei S&P 500, tačiau *ProbSR* nepasiekia 0.95 nei vienu atveju. Taip pat, visų strategijų *MaxDD* ir *VaR95* rodikliai geresni. Prasčiausiai atrodo strategija naudojant SVC prognozes, geriausiai DNN ir RF, tačiau jų grąža nėra žymiai didesnė nei S&P 500 indekso. Lyginant strategijas su atitinkamais lyginamaisiais indeksais, matoma, kad neretai rizika sverto indekso (kartais ir laiku sverto indekso) grąža didesnė. Nagrinėjant *Foresight* strategiją (žr. 13 lentelę), matoma, kad kai kurių strategijų (LR_L1, LR_L2) grąžos beveik dvigubai didesnės nei S&P 500, kiti rodikliai taip pat atrodo žymiai geriau. Įspūdingai atrodo *Regime_backtest* rezultatai, nes čia yra 1 mėnesio pažiūrėjimas į ateitį naudojant S&P 500 grąžas. *Foresight* strategijos atveju prasčiausiai atrodo LR_enet ir SVC, o DNN atveju grąža nėra žymiai geresnė nei S&P 500 indekso.

Toliau prieduose pateikiamos analogiškos lentelės investuojant savaitiniu horizontu. Nagrinėjant *Lag* strategijos rezultatus pateikiamus 14 lentelėje, didesne grąža išsiskiria XGBM ir DNN. Kaip ir investuojant mėnesiniu horizontu, didesnės dalies strategijų *Stdev*, *MaxDD*, *VaR95* rodikliai yra geresni nei S&P 500. *TreynorR* ir *SharpeR* rodikliai taip pat dažniau yra geresni, bet *ProbSR* nepasiekia 0.95. Be to, yra keli atvejai, kai lyginamųjų indeksų grąžos geresnės nei atitinkamų strategijų. Prasčiausiai šioje lentelėje atrodo SVC, LR_enet ir GBM. Nagrinėjant *Foresight* strategiją (žr. 15 lentelę) prasčiausiai atrodo tie patys modeliai. Žiūrint bendrai didesnė dalis strategijų rodiklių atrodo geriau.

Prieš tai minėtoje imtyje taip pat buvo skaičiuoti modeliai, kai naudojama dalis makroekonominių rodiklių. Šie rodikliai aprašyti 2.4 skyriuje. Mėnesinio horizonto atveju su *Lag* strategija (žr. 16 lentelę), kai naudojama dalis makroekonominių rodiklių, gauti geresni rezultatai naudojant LR_L1, RF, GBM, SVC ir Median modelių prognozes, nei modelių apmokymui imant visus makroekonominius duomenis. Rezultatų lentelėje (žr. 16 lentelę) matoma, kad 3 strategijų atvejais vidutinė metinė grąža buvo didesnė nei S&P 500 indekso. Rizikos rodikliai *Stdev*, *MaxDD* ir *VaR95* visais atvejais geresni nei S&P 500. DNN atveju *ProbSR* viršija 0.95, bet šio modelio vidutinė grąža beveik 1% mažesnė nei S&P 500 indekso. Vertinant

bendrą vaizdą, strategijų rezultatai šiek tiek geresni nei modelių, kurių apmokymui buvo naudojami visi makroekonominiai duomenys. *Foresight* strategijos atveju (žr. 17 lentelę) visos vidutinės grąžos didesnės nei S&P 500 indekso. Geriausi rezultatai (be Regime_backtest) yra LR_L1, LR_L2 atvejais, prasčiausi - LR_enet ir DNN.

Savaitinio horizonto modelių atveju nagrinėjant *Lag* strategijos rezultatų 18 lentelę, geriausi rezultatai matomi LR_L1 ir LR_L2 modelių atvejais, prasčiausi - LR_enet ir SVC. Čia penkiais atvejais vidutinė metinė grąža buvo didesnė arba sutapo su S&P 500 indekso. Kaip ir prieš tai, didesnė dalis rizikos rodiklių yra geresni nei S&P 500 indekso. Strategijų *SharpeR* rodikliai dažniausiai taip pat geresni, tačiau nei vienu atveju *ProbSR* nėra didesnis nei 0.95. Lyginant strategijų rezultatus su modelių, kuriems buvo naudojami visi makroekonominiai duomenys, bendrai rezultatai atrodo geriau. *Foresight* strategijos atveju (žr. 19 lentelę), priešingai nei su mėnesinio horizonto modeliais, ne visų strategijų vidutinė metinė grąža didesnė nei S&P 500 indekso. Čia prasčiausiai atrodo LR_enet ir SVC rodikliai, o geriausiai LR_L1 bei LR_L2. Lyginant su savaitinių modelių, kurie apmokyti naudojant visus makroekonominis duomenis, rezultatais, gaunama, kad dalies strategijų vidutinės grąžos padidėjo (LR_L1, LR_L2, GBM, XGBM, SVC), dalies pablogėjo (LR_enet, RF, DNN), o Median nepakito.

Toliau bus aptariami modelių tikslumo ir investavimo strategijų rezultatai nuo 2015 sausio iki 2022 metų gegužės. Atsižvelgus į prastus investavimo rezultatus prieš tai nagrinėjote imtyje, toliau nenaudojami LR_enet ir SVC modeliai. Nagrinėjant mėnesinių modelių (žr. 20 lentelę) ir savaitinių modelių tikslumo rezultatus (žr. 22 lentelę), kai apmokymui naudojami visi makroekonominiai rodikliai, matosi, kad šioje imtyje kriziniai režimai atpažįstami labai prastai ir jų buvo ganėtinais nedaug. Naudojant tuos pačius modelius mėnesiniam (žr. 24 lentelę) ir savaitiniam (žr. 26 lentelę) prognozavimui, bet apmokant juos su dalimi makroekonominių kintamųjų, kai kurių modelių atvejais padažnėja krizinio režimo teisingas prognozavimas, bet labai nežymiai.

Nagrinėjant investavimo rezultatus naudojant mėnesinio horizonto ir savaitinio horizonto modelių prognozes, strategijų rezultatai yra prastesni nei investavus į S&P 500 indeksą nekeičiant pozicijos. Mėnesinio ir savaitinio horizonto modelių, kurie apmokyti imant visus makroekonominis rodiklius, *Lag* investavimo strategijų rezultatai pateikiami priedų skyriaus 28 ir 30 lentelėse, o *Foresight* investavimo strategijų rezultatai pateikiami 29 ir 31 lentelėse. Mėnesinio ir savaitinio horizonto

modelių, kurie apmokyti naudojant dalį makroekonominių rodiklių, *Lag* investavimo strategijų rezultatus galima rasti 32 ir 34 lentelėse, o *Foresight* investavimo strategijų rezultatai yra 33 ir 35 lentelėse.

2.6 Palyginimas su režimų pasikeitimo modeliu

Šiame skyrelyje pristatomas režimų pasikeitimo modelio praktinis panaudojimas ir gauti rezultatai palyginami su mašininio mokymosi modelių rezultatais toje pačioje imtyje.

Lietuvos banko naudojamas režimų pasikeitimo modelis yra skaičiuojamas 4 kartus per mėnesį: 4, 11, 18 ir 25 dienomis. Jei nurodyta diena yra nedarbo, tai naudojami praeitos darbo dienos duomenys. Skaičiuojant modelį nurodytomis dienomis yra naudojami atitinkamų dienų mėnesiniai duomenys, tai yra skaičiuojant 4 dieną, naudojamas mėnesinių duomenų rinkinys, kur mėnesių duomenys yra nuo vieno mėnesio 4 dienos iki kito mėnesio 4 dienos, skaičiuojant 11 dieną mėnesiniai duomenys yra nuo vieno mėnesio 11 dienos iki kito mėnesio 11 dienos ir t.t. Režimų pasikeitimo modelyje naudojami makroekonominiai ir rinkos duomenys aprašyti 2.1 skyrelyje. Dėl duomenų trūkumo, šio modelio testavimo imtis prasideda nuo 1986-03-04 ir baigiasi 2022-05-04, todėl mašininio mokymosi modelių rezultatai bus lyginami toje pat imtyje. Mėnesinių modelių atveju buvo imama imtis nuo 1986-03-11 iki 2022-05-11. Kadangi imtyje nuo 2015 iki 2022 metų nenagrinėjami LR_enet ir SVC modelių rezultatai, tai jie nebuvo įtraukti į šį palyginimą. Mašininio mokymosi modelių palyginimui su režimų pasikeitimo modeliu naudojami iš *Lag* strategijos gauti investavimo rezultatai.

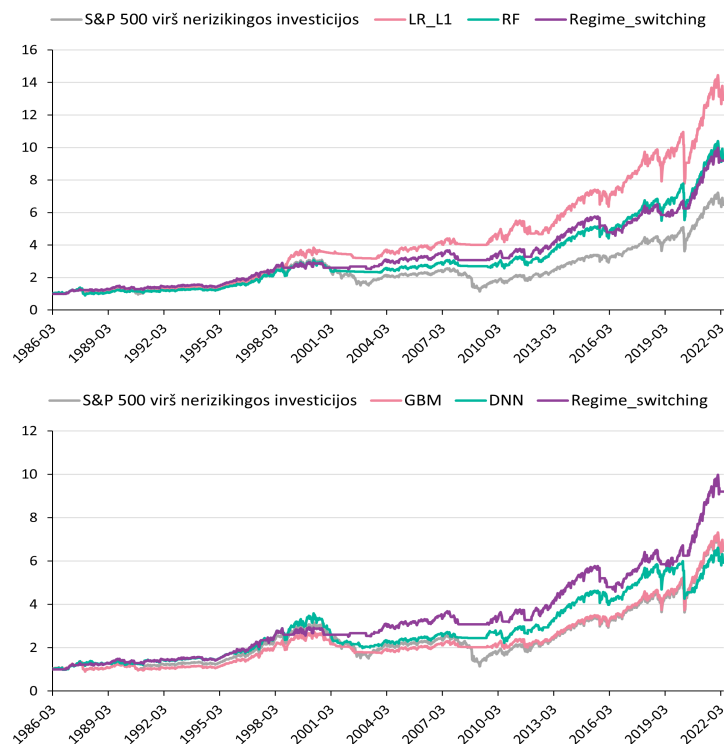
Investavimo sprendimai, turint krizinio režimo tikimybes iš režimų pasikeitimo modelio, priimami taip:

Akcijos perkamos, kai 2 savaites iš eilės krizinio režimo tikimybė neviršija 20%,
Akcijos parduodamos, kai 2 iš 3 paskutinių savaičių rodo, kad krizinio režimo tikimybė viršijo 50%.

Modelių rezultatų palyginimui skirtos lentelės pateikiamos priedų skyriuje. Jose režimų pasikeitimo modelio rezultatai vadinami **Regime_switching**. Lyginant mėnesinio (žr. 36 lentelę) ir savaitinio (žr. 37 lentelę) horizonto investavimo rezultatus, naudojant mašininio mokymosi modelius, kurie apmokyti imant visus makroekonominius rodiklius, ir režimo pasikeitimo modelį, visais atvejais vidutinė metinė grąža

buvo didesnė arba lygi S&P 500 indekso vidutinei gražai, rizikos rodikliai taip pat geresni, aktyvaus valdymo strategijų gražos didesnės nei atitinkamų lyginamųjų indeksų, tačiau bendrai geresni rezultatai buvo gauti naudojant režimų pasikeitimo modelį. Nagrinėjant tų pačių mašininio mokymosi modelių rezultatus (mėnesinio modelio 38 lentelė, savaitinio modelio 39 lentelė), kai apmokymui panaudota dalis duomenų, pastebimos šiek tiek geresnės vidutinės metinės gražos. Mėnesinių modelių LR_L1 ir Median vidutinės gražos didesnės nei Regime_switching, o savaitinių modelių didesnės gražos yra LR_L1, LR_L2 ir RF atveju, tačiau žiūrint bendrai režimų pasikeitimo rodikliai yra geresni nei kitų strategijų.

Apačioje pateikiami savaitinio horizonto modelių, naudojant dalį rodiklių apmokymui, geriausiai ir blogiausiai veikusių strategijų investiciniai vienetai. Matoma, kad pirmame grafike esančios strategijos veikė pakankamai gerai iki 2020 metų, tačiau priešingai nei režimų pasikeitimo modelio atveju, nebuvo laiku uždaryta 2020 kovo mėnesį (COVID-19 krizės pradžia) turėta pozicija ir patirti dideli nuostoliai. Antrame grafike mašininio mokymosi modeliai laiku neatpažino „Dot-com“ burbulo sukeltos krizės, o tuo pačiu laiku neuždaryta 2020 kovo mėnesį turėta pozicija, todėl galutiniai rezultatai gerokai prastesni nei režimų pasikeitimo modelio.



2.1 pav.: Geriausiai ir blogiausiai veikusios mašininio mokymosi modelių strategijos

Išvados

Naudojant mašininio mokymosi modelius, kurių apmokymui naudojami makroekonominiai rodikliai, nepavyksta reikšmingai pagerinti investavimo rezultatų juos lyginant su režimų pasikeitimo modeliu. Makroekonominių rodiklių duomenys vėluoja bent 1 mėnesiu, todėl režimų prognozavimas akcijų rinkoje nėra efektyvus. Rezultatus šiek tiek pagerina akcijų rinkos signalų įtraukimas, makroekonominių rodiklių imties sumažinimas ir modelių skaičiavimo dažnio padidinimas iš mėnesinio į savaitinį horizontą. Tačiau mašininio mokymosi modelių atveju makroekonominiai duomenys turi ganėtinai didelį svorį, taigi krizinis režimas, kuris įvyksta dėl egzogeninių šokų, neatpažįstamas pakankamai greitai, nes minėtieji šokai dar neatsispindi makroekonominiuose duomenyse.

Viena iš galimų ateities tyrimų kryptių galėtų būti didesnio svorio suteikimas rinkos duomenims ir/arba makroekonominių duomenų dabarties prognozių (angl. *nowcasting*) naudojimas tam, kad būtų sumažintos problemos kylančios dėl jų vėlavimo. Taip pat būtų galima atskirti makroekonominius ir rinkos duomenis bei sukurti du modelius, kurie prognozuotų rinkos situaciją tik iš makroekonominių ir tik iš rinkos duomenų. Esant modelių prognozių išsiskyrimui būtų atliekama papildoma analizė ar reikėtų keisti pozicijas akcijose.

