



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Daugiamačių laiko eilučių prognozavimas mažmeninėje prekyboje

Baigiamasis magistro studijų projektas

Tauras Gaulia

Projekto autorius

Prof. dr. Evaldas Vaičiukynas

Vadovas

Doc. dr. Aistė Dovalienė

Vadovė

Kaunas, 2025



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Daugiamatį laiko eilučių prognozavimas mažmeninėje prekyboje

Baigiamasis magistro studijų projektas

Didžiųjų verslo duomenų analitika (6213AX001)

Tauras Gaulia

Projekto autorius

Prof. dr. Evaldas Vaičiukynas

Vadovas

Doc. dr. Aistė Dovalienė

Vadovė

Doc. dr. Mindaugas Kavaliauskas

Recenzentas

Prof. dr. Mantas Vilkas

Recenzentas

Kaunas, 2025



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Tauras Gaulia

Daugiamačių laiko eilučių prognozavimas mažmeninėje prekyboje

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;
2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąrašė;
3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);
4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Tauras Gaulia

Patvirtinta elektroniniu būdu

Gaulia, Tauras. Daugiamačių laiko eilučių prognozavimas mažmeninėje prekyboje. Magistro studijų baigiamasis projektas / vadovas prof. dr. Evaldas Vaičiukynas, vadovė doc. dr. Aistė Dovalienė; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

Studijų kryptis ir sritis (studijų krypties grupė): Taikomoji matematika (Matematikos mokslai).

Reikšminiai žodžiai: laiko eilučių analizė, hierarchinis prognozavimas, *ARIMAX*, išoriniai regresoriai, *M5* duomenys, *WRMSSE*.

Kaunas, 2025. 54 p.

Santrauka

Šiame magistro baigiamajame projekte analizuojama daugiamačių laiko eilučių prognozavimo problema, pasitelkiant „*M5 Forecasting*“ konkurso duomenų rinkinį, kuriame pateikiami daugiau nei 30 tūkstančių prekių pardavimų įrašai 10 skirtingų „*Walmart*“ parduotuvių, esančių JAV trijose valstijose. Šie duomenys yra struktūrizuoti hierarchiškai – nuo visų parduotuvių bendros apyvartos iki atskiros prekės kiekvienoje parduotuvėje – iš viso formuojant 12 hierarchijos lygių. Todėl šis rinkinys itin tinkamas hierarchinių prognozavimo metodų tikslumo ir efektyvumo analizei.

Šio tyrimo aktualumą lemia vis spartėjantis duomenų kiekio augimas mažmeninėje prekyboje ir tuo pačiu didėjantis poreikis automatizuotai bei tiksliai prognozuoti skirtingo lygmens paklausą. Tikslus prognozavimas leidžia priimti pagrįstus sprendimus dėl atsargų planavimo, tiekimo grandinės valdymo ir reklaminių kampanijų planavimo. Tinkamai parinkti prognozavimo modeliai gali tiesiogiai prisidėti prie įmonės veiklos efektyvumo ir pelno augimo.

Darbo tikslas – ištirti skirtingus daugiamačių laiko eilučių prognozavimo metodus, jų taikymo galimybes ir ribojimus bei įvertinti jų tikslumą skirtinguose hierarchijos lygiuose. Tyrimo metu buvo pritaikyti keliolika etaloninių (*benchmark*) modelių: nuo paprastų metodų, tokių kaip *Naive* ar *Seasonal Naive*, iki sudėtingesnių statistinių (*ETS*, *ARIMA*), mašininio mokymosi (*Random Forest*, *MLP*) ir giluminio mokymosi metodų. Visi modeliai buvo įvertinti naudojant *WRMSSE* metriką, kuri atsižvelgia į kiekvienos prognozės svarbą pagal atitinkamos serijos apyvartą.

Tyrimo metu taip pat sukurtas pažangus *ARIMAX* modelis su 7 išoriniais regresoriais, kuriuose integruota papildoma informacija apie šventines dienas, departamento, kategorijos, valstijos, parduotuvės bei valstijos–kategorijos pardavimus, taip pat kainų dinamika. Tokie regresoriai buvo atrinkti remiantis duomenų hierarchija ir ekonomine logika, siekiant kuo geriau išnaudoti turimą informaciją ir sumažinti prognozės paklaidą. Siekiant tikslumo ir lankstumo, pasirinkta *Middle-out* prognozavimo strategija – prognozės buvo atliekamos viduriniame, parduotuvės–departamento lygyje, iš kurio jos buvo agreguojamos į aukštesnius lygius ir proporcingai paskirstytos žemyn.

ARIMAX modelis parduotuvės–departamento lygyje pasiekė vieną geriausių rezultatų tarp visų testuotų metodų – $WRMSSE = 0.5404$. Didesniuose lygiuose šis modelis taip pat konkuravo su geriausiais etaloniniais metodais. Vis dėlto, žemesniuose (produktų ar produktų–parduotuvės lygmenys) lygiuose *WRMSSE* reikšmės viršijo 1, atskleidžiamos modeliavimo ribotumas labai diskrečių laiko eilučių atveju. Tyrimas patvirtino, kad suderinta regresorių atranka ir prognozavimo strategija gali ženkliai pagerinti rezultatų tikslumą bei būti naudinga sprendžiant realaus pasaulio prognozavimo uždavinius mažmeninėje prekyboje.

Gaulia, Tauras. Multivariate Time Series Forecasting for Retail Sales. Master's Final Degree Project / supervisor prof. dr. Evaldas Vaičiukynas, supervisor assoc. prof. dr. Aistė Dovalienė; Faculty of Mathematics and Natural Sciences, Kaunas University of Technology.

Study field and area (study field group): Applied Mathematics (Mathematical Sciences).

Keywords: time series analysis, hierarchical forecasting, *ARIMAX*, exogenous regressors, *WRMSSE*, *M5* dataset.

Kaunas, 2025. 54 pages.

Summary

This master's thesis analyzes the problem of multivariate time series forecasting using the *M5 Forecasting* competition dataset, which contains over 30,000 product sales records from 10 different *Walmart* stores located in three U.S. states. The data is hierarchically structured – from total revenue across all stores to the sales of individual products in each store – forming a 12-level hierarchy. This makes the dataset highly suitable for evaluating the accuracy and efficiency of hierarchical forecasting methods.