Literatūra

- [1] Ang, A., & Bekaert, G. (2002, July). International Asset Allocation With Regime Shifts. *The Review of Financial Studies*, 15(4), 1137-1187.
- [2] Ang, A., & Bekaert, G. (2004). How Regimes Affect Asset Allocation. *Financial Analyst Journal*, 60(2), 86-99.
- [3] Bailey, D. H., & De Prado, M. L. (2012). The Sharpe Ratio Efficient Frontier. *Journal of Risk*, 15(2), 3-44.
- [4] Boero, G., Smith J., & Wallis K. F. (2011). Scoring rules and survey density forecasts. *International Journal of Forecasting*, 27(2), 379-393.
- [5] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.
- [6] Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
- [7] Chicco, D., & Jurman, G. (2020, January). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1).
- [8] De Angelis, L., & Paas, L. J. (2013). A dynamic analysis of stock markets using a hidden Markov model. *Journal of Applied Statistics*, 40(8), 1682-1700.
- [9] De Prado, M. L., & Lewis, M. J. (2019). Confidence and Power of the Sharpe Ratio under Multiple Testing. *Social Science Research Network*, 2019.
<https://ssrn.com/abstract=3193697>
- [10] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874.
- [11] Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.
- [12] Gu, J., & Mulvey, J. M. (2021). Factor Momentum and Regime-Switching Overlay Strategy. *The Journal of Financial Data Science*, 3(4), 101-129.

- [13] Hamilton, J. D. (1989, March). A New Approach to the Economic Analysis of Nonstationary Time Series and Business Cycle. *Econometrica*, 57(2), 357-384.
- [14] Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.
- [15] Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques*. 3rd Edition. Morgan Kaufmann Publishers.
- [16] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction*. 2nd Edition. Springer.
- [17] James, A., Abu-Mostafa, Y. S., & Qiao, X. (2019). Machine Learning for Recession Prediction and Dynamic Asset Allocation. *The Journal of Financial Data Science*, 1(3), 41-56.
- [18] Jorion, P. (2007). *Value at Risk: The New Benchmark for managing Financial Risk*. 3rd Edition. McGraw-Hill.
- [19] Jorion, P. (2011). *Financial Risk Manager Handbook*. 6th Edition. Wiley Finance.
- [20] Karatzoglou, A., Meyer, D., & Hornik, K. (2006). Support Vector Machines in R. *Journal of Statistical Software*, 15(9), 1-28.
- [21] Kim, M. S., Koh, K., Boyd, S., & Gorinevsky, D. (2009). ℓ_1 Trend Filtering. *SIAM Review*, 51(2), 339-360.
<https://doi.org/10.1137/070690274>
- [22] Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the Artificial Neural Networks.
- [23] McCracken, M. W., & Ng, S. (2016). FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, 34(4), 574-589.
<https://doi.org/10.20955/wp.2015.012>
- [24] Mulvey, J. M., Hao, H., & Li, N. (2018). Machine learning, economic regimes and portfolio optimisation. *International Journal of Financial Engineering and Risk Management*, 2(4), 260-282.

- [25] Nielsen, A. (2014). Hide and seek: Profitting by hidden states of financial markets in real time. *GOAL - Global Strategy Paper*, (14).
- [26] Nguyen, N. (2018). Hidden Markov Model for Stock Trading. *International Journal of Financial Studies*, 6(2).
- [27] Perez-Quiros, G., & Timmermann, A. (2000). Firm size and Cyclical Variations in Stock Returns. *The Journal of Finance*, 55(3), 1229-1262.
- [28] Perlin, M. (2015, April). The MATLAB Package for Markov Regime Switching Models. *Social Science Research Network*.
<http://ssrn.com/abstract=1714016>
- [29] Reus, L., & Mulvey, J. M. (2016). Dynamic Allocation for Currency Futures under Switching Regimes Signals. *European Journal of Operational Research*, 253(1), 85-93.
- [30] Sharpe, W. F. (1994). The Sharpe ratio. *The Journal of Portfolio Management*, 21(1), 49-58.
- [31] Simonian, J., & Wu, C. (2019). Minsky vs. Machine: New Foundations for Quant-Macro Investing. *The Journal of Financial Data Science*, 1(2), 94-110.
- [32] Tibshirani, R. J. (2014). Adaptive piecewise polynomial estimation via trend filtering. *The Annals of Statistics*, 42(1), 285-323.
<https://doi.org/10.1214/13-AOS1189>
- [33] Tibshirani, R. J., & Taylor, J. (2011). The solution path of the generalized lasso. *The Annals of Statistics*, 39(3), 1335-1371.
<https://doi.org/10.1214/11-AOS878>
- [34] Uysal, A. S., & Mulvey, J. M. (2021). A Machine Learning Approach in Regime-Switching Risk Parity Portfolios. *The Journal of Financial Data Science*, 3(2), 87-108.
- [35] Wu, D. J., Feng, T., Naehrig, M., & Lauter, K. (2016). Privately Evaluating Decision Trees and Random Forests. *Proceedings on Privacy Enhancing Technologies*, 2016(4), 335-355.

- [36] Yazdani, A. (2020). Machine Learning Prediction of Recessions: An Imbalanced Classification Approach. *The Journal of Financial Data Science*, 2(4), 21-32.

Priedai

Makroekonominių duomenų sąrašas

Duomenų transformacijos nr.: 1. nekeičiama, 2. Δx_t , 3. $\Delta^2 x_t$, 4. $\log(x_t)$, 5. $\Delta \log(x_t)$, 6. $\Delta^2 \log(x_t)$, 7. $\Delta(x_t/x_{t-1} - 1)$.

Darbo rinka (angl. Labor Market)	Transformacijos nr.	Kainų indeksai (angl. Prices)	Transformacijos nr.
Help-Wanted Index for United States	2	PPI: Finished Goods	6
Ratio of Help Wanted/No. Unemployed	2	PPI: Finished Consumer Goods	6
Civilian Labor Force	5	PPI: Intermediate Materials	6
Civilian Employment	5	PPI: Crude Materials	6
Civilian Unemployment Rate	2	Crude Oil, spliced WTI and Cushing	6
Average Duration of Unemployment (Weeks)	2	PPI: Metals and metal products:	6
Civilians Unemployed - Less Than 5 Weeks	5	CPI: All Items	6
Civilians Unemployed for 5-14 Weeks	5	CPI: Apparel	6
Civilians Unemployed - 15 Weeks & Over	5	CPI: Transportation	6
Civilians Unemployed for 15-26 Weeks	5	CPI: Medical Care	6
Civilians Unemployed for 27 Weeks and Over	5	CPI: Commodities	6
Initial Claims	5	CPI: Durables	6
All Employees: Total nonfarm	5	CPI: Services	6
All Employees: Goods-Producing Industries	5	CPI: All Items Less Food	6
All Employees: Mining and Logging: Mining	5	CPI: All items less shelter	6
All Employees: Construction	5	CPI: All items less medical care	6
All Employees: Manufacturing	5	Personal Cons. Expend.: Chain Index	6
All Employees: Durable goods	5	Personal Cons. Exp: Durable goods	6
All Employees: Nondurable goods	5	Personal Cons. Exp: Nondurable goods	6
All Employees: Service-Providing Industries	5	Personal Cons. Exp: Services	6
All Employees: Trade, Transportation & Utilities	5	Produkcija ir pajamos (angl. Output and Income)	
All Employees: Wholesale Trade	5	Real Personal Income	5
All Employees: Retail Trade	5	Real personal income ex transfer receipts	5
All Employees: Financial Activities	5	IP Index	5
All Employees: Government	5	IP: Final Products and Nonindustrial Supplies	5
Avg Weekly Hours : Goods-Producing	1	IP: Final Products (Market Group)	5
Avg Weekly Overtime Hours : Manufacturing	2	IP: Consumer Goods	5
Avg Weekly Hours : Manufacturing	1	IP: Durable Consumer Goods	5
Avg Hourly Earnings : Goods-Producing	6	IP: Nondurable Consumer Goods	5
Avg Hourly Earnings : Construction	6	IP: Business Equipment	5
Avg Hourly Earnings : Manufacturing	6	IP: Materials	5
Akcijų rinka (angl. Stock Market)		IP: Durable Materials	5
S&P's Common Stock Price Index: Composite	5	IP: Nondurable Materials	5
S&P's Common Stock Price Index: Industrials	5	IP: Manufacturing (SIC)	5
S&P's Composite Common Stock: Dividend Yield	2	IP: Residential Utilities	5
S&P's Composite Common Stock: Price-Earnings Ratio	5	IP: Fuels	5
VIX	1	Capacity Utilization: Manufacturing	2

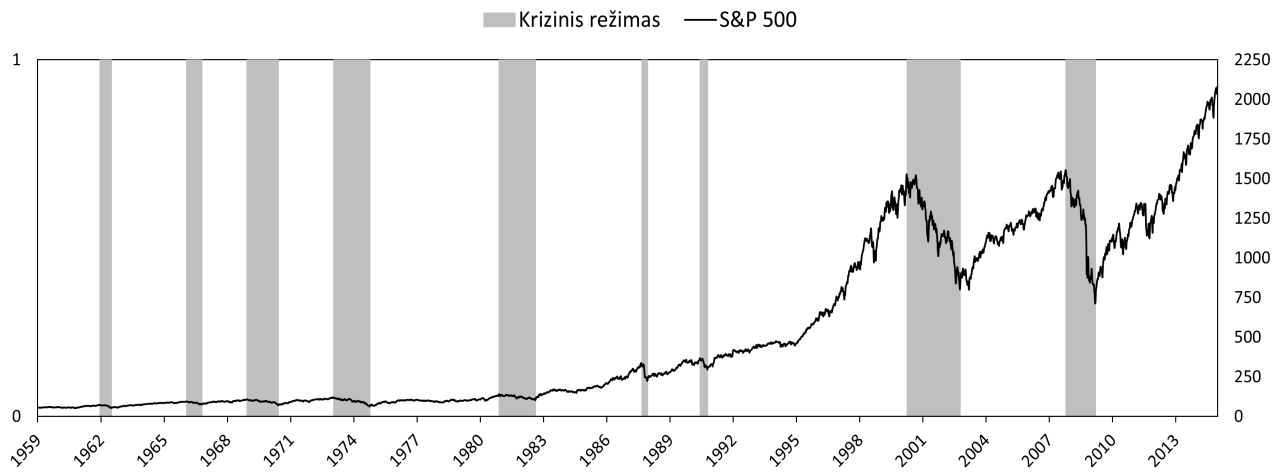
1 lentelė: Makroekonominiai rodikliai (1/2)

Palūkanų normos ir valiutų kursai (angl. Interest rates and Exchange Rates)	Transformacijos nr.	Vartojimas ir užsakymai (angl. Consumption and Orders)	Transformacijos nr.
Effective Federal Funds Rate	2	Housing Starts: Total New Privately Owned	4
3-Month AA Financial Commercial Paper Rate	2	Housing Starts, Northeast	4
3-Month Treasury Bill:	2	Housing Starts, Midwest	4
6-Month Treasury Bill:	2	Housing Starts, South	4
1-Year Treasury Rate	2	Housing Starts, West	4
5-Year Treasury Rate	2	New Private Housing Permits (SAAR)	4
10-Year Treasury Rate	2	New Private Housing Permits, Northeast (SAAR)	4
Moody's Seasoned Aaa Corporate Bond Yield	2	New Private Housing Permits, Midwest (SAAR)	4
Moody's Seasoned Baa Corporate Bond Yield	2	New Private Housing Permits, South (SAAR)	4
3-Month Commercial Paper Minus FEDFUNDS	1	New Private Housing Permits, West (SAAR)	4
3-Month Treasury C Minus FEDFUNDS	1	Pinigai ir kreditas (angl. Money and Credit)	
6-Month Treasury C Minus FEDFUNDS	1	M1 Money Stock	6
1-Year Treasury C Minus FEDFUNDS	1	M2 Money Stock	6
5-Year Treasury C Minus FEDFUNDS	1	Real M2 Money Stock	5
10-Year Treasury C Minus FEDFUNDS	1	Monetary Base	6
Moody's Aaa Corporate Bond Minus FEDFUNDS	1	Total Reserves of Depository Institutions	6
Moody's Baa Corporate Bond Minus FEDFUNDS	1	Reserves Of Depository Institutions	7
Trade Weighted U.S. Dollar Index	5	Commercial and Industrial Loans	6
Switzerland / U.S. Foreign Exchange Rate	5	Real Estate Loans at All Commercial Banks	6
Japan / U.S. Foreign Exchange Rate	5	Total Nonrevolving Credit	6
U.S. / U.K. Foreign Exchange Rate	5	Nonrevolving consumer credit to Personal Income	2
Canada / U.S. Foreign Exchange Rate	5	Consumer Motor Vehicle Loans Outstanding	6
Užsakymai ir atsargos (angl. Orders and Inventories)		Total Consumer Loans and Leases Outstanding	6
Real personal consumption expenditures	5	Securities in Bank Credit at All Commercial Banks	6
Real Manu. and Trade Industries Sales	5		
Retail and Food Services Sales	5		
New Orders for Consumer Goods	5		
New Orders for Durable Goods	5		
New Orders for Nondefense Capital Goods	5		
Unfilled Orders for Durable Goods	5		
Total Business Inventories	5		
Total Business: Inventories to Sales Ratio	2		
Consumer Sentiment Index	2		

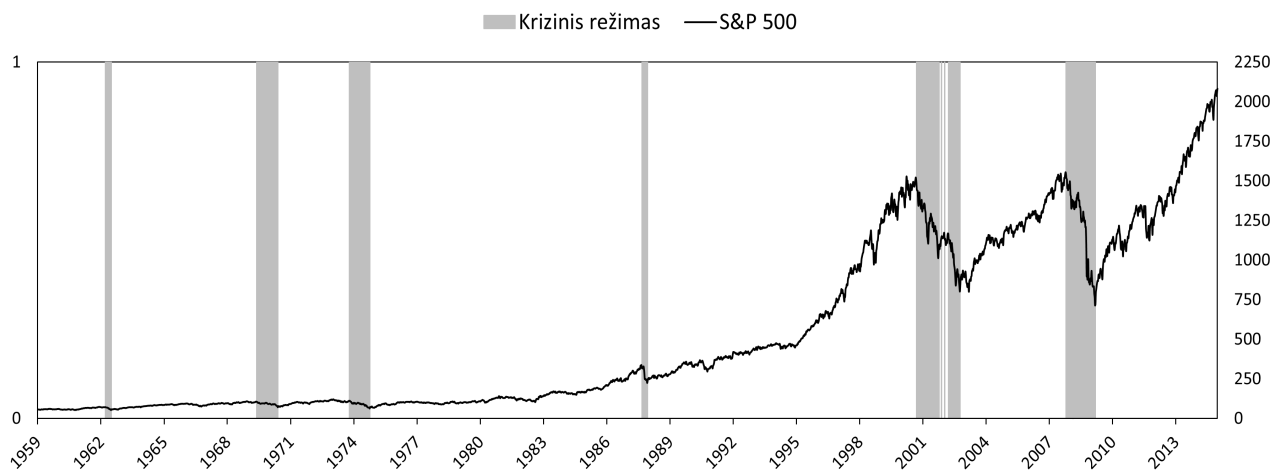
2 lentelė: Makroekonominiai rodikliai (2/2)

Rodiklis	Paaiškinimas
DGS10	JAV vyriausybės 10 metų obligacijos palūkanų norma
DGS5	JAV vyriausybės 5 metų obligacijos palūkanų norma
DGS2	JAV vyriausybės 2 metų obligacijos palūkanų norma
DTB3	JAV vyriausybės 3 mėnesių obligacijos palūkanų norma
CFNAIMA3 index	Čikagos Fed nacionalinio aktyvumo indekso 3 mėnesių slenkantis vidurkis (angl. Chicago Fed National Activity Index: Three Month Moving Average)
NAPMPMI Index	Čikagos pramonės pasitikėjimo rodiklis (angl. Chicago Purchasing Managers Index (PMI))
USURTOT Index	JAV nedarbo lygis (angl. US Unemployment Rate)
MOODCBAA	BAA reitingą turinčių korporacijų obligacijų palūkanų norma (angl. Moody's Seasoned Baa Corporate Bond Yield)
MOODCAAA	AAA reitingą turinčių korporacijų obligacijų palūkanų norma (angl. Moody's Seasoned Aaa Corporate Bond Yield)
CONCCONF Index	JAV vartotojų pasitikėjimo rodiklis (angl. US Consumer confidence index)
IP YoY Index	Pramonės produkcija (angl. Industrial production)
Returns of G0Q0	JAV vyriausybės 1+ metų obligacijų indekso grąžos
Returns of SPX	S&P 500 indekso grąžos
Returns of DTB3	JAV vyriausybės 3 mėnesių obligacijos grąžos

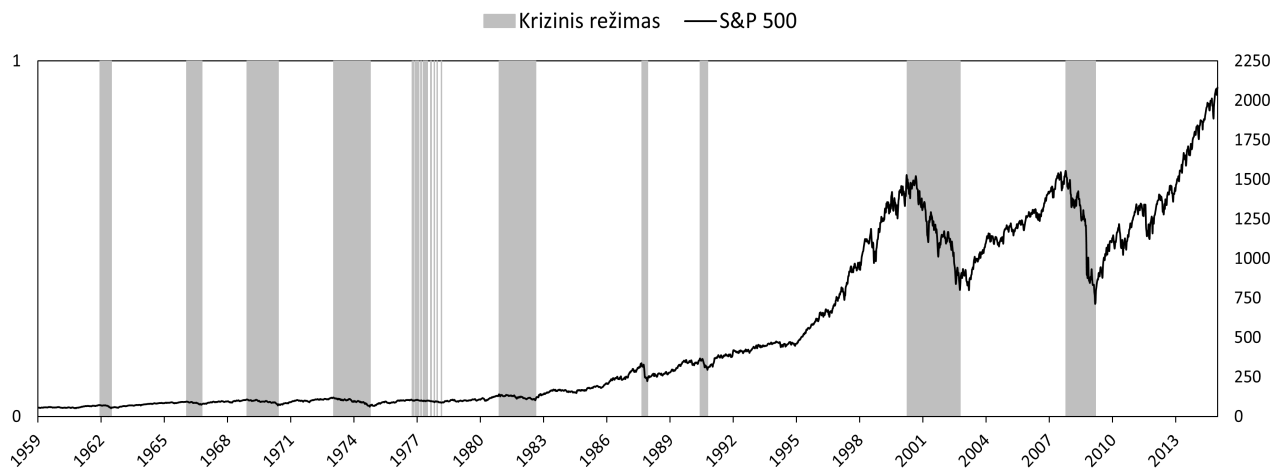
3 lentelė: Režimų pasikeitimo modelio rodikliai



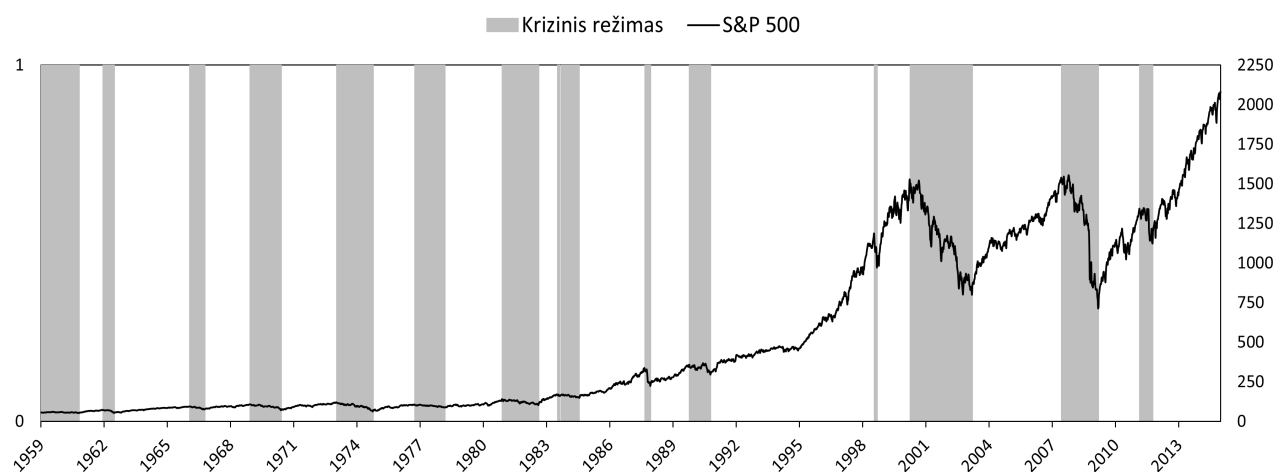
2 pav.: Režimai, kai slenkstis yra 0



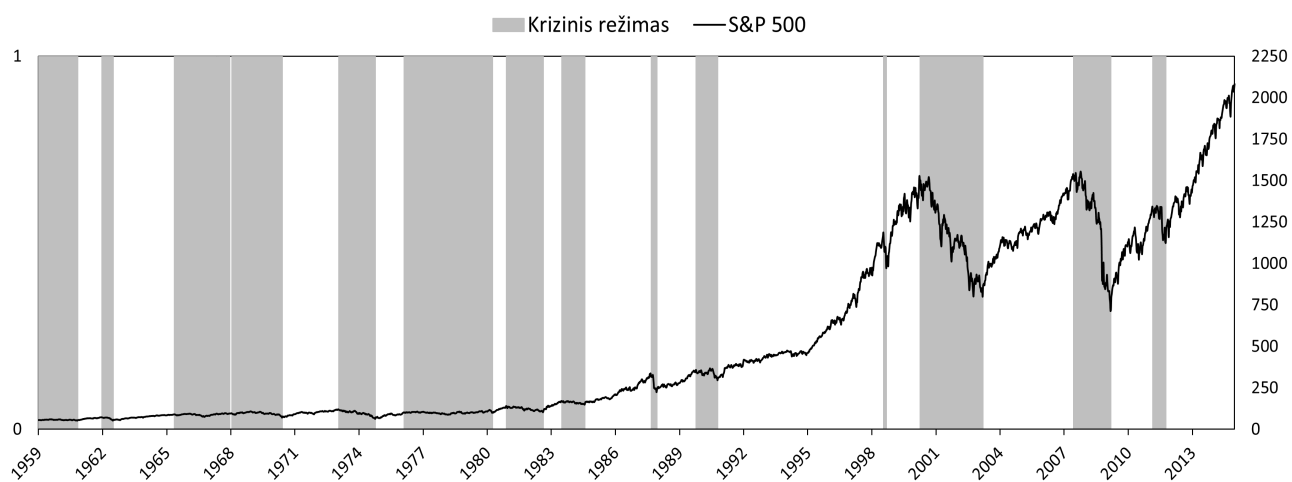
3 pav.: Režimai, kai slenkstis yra 10-tas procentilis



4 pav.: Režimai, kai slenkstis yra 20-tas procentilis



5 pav.: Režimai, kai slenkstis yra 30-tas procentilis



6 pav.: Režimai, kai slenkstis yra 40-tas procentilis

Mašininio mokymosi modelių tikslumo rodikliai (1978-2014 imtyje), kai apmokymui naudojami visi makroekonominiai duomenys

	ACC	MCC	QPS	AUC
LR-elastic net	0.701	0.067	0.202	0.579
LR-I1	0.828	0.559	0.143	0.845
LR-I2	0.801	0.492	0.158	0.824
RF	0.803	0.473	0.141	0.837
GBM	0.776	0.385	0.149	0.813
XGBM	0.799	0.467	0.144	0.837
SVC	0.758	0.424	0.171	0.738
DNN	0.692	0.258	0.237	0.713
MEDIAN	0.803	0.471	0.137	0.845

4 lentelė: Mėnesinio horizonto modelių tikslumo rodikliai

	LR-elastic net		LR-I1		LR-I2	
Tikras režimas		0	1		0	1
	0	292	35	0	287	40
	1	97	18	1	36	79
		Režimo prognozė			Režimo prognozė	
		RF			GBM	
Tikras režimas		0	1		0	1
	0	289	38	0	288	39
	1	49	66	1	60	55
		Režimo prognozė			Režimo prognozė	
		XGBM			DNN	
Tikras režimas		0	1		0	1
	0	286	41	0	246	81
	1	48	67	1	55	60
		Režimo prognozė			Režimo prognozė	
		SVC			Median	
Tikras režimas		0	1		0	1
	0	259	68	0	290	37
	1	39	76	1	50	65
		Režimo prognozė			Režimo prognozė	

5 lentelė: Mėnesinio horizonto modelių maišaties matricos

Mašininio mokymosi modelių tikslumo rodikliai (1978-2014 imtyje), kai apmokymui naudojami visi makroekonominiai duomenys

	ACC	MCC	QPS	AUC
LR-elastic net	0.712	0.091	0.200	0.584
LR-I1	0.816	0.529	0.148	0.837
LR-I2	0.804	0.499	0.161	0.819
RF	0.796	0.446	0.143	0.831
GBM	0.768	0.356	0.151	0.805
XGBM	0.789	0.436	0.150	0.826
SVC	0.741	0.370	0.176	0.719
DNN	0.707	0.278	0.238	0.696
MEDIAN	0.794	0.436	0.140	0.834

6 lentelė: Savaitinio horizonto modelių tikslumo rodikliai

		LR-elastic net		LR-I1		LR-I2	
Tikras režimas		0	1	0	1	0	1
	0	1183	129	1134	178	1125	187
1	381	75	148	308	159	297	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		RF		GBM		XGBM	
Tikras režimas		0	1	0	1	0	1
	0	1159	153	1152	160	1140	172
1	208	248	250	206	202	254	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		SVC		DNN		Median	
Tikras režimas		0	1	0	1	0	1
	0	1035	277	1013	299	1163	149
1	181	275	219	237	216	240	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	

7 lentelė: Savaitinio horizonto modelių maišaties matricos

Mašininio mokymosi modelių tikslumo rodikliai (1978-2014 imtyje), kai apmokymui naudojama dalis makroekonominių duomenų

	ACC	MCC	QPS	AUC
LR-elastic net	0.762	0.340	0.171	0.761
LR-I1	0.787	0.441	0.156	0.805
LR-I2	0.790	0.449	0.155	0.802
RF	0.812	0.508	0.145	0.837
GBM	0.790	0.437	0.154	0.814
XGBM	0.803	0.490	0.145	0.826
SVC	0.729	0.351	0.168	0.752
DNN	0.715	0.297	0.234	0.673
MEDIAN	0.801	0.463	0.148	0.828

8 lentelė: Mėnesinio horizonto modelių tikslumo rodikliai

		LR-elastic net		LR-I1		LR-I2	
Tikras režimas		0	1	0	1	0	1
0	287	40	282	45	282	45	
1	65	50	49	66	48	67	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		RF		GBM		XGBM	
Tikras režimas		0	1	0	1	0	1
0	287	40	286	41	283	44	
1	43	72	52	63	43	72	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		SVC		DNN		Median	
Tikras režimas		0	1	0	1	0	1
0	253	74	255	72	290	37	
1	46	69	54	61	51	64	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	

9 lentelė: Mėnesinio horizonto modelių maišaties matricos

Mašininio mokymosi modelių tikslumo rodikliai (1978-2014 imtyje), kai apmokymui naudojama dalis makroekonominių duomenų

	ACC	MCC	QPS	AUC
LR-elastic net	0.762	0.341	0.172	0.735
LR-I1	0.798	0.480	0.150	0.801
LR-I2	0.803	0.479	0.149	0.796
RF	0.780	0.422	0.152	0.816
GBM	0.763	0.354	0.163	0.784
XGBM	0.776	0.411	0.157	0.802
SVC	0.762	0.432	0.161	0.771
DNN	0.727	0.329	0.219	0.715
MEDIAN	0.804	0.469	0.147	0.810

10 lentelė: Savaitinio horizonto modelių tikslumo rodikliai

		LR-elastic net		LR-I1		LR-I2	
Tikras režimas		0	1	0	1	0	1
	0	1146	166	1124	188	1145	167
1	254	202	169	287	182	274	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		RF		GBM		XGBM	
Tikras režimas		0	1	0	1	0	1
	0	1121	191	1133	179	1117	195
1	198	258	240	216	202	254	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		SVC		DNN		Median	
Tikras režimas		0	1	0	1	0	1
	0	1044	268	1030	282	1165	147
1	152	304	200	256	200	256	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	

11 lentelė: Savaitinio horizonto modelių maišaties matricos

Toliau pateikiamos investavimo strategijų rezultatų lentelės. Lentelių pavadinimų šifravimas yra toks: „Mėn. (Sav.), visi (dalis), *Lag (Foresight)*, 1978-2014 (2015-2022)“ - reiškia, kad strategijoms buvo naudojami mėnesinio (savaitinio) horizonto mašininio mokymosi modeliai, kurie buvo apmokyti naudojant visus (dalį) makroekonominius rodiklius, investavimo strategija yra *Lag (Foresight)*, rezultatai skaičiuoti imtyje nuo 1978 vasario mėnesio iki 2014 gruodžio mėnesio (nuo 2015 iki 2022 metų).