The relevance of this research stems from the rapid growth of data volumes in the retail sector and the increasing demand for automated and accurate demand forecasting across multiple levels. Accurate forecasts enable informed decisions on inventory planning, supply chain management, and promotional campaign scheduling. Appropriately chosen forecasting models can directly contribute to improved operational efficiency and increased profitability.

The main objective of this thesis is to investigate various multivariate time series forecasting methods, assess their practical applicability and limitations, and evaluate their accuracy at different hierarchy levels. The study applied a wide range of benchmark models – from simple methods like *Naive* and *Seasonal Naive* to more complex statistical (*ETS*, *ARIMA*), machine learning (*Random Forest*, *MLP*), and deep learning approaches. All models were evaluated using the *WRMSSE* metric, which weights each forecast according to the importance of the corresponding time series in terms of sales value.

In addition, an advanced *ARIMAX* model was developed using seven external regressors that incorporated supplementary information on holidays, department-level sales, category-level sales, state- and store-level sales, state–category sales, and price dynamics. These regressors were selected based on the data hierarchy and economic rationale to better leverage available information and reduce forecast error. To ensure both accuracy and flexibility, a *Middle-out* forecasting strategy was employed – forecasts were generated at the intermediate store–department level, then aggregated upward and proportionally disaggregated to lower levels.

At the store–department level, the *ARIMAX* model achieved one of the best results among all tested methods (*WRMSSE* = 0.5404). It also performed competitively at higher levels. However, at lower levels (such as individual products or product–store combinations), *WRMSSE* exceeded 1, highlighting the modeling challenges associated with highly sparse and discrete time series. The results confirm that well-designed external regressors and a thoughtfully chosen forecasting strategy can significantly improve forecast accuracy and provide practical value in real-world retail forecasting scenarios.

Turinys

Lentelių sąrašas	8
Paveikslų sąrašas	9
Įvadas.....	10
1. Literatūros apžvalga	11
1.1. Tyrimų tikslai ir aktualumas.....	12
1.2. Naudoti duomenų rinkiniai.....	13
1.3. Naudoti metodai	17
1.3.1. Hierarchinės prognozavimo strategijos: <i>Bottom-up</i> , <i>Top-down</i> ir <i>Middle-out</i>	18
1.3.2. <i>TsBench</i> : Standartizuota laiko eilučių prognozavimo metodų vertinimo aplinka.....	20
1.3.3. Prognozavimo metodų palyginimo aspektai	21
1.3.4. Išoriniai kintamieji laiko eilučių prognozavimo modeliuose	21
1.4. Pagrindiniai rezultatai ir išvalgos	22
1.4.1. Giluminio mokymosi metodai.....	23
1.4.2. Tikimybiniai modeliai	24
1.4.3. Autoregresiniai modeliai	24
1.4.4. Hibridiniai metodai.....	24
1.4.5. Srautų normalizavimo metodai.....	25
1.5. Baigiamojo magistro projekto temos ir uždavinių pagrindimas.....	25
2. Tyrimo metodai	26
2.1. Naudoti prognozavimo metodai	26
2.1.1. Etaloniniai varžybų (<i>benchmark</i>) modeliai	26
2.1.2. <i>ARIMAX</i> su išoriniais regresoriais.....	28
2.2. Duomenų paruošimo strategijos	31
2.2.1. Šventinių dienų identifikavimas	31
2.2.2. Laiko eilučių transformacijos	31
2.2.3. Išorinių regresorių formavimas	32
2.2.4. Laiko eilučių stacionarumo įvertinimas	32
2.2.5. Papildomos pastabos	33
2.3. Modelių vertinimo metrikos	33
2.3.1. Root Mean Squared Error (<i>RMSE</i>).....	33
2.3.2. Median Absolute Error (<i>MDAE</i>)	33
2.3.3. Mean Absolute Scaled Error (<i>MASE</i>).....	34
2.3.4. Symmetric Mean Absolute Percentage Error (<i>SMAPE</i>).....	34
2.3.5. Mean Arctangent Absolute Percentage Error (<i>MAAPE</i>)	34
2.3.6. Weighted Root Mean Squared Scaled Error (<i>WRMSSE</i>)	34
2.3.7. Modelių reitingavimo strategija	35
2.4. Skaičiavimų aplinka ir techniniai sprendimai	36
3. Tyrimo rezultatai.....	37
3.1. <i>M5</i> duomenų rinkinio aprašymas	37
3.1.1. Duomenų failų struktūra.....	37
3.1.2. Duomenų hierarchija	38
3.1.3. Tyrime naudotų duomenų specifikacija	38
3.1.4. Duomenų savybės.....	39
3.1.5. Duomenų išskirtinumai	39

3.2.	Pirminė duomenų analizė	39
3.2.1.	Pardavimų dinamika laike	39
3.2.2.	Pardavimų pasiskirstymas pagal savaitės dieną	40
3.2.3.	Šventinių dienų įtaka	40
3.2.4.	Pardavimai pagal departamentus	41
3.3.	Etaloninių modelių rezultatų analizė	42
3.4.	Modelių taikymo praktikoje ir resursų poreikio analizė.....	44
3.4.1.	Modelių palyginimas pagal treniravimo sudėtingumą ir resursus.....	44
3.5.	Sukurto <i>ARIMAX</i> modelio aprašymas ir taikymo principas	45
3.6.	Etaloninių ir <i>ARIMAX</i> modelių palyginimas	47
3.7.	Hierarchinių prognozių nauda mažmeninėje prekyboje	48
3.8.	Rezultatų interpretacija ir tyrimo ribojimai	50
	Išvados	52
	Literatūros sąrašas	53
	Priedai.....	55
1	priedas. Likusiems išoriniams regresoriams atrinktų modelių klaidų lentelė	55
2	priedas. Alternatyvus regresorių prognozės agregavimo metodas	55