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample	7.2%	11.2%	0.65	1.000	24.6%	-15.1%	0.142	0.51
Regime_full_sample TWBmk	2.1%	8.0%	0.26	0.682	37.3%	-12.1%	0.042	0.51
Regime_full_sample RiskBmk	2.8%	11.2%	0.25	0.672	48.7%	-16.7%	0.039	0.71
Regime_backtest	2.3%	12.0%	0.19	0.622	51.3%	-18.5%	0.039	0.57
Regime_backtest TWBmk	2.3%	9.0%	0.26	0.679	41.1%	-13.6%	0.041	0.57
Regime_backtest RiskBmk	2.9%	12.0%	0.24	0.669	51.6%	-18.1%	0.039	0.76
LR_enet	3.1%	14.5%	0.22	0.638	53.8%	-21.9%	0.037	0.84
LR_enet TWBmk	3.2%	13.3%	0.24	0.665	55.7%	-20.0%	0.038	0.84
LR_enet RiskBmk	3.4%	14.5%	0.23	0.661	59.3%	-21.8%	0.037	0.92
LR_L1	3.4%	11.9%	0.28	0.715	46.9%	-17.1%	0.060	0.56
LR_L1 TWBmk	2.3%	8.9%	0.26	0.680	40.6%	-13.4%	0.041	0.56
LR_L1 RiskBmk	2.9%	11.9%	0.24	0.670	51.2%	-17.9%	0.039	0.75
LR_L2	3.4%	12.3%	0.28	0.720	47.6%	-18.2%	0.057	0.60
LR_L2 TWBmk	2.4%	9.5%	0.26	0.678	42.7%	-14.2%	0.041	0.60
LR_L2 RiskBmk	3.0%	12.3%	0.24	0.668	52.4%	-18.4%	0.038	0.78
RF	3.6%	11.6%	0.31	0.711	36.2%	-16.8%	0.068	0.53
RF TWBmk	2.2%	8.4%	0.26	0.681	38.8%	-12.7%	0.041	0.53
RF RiskBmk	2.9%	11.6%	0.25	0.671	50.1%	-17.4%	0.039	0.73
GBM	2.3%	11.8%	0.20	0.627	44.0%	-18.2%	0.042	0.55
GBM TWBmk	2.3%	8.6%	0.26	0.680	39.6%	-13.0%	0.041	0.55
GBM RiskBmk	2.9%	11.8%	0.25	0.670	50.7%	-17.6%	0.039	0.74
XGBM	3.4%	11.7%	0.29	0.723	43.9%	-16.9%	0.062	0.54
XGBM TWBmk	2.2%	8.6%	0.26	0.681	39.4%	-12.9%	0.041	0.54
XGBM RiskBmk	2.9%	11.7%	0.25	0.670	50.5%	-17.6%	0.039	0.74
SVC	1.2%	11.3%	0.11	0.574	38.0%	-17.0%	0.024	0.51
SVC TWBmk	2.1%	8.0%	0.26	0.682	37.2%	-12.0%	0.042	0.51
SVC RiskBmk	2.8%	11.3%	0.25	0.672	49.2%	-17.0%	0.039	0.72
DNN	3.7%	12.0%	0.31	0.807	35.1%	-16.9%	0.065	0.57
DNN TWBmk	2.3%	9.1%	0.26	0.679	41.2%	-13.6%	0.041	0.57
DNN RiskBmk	2.9%	12.0%	0.24	0.669	51.5%	-18.0%	0.039	0.76
Median	2.5%	11.9%	0.21	0.646	50.4%	-17.5%	0.045	0.56
Median TWBmk	2.3%	8.8%	0.26	0.680	40.3%	-13.2%	0.041	0.56
Median RiskBmk	2.9%	11.9%	0.24	0.670	51.1%	-17.8%	0.039	0.75
S&P 500	3.6%	15.8%	0.22	0.656	62.9%	-23.8%	0.036	1

12 lentelė: Mėn., visi, *Lag*, 1978-2014

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample	9.5%	11.6%	0.81	1	25.5%	-14.5%	0.168	0.56
Regime_full_sample TWBmk	2.3%	8.9%	0.26	0.680	40.6%	-13.4%	0.041	0.56
Regime_full_sample RiskBmk	2.9%	11.6%	0.25	0.671	50.2%	-17.4%	0.039	0.73
Regime_backtest	11.8%	11.7%	1.01	1	17.3%	-14.6%	0.201	0.59
Regime_backtest TWBmk	2.4%	9.3%	0.26	0.678	42.1%	-14.0%	0.041	0.59
Regime_backtest RiskBmk	2.9%	11.7%	0.25	0.670	50.5%	-17.6%	0.039	0.74
LR_enet	2.8%	14.4%	0.19	0.622	57.6%	-21.9%	0.034	0.82
LR_enet TWBmk	3.1%	13.0%	0.24	0.666	54.7%	-19.5%	0.038	0.82
LR_enet RiskBmk	3.3%	14.4%	0.23	0.661	58.8%	-21.6%	0.037	0.91
LR_L1	7.6%	11.3%	0.67	1	21.5%	-15.5%	0.144	0.53
LR_L1 TWBmk	2.2%	8.3%	0.26	0.681	38.4%	-12.5%	0.042	0.53
LR_L1 RiskBmk	2.8%	11.3%	0.25	0.672	49.3%	-17.0%	0.039	0.72
LR_L2	6.3%	11.3%	0.56	1	21.5%	-15.6%	0.122	0.52
LR_L2 TWBmk	2.1%	8.2%	0.26	0.682	37.7%	-12.3%	0.042	0.52
LR_L2 RiskBmk	2.8%	11.3%	0.25	0.672	49.1%	-16.9%	0.039	0.71
RF	4.3%	11.4%	0.37	0.750	31.3%	-16.2%	0.082	0.52
RF TWBmk	2.2%	8.2%	0.26	0.682	37.9%	-12.3%	0.042	0.52
RF RiskBmk	2.8%	11.4%	0.25	0.671	49.5%	-17.1%	0.039	0.72
GBM	4.3%	12.2%	0.35	0.804	34.4%	-16.9%	0.072	0.59
GBM TWBmk	2.4%	9.4%	0.26	0.678	42.5%	-14.1%	0.041	0.59
GBM RiskBmk	3.0%	12.2%	0.24	0.669	52.1%	-18.3%	0.038	0.77
XGBM	4.6%	12.0%	0.38	0.804	32.2%	-16.9%	0.080	0.57
XGBM TWBmk	2.4%	9.1%	0.26	0.679	41.3%	-13.6%	0.041	0.57
XGBM RiskBmk	2.9%	12.0%	0.24	0.669	51.5%	-18.0%	0.039	0.76
SVC	3.3%	10.8%	0.30	0.839	52.6%	-16.1%	0.071	0.46
SVC TWBmk	2.0%	7.3%	0.27	0.685	34.4%	-11.0%	0.042	0.46
SVC RiskBmk	2.7%	10.8%	0.25	0.673	47.5%	-16.2%	0.040	0.68
DNN	3.7%	11.4%	0.32	0.999	43.3%	-17.0%	0.070	0.52
DNN TWBmk	2.2%	8.3%	0.26	0.682	38.1%	-12.4%	0.042	0.52
DNN RiskBmk	2.8%	11.4%	0.25	0.671	49.6%	-17.2%	0.039	0.72
Median	5.1%	12.1%	0.42	0.852	30.9%	-16.2%	0.086	0.59
Median TWBmk	2.4%	9.3%	0.26	0.678	42.2%	-14.0%	0.041	0.59
Median RiskBmk	2.9%	12.1%	0.24	0.669	51.8%	-18.2%	0.039	0.76
S&P 500	3.6%	15.8%	0.22	0.656	62.9%	-23.8%	0.036	1

13 lentelė: Mėn., visi, *Foresight*, 1978-2014

Strategija	Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample		7.6%	11.9%	0.64	0.971	25.0%	-18.3%	0.140	0.54
Regime_full_sample TWBmk		2.3%	8.8%	0.26	0.719	40.2%	-13.8%	0.043	0.54
Regime_full_sample RiskBmk		3.0%	11.9%	0.25	0.707	51.3%	-18.8%	0.040	0.73
Regime_backtest		2.7%	12.5%	0.22	0.684	41.2%	-20.0%	0.046	0.58
Regime_backtest TWBmk		2.5%	9.5%	0.26	0.716	42.8%	-14.9%	0.042	0.58
Regime_backtest RiskBmk		3.1%	12.5%	0.24	0.705	52.9%	-19.6%	0.040	0.77
LR_enet		3.2%	15.2%	0.21	0.668	57.6%	-24.0%	0.037	0.87
LR_enet TWBmk		3.3%	14.1%	0.24	0.698	58.0%	-22.1%	0.039	0.87
LR_enet RiskBmk		3.5%	15.2%	0.23	0.694	61.0%	-23.8%	0.038	0.93
LR_L1		3.6%	12.5%	0.29	0.738	40.5%	-19.4%	0.062	0.59
LR_L1 TWBmk		2.5%	9.6%	0.26	0.716	43.0%	-15.0%	0.042	0.59
LR_L1 RiskBmk		3.1%	12.5%	0.24	0.705	53.0%	-19.6%	0.040	0.77
LR_L2		3.2%	12.5%	0.26	0.714	41.3%	-20.0%	0.055	0.59
LR_L2 TWBmk		2.5%	9.6%	0.26	0.716	43.0%	-15.0%	0.042	0.59
LR_L2 RiskBmk		3.1%	12.5%	0.24	0.705	53.0%	-19.6%	0.040	0.77
RF		3.5%	12.7%	0.28	0.709	37.2%	-19.8%	0.058	0.61
RF TWBmk		2.5%	9.9%	0.26	0.715	44.1%	-15.5%	0.042	0.61
RF RiskBmk		3.1%	12.7%	0.24	0.704	53.6%	-19.9%	0.040	0.78
GBM		2.7%	12.9%	0.21	0.665	43.4%	-20.6%	0.043	0.63
GBM TWBmk		2.6%	10.2%	0.26	0.714	45.3%	-16.0%	0.042	0.63
GBM RiskBmk		3.1%	12.9%	0.24	0.703	54.3%	-20.2%	0.040	0.79
XGBM		3.9%	12.8%	0.31	0.742	36.7%	-19.6%	0.064	0.61
XGBM TWBmk		2.6%	10.0%	0.26	0.714	44.5%	-15.7%	0.042	0.61
XGBM RiskBmk		3.1%	12.8%	0.24	0.704	53.9%	-20.0%	0.040	0.78
SVC		1.4%	12.1%	0.12	0.600	49.6%	-19.0%	0.026	0.55
SVC TWBmk		2.4%	9.0%	0.26	0.718	40.8%	-14.1%	0.043	0.55
SVC RiskBmk		3.0%	12.1%	0.25	0.706	51.8%	-19.0%	0.040	0.74
DNN		3.9%	12.4%	0.31	0.770	35.2%	-18.9%	0.066	0.58
DNN TWBmk		2.5%	9.5%	0.26	0.716	42.8%	-14.9%	0.042	0.58
DNN RiskBmk		3.0%	12.4%	0.24	0.705	52.9%	-19.5%	0.040	0.76
Median		3.5%	12.9%	0.27	0.727	40.7%	-20.5%	0.056	0.63
Median TWBmk		2.6%	10.3%	0.26	0.713	45.5%	-16.1%	0.042	0.63
Median RiskBmk		3.1%	12.9%	0.24	0.703	54.4%	-20.3%	0.039	0.79
S&P 500		3.7%	16.3%	0.23	0.690	64.0%	-25.5%	0.037	1

14 lentelė: Sav., visi, Lag, 1978-2014

Strategija	Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample		8.4%	11.9%	0.71	0.995	24.8%	-18.3%	0.156	0.53
Regime_full_sample TWBmk		2.3%	8.7%	0.26	0.719	39.8%	-13.7%	0.043	0.53
Regime_full_sample RiskBmk		2.9%	11.9%	0.25	0.707	51.0%	-18.6%	0.040	0.73
Regime_backtest		13.2%	12.3%	1.08	1	19.7%	-16.9%	0.227	0.58
Regime_backtest TWBmk		2.5%	9.5%	0.26	0.716	42.8%	-14.9%	0.042	0.58
Regime_backtest RiskBmk		3.0%	12.3%	0.25	0.705	52.5%	-19.3%	0.040	0.76
LR_enet		3.8%	15.1%	0.25	0.710	55.2%	-23.9%	0.044	0.86
LR_enet TWBmk		3.3%	13.9%	0.24	0.699	57.5%	-21.9%	0.039	0.86
LR_enet RiskBmk		3.5%	15.1%	0.23	0.695	60.7%	-23.6%	0.038	0.93
LR_L1		4.4%	12.5%	0.35	0.780	35.5%	-19.0%	0.075	0.59
LR_L1 TWBmk		2.5%	9.5%	0.26	0.716	42.9%	-15.0%	0.042	0.59
LR_L1 RiskBmk		3.1%	12.5%	0.24	0.705	52.9%	-19.6%	0.040	0.77
LR_L2		4.1%	12.6%	0.33	0.770	37.6%	-19.2%	0.069	0.60
LR_L2 TWBmk		2.5%	9.7%	0.26	0.715	43.6%	-15.3%	0.042	0.60
LR_L2 RiskBmk		3.1%	12.6%	0.24	0.704	53.3%	-19.8%	0.040	0.77
RF		4.3%	12.7%	0.34	0.750	34.6%	-19.7%	0.071	0.61
RF TWBmk		2.5%	9.9%	0.26	0.715	44.1%	-15.5%	0.042	0.61
RF RiskBmk		3.1%	12.7%	0.24	0.704	53.6%	-19.9%	0.040	0.78
GBM		3.0%	12.9%	0.23	0.677	41.8%	-20.6%	0.047	0.63
GBM TWBmk		2.6%	10.2%	0.26	0.714	45.3%	-16.0%	0.042	0.63
GBM RiskBmk		3.1%	12.9%	0.24	0.703	54.3%	-20.2%	0.040	0.79
XGBM		4.3%	12.7%	0.34	0.763	41.9%	-19.8%	0.071	0.61
XGBM TWBmk		2.6%	10.0%	0.26	0.714	44.5%	-15.7%	0.042	0.61
XGBM RiskBmk		3.1%	12.7%	0.24	0.704	53.8%	-20.0%	0.040	0.78
SVC		1.5%	12.1%	0.13	0.607	44.0%	-19.0%	0.028	0.55
SVC TWBmk		2.3%	9.0%	0.26	0.718	40.7%	-14.1%	0.043	0.55
SVC RiskBmk		3.0%	12.1%	0.25	0.706	51.7%	-19.0%	0.040	0.74
DNN		4.7%	12.5%	0.38	0.813	37.8%	-19.2%	0.080	0.59
DNN TWBmk		2.5%	9.7%	0.26	0.716	43.3%	-15.2%	0.042	0.59
DNN RiskBmk		3.1%	12.5%	0.24	0.705	53.2%	-19.7%	0.040	0.77
Median		4.4%	12.9%	0.34	0.766	41.6%	-20.3%	0.070	0.63
Median TWBmk		2.6%	10.3%	0.26	0.713	45.5%	-16.1%	0.042	0.63
Median RiskBmk		3.1%	12.9%	0.24	0.703	54.4%	-20.3%	0.039	0.79
S&P 500		3.7%	16.3%	0.23	0.690	64.0%	-25.5%	0.037	1

15 lentelė: Sav., visi, Foresight, 1978-2014

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample	6.4%	11.1%	0.57	0.997	26.7%	-15.6%	0.128	0.50
Regime_full_sample TWBmk	2.1%	7.9%	0.26	0.683	36.7%	-11.9%	0.042	0.50
Regime_full_sample RiskBmk	2.8%	11.1%	0.25	0.672	48.5%	-16.7%	0.039	0.70
Regime_backtest	2.3%	12.0%	0.19	0.622	51.3%	-18.5%	0.039	0.57
Regime_backtest TWBmk	2.3%	9.0%	0.26	0.679	41.1%	-13.6%	0.041	0.57
Regime_backtest RiskBmk	2.9%	12.0%	0.24	0.669	51.6%	-18.1%	0.039	0.76
LR_enet	3.1%	12.6%	0.24	0.724	40.6%	-19.1%	0.049	0.63
LR_enet TWBmk	2.5%	10.0%	0.25	0.676	44.7%	-15.0%	0.040	0.63
LR_enet RiskBmk	3.0%	12.6%	0.24	0.667	53.5%	-18.9%	0.038	0.80
LR_L1	4.2%	12.1%	0.35	0.814	39.0%	-16.9%	0.073	0.58
LR_L1 TWBmk	2.4%	9.2%	0.26	0.679	41.6%	-13.8%	0.041	0.58
LR_L1 RiskBmk	2.9%	12.1%	0.24	0.669	51.7%	-18.1%	0.039	0.76
LR_L2	2.8%	12.0%	0.23	0.709	46.3%	-17.7%	0.049	0.57
LR_L2 TWBmk	2.3%	9.0%	0.26	0.679	40.9%	-13.5%	0.041	0.57
LR_L2 RiskBmk	2.9%	12.0%	0.24	0.669	51.4%	-18.0%	0.039	0.76
RF	3.9%	11.7%	0.34	0.770	39.1%	-16.9%	0.072	0.54
RF TWBmk	2.2%	8.6%	0.26	0.681	39.5%	-12.9%	0.041	0.54
RF RiskBmk	2.9%	11.7%	0.25	0.670	50.4%	-17.5%	0.039	0.74
GBM	3.5%	11.9%	0.29	0.740	43.8%	-17.0%	0.062	0.57
GBM TWBmk	2.3%	9.0%	0.26	0.679	40.8%	-13.4%	0.041	0.57
GBM RiskBmk	2.9%	11.9%	0.24	0.669	51.3%	-17.9%	0.039	0.75
XGBM	3.3%	11.7%	0.28	0.731	43.5%	-17.0%	0.060	0.55
XGBM TWBmk	2.3%	8.7%	0.26	0.680	39.7%	-13.0%	0.041	0.55
XGBM RiskBmk	2.9%	11.7%	0.25	0.670	50.6%	-17.6%	0.039	0.74
SVC	2.2%	11.0%	0.20	0.824	43.9%	-17.1%	0.046	0.47
SVC TWBmk	2.0%	7.5%	0.27	0.684	35.2%	-11.3%	0.042	0.47
SVC RiskBmk	2.7%	11.0%	0.25	0.673	48.0%	-16.5%	0.039	0.69
DNN	2.7%	11.6%	0.24	0.983	35.6%	-16.9%	0.051	0.53
DNN TWBmk	2.2%	8.4%	0.26	0.681	38.7%	-12.6%	0.041	0.53
DNN RiskBmk	2.9%	11.6%	0.25	0.671	50.1%	-17.4%	0.039	0.73
Median	4.2%	12.2%	0.34	0.801	40.8%	-17.6%	0.070	0.60
Median TWBmk	2.4%	9.5%	0.26	0.678	42.7%	-14.2%	0.041	0.60
Median RiskBmk	3.0%	12.2%	0.24	0.668	52.3%	-18.4%	0.038	0.77
S&P 500	3.6%	15.8%	0.22	0.656	62.9%	-23.8%	0.036	1

16 lentelė: Mėn., dalis, Lag, 1978-2014

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample	8.3%	11.3%	0.73	1	20.9%	-14.8%	0.157	0.53
Regime_full_sample TWBmk	2.2%	8.4%	0.26	0.681	38.6%	-12.6%	0.041	0.53
Regime_full_sample RiskBmk	2.8%	11.3%	0.25	0.672	49.3%	-17.0%	0.039	0.72
Regime_backtest	11.8%	11.7%	1.01	1	17.3%	-14.6%	0.201	0.59
Regime_backtest TWBmk	2.4%	9.3%	0.26	0.678	42.1%	-14.0%	0.041	0.59
Regime_backtest RiskBmk	2.9%	11.7%	0.25	0.670	50.5%	-17.6%	0.039	0.74
LR_enet	3.9%	12.7%	0.30	0.764	51.6%	-19.1%	0.060	0.65
LR_enet TWBmk	2.6%	10.2%	0.25	0.675	45.5%	-15.4%	0.040	0.65
LR_enet RiskBmk	3.1%	12.7%	0.24	0.667	53.9%	-19.1%	0.038	0.80
LR_L1	8.0%	11.8%	0.68	1	46.6%	-16.1%	0.140	0.57
LR_L1 TWBmk	2.3%	9.0%	0.26	0.679	41.0%	-13.5%	0.041	0.57
LR_L1 RiskBmk	2.9%	11.8%	0.25	0.670	50.8%	-17.7%	0.039	0.74
LR_L2	8.0%	11.8%	0.68	1	47.3%	-16.1%	0.140	0.57
LR_L2 TWBmk	2.3%	9.1%	0.26	0.679	41.2%	-13.6%	0.041	0.57
LR_L2 RiskBmk	2.9%	11.8%	0.25	0.670	50.9%	-17.7%	0.039	0.75
RF	6.1%	12.2%	0.50	0.896	30.9%	-16.4%	0.101	0.60
RF TWBmk	2.4%	9.5%	0.26	0.678	42.9%	-14.3%	0.041	0.60
RF RiskBmk	3.0%	12.2%	0.24	0.669	52.1%	-18.3%	0.038	0.77
GBM	5.5%	12.0%	0.46	0.837	30.9%	-16.9%	0.095	0.58
GBM TWBmk	2.4%	9.2%	0.26	0.678	41.9%	-13.9%	0.041	0.58
GBM RiskBmk	2.9%	12.0%	0.24	0.669	51.6%	-18.1%	0.039	0.76
XGBM	6.3%	12.2%	0.52	0.901	30.9%	-16.8%	0.105	0.60
XGBM TWBmk	2.4%	9.5%	0.26	0.678	42.8%	-14.3%	0.041	0.60
XGBM RiskBmk	3.0%	12.2%	0.24	0.669	52.1%	-18.3%	0.039	0.77
SVC	6.2%	11.9%	0.52	0.825	31.2%	-16.0%	0.107	0.58
SVC TWBmk	2.4%	9.1%	0.26	0.679	41.4%	-13.7%	0.041	0.58
SVC RiskBmk	2.9%	11.9%	0.24	0.670	51.2%	-17.9%	0.039	0.75
DNN	4.8%	12.1%	0.40	0.851	37.1%	-17.1%	0.083	0.58
DNN TWBmk	2.4%	9.2%	0.26	0.679	41.7%	-13.8%	0.041	0.58
DNN RiskBmk	2.9%	12.1%	0.24	0.669	51.7%	-18.1%	0.039	0.76
Median	6.7%	12.4%	0.54	0.908	36.1%	-17.0%	0.106	0.63
Median TWBmk	2.5%	9.9%	0.25	0.676	44.4%	-14.9%	0.040	0.63
Median RiskBmk	3.0%	12.4%	0.24	0.668	52.9%	-18.7%	0.038	0.79
S&P 500	3.6%	15.8%	0.22	0.656	62.9%	-23.8%	0.036	1

17 lentelė: Mėn., dalis, Foresight, 1978-2014

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample	7.6%	11.9%	0.64	0.971	25.0%	-18.3%	0.140	0.54
Regime_full_sample TWBmk	2.3%	8.8%	0.26	0.719	40.2%	-13.8%	0.043	0.54
Regime_full_sample RiskBmk	3.0%	11.9%	0.25	0.707	51.3%	-18.8%	0.040	0.73
Regime_backtest	2.7%	12.5%	0.22	0.684	41.2%	-20.0%	0.046	0.58
Regime_backtest TWBmk	2.5%	9.5%	0.26	0.716	42.8%	-14.9%	0.042	0.58
Regime_backtest RiskBmk	3.1%	12.5%	0.24	0.705	52.9%	-19.6%	0.040	0.77
LR_enet	2.8%	13.3%	0.21	0.683	42.2%	-21.8%	0.043	0.66
LR_enet TWBmk	2.7%	10.8%	0.25	0.711	47.3%	-16.9%	0.041	0.66
LR_enet RiskBmk	3.2%	13.3%	0.24	0.702	55.4%	-20.8%	0.039	0.81
LR_L1	5.0%	12.4%	0.40	0.847	38.5%	-19.0%	0.086	0.58
LR_L1 TWBmk	2.5%	9.5%	0.26	0.716	42.8%	-14.9%	0.042	0.58
LR_L1 RiskBmk	3.0%	12.4%	0.24	0.705	52.9%	-19.5%	0.040	0.76
LR_L2	4.7%	12.7%	0.37	0.820	36.4%	-19.8%	0.078	0.61
LR_L2 TWBmk	2.5%	9.9%	0.26	0.715	44.0%	-15.5%	0.042	0.61
LR_L2 RiskBmk	3.1%	12.7%	0.24	0.704	53.6%	-19.9%	0.040	0.78
RF	4.0%	12.7%	0.32	0.763	39.4%	-18.9%	0.067	0.61
RF TWBmk	2.5%	9.9%	0.26	0.715	44.2%	-15.5%	0.042	0.61
RF RiskBmk	3.1%	12.7%	0.24	0.704	53.7%	-19.9%	0.040	0.78
GBM	2.9%	13.0%	0.23	0.694	37.0%	-20.3%	0.046	0.63
GBM TWBmk	2.6%	10.3%	0.26	0.713	45.7%	-16.2%	0.042	0.63
GBM RiskBmk	3.1%	13.0%	0.24	0.703	54.5%	-20.4%	0.039	0.80
XGBM	3.6%	12.7%	0.28	0.737	37.7%	-19.0%	0.059	0.60
XGBM TWBmk	2.5%	9.8%	0.26	0.715	44.0%	-15.5%	0.042	0.60
XGBM RiskBmk	3.1%	12.7%	0.24	0.704	53.6%	-19.9%	0.040	0.78
SVC	1.5%	11.9%	0.12	0.599	48.8%	-18.8%	0.027	0.53
SVC TWBmk	2.3%	8.7%	0.26	0.719	39.6%	-13.6%	0.043	0.53
SVC RiskBmk	2.9%	11.9%	0.25	0.707	51.1%	-18.7%	0.040	0.73
DNN	3.7%	12.2%	0.30	0.794	44.6%	-19.0%	0.066	0.56
DNN TWBmk	2.4%	9.2%	0.26	0.717	41.6%	-14.4%	0.042	0.56
DNN RiskBmk	3.0%	12.2%	0.25	0.706	52.2%	-19.2%	0.040	0.75
Median	3.7%	13.0%	0.29	0.740	42.0%	-20.4%	0.059	0.64
Median TWBmk	2.6%	10.3%	0.26	0.713	45.8%	-16.2%	0.042	0.64
Median RiskBmk	3.1%	13.0%	0.24	0.703	54.6%	-20.4%	0.039	0.80
S&P 500	3.7%	16.3%	0.23	0.690	64.0%	-25.5%	0.037	1

18 lentelė: Sav., dalis, *Lag*, 1978-2014

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_full_sample	8.4%	11.9%	0.71	0.995	24.8%	-18.3%	0.156	0.53
Regime_full_sample TWBmk	2.3%	8.7%	0.26	0.719	39.8%	-13.7%	0.043	0.53
Regime_full_sample RiskBmk	2.9%	11.9%	0.25	0.707	51.0%	-18.6%	0.040	0.73
Regime_backtest	13.2%	12.3%	1.08	1	19.7%	-16.9%	0.227	0.58
Regime_backtest TWBmk	2.5%	9.5%	0.26	0.716	42.8%	-14.9%	0.042	0.58
Regime_backtest RiskBmk	3.0%	12.3%	0.25	0.705	52.5%	-19.3%	0.040	0.76
LR_enet	2.9%	13.2%	0.22	0.687	46.9%	-21.3%	0.045	0.65
LR_enet TWBmk	2.7%	10.6%	0.25	0.712	46.8%	-16.7%	0.041	0.65
LR_enet RiskBmk	3.2%	13.2%	0.24	0.702	55.1%	-20.7%	0.039	0.81
LR_L1	5.3%	12.5%	0.43	0.860	41.1%	-19.0%	0.090	0.59
LR_L1 TWBmk	2.5%	9.6%	0.26	0.716	43.2%	-15.1%	0.042	0.59
LR_L1 RiskBmk	3.1%	12.5%	0.24	0.705	53.1%	-19.6%	0.040	0.77
LR_L2	5.3%	12.6%	0.42	0.857	41.6%	-19.2%	0.088	0.60
LR_L2 TWBmk	2.5%	9.8%	0.26	0.715	43.8%	-15.3%	0.042	0.60
LR_L2 RiskBmk	3.1%	12.6%	0.24	0.704	53.4%	-19.8%	0.040	0.77
RF	4.1%	12.6%	0.32	0.760	36.7%	-18.9%	0.068	0.60
RF TWBmk	2.5%	9.8%	0.26	0.715	43.8%	-15.3%	0.042	0.60
RF RiskBmk	3.1%	12.6%	0.24	0.704	53.4%	-19.8%	0.040	0.78
GBM	3.9%	12.9%	0.30	0.749	36.3%	-20.0%	0.062	0.63
GBM TWBmk	2.6%	10.3%	0.26	0.713	45.6%	-16.2%	0.042	0.63
GBM RiskBmk	3.1%	12.9%	0.24	0.703	54.5%	-20.3%	0.039	0.80
XGBM	4.7%	12.6%	0.37	0.790	34.6%	-18.9%	0.078	0.60
XGBM TWBmk	2.5%	9.8%	0.26	0.715	43.7%	-15.3%	0.042	0.60
XGBM RiskBmk	3.1%	12.6%	0.24	0.704	53.4%	-19.8%	0.040	0.77
SVC	3.1%	12.0%	0.26	0.727	42.6%	-18.7%	0.057	0.54
SVC TWBmk	2.3%	8.8%	0.26	0.719	40.2%	-13.8%	0.043	0.54
SVC RiskBmk	3.0%	12.0%	0.25	0.707	51.4%	-18.8%	0.040	0.74
DNN	3.9%	12.1%	0.32	0.839	34.8%	-18.8%	0.070	0.55
DNN TWBmk	2.3%	9.0%	0.26	0.718	40.7%	-14.1%	0.043	0.55
DNN RiskBmk	3.0%	12.1%	0.25	0.706	51.7%	-19.0%	0.040	0.74
Median	4.4%	12.9%	0.34	0.772	46.8%	-20.0%	0.069	0.63
Median TWBmk	2.6%	10.3%	0.26	0.713	45.5%	-16.1%	0.042	0.63
Median RiskBmk	3.1%	12.9%	0.24	0.703	54.4%	-20.3%	0.039	0.79
S&P 500	3.7%	16.3%	0.23	0.690	64.0%	-25.5%	0.037	1