Lentelių sąrašas

1 lentelė. Populiariausių duomenų rinkinių savybės.....	15
2 lentelė. Hierarchinių prognozavimo strategijų palyginimas	19
3 lentelė. Prognozavimo metodų palyginimas	22
4 lentelė. Modelių palyginimo lentelė.....	28
5 lentelė. Trijų pagrindinių <i>M5</i> rinkinio <i>CSV</i> failų struktūra.....	37
6 lentelė. <i>Benchmark</i> modelių <i>WRMSSE</i> rezultatai	43
7 lentelė. Išoriniams regresoriams atrinktų modelių klaidos	46
8 lentelė. <i>ARIMAX</i> ir <i>Bottom-up Exponential Smoothing</i> modelių lyginimas.....	47
9 lentelė. Likusiems išoriniams regresoriams atrinktų modelių klaidos.....	55

Paveikslų sąrašas

1 pav. Kiekvienos savaitės pardavimų grafikas	39
2 pav. Vidutinių kategorijų pardavimų pagal savaitės dieną grafikas	40
3 pav. Vidutinių kategorijų pardavimų pasiskirstymo pagal šventes grafikas.....	41
4 pav. Departamentų pardavimų grafikas	42
5 pav. Modelių vidutinio treniravimo laiko vienai laiko eilutei palyginimo grafikas	45
6 pav. <i>ARIMAX</i> ir <i>Bottom-up Exponential Smoothing</i> modelių prognozavimo palyginimo grafikas	48
7 pav. Kaupiamieji prognozavimo kaštai pagal <i>ARIMAX</i> ir <i>ETS_bu</i> modelius (vienai parduotuvės–departamento serijai)	50

Ivadas

Daugiamačių laiko eilučių prognozavimas yra viena iš sudėtingiausių ir aktualiausių duomenų analizės sričių, kuri randa pritaikymą tokiose srityse kaip finansai, energetika, logistika ir mažmeninė prekyba. Nepaisant reikšmingos pažangos, šioje srityje išlieka daug neišspręstų klausimų, susijusių su skirtingų metodų efektyvumo palyginimu, jų pritaikymo specifika ir iššūkiais, kylančiais dėl sudėtingų duomenų struktūrų. Populiarėjant giluminio mokymosi metodams, tokiems kaip *DeepGLO*, *LSTNet*, *DeepAR* ir *N-BEATS*, kyla diskusijų apie jų pranašumus prieš paprastesnius statistinius ar tradicinio mašininio mokymosi metodus. Tačiau nėra aišku, kurie iš šių metodų yra tinkamiausi įvairioms daugiamačių laiko eilučių prognozavimo užduotims spręsti.

Viena iš pagrindinių problemų – pasirinkimas tarp universalios, visoms eilutėms skirtos modelio, ir individualiai kiekvienai eilei apmokytų daugiamatės prognozės modelių. Nors individualiai apmokyti modeliai teoriškai gali pasiūlyti didesnę tikslumą, praktinis jų taikymas dažnai ribojamas augančių skaičiavimo resursų poreikių. Taip pat svarbu apsispręsti, kaip modeliai turėtų būti pritaikomi hierarchinėms laiko eilutėms. Ar efektyviau prognozuoti žemiausio lygio eilutes ir iš jų rekonstruoti aukštesnio agreguotumo lygmenis, ar atvirkščiai – pradėti nuo aukštesnių lygmenų ir juos suskaidyti? Šiame kontekste aktualus tampa suderinimo (angl. forecast reconciliation) uždavinys, kai pradinės prognozės koreguojamos taip, kad jos atitiktų hierarchijos struktūrą.

Ižvalgos iš *M5* varžybų rodo, kad nėra vieno universalios metodo, kuris užtikrintų maksimalų tikslumą visais hierarchijos lygmenimis. Tai pabrėžia, kad skirtingiems duomenų lygmenims būtina taikyti specifinius metodus. Be to, diskusijos apie giluminio mokymosi metodų efektyvumą, lyginant su paprastesniais *GBM*-tipo mašininio mokymosi algoritmais, rodo, jog giluminio mokymosi privalumai dar nėra iki galo įrodyti. Ši neapibrėžtis ir technologijų įvairovė skatina ieškoti naujų sprendimų, kurie leistų užtikrinti aukštą prognozių tikslumą ir praktišką jų pritaikymą.

Šio **magistro projekto tikslas** – daugiamačių laiko eilučių prognozavimo metodų tinkamumo vertinimas, atsižvelgiant į jų efektyvumą, tikslumą ir pritaikomumą skirtingoms užduotims bei hierarchijos lygmenims.

Baigiamojo projekto uždaviniai:

1. Apžvelgti ir įvertinti naujausius daugiamačių laiko eilučių prognozavimo metodus, įskaitant giluminio mokymosi, tradicinio mašininio mokymosi ir statistinius modelius.
2. Eksperimentiniu būdu palyginti šių metodų prognozavimo tikslumą ir efektyvumą pasirinktame duomenų rinkinyje.
3. Išanalizuoti skirtingas hierarchinių laiko eilučių prognozavimo strategijas (*Bottom-up*, *Top-down*, *Middle-out*) ir įvertinti jų taikymo poveikį prognozių tikslumui skirtinguose hierarchijos lygiuose, atsižvelgiant į prognozavimo lygmens pasirinkimo svarbą bendrai modelio veiksmingumui.
4. Ištirti daugiamačių laiko eilučių prognozavimo metodų efektyvumą didelio masto duomenų rinkiniuose, atsižvelgiant į skaičiavimo resursų poreikį ir modelių pritaikymo galimybes praktinėms užduotims.
5. Pateikti rekomendacijas dėl daugiamačių laiko eilučių prognozavimo metodų pasirinkimo priklausomai nuo taikymo srities ir duomenų struktūros.

1. Literatūros apžvalga

Planavimas yra vienas pagrindinių organizacijų veiklos procesų, kuris ne tik leidžia optimaliai panaudoti turimus išteklius, bet ir padeda užtikrinti ilgalaikį organizacijos stabilumą. Efektyvus planavimas suteikia galimybę sumažinti veiklos kaštus, padidinti pelningumą ir geriau valdyti rizikas, o tai ypač svarbu sparčiai besikeičiant ekonominei aplinkai. Ekonomikos sektoriuje planavimas yra nepakeičiama priemonė, leidžianti priimti strateginius sprendimus dėl rinkos paklausos prognozavimo, atsargų valdymo, gamybos apimčių planavimo bei transportavimo procesų optimizavimo. Šie procesai prisideda prie konkurencinio pranašumo stiprinimo, nes užtikrina veiklos lankstumą ir galimybę greitai prisitaikyti prie rinkos pokyčių.