19 lentelė: Sav., dalis, *Foresight*, 1978-2014

Mašininio mokymosi modelių tikslumo rodikliai (2015-2022 imtyje), kai apmokymui naudojami visi makroekonominiai duomenys

	ACC	MCC	QPS	AUC
LR-I1	0.865	-0.040	0.112	0.676
LR-I2	0.832	0.191	0.154	0.646
RF	0.854	-0.057	0.118	0.669
GBM	0.865	-0.040	0.119	0.576
XGBM	0.843	-0.070	0.138	0.604
DNN	0.809	0.152	0.182	0.624
MEDIAN	0.865	-0.040	0.117	0.648

20 lentelė: Mėnesinio horizonto modelių tikslumo rodikliai

	LR-I1		LR-I2		RF	
Tikras režimas		0	1		0	1
	0	77	1	0	71	7
	1	11	0	1	8	3
		Režimo prognozė			Režimo prognozė	
	GBM		XGBM		DNN	
Tikras režimas		0	1		0	1
	0	77	1	0	75	3
	1	11	0	1	11	0
		Režimo prognozė			Režimo prognozė	
	Median					
Tikras režimas		0	1		0	1
	0	77	1	0	69	9
	1	11	0	1	8	3
		Režimo prognozė			Režimo prognozė	

21 lentelė: Mėnesinio horizonto modelių maišaties matricos

Mašininio mokymosi modelių tikslumo rodikliai (2015-2022 imtyje), kai apmokymui naudojami visi makroekonominiai duomenys

	ACC	MCC	QPS	AUC
LR-I1	0.854	0.012	0.125	0.581
LR-I2	0.798	0.012	0.172	0.601
RF	0.843	-0.060	0.128	0.622
GBM	0.854	-0.042	0.126	0.587
XGBM	0.832	-0.074	0.145	0.562
DNN	0.787	0.034	0.191	0.553
MEDIAN	0.851	-0.047	0.125	0.595

22 lentelė: Savaitinio horizonto modelių tikslumo rodikliai

		LR-I1		LR-I2		RF	
Tikras režimas		0	1	0	1	0	1
	0	303	5	279	29	300	8
1	47	1	43	5	48	0	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	

		GBM		XGBM		DNN	
Tikras režimas		0	1	0	1	0	1
	0	304	4	296	12	273	35
1	48	0	48	0	41	7	
		Režimo prognozė		Režimo prognozė		Režimo prognozė	

		Median	
Tikras režimas		0	1
	0	303	5
1	48	0	
		Režimo prognozė	

23 lentelė: Savaitinio horizonto modelių maišaties matricos

Mašininio mokymosi modelių tikslumo rodikliai (2015-2022 imtyje), kai apmokymui naudojama dalis makroekonominių duomenų

	ACC	MCC	QPS	AUC
LR-I1	0.876	0.173	0.107	0.680
LR-I2	0.820	0.101	0.123	0.615
RF	0.854	-0.057	0.110	0.732
GBM	0.865	-0.040	0.111	0.677
XGBM	0.854	-0.057	0.135	0.740
DNN	0.798	0.066	0.183	0.626
MEDIAN	0.865	0.119	0.109	0.702

24 lentelė: Mėnesinio horizonto modelių tikslumo rodikliai

	LR-I1		LR-I2		RF	
Tikras režimas		0	1		0	1
	0	77	1	0	71	7
	1	10	1	1	9	2
		Režimo prognozė			Režimo prognozė	
	GBM		XGBM		DNN	
Tikras režimas		0	1		0	1
	0	77	1	0	76	2
	1	11	0	1	11	0
		Režimo prognozė			Režimo prognozė	
	Median					
Tikras režimas		0	1		0	1
	0	76	2	0	69	9
	1	10	1	1	9	2
		Režimo prognozė			Režimo prognozė	

25 lentelė: Mėnesinio horizonto modelių maišaties matricos

Mašininio mokymosi modelių tikslumo rodikliai (2015-2022 imtyje), kai apmokymui naudojama dalis makroekonominių duomenų

	ACC	MCC	QPS	AUC
LR-I1	0.857	0.094	0.115	0.704
LR-I2	0.801	0.072	0.143	0.622
RF	0.846	-0.011	0.117	0.728
GBM	0.854	-0.042	0.118	0.681
XGBM	0.846	-0.056	0.140	0.717
DNN	0.770	0.013	0.197	0.581
MEDIAN	0.848	-0.004	0.117	0.719

26 lentelė: Savaitinio horizonto modelių tikslumo rodikliai

		LR-I1		LR-I2		RF	
Tikras režimas		0	1	0	1	0	1
	0	302	6	277	31	300	8
	1	45	3	40	8	47	1
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		GBM		XGBM		DNN	
Tikras režimas		0	1	0	1	0	1
	0	304	4	301	7	267	41
	1	48	0	48	0	41	7
		Režimo prognozė		Režimo prognozė		Režimo prognozė	
		Median					
Tikras režimas		0	1				
	0	301	7				
	1	47	1				
		Režimo prognozė					

27 lentelė: Savaitinio horizonto modelių maišaties matricos

Strategija / Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample	6.7%	12.0%	0.56	0.591	18.5%	-18.9%	0.087	0.77
Regime_bt_sample TWBmk	6.6%	10.5%	0.63	0.619	14.3%	-15.2%	0.086	0.77
Regime_bt_sample RiskBmk	7.5%	12.0%	0.63	0.618	16.3%	-17.4%	0.085	0.88
Regime_backtest	8.0%	11.8%	0.68	0.603	18.5%	-18.9%	0.105	0.76
Regime_backtest TWBmk	6.5%	10.3%	0.64	0.619	14.0%	-14.9%	0.086	0.76
Regime_backtest RiskBmk	7.4%	11.8%	0.63	0.618	16.1%	-17.2%	0.085	0.87
LR_L1	8.1%	13.5%	0.60	0.612	18.5%	-19.7%	0.081	1.00
LR_L1 TWBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
LR_L1 RiskBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
LR_L2	7.7%	12.5%	0.61	0.612	18.5%	-18.9%	0.090	0.85
LR_L2 TWBmk	7.3%	11.6%	0.63	0.618	15.8%	-16.9%	0.085	0.85
LR_L2 RiskBmk	7.8%	12.5%	0.63	0.617	17.1%	-18.3%	0.085	0.93
RF	7.2%	13.4%	0.54	0.601	18.5%	-19.7%	0.074	0.97
RF TWBmk	8.2%	13.2%	0.62	0.616	18.0%	-19.2%	0.084	0.97
RF RiskBmk	8.3%	13.4%	0.62	0.616	18.2%	-19.5%	0.084	0.99
GBM	8.1%	13.5%	0.60	0.612	18.5%	-19.7%	0.081	1.00
GBM TWBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
GBM RiskBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
XGBM	7.1%	13.5%	0.53	0.602	18.5%	-19.7%	0.072	0.98
XGBM TWBmk	8.3%	13.3%	0.62	0.616	18.2%	-19.4%	0.084	0.98
XGBM RiskBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.6%	0.084	0.99
DNN	6.0%	13.1%	0.46	0.588	18.7%	-18.9%	0.065	0.92
DNN TWBmk	7.8%	12.5%	0.63	0.617	17.1%	-18.3%	0.085	0.92
DNN RiskBmk	8.2%	13.1%	0.62	0.616	17.8%	-19.1%	0.084	0.97
Median	8.1%	13.5%	0.60	0.612	18.5%	-19.7%	0.081	1.00
Median TWBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
Median RiskBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
S&P 500	8.4%	13.5%	0.62	0.616	18.5%	-19.7%	0.084	1

28 lentelė: Mėn., visi, *Lag*, 2015-2022

Strategija / Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample	8.0%	11.8%	0.68	0.603	18.5%	-18.9%	0.105	0.76
Regime_bt_sample TWBmk	6.5%	10.3%	0.64	0.619	14.0%	-14.9%	0.086	0.76
Regime_bt_sample RiskBmk	7.4%	11.8%	0.63	0.618	16.1%	-17.2%	0.085	0.87
Regime_backtest	15.4%	8.4%	1.83	0.892	6.4%	-10.4%	0.359	0.43
Regime_backtest TWBmk	3.8%	5.8%	0.65	0.623	7.9%	-8.5%	0.088	0.43
Regime_backtest RiskBmk	5.4%	8.4%	0.64	0.621	11.5%	-12.3%	0.087	0.62
LR_L1	7.7%	13.5%	0.57	0.608	18.5%	-19.7%	0.078	0.98
LR_L1 TWBmk	8.3%	13.3%	0.62	0.616	18.2%	-19.4%	0.084	0.98
LR_L1 RiskBmk	8.4%	13.5%	0.62	0.616	18.3%	-19.6%	0.084	0.99
LR_L2	6.6%	12.4%	0.54	0.598	18.5%	-18.9%	0.080	0.83
LR_L2 TWBmk	7.1%	11.2%	0.63	0.618	15.3%	-16.3%	0.086	0.83
LR_L2 RiskBmk	7.8%	12.4%	0.63	0.617	16.9%	-18.0%	0.085	0.91
RF	7.4%	13.4%	0.55	0.604	18.5%	-19.7%	0.075	0.98
RF TWBmk	8.3%	13.3%	0.62	0.616	18.1%	-19.4%	0.084	0.98
RF RiskBmk	8.4%	13.4%	0.62	0.616	18.3%	-19.6%	0.084	0.99
GBM	7.7%	13.5%	0.57	0.608	18.5%	-19.7%	0.078	0.98
GBM TWBmk	8.3%	13.3%	0.62	0.616	18.2%	-19.4%	0.084	0.98
GBM RiskBmk	8.4%	13.5%	0.62	0.616	18.3%	-19.6%	0.084	0.99
XGBM	7.9%	13.4%	0.59	0.609	18.5%	-19.7%	0.080	0.98
XGBM TWBmk	8.2%	13.2%	0.62	0.616	18.0%	-19.3%	0.084	0.98
XGBM RiskBmk	8.3%	13.4%	0.62	0.616	18.3%	-19.5%	0.084	0.99
DNN	5.4%	11.8%	0.46	0.574	18.5%	-18.9%	0.072	0.75
DNN TWBmk	6.5%	10.2%	0.64	0.619	13.9%	-14.9%	0.086	0.75
DNN RiskBmk	7.4%	11.8%	0.63	0.618	16.1%	-17.3%	0.085	0.87
Median	7.7%	13.5%	0.57	0.608	18.5%	-19.7%	0.078	0.98
Median TWBmk	8.3%	13.3%	0.62	0.616	18.2%	-19.4%	0.084	0.98
Median RiskBmk	8.4%	13.5%	0.62	0.616	18.3%	-19.6%	0.084	0.99
S&P 500	8.4%	13.5%	0.62	0.616	18.5%	-19.7%	0.084	1

29 lentelė: Mėn., visi, *Foresight*, 2015-2022

Strategija	Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample		5.4%	14.9%	0.37	0.581	28.9%	-19.2%	0.065	0.84
Regime_bt_sample TWBmk		7.4%	13.7%	0.54	0.628	24.8%	-20.3%	0.087	0.84
Regime_bt_sample RiskBmk		8.0%	14.9%	0.53	0.627	26.8%	-22.1%	0.087	0.92
Regime_backtest		5.2%	12.5%	0.41	0.574	20.1%	-17.3%	0.087	0.59
Regime_backtest TWBmk		5.3%	9.6%	0.56	0.633	17.9%	-14.3%	0.090	0.59
Regime_backtest RiskBmk		6.8%	12.5%	0.54	0.630	22.8%	-18.6%	0.088	0.77
LR_L1		7.8%	16.0%	0.49	0.613	28.9%	-23.2%	0.080	0.97
LR_L1 TWBmk		8.4%	15.7%	0.53	0.626	28.2%	-23.4%	0.086	0.97
LR_L1 RiskBmk		8.5%	16.0%	0.53	0.625	28.6%	-23.8%	0.086	0.99
LR_L2		9.1%	15.5%	0.59	0.636	28.9%	-20.0%	0.099	0.92
LR_L2 TWBmk		7.9%	14.8%	0.53	0.627	26.7%	-22.1%	0.087	0.92
LR_L2 RiskBmk		8.2%	15.5%	0.53	0.626	27.8%	-23.1%	0.086	0.96
RF		7.3%	15.9%	0.46	0.607	28.9%	-23.2%	0.075	0.97
RF TWBmk		8.3%	15.7%	0.53	0.626	28.1%	-23.3%	0.086	0.97
RF RiskBmk		8.4%	15.9%	0.53	0.625	28.5%	-23.7%	0.086	0.98
GBM		7.8%	16.0%	0.49	0.613	28.9%	-23.2%	0.080	0.97
GBM TWBmk		8.4%	15.7%	0.53	0.626	28.2%	-23.4%	0.086	0.97
GBM RiskBmk		8.5%	16.0%	0.53	0.625	28.6%	-23.8%	0.086	0.99
XGBM		7.3%	15.9%	0.46	0.606	28.9%	-23.2%	0.076	0.96
XGBM TWBmk		8.3%	15.6%	0.53	0.626	27.9%	-23.2%	0.086	0.96
XGBM RiskBmk		8.4%	15.9%	0.53	0.625	28.4%	-23.6%	0.086	0.98
DNN		6.7%	15.8%	0.43	0.600	28.9%	-20.0%	0.071	0.96
DNN TWBmk		8.2%	15.5%	0.53	0.626	27.7%	-23.0%	0.086	0.96
DNN RiskBmk		8.4%	15.8%	0.53	0.625	28.3%	-23.6%	0.086	0.98
Median		7.7%	16.0%	0.48	0.612	28.9%	-23.2%	0.079	0.97
Median TWBmk		8.3%	15.7%	0.53	0.626	28.1%	-23.4%	0.086	0.97
Median RiskBmk		8.5%	16.0%	0.53	0.625	28.5%	-23.7%	0.086	0.99
S&P 500		8.6%	16.2%	0.53	0.625	28.9%	-24.1%	0.086	1

30 lentelė: Sav., visi, *Lag*, 2015-2022

Strategija	Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample		5.6%	14.9%	0.38	0.583	28.9%	-19.2%	0.067	0.84
Regime_bt_sample TWBmk		7.3%	13.6%	0.54	0.628	24.7%	-20.3%	0.087	0.84
Regime_bt_sample RiskBmk		7.9%	14.9%	0.53	0.627	26.8%	-22.1%	0.087	0.92
Regime_backtest		21.8%	11.5%	1.90	1.000	7.8%	-14.2%	0.421	0.52
Regime_backtest TWBmk		4.7%	8.4%	0.56	0.634	15.8%	-12.5%	0.091	0.52
Regime_backtest RiskBmk		6.3%	11.5%	0.55	0.631	21.1%	-17.0%	0.089	0.71
LR_L1		8.2%	16.0%	0.51	0.619	28.9%	-23.2%	0.084	0.98
LR_L1 TWBmk		8.4%	15.9%	0.53	0.625	28.4%	-23.6%	0.086	0.98
LR_L1 RiskBmk		8.5%	16.0%	0.53	0.625	28.6%	-23.8%	0.086	0.99
LR_L2		8.4%	15.5%	0.54	0.627	28.9%	-20.0%	0.091	0.92
LR_L2 TWBmk		8.0%	14.9%	0.53	0.627	26.8%	-22.1%	0.087	0.92
LR_L2 RiskBmk		8.3%	15.5%	0.53	0.626	27.8%	-23.1%	0.086	0.96
RF		7.3%	15.9%	0.46	0.607	28.9%	-23.2%	0.076	0.97
RF TWBmk		8.3%	15.7%	0.53	0.626	28.1%	-23.3%	0.086	0.97
RF RiskBmk		8.4%	15.9%	0.53	0.625	28.5%	-23.7%	0.086	0.98
GBM		8.2%	16.0%	0.51	0.619	28.9%	-23.2%	0.084	0.98
GBM TWBmk		8.4%	15.9%	0.53	0.625	28.4%	-23.6%	0.086	0.98
GBM RiskBmk		8.5%	16.0%	0.53	0.625	28.6%	-23.8%	0.086	0.99
XGBM		7.5%	15.9%	0.47	0.610	28.9%	-23.2%	0.078	0.97
XGBM TWBmk		8.3%	15.7%	0.53	0.626	28.1%	-23.3%	0.086	0.97
XGBM RiskBmk		8.4%	15.9%	0.53	0.625	28.5%	-23.7%	0.086	0.98
DNN		5.9%	15.9%	0.37	0.590	28.9%	-23.2%	0.062	0.96
DNN TWBmk		8.3%	15.5%	0.53	0.626	27.8%	-23.1%	0.086	0.96
DNN RiskBmk		8.4%	15.9%	0.53	0.625	28.4%	-23.6%	0.086	0.98
Median		7.8%	16.0%	0.49	0.613	28.9%	-23.2%	0.080	0.97
Median TWBmk		8.3%	15.7%	0.53	0.626	28.2%	-23.4%	0.086	0.97
Median RiskBmk		8.5%	16.0%	0.53	0.625	28.5%	-23.8%	0.086	0.99
S&P 500		8.6%	16.2%	0.53	0.625	28.9%	-24.1%	0.086	1

31 lentelė: Sav., visi, *Foresight*, 2015-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample	6.7%	12.0%	0.56	0.591	18.5%	-18.9%	0.087	0.77
Regime_bt_sample TWBmk	6.6%	10.5%	0.63	0.619	14.3%	-15.2%	0.086	0.77
Regime_bt_sample RiskBmk	7.5%	12.0%	0.63	0.618	16.3%	-17.4%	0.085	0.88
Regime_backtest	8.0%	11.8%	0.68	0.603	18.5%	-18.9%	0.105	0.76
Regime_backtest TWBmk	6.5%	10.3%	0.64	0.619	14.0%	-14.9%	0.086	0.76
Regime_backtest RiskBmk	7.4%	11.8%	0.63	0.618	16.1%	-17.2%	0.085	0.87
LR_L1	7.4%	13.4%	0.55	0.604	18.5%	-19.7%	0.075	0.98
LR_L1 TWBmk	8.3%	13.3%	0.62	0.616	18.1%	-19.4%	0.084	0.98
LR_L1 RiskBmk	8.4%	13.4%	0.62	0.616	18.3%	-19.6%	0.084	0.99
LR_L2	6.0%	13.3%	0.45	0.588	18.5%	-19.7%	0.062	0.96
LR_L2 TWBmk	8.1%	13.0%	0.62	0.616	17.8%	-19.0%	0.085	0.96
LR_L2 RiskBmk	8.3%	13.3%	0.62	0.616	18.2%	-19.5%	0.084	0.99
RF	8.2%	13.5%	0.61	0.614	18.5%	-19.7%	0.083	1.00
RF TWBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
RF RiskBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
GBM	8.1%	13.5%	0.60	0.612	18.5%	-19.7%	0.081	1.00
GBM TWBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
GBM RiskBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
XGBM	8.2%	13.5%	0.61	0.614	18.5%	-19.7%	0.083	1.00
XGBM TWBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
XGBM RiskBmk	8.4%	13.5%	0.62	0.616	18.4%	-19.7%	0.084	1.00
DNN	4.0%	12.7%	0.32	0.556	18.7%	-19.7%	0.046	0.87
DNN TWBmk	7.4%	11.8%	0.63	0.618	16.1%	-17.2%	0.085	0.87
DNN RiskBmk	8.0%	12.7%	0.63	0.617	17.4%	-18.6%	0.085	0.94
Median	7.5%	13.4%	0.56	0.605	18.5%	-19.7%	0.077	0.98
Median TWBmk	8.3%	13.3%	0.62	0.616	18.1%	-19.3%	0.084	0.98
Median RiskBmk	8.4%	13.4%	0.62	0.616	18.3%	-19.6%	0.084	0.99
S&P 500	8.4%	13.5%	0.62	0.616	18.5%	-19.7%	0.084	1

32 lentelė: Mėn., dalis, Lag, 2015-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample	8.0%	11.8%	0.68	0.603	18.5%	-18.9%	0.105	0.76
Regime_bt_sample TWBmk	6.5%	10.3%	0.64	0.619	14.0%	-14.9%	0.086	0.76
Regime_bt_sample RiskBmk	7.4%	11.8%	0.63	0.618	16.1%	-17.2%	0.085	0.87
Regime_backtest	15.4%	8.4%	1.83	0.892	6.4%	-10.4%	0.359	0.43
Regime_backtest TWBmk	3.8%	5.8%	0.65	0.623	7.9%	-8.5%	0.088	0.43
Regime_backtest RiskBmk	5.4%	8.4%	0.64	0.621	11.5%	-12.3%	0.087	0.62
LR_L1	7.5%	13.4%	0.55	0.605	18.5%	-19.7%	0.076	0.98
LR_L1 TWBmk	8.3%	13.3%	0.62	0.616	18.1%	-19.4%	0.084	0.98
LR_L1 RiskBmk	8.4%	13.4%	0.62	0.616	18.3%	-19.6%	0.084	0.99
LR_L2	12.6%	10.2%	1.23	0.875	13.3%	-18.7%	0.213	0.59
LR_L2 TWBmk	5.2%	8.0%	0.64	0.621	10.9%	-11.7%	0.087	0.59
LR_L2 RiskBmk	6.5%	10.2%	0.64	0.619	13.9%	-14.9%	0.086	0.76
RF	7.3%	13.4%	0.54	0.603	18.5%	-19.7%	0.075	0.98
RF TWBmk	8.3%	13.3%	0.62	0.616	18.1%	-19.4%	0.084	0.98
RF RiskBmk	8.4%	13.4%	0.62	0.616	18.3%	-19.6%	0.084	0.99
GBM	7.7%	13.5%	0.57	0.608	18.5%	-19.7%	0.078	0.98
GBM TWBmk	8.3%	13.3%	0.62	0.616	18.2%	-19.4%	0.084	0.98
GBM RiskBmk	8.4%	13.5%	0.62	0.616	18.3%	-19.6%	0.084	0.99
XGBM	7.3%	13.4%	0.54	0.603	18.5%	-19.7%	0.075	0.98
XGBM TWBmk	8.3%	13.3%	0.62	0.616	18.1%	-19.4%	0.084	0.98
XGBM RiskBmk	8.4%	13.4%	0.62	0.616	18.3%	-19.6%	0.084	0.99
DNN	7.3%	13.0%	0.56	0.603	18.5%	-19.7%	0.080	0.92
DNN TWBmk	7.8%	12.5%	0.63	0.617	17.0%	-18.2%	0.085	0.92
DNN RiskBmk	8.1%	13.0%	0.62	0.617	17.7%	-19.0%	0.085	0.96
Median	7.1%	13.4%	0.53	0.600	18.5%	-19.7%	0.072	0.98
Median TWBmk	8.3%	13.3%	0.62	0.616	18.1%	-19.3%	0.084	0.98
Median RiskBmk	8.4%	13.4%	0.62	0.616	18.3%	-19.6%	0.084	0.99
S&P 500	8.4%	13.5%	0.62	0.616	18.5%	-19.7%	0.084	1

33 lentelė: Mėn., dalis, Foresight, 2015-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample	5.4%	14.9%	0.37	0.581	28.9%	-19.2%	0.065	0.84
Regime_bt_sample TWBmk	7.4%	13.7%	0.54	0.628	24.8%	-20.3%	0.087	0.84
Regime_bt_sample RiskBmk	8.0%	14.9%	0.53	0.627	26.8%	-22.1%	0.087	0.92
Regime_backtest	5.2%	12.5%	0.41	0.574	20.1%	-17.3%	0.087	0.59
Regime_backtest TWBmk	5.3%	9.6%	0.56	0.633	17.9%	-14.3%	0.090	0.59
Regime_backtest RiskBmk	6.8%	12.5%	0.54	0.630	22.8%	-18.6%	0.088	0.77
LR_L1	7.3%	15.9%	0.46	0.607	28.9%	-23.2%	0.075	0.97
LR_L1 TWBmk	8.3%	15.7%	0.53	0.626	28.1%	-23.3%	0.086	0.97
LR_L1 RiskBmk	8.4%	15.9%	0.53	0.625	28.5%	-23.7%	0.086	0.98
LR_L2	6.6%	14.7%	0.45	0.583	28.9%	-19.8%	0.079	0.83
LR_L2 TWBmk	7.2%	13.4%	0.54	0.628	24.3%	-19.9%	0.088	0.83
LR_L2 RiskBmk	7.9%	14.7%	0.53	0.627	26.6%	-21.9%	0.087	0.91
RF	7.8%	16.0%	0.49	0.614	28.9%	-23.2%	0.081	0.97
RF TWBmk	8.3%	15.7%	0.53	0.626	28.1%	-23.4%	0.086	0.97
RF RiskBmk	8.5%	16.0%	0.53	0.625	28.5%	-23.7%	0.086	0.99
GBM	8.3%	16.0%	0.52	0.620	28.9%	-23.2%	0.085	0.98
GBM TWBmk	8.4%	15.9%	0.53	0.625	28.4%	-23.6%	0.086	0.98
GBM RiskBmk	8.5%	16.0%	0.53	0.625	28.6%	-23.8%	0.086	0.99
XGBM	8.0%	16.0%	0.50	0.615	28.9%	-23.2%	0.082	0.97
XGBM TWBmk	8.3%	15.7%	0.53	0.626	28.2%	-23.4%	0.086	0.97
XGBM RiskBmk	8.5%	16.0%	0.53	0.625	28.5%	-23.7%	0.086	0.99
DNN	2.9%	15.0%	0.19	0.536	28.9%	-23.2%	0.034	0.85
DNN TWBmk	7.4%	13.8%	0.54	0.628	25.0%	-20.5%	0.087	0.85
DNN RiskBmk	8.0%	15.0%	0.53	0.627	26.9%	-22.3%	0.086	0.92
Median	7.3%	15.9%	0.46	0.607	28.9%	-23.2%	0.075	0.97
Median TWBmk	8.3%	15.7%	0.53	0.626	28.1%	-23.3%	0.086	0.97
Median RiskBmk	8.4%	15.9%	0.53	0.625	28.5%	-23.7%	0.086	0.98
S&P 500	8.6%	16.2%	0.53	0.625	28.9%	-24.1%	0.086	1

34 lentelė: Sav., dalis, *Lag*, 2015-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_bt_sample	5.6%	14.9%	0.38	0.583	28.9%	-19.2%	0.067	0.84
Regime_bt_sample TWBmk	7.3%	13.6%	0.54	0.628	24.7%	-20.3%	0.087	0.84
Regime_bt_sample RiskBmk	7.9%	14.9%	0.53	0.627	26.8%	-22.1%	0.087	0.92
Regime_backtest	21.8%	11.5%	1.90	1.000	7.8%	-14.2%	0.421	0.52
Regime_backtest TWBmk	4.7%	8.4%	0.56	0.634	15.8%	-12.5%	0.091	0.52
Regime_backtest RiskBmk	6.3%	11.5%	0.55	0.631	21.1%	-17.0%	0.089	0.71
LR_L1	5.1%	15.4%	0.33	0.564	28.9%	-24.1%	0.057	0.90
LR_L1 TWBmk	7.8%	14.6%	0.54	0.627	26.3%	-21.7%	0.087	0.90
LR_L1 RiskBmk	8.2%	15.4%	0.53	0.626	27.6%	-22.9%	0.086	0.95
LR_L2	8.0%	14.0%	0.57	0.619	21.6%	-19.8%	0.107	0.75
LR_L2 TWBmk	6.6%	12.1%	0.55	0.630	22.3%	-18.1%	0.088	0.75
LR_L2 RiskBmk	7.5%	14.0%	0.54	0.628	25.4%	-20.9%	0.087	0.87
RF	8.0%	16.0%	0.50	0.616	28.9%	-23.2%	0.082	0.97
RF TWBmk	8.4%	15.8%	0.53	0.626	28.2%	-23.4%	0.086	0.97
RF RiskBmk	8.5%	16.0%	0.53	0.625	28.6%	-23.8%	0.086	0.99
GBM	7.7%	16.1%	0.48	0.612	28.9%	-24.1%	0.078	0.99
GBM TWBmk	8.5%	16.0%	0.53	0.625	28.5%	-23.8%	0.086	0.99
GBM RiskBmk	8.5%	16.1%	0.53	0.625	28.7%	-23.9%	0.086	0.99
XGBM	8.4%	16.0%	0.53	0.622	28.9%	-23.2%	0.086	0.98
XGBM TWBmk	8.4%	15.8%	0.53	0.625	28.3%	-23.6%	0.086	0.98
XGBM RiskBmk	8.5%	16.0%	0.53	0.625	28.6%	-23.8%	0.086	0.99
DNN	2.3%	14.8%	0.16	0.529	33.0%	-20.0%	0.028	0.83
DNN TWBmk	7.3%	13.5%	0.54	0.628	24.5%	-20.1%	0.087	0.83
DNN RiskBmk	7.9%	14.8%	0.53	0.627	26.7%	-22.0%	0.087	0.91
Median	5.1%	15.4%	0.33	0.564	28.9%	-24.1%	0.057	0.90
Median TWBmk	7.8%	14.6%	0.54	0.627	26.3%	-21.7%	0.087	0.90
Median RiskBmk	8.2%	15.4%	0.53	0.626	27.6%	-22.9%	0.086	0.95
S&P 500	8.6%	16.2%	0.53	0.625	28.9%	-24.1%	0.086	1