Ekonomikos planavimas yra reikšmingas ne tik įmonėms, bet ir platesnei makroekonominiai aplinkai. Nacionaliniai planavimo procesai padeda subalansuoti biudžetą, nustatyti prioritetus infrastruktūrai vystyti ir užtikrinti stabilų ekonominį augimą. Pavyzdžiui, tinkamai suplanuotos investicijos į viešąją infrastruktūrą, tokią kaip kelių tinklai, energetikos sistemos ar informacinių technologijų plėtra, gali skatinti ilgalaikį produktyvumą ir darbo vietų kūrimą. Planavimas taip pat leidžia valdyti ekonomines krizes, pvz., didėjant infliacijai ar kylant nedarbo lygiui, vyriausybės gali įgyvendinti fiskalines ar monetarines priemones, pagrįstas kruopščia situacijos analize.

Tuo tarpu privataus sektoriaus organizacijose planavimas apima gamybos procesų, tiekimo grandinės, personalo valdymo ir finansų planavimo sprendimus. Vienas svarbiausių aspektų yra rinkos paklausos prognozė, kuri leidžia įmonėms tiksliai nustatyti reikalingus gamybos pajėgumus ir efektyviai valdyti atsargas. Pavyzdžiui, mažmeninės prekybos sektoriuje planavimas padeda išvengti situacijų, kai atsargose trūksta populiarių prekių arba, atvirkščiai, susidaro perteklinis prekių kiekis, sukeliantis nuostolius.

Be to, planavimas yra nepakeičiamas siekiant optimizuoti logistikos procesus. Efektyvus tiekimo grandinės valdymas padeda sumažinti transportavimo kaštus, sumažinti pristatymo laiką ir pagerinti klientų aptarnavimo kokybę. Tarptautinėje prekyboje tinkamai suplanuoti logistikos maršrutai leidžia sumažinti tranzito laiką ir transportavimo riziką, o tai ypač svarbu, kai prekiaujama jautriomis ar greitai gendančiomis prekėmis.

Taip pat svarbu paminėti, kad planavimas šiandien vis dažniau remiasi duomenų analitika ir prognozavimo modeliais. Šiuolaikinėse technologijose, tokiose kaip dirbtinis intelektas ar didžiųjų duomenų analitika, slypi milžiniškas potencialas pagerinti planavimo procesų tikslumą. Pavyzdžiui, finansų sektoriuje įmonės naudoja analitinius modelius, leidžiančius prognozuoti valiutų kursų pokyčius, palūkanų normas ar kitus makroekonominis rodiklius. Tuo tarpu energetikos sektoriuje planavimo priemonės padeda optimizuoti energijos paskirstymą ir sumažinti perprodukcijos nuostolius.

Planavimas tampa ypač aktualus greitai kintančioje ekonominėje aplinkoje, kurioje reikia priimti operatyvius ir informuotus sprendimus. Pavyzdžiui, COVID-19 pandemijos metu daugelis įmonių ir vyriausybių susidūrė su precedento neturinčiomis planavimo problemomis, tokiomis kaip sveikatos priežiūros sistemos apkrovos valdymas, skiepijimo programų logistikos planavimas ar ekonomikos atsigavimo strategijos sudarymas. Tokios situacijos parodo, kaip svarbu turėti lanksčius planavimo modelius, kurie galėtų adaptuotis prie netikėtų aplinkybių.

COVID-19 pandemija atskleidė planavimo svarbą tiek nacionaliniu, tiek tarptautiniu mastu. Pandemijos metu daugelis šalių susidūrė su ekonominiais iššūkiais, kuriems spręsti reikėjo efektyvaus planavimo ir koordinavimo. Pavyzdžiui, Europos Audito Rūmai¹ pabrėžė suderinto ekonominio atsako į COVID-19 protrūkį svarbą, siekiant užtikrinti likvidumą ir stabilumą rinkose.

Apibendrinant galima teigti, kad planavimas yra kompleksinis procesas, apimantis daugelį ekonominės veiklos aspektų – nuo strateginio valdymo iki kasdienių operatyvių sprendimų. Siekiant užtikrinti efektyvų planavimą, svarbu ne tik analizuoti esamus duomenis, bet ir pasitelkti modernius prognozavimo metodus, kurie leistų numatyti būsimus pokyčius bei pasiruošti galimiems iššūkiams.

1.1. Tyrimų tikslai ir aktualumas

Planavimas yra esminis veiksnys, leidžiantis efektyviai paskirstyti išteklius, numatyti galimus iššūkius ir priimti pagrįstus sprendimus. Šiuolaikiniame dinamiškame pasaulyje, kuriame dauguma procesų yra sudėtingi ir tarpusavyje susiję, planavimas vis dažniau remiasi duomenų analizės metodais, tokiais kaip daugiamatės laiko eilutės. Šie metodai suteikia galimybę ne tik suprasti įvairių veiksmų sąveiką per laiką, bet ir tiksliau prognozuoti jų raidą ateityje, sudarant pagrindą geresniems planavimo sprendimams.

Daugiamatės laiko eilutės yra svarbios sprendžiant sudėtingas realaus pasaulio problemas, kuriose skirtingi kintamieji yra tarpusavyje susiję ir daro įtaką vieni kitiems per laiką. Šie metodai plačiai taikomi įvairiose srityse, įskaitant ekonomiką, kur jie naudojami finansinių rodiklių prognozėms; energetiką, siekiant numatyti elektros paklausos ar gamybos dinamiką; mediciną, kurioje jie padeda analizuoti paciento sveikatos rodiklių kitimą; ir klimato mokslus, kur taikomi orų, klimato pokyčiams ar ekosistemoms modeliuoti. Pagrindinis šių metodų tikslas – išnaudoti turimus duomenis, siekiant sukurti tikslesnius prognozavimo modelius, kurie leistų priimti efektyvesnius sprendimus.

Apžvelgiant mokslinius straipsnius, kuriuose naudojamos daugiamatės laiko eilutės, pastebėta, jog visi jie orientuoti į daugiamatį laiko eilučių prognozavimo metodų tobulinimą, siekiant spręsti realaus pasaulio problemas, susijusias su kompleksiniais ir tarpusavyje sąveikaujančiais kintamaisiais. Šie tyrimai nagrinėja įvairius metodus, tokius kaip neuroninių tinklų taikymas, hibridinių modelių kūrimas bei sudėtingų sąveikų tarp kintamųjų identifikavimas. Kiekvienas straipsnis pateikia unikalių įžvalgų – nuo efektyvesnių algoritmų pasiūlymų iki naujų požiūrių į duomenų normalizavimą ir reprezentavimą. Šios įžvalgos ne tik prisideda prie gilesnio šios srities supratimo, bet ir padeda plėtoti praktinius sprendimus, kurie gali būti tiesiogiai pritaikomi sprendžiant aktualias problemas, tokias kaip finansų rinkų stabilumo analizė, sveikatos priežiūros sistemų optimizavimas ar tvaraus energijos vartojimo užtikrinimas.

Vienas iš pažangių sprendimų, skirtų daugiamatėms laiko eilutėms prognozuoti, yra hibridinių modelių kūrimas, siekiant išnaudoti tiek globalius duomenų dėsningumus, tiek lokalias savybes. Tokie modeliai leidžia pasiekti tikslesnes prognozes, ypač sprendžiant kompleksines realaus pasaulio problemas, kur duomenų mastas ir įvairovė yra itin dideli.

¹ Europos Audito Rūmai. (2020). Risks, challenges and opportunities in the EU's economic policy response to the COVID 19 crisis.

https://www.eca.europa.eu/lists/ecadocuments/rw20_06/rw_economic_response_to_covid19_lt.pdf?

Seno, Yu ir Dhillion'o tyrime [1] siūlomas hibridinis modelis *DeepGLO*, kuris derina globalų matricių faktorizavimą su lokaliu laiko eilučių kalibravimu. Autoriai pabrėžia, kad svarbu išnaudoti globalius duomenų dėsninumus ir juos derinti su lokaliai reikšmingomis savybėmis. Pritaikymas tokiose srityse kaip prekių paklausos prognozės arba finansiniai rodikliai atskleidžia *DeepGLO* universalumą ir efektyvumą didelio masto prognozėms. Šis modelis leidžia vienu metu valdyti ir globalius, ir lokalius laiko eilučių ypatumus, kas ypač aktualu pramonėje, kur duomenų kiekis ir įvairovė nuolat auga. Autoriai taip pat pabrėžia modelio lankstumą, kuris leidžia jį pritaikyti įvairiose srityse, nuo energetikos iki elektroninės prekybos. Be to, šis metodas pasižymi skaičiavimo efektyvumu ir gebėjimu apdoroti didelio masto duomenis, net jei jie yra labai nevienalyčiai.

2020-tųjų metų Shchuro, Bilošo ir Günnemanno tyrime [2] nagrinėjamas alternatyvus požiūris į taškinius procesus, pakeičiant intensyvumo funkcijų parametrizavimą tiesioginiu sąlyginio pasiskirstymo modeliavimu. Autoriai pristato normalizuotas srauto funkcijas, kurios leidžia tiksliai modeliuoti procesus, vykstančius nereguliais intervalais. Tokia metodologija ypač naudinga sveikatos priežiūros sistemose, kur pacientų apsilankymai, tyrimų rezultatai ar gydymo procesai vyksta dinamiškai ir neprognozuojamai. Be to, šis metodas taikomas socialinių tinklų analizei, padedant analizuoti interaktyvias vartotojų veiklas realiu laiku. Straipsnyje taip pat akcentuojama, kad siūlomas požiūris gerokai sumažina prognozavimo kaštus, palyginti su tradiciniais metodais, ir leidžia patikimiau analizuoti kompleksines sistemas.

Priežasties ir pasekmės analizė tampa itin reikšminga makroekonominiuose tyrimuose, kur globalūs ekonominiai procesai lemia daugybę lokalių pokyčių. Naudojant Gauso procesus ir giluminio mokymosi metodus, galima nustatyti sudėtingas sąsajas tarp tokių rodiklių kaip infliacija, nedarbo lygis ar valiutų kursai. Toks požiūris suteikia galimybę ne tik geriau suprasti esamas tendencijas, bet ir numatyti būsimus pokyčius, kurie gali turėti reikšmingą poveikį tiek įmonėms, tiek visai ekonomikai [3].

Kitas svarbus aspektas yra mastelio pritaikymo galimybė. Hibridiniai metodai, jungiantys tradicinius statistinius modelius su neuroniniais tinklais, leidžia dirbti tiek su mažais, tiek su dideliais duomenų rinkiniais. Tai leidžia kurti universalius sprendimus, kurie gali būti pritaikomi labai skirtinguose kontekstuose – nuo prekių paklausos prognozavimo iki klimato kaitos tyrimų. Tokių metodų lankstumas tampa ypač aktualus sprendžiant problemas, kur duomenų struktūra ir kokybė gali labai skirtis [4].

Apibendrinant, daugiamačių laiko eilučių prognozavimo tyrimai pabrėžia integruoto požiūrio svarbą. Globalūs dėsninukai, lokaliai savybės ir tikimybinė analizė suteikia galimybę ne tik tiksliai prognozuoti, bet ir suprasti sudėtingą sąsajų tinklą tarp kintamųjų. Tai leidžia ne tik geriau planuoti ir valdyti resursus, bet ir efektyviau spręsti problemas dinamiškame pasaulyje.

1.2. Naudoti duomenų rinkiniai

Daugiamačių laiko eilučių prognozavimo tyrimuose naudojami duomenų rinkiniai yra labai įvairūs, atspindintys tiek realaus pasaulio problemas, tiek imituotus (modeliuotus) scenarijus. Daugelyje atvejų tyrimai remiasi viešai prieinamais duomenų rinkiniais, tokiais kaip finansinių rodiklių ar energetikos suvartojimo laiko eilutės, kurios yra gerai dokumentuotos ir plačiai naudojamos prognozavimo modeliams vertinti. Taip pat dažnai naudojami specifiniai duomenys, surinkti iš pramonės ar socialinių tinklų, siekiant analizuoti unikalius dėsninumus ir tendencijas.

Pavyzdžiui, energetikos srityje naudojami duomenys apie elektros suvartojimą skirtingose geografinėse vietovėse ir skirtingais laikotarpiais. Tokie rinkiniai apima sezoninius svyravimus, darbo dienų ir savaitgalių skirtumus bei įvairius ekstremalius įvykius, tokius kaip karščio bangos ar šalčių laikotarpiai. Finansų rinkose dažnai naudojami akcijų kainų ir prekybos apimčių duomenys, kurie leidžia modeliuoti kintamųjų sąveikas ir tiksliai prognozuoti rinkos tendencijas [5].

Socialinių tinklų analizės srityje naudojami duomenys apie vartotojų interakcijas, tokias kaip pranešimų siuntimo dažnis, reakcijų kiekis ar vartotojų elgsenos pokyčiai laike. Tokie rinkiniai ypač aktualūs, kai reikia modeliuoti procesus, vykstančius nereguliariais intervalais, ir numatyti galimus elgsenos pokyčius [2].

Moksliniuose straipsniuose daugiamatėms laiko eilutėms prognozuoti dažnai pasitelkiami trys tie patys duomenų rinkiniai. Pirmasis elektros energijos duomenų rinkinys apima valandinius elektros energijos suvartojimo duomenis, surinktus iš 370 skirtingų klientų. Antrasis duomenų rinkinys skirtas analizuoti ir prognozuoti daugiamatį laiko eilučių duomenis yra susijęs su transporto eismu. Šis rinkinys apima valandinius transporto srautų duomenis, surinktus iš kelių šimtų kelių jutiklių, įrengtų San Francisko miesto ir aplinkinių rajonų keliuose. Trečiasis duomenų rinkinys - *Wikipedia* svetainės statistika. Šis duomenų rinkinys yra populiarus pasirinkimas laiko eilučių prognozavimo tyrimuose, nes jis atspindi realaus pasaulio vartotojų elgsenos modelius internete. Rinkinys apima kasdienius įvairių *Wikipedia* straipsnių puslapių peržiūrų skaičius [1, 3, 6].

Kai kuriuose duomenų rinkiniuose laiko eilutės yra hierarchinės. Hierarchinės laiko eilutės yra specifinė duomenų rinkinių rūšis, kuriose duomenys organizuojami įvairiuose hierarchijos lygiuose. Šios eilutės pasižymi tuo, kad aukštesnio lygio duomenys yra agreguoti iš žemesnio lygio duomenų. Vienas iš svarbių hierarchinių duomenų rinkinių pavyzdžių yra jau anksčiau minėti *Wikipedia* duomenys, kurie apima vartotojų veiklos stebėjimą skirtinguose hierarchijos lygiuose, pavyzdžiui, visos svetainės, specifinių straipsnių kategorijų ar individualių puslapių analizę. Taip pat verta paminėti ir *M5* konkurso duomenis. Šie duomenys apima mažmeninės prekybos pardavimų ir atsargų duomenis, surinktus iš daugiau nei 30 000 skirtingų produktų, suskirstytų pagal hierarchines kategorijas (pavyzdžiui, produktai → kategorijos → parduotuvės → regionai). Duomenų rinkinys pateikia kasdienius pardavimų rodiklius 10 skirtingų parduotuvių, esančių 3 JAV regionuose, ir yra naudojamas sudėtingų hierarchinių laiko eilučių prognozavimo uždaviniams. Hierarchinės laiko eilutės išsiskiria tuo, kad aukštesnio lygio duomenys yra agreguoti iš žemesnio lygio, todėl prognozavimo metu būtina užtikrinti suderinamumą tarp lygių. Tai suteikia galimybę modeliuoti tiek bendras tendencijas, tiek specifinius pokyčius atskirose hierarchijos dalyse. Tokie duomenys leidžia geriau suprasti vartotojų elgesį globaliame ir lokaliame kontekstuose. Be to, jog kiekvienas hierarchijos lygis gali turėti skirtingus sezoninius ir struktūrinius ypatumus, hierarchinės laiko eilutės taip pat pasunkina ir skaičiavimus. Hierarchiniai duomenys dažnai yra dideli ir reikalauja efektyvių algoritmų, kad būtų galima tiksliai prognozuoti tiek aukštame, tiek žemame lygmenyje. Vis dėlto hierarchinių eilučių detalumas ir tai, jog efektyviai integruojant hierarchijos lygius, galima gauti tiksliau suderintas prognozes, padaro jas labai naudingas interpretuojant rezultatus ir darant išvadas. Pavyzdžiui, *M5* konkurso metu dalyviai naudojo įvairius metodus, įskaitant giluminio mokymosi algoritmus, statistinius modelius ir hibridinius metodus, kad pateiktų tikslias prognozes skirtingais hierarchijos lygmenimis. Konkursas pabrėžė, kad suderinamumas tarp skirtingų lygmenų yra esminis veiksnys norint užtikrinti tikslias prognozes. Tai ypač aktualu mažmeninės prekybos srityje, kur hierarchinė struktūra leidžia modeliuoti tiek lokalias pardavimų tendencijas, tiek globalias regionines dinamikas [4, 6-8].

1 lentelė. Populiariausių duomenų rinkinių savybės

Duomenų rinkinys	Savybės
<i>Electricity</i> [1, 3, 6, 9, 10]	<ul style="list-style-type: none"> • Didelis mastas: apima daug subjektų (klientų), tai leidžia tirti sudėtingus globalius ir lokalius dėsningumus. • Aukštas laiko dažnis: valandiniai matavimai suteikia detalią informaciją apie suvartojimo dinaminį pokyčius. • Daugialypiškumas: klientų duomenys atspindi skirtingas vartotojų grupes (pvz., komercinius ir buitinius vartotojus), o tai leidžia modeliams prisitaikyti prie heterogeniškų duomenų. • Praktinis pritaikymas: duomenys yra naudingi prognozuojant energijos paklausą, optimizuojant energijos paskirstymą ir tiriant vartotojų elgseną.
<i>Traffic</i> [1, 3, 4, 6, 9, 10]	<ul style="list-style-type: none"> • Jutiklių įvairovė: duomenys renkami iš daugiau nei 300 jutiklių, kurie apima skirtingas kelio atkarpas, leidžiančias analizuoti eismo srautų heterogeniškumą. • Aukštas laiko dažnis: valandiniai matavimai suteikia detalią informaciją apie eismo dinamiką per parą, padedant nustatyti piko laikotarpius ar neįprastus srautus. • Priklausomybė nuo aplinkos: transporto srautai stipriai priklauso nuo tokių veiksnių kaip dienos laikas, savaitės diena, šventės ar oro sąlygos, todėl šie duomenys ypač tinka modeliams, gebantiems analizuoti sezonines ir nenumatytas anomalijas, testuoti. • Praktinis pritaikymas: tinkamas modeliams, sprendžiantiems tokias problemas kaip eismo valdymas, transporto srautų prognozė ar infrastruktūros optimizavimas.
<i>Wikipedia</i> [1-4, 6]	<ul style="list-style-type: none"> • Didelis subjektų skaičius: apima tūkstančius skirtingų straipsnių, kurie reprezentuoja įvairias temas - nuo populiarios kultūros iki techninių žinių. • Kasdieniai matavimai: užtikrina nuoseklų laiką tarp stebėjimų, leidžiant analizuoti ilgalaikes tendencijas, sezoniškumą ir trumpalaikes anomalijas. • Elgsenos dinamika: duomenys dažnai atspindi vartotojų interesų pokyčius, susijusius su įvykiais, sezoninėmis temomis ar informacijos poreikio piko laikotarpiais (pvz., susijusiais su naujienomis ar šventėmis). • Heterogeniškumas: skirtingų straipsnių peržiūrų dinamika gali labai skirtis – nuo stabilios iki itin kintančios, todėl tai puikus išbandymas modeliams, siekiantiems prognozuoti įvairių tipų laiko eilutes. • Praktinis pritaikymas: naudojamas prognozuojant turinio populiarumą, optimizuojant serverių išteklius ar analizuojant informacijos sklaidos tendencijas internete.
<i>M5</i> [7, 8]	<ul style="list-style-type: none"> • Hierarchinė struktūra: duomenys yra organizuoti hierarchiškai (1 lygis – regionai, 2 lygis – parduotuvės, 3 lygis – produktų kategorijos, 4 lygis – produktų subkategorijos, 5 lygis – produktai). Ši struktūra leidžia analizuoti pardavimų tendencijas skirtingais detalumo lygmenimis. • Kasdieniai matavimai: pateikiami kasdieniai pardavimų rodikliai 1942 dienų laikotarpiui. • Praktinis pritaikymas: naudojamas mažmeninės prekybos planavimui, atsargų valdymui ir pardavimų prognozėms. Testuoja skirtingų prognozavimo metodų gebėjimą tvarkytis su hierarchiniais duomenimis.
<i>Global Superstore Sales (Tableau, 2018)</i>	<ul style="list-style-type: none"> • Hierarchinė struktūra: duomenys apima pardavimus įvairiais hierarchijos lygiais: produktų kategorijas, regionus ir skirtingus pardavimo kanalus. • Daugiamatis pobūdis: pateikiami pardavimų, pelno, nuolaidų ir užsakymų apimtys duomenys, leidžiantys modeliuoti kompleksinius ryšius tarp kintamųjų. • Plati geografinė aprėptis: duomenys surinkti iš skirtingų pasaulio regionų, todėl tinkami analizuoti prekybos tendencijas skirtingose rinkose.
<i>Brazilian E-Commerce Olist</i>	<ul style="list-style-type: none"> • Detalus elgsenos duomenys: apima užsakymų informaciją, produktų kategorijas, pristatymo laikus ir klientų atsiliepimus, leidžiančius modeliuoti

(Kaggle, 2018)	<p>klientų elgseną.</p> <ul style="list-style-type: none"> • E. prekybos specifika: duomenys yra iš e. prekybos platformos, todėl ypač naudingi prognozuojant pardavimus internetu, vertinant sezoniskumą ir paklausą. • Išoriniai veiksniai: galima analizuoti, kaip pristatymo laikas, logistikos problemos ar klientų atsiliepimai veikia pardavimų sėkmę.
TsBench (44 heterogenous datasets)	<ul style="list-style-type: none"> • Plati duomenų įvairovė: apima 44 skirtingus duomenų rinkinius iš įvairių sričių (finansai, energetika, mažmeninė prekyba, transportas ir kt.). • Standartizuotas testavimo rinkinys: naudojamas palyginti prognozavimo metodus kontroliuojamoje aplinkoje, užtikrinant objektyvią analizę. • Diferencijuoti sudėtingumo lygiai: duomenų rinkiniai skiriasi pagal sezoniskumą, triukšmą, tendencijas ir hierarchinę struktūrą, leidžiant įvertinti modelių lankstumą.

Simuliuoti duomenų rinkiniai taip pat plačiai naudojami, siekiant patikrinti naujų metodų efektyvumą ir išvengti realių duomenų apribojimų, tokių kaip trūkstamos reikšmės ar triukšmas. Simuliacijos leidžia kontroliuoti pagrindinius modelio parametrus ir tiksliai įvertinti jo veikimą esant įvairioms sąlygoms. Straipsnyje "*Intensity-Free Learning of Temporal Point Processes*" [2] sintetiniai duomenys buvo naudojami siekiant patikrinti pasiūlyto metodo efektyvumą ir įvertinti jo tikslumą esant įvairioms sąlygoms. Pagrindinis tikslas buvo sukurti kontroliuojamą aplinką, kurioje būtų galima modeliuoti ir analizuoti sudėtingus taškinis procesus be realių duomenų apribojimų, tokių kaip trūkstamos reikšmės, triukšmas ar duomenų surinkimo nenuoseklumas.

Eksperimentuose buvo naudojami sintetiniai duomenų rinkiniai, kurie apėmė įvairius laiko intervalus ir įvykių tankius. Šie duomenys buvo generuojami pagal specifines taisykles, leidžiančias tiksliai kontroliuoti įvykių pasiskirstymą bei dinaminis pokyčius laike. Tai leido efektyviai analizuoti metodo gebėjimą prisitaikyti prie skirtingų sąlygų ir įvertinti jo stabilumą esant netipinėms situacijoms.

Taikytas metodas, pagrįstas srautų normalizavimu (angl. *normalized flows*), leido išvengti tradicinių intensyvumo funkcijų parametrizavimo. Vietoje to buvo naudojami sąlyginio pasiskirstymo modeliai, kurie suteikė galimybę lanksčiau apdoroti duomenis. Šis požiūris leido tiksliau modeliuoti įvykių laiką ir dažnį, kartu sumažinant skaičiavimo kaštus. Sintetinių duomenų pagalba buvo įrodyta, kad šis metodas gali efektyviai veikti įvairiuose scenarijuose, įskaitant tuos, kur tradiciniai metodai susiduria su reikšmingais apribojimais.

Rezultatai parodė, kad sintetiniai duomenys buvo esminiai vertinant metodo pritaikomumą ir efektyvumą. Jie leido ne tik pademonstruoti modelio lankstumą ir tikslumą, bet ir atskleisti jo ribas bei galimas optimizavimo kryptis. Tokia eksperimentinė aplinka suteikė galimybę detalai ištirti modelio elgseną ir geriau suprasti, kaip jis gali būti pritaikytas realaus pasaulio problemoms spręsti [2].

Apibendrinant galima teigti, kad daugiamačių laiko eilučių prognozavimo tyrimuose naudojami duomenys yra ne tik įvairūs, bet ir atspindintys realius taikymo scenarijus. Tai leidžia modeliuoti sudėtingus procesus ir kurti sprendimus, pritaikomus tiek pramonėje, tiek akademiniuose tyrimuose.

1.3. Naudoti metodai

Daugiamačių laiko eilučių prognozavimo tyrimuose buvo taikoma plati metodų įvairovė, siekiant efektyviai modeliuoti sudėtingas priklausomybes tarp kintamųjų ir užtikrinti prognozių tikslumą. Kiekvienas metodas buvo pasirinktas pagal specifinius duomenų pobūdžio, tyrimų tikslų ir taikymo sričių reikalavimus. Toliau apžvelgiami svarbiausi iš naudotų metodų, paminėti tyrimuose.

Giluminio mokymosi algoritmai buvo vieni svarbiausių metodų, naudojamų analizuotuose straipsniuose. Jie leido automatiškai identifikuoti paslėptus dėsningumus duomenyse ir modeliuoti sudėtingas priklausomybes tarp kintamųjų. Giluminiai neuroniniai tinklai, tokie kaip *LSTM* (*long short-term memory*) ir *GRU* (*gated recurrent unit*), buvo naudojami dėl jų gebėjimo apdoroti laiko eilutes su ilgalaikėmis priklausomybėmis. Be to, kai kuriuose tyrimuose giluminiai tinklai buvo derinami su globaliais matricų faktorizavimo metodais, siekiant išlaikyti tiek globalias, tiek lokalias duomenų savybes. Pavyzdžiui, toks metodas buvo sėkmingai pritaikytas elektros suvartojimo prognozėse, kur globalūs sezoniniai modeliai buvo papildyti lokaliais regionų ypatumais [9].

Tikimybiniai metodai taip pat buvo plačiai naudojami. Šie modeliai leido pateikti ne tik vidutinę prognozę, bet ir įvertinti jos patikimumo intervalus, kas ypač svarbu priimant sprendimus esant neapibrėžtumui. Pavyzdžiui, Bayeso metodai suteikė galimybę generuoti alternatyvius galimus scenarijus ir įvertinti jų tikimybę. Tokie metodai buvo naudingi prognozuojant finansinius rodiklius ir vertinant klimato kaitos poveikį. Tikimybiniai modeliai padėjo išsamiau suprasti duomenų pasiskirstymą ir rizikas [6].

Autoregresiniai modeliai, tokie kaip *SARIMAX* (*Seasonal Autoregressive Integrated Moving Average with Exogenous Variables*) ir *VAR* (*Vector Autoregression*), buvo naudojami analizuoti sezonines ir struktūrines priklausomybes. *SARIMAX* leido įtraukti išorinius kintamuosius, tokius kaip oro sąlygos ar makroekonominiai rodikliai, siekiant pagerinti prognozių tikslumą. *VAR* modeliai buvo naudingi tiriant tarpusavio priklausomybes tarp daugelio kintamųjų, pavyzdžiui, energetikos ir ekonominių rodiklių. Šie modeliai ypač pasitarnavo prognozuojant tendencijas srityse, kur reikšmingos abipusės įtakos tarp kintamųjų [5].

Srautų normalizavimo metodai (*normalized flows*) buvo taikomi sprendžiant problemas, susijusias su taškinių procesų modeliavimu. Šis metodas leido išvengti tradicinių intensyvumo funkcijų parametrizavimo, naudojant sąlyginio pasiskirstymo modelius. Tokia metodologija buvo ypač naudinga analizuojant duomenis su netipinėmis struktūromis arba nereguliariais laiko intervalais, pavyzdžiui, sveikatos priežiūros ir socialinių tinklų analizėje [2, 10].

Hibridiniai metodai, jungiantys tradicinius statistinius modelius su giluminio mokymosi algoritmais, išsiskyrė savo lankstumu. Šie metodai buvo taikomi prognozuojant hierarchines laiko eilutes, kur duomenys buvo organizuojami įvairiais lygmenimis (pavyzdžiui, regionai → miestai → vietovės). Derinant šiuos metodus, buvo užtikrintas prognozių suderinamumas tarp skirtingų hierarchijos lygių, kas ypač svarbu logistikos ir energetikos srityse [3, 4].

Apibendrinant, kiekvienas metodas buvo