35 lentelė: Sav., dalis, *Foresight*, 2015-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_model	8.7%	11.3%	0.77	0.976	18.5%	-15.8%	0.158	0.55
Regime_model TWBmk	3.1%	8.5%	0.37	0.700	39.8%	-12.4%	0.057	0.55
Regime_model RiskBmk	4.0%	11.3%	0.35	0.693	50.1%	-16.5%	0.054	0.73
Regime_backtest	5.3%	12.0%	0.44	0.716	34.0%	-18.9%	0.088	0.61
Regime_backtest TWBmk	3.4%	9.4%	0.36	0.698	43.2%	-13.7%	0.056	0.61
Regime_backtest RiskBmk	4.2%	12.0%	0.35	0.691	52.6%	-17.6%	0.054	0.78
LR_L1	6.4%	12.4%	0.51	0.783	29.3%	-18.4%	0.098	0.65
LR_L1 TWBmk	3.6%	10.1%	0.36	0.696	45.9%	-14.7%	0.055	0.65
LR_L1 RiskBmk	4.3%	12.4%	0.35	0.690	53.9%	-18.2%	0.054	0.81
LR_L2	6.2%	12.6%	0.49	0.788	34.1%	-18.6%	0.093	0.67
LR_L2 TWBmk	3.7%	10.3%	0.36	0.695	46.6%	-15.0%	0.055	0.67
LR_L2 RiskBmk	4.4%	12.6%	0.35	0.689	54.4%	-18.4%	0.054	0.82
RF	6.0%	12.2%	0.49	0.751	33.7%	-18.4%	0.095	0.63
RF TWBmk	3.5%	9.7%	0.36	0.697	44.6%	-14.2%	0.056	0.63
RF RiskBmk	4.3%	12.2%	0.35	0.690	53.3%	-17.9%	0.054	0.79
GBM	5.1%	12.4%	0.41	0.709	42.5%	-19.2%	0.080	0.64
GBM TWBmk	3.6%	9.9%	0.36	0.696	45.3%	-14.5%	0.056	0.64
GBM RiskBmk	4.3%	12.4%	0.35	0.690	53.8%	-18.1%	0.054	0.80
XGBM	6.0%	12.2%	0.49	0.750	33.2%	-18.4%	0.095	0.63
XGBM TWBmk	3.5%	9.7%	0.36	0.697	44.6%	-14.2%	0.056	0.63
XGBM RiskBmk	4.3%	12.2%	0.35	0.690	53.3%	-17.9%	0.054	0.79
DNN	5.6%	12.1%	0.46	0.774	33.1%	-18.0%	0.090	0.62
DNN TWBmk	3.5%	9.5%	0.36	0.697	43.9%	-13.9%	0.056	0.62
DNN RiskBmk	4.2%	12.1%	0.35	0.691	52.9%	-17.7%	0.054	0.79
Median	5.6%	12.5%	0.45	0.733	34.2%	-18.8%	0.086	0.65
Median TWBmk	3.6%	10.1%	0.36	0.696	45.8%	-14.7%	0.055	0.65
Median RiskBmk	4.3%	12.5%	0.35	0.690	54.0%	-18.2%	0.054	0.81
Regime_switching	6.3%	11.6%	0.55	0.850	20.6%	-18.4%	0.138	0.46
Regime_switching_TW Bmk	2.8%	7.8%	0.36	0.688	34.8%	-11.6%	0.061	0.46
Regime_switching_RiskBmk	3.9%	11.6%	0.34	0.678	48.0%	-17.1%	0.058	0.68
S&P 500	5.1%	15.4%	0.33	0.682	62.9%	-22.5%	0.051	1

36 lentelė: Palyginimas su režimų pasikeitimo modeliu. Mėn, visi, *Lag*, 1986-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_model	8.6%	12.6%	0.69	0.905	28.9%	-18.8%	0.147	0.59
Regime_model TWBmk	3.4%	9.6%	0.35	0.740	42.9%	-15.0%	0.058	0.59
Regime_model RiskBmk	4.3%	12.6%	0.34	0.731	52.8%	-19.6%	0.056	0.76
Regime_backtest	4.3%	12.6%	0.34	0.704	41.2%	-19.5%	0.073	0.58
Regime_backtest TWBmk	3.4%	9.6%	0.35	0.740	42.8%	-15.0%	0.058	0.58
Regime_backtest RiskBmk	4.3%	12.6%	0.34	0.731	52.9%	-19.6%	0.056	0.77
LR_L1	6.0%	13.4%	0.45	0.769	31.1%	-20.0%	0.090	0.67
LR_L1 TWBmk	3.8%	11.0%	0.35	0.736	47.6%	-17.1%	0.057	0.67
LR_L1 RiskBmk	4.5%	13.4%	0.34	0.728	55.6%	-21.0%	0.055	0.82
LR_L2	5.5%	13.4%	0.41	0.753	35.8%	-20.7%	0.084	0.66
LR_L2 TWBmk	3.8%	10.9%	0.35	0.736	47.4%	-17.0%	0.057	0.66
LR_L2 RiskBmk	4.5%	13.4%	0.34	0.728	55.4%	-20.9%	0.055	0.81
RF	5.9%	13.8%	0.42	0.760	37.0%	-21.1%	0.083	0.71
RF TWBmk	4.0%	11.6%	0.35	0.734	49.9%	-18.2%	0.057	0.71
RF RiskBmk	4.6%	13.8%	0.34	0.727	56.8%	-21.6%	0.055	0.84
GBM	5.4%	14.0%	0.39	0.745	43.4%	-21.5%	0.075	0.73
GBM TWBmk	4.1%	12.0%	0.34	0.733	50.9%	-18.7%	0.057	0.73
GBM RiskBmk	4.7%	14.0%	0.33	0.726	57.3%	-21.9%	0.055	0.85
XGBM	5.9%	13.8%	0.43	0.764	35.6%	-21.1%	0.084	0.70
XGBM TWBmk	4.0%	11.6%	0.35	0.734	49.7%	-18.1%	0.057	0.70
XGBM RiskBmk	4.6%	13.8%	0.34	0.727	56.7%	-21.5%	0.055	0.84
DNN	5.9%	13.5%	0.43	0.774	35.2%	-20.2%	0.087	0.68
DNN TWBmk	3.9%	11.1%	0.35	0.735	48.0%	-17.3%	0.057	0.68
DNN RiskBmk	4.5%	13.5%	0.34	0.728	55.8%	-21.1%	0.055	0.82
Median	5.9%	13.9%	0.43	0.763	40.7%	-21.3%	0.083	0.72
Median TWBmk	4.1%	11.8%	0.34	0.733	50.3%	-18.4%	0.057	0.72
Median RiskBmk	4.7%	13.9%	0.33	0.727	57.0%	-21.7%	0.055	0.85
Regime_switching	6.3%	11.6%	0.55	0.850	20.6%	-18.4%	0.138	0.46
Regime_switching_TW Bmk	2.8%	7.8%	0.36	0.688	34.8%	-11.6%	0.061	0.46
Regime_switching_RiskBmk	3.9%	11.6%	0.34	0.678	48.0%	-17.1%	0.058	0.68
S&P 500	5.3%	16.4%	0.32	0.719	64.0%	-25.6%	0.053	1

37 lentelė: Palyginimas su režimų pasikeitimo modeliu. Sav., visi, *Lag*, 1986-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_model	8.1%	11.5%	0.71	0.945	24.8%	-16.9%	0.144	0.56
Regime_model TWBmk	3.2%	8.7%	0.37	0.700	40.5%	-12.6%	0.056	0.56
Regime_model RiskBmk	4.0%	11.5%	0.35	0.692	50.7%	-16.7%	0.054	0.74
Regime_backtest	5.3%	12.0%	0.44	0.716	34.0%	-18.9%	0.088	0.61
Regime_backtest TWBmk	3.4%	9.4%	0.36	0.698	43.2%	-13.7%	0.056	0.61
Regime_backtest RiskBmk	4.2%	12.0%	0.35	0.691	52.6%	-17.6%	0.054	0.78
LR_L1	6.6%	12.4%	0.53	0.795	39.0%	-18.0%	0.102	0.65
LR_L1 TWBmk	3.6%	10.0%	0.36	0.696	45.7%	-14.6%	0.055	0.65
LR_L1 RiskBmk	4.3%	12.4%	0.35	0.690	53.8%	-18.1%	0.054	0.80
LR_L2	5.1%	12.3%	0.41	0.731	46.0%	-18.7%	0.080	0.64
LR_L2 TWBmk	3.5%	9.8%	0.36	0.697	44.9%	-14.3%	0.056	0.64
LR_L2 RiskBmk	4.3%	12.3%	0.35	0.690	53.6%	-18.0%	0.054	0.80
RF	5.9%	12.3%	0.49	0.748	30.9%	-18.7%	0.094	0.63
RF TWBmk	3.5%	9.8%	0.36	0.697	44.7%	-14.3%	0.056	0.63
RF RiskBmk	4.3%	12.3%	0.35	0.690	53.3%	-17.9%	0.054	0.79
GBM	5.8%	12.4%	0.47	0.743	30.9%	-18.7%	0.090	0.65
GBM TWBmk	3.6%	10.0%	0.36	0.696	45.4%	-14.6%	0.055	0.65
GBM RiskBmk	4.3%	12.4%	0.35	0.690	53.8%	-18.1%	0.054	0.80
XGBM	5.8%	12.3%	0.47	0.745	30.9%	-18.7%	0.091	0.64
XGBM TWBmk	3.5%	9.8%	0.36	0.697	44.9%	-14.4%	0.056	0.64
XGBM RiskBmk	4.3%	12.3%	0.35	0.690	53.5%	-18.0%	0.054	0.80
DNN	4.2%	11.9%	0.35	0.807	34.8%	-18.8%	0.071	0.60
DNN TWBmk	3.3%	9.2%	0.36	0.698	42.6%	-13.4%	0.056	0.60
DNN RiskBmk	4.2%	11.9%	0.35	0.691	52.3%	-17.4%	0.054	0.77
Median	6.4%	12.6%	0.50	0.781	30.9%	-18.7%	0.095	0.67
Median TWBmk	3.7%	10.4%	0.36	0.695	46.9%	-15.1%	0.055	0.67
Median RiskBmk	4.4%	12.6%	0.35	0.689	54.5%	-18.4%	0.053	0.82
Regime_switching	6.3%	11.6%	0.55	0.850	20.6%	-18.4%	0.138	0.46
Regime_switching_TW Bmk	2.8%	7.8%	0.36	0.688	34.8%	-11.6%	0.061	0.46
Regime_switching_RiskBmk	3.9%	11.6%	0.34	0.678	48.0%	-17.1%	0.058	0.68
S&P 500	5.1%	15.4%	0.33	0.682	62.9%	-22.5%	0.051	1

38 lentelė: Palyginimas su režimų pasikeitimo modeliu. Mėn, dalis, *Lag*, 1986-2022

Strategija \ Rodikliai	Mean	Stdev	SharpeR	ProbSR	MaxDD	VaR95	TreynorR	Beta
Regime_model	8.6%	12.6%	0.69	0.905	28.9%	-18.8%	0.147	0.59
Regime_model TWBmk	3.4%	9.6%	0.35	0.740	42.9%	-15.0%	0.058	0.59
Regime_model RiskBmk	4.3%	12.6%	0.34	0.731	52.8%	-19.6%	0.056	0.76
Regime_backtest	4.3%	12.6%	0.34	0.704	41.2%	-19.5%	0.073	0.58
Regime_backtest TWBmk	3.4%	9.6%	0.35	0.740	42.8%	-15.0%	0.058	0.58
Regime_backtest RiskBmk	4.3%	12.6%	0.34	0.731	52.9%	-19.6%	0.056	0.77
LR_L1	7.3%	13.2%	0.56	0.833	28.9%	-19.2%	0.114	0.65
LR_L1 TWBmk	3.7%	10.6%	0.35	0.737	46.4%	-16.6%	0.058	0.65
LR_L1 RiskBmk	4.5%	13.2%	0.34	0.729	54.8%	-20.6%	0.056	0.80
LR_L2	6.5%	13.2%	0.49	0.783	31.6%	-19.9%	0.100	0.65
LR_L2 TWBmk	3.7%	10.6%	0.35	0.737	46.4%	-16.6%	0.058	0.65
LR_L2 RiskBmk	4.5%	13.2%	0.34	0.729	54.9%	-20.6%	0.056	0.80
RF	6.4%	13.6%	0.47	0.778	34.6%	-20.0%	0.093	0.69
RF TWBmk	3.9%	11.3%	0.35	0.735	48.6%	-17.6%	0.057	0.69
RF RiskBmk	4.6%	13.6%	0.34	0.728	56.1%	-21.2%	0.055	0.83
GBM	5.3%	13.8%	0.39	0.739	37.0%	-21.2%	0.075	0.71
GBM TWBmk	4.0%	11.7%	0.35	0.734	49.9%	-18.2%	0.057	0.71
GBM RiskBmk	4.6%	13.8%	0.34	0.727	56.8%	-21.6%	0.055	0.84
XGBM	5.9%	13.6%	0.43	0.762	35.0%	-20.1%	0.086	0.69
XGBM TWBmk	3.9%	11.3%	0.35	0.735	48.6%	-17.6%	0.057	0.69
XGBM RiskBmk	4.6%	13.6%	0.34	0.728	56.1%	-21.2%	0.055	0.83
DNN	5.0%	12.9%	0.39	0.737	44.6%	-20.0%	0.082	0.62
DNN TWBmk	3.6%	10.1%	0.35	0.738	44.7%	-15.8%	0.058	0.62
DNN RiskBmk	4.4%	12.9%	0.34	0.730	54.0%	-20.1%	0.056	0.79
Median	6.1%	13.7%	0.45	0.771	34.6%	-20.5%	0.088	0.70
Median TWBmk	4.0%	11.5%	0.35	0.734	49.4%	-17.9%	0.057	0.70
Median RiskBmk	4.6%	13.7%	0.34	0.727	56.5%	-21.4%	0.055	0.84
Regime_switching	6.3%	11.6%	0.55	0.850	20.6%	-18.4%	0.138	0.46
Regime_switching_TW Bmk	2.8%	7.8%	0.36	0.688	34.8%	-11.6%	0.061	0.46
Regime_switching_RiskBmk	3.9%	11.6%	0.34	0.678	48.0%	-17.1%	0.058	0.68
S&P 500	5.3%	16.4%	0.32	0.719	64.0%	-25.6%	0.053	1

39 lentelė: Palyginimas su režimų pasikeitimo modeliu. Sav., dalis, *Lag*, 1986-2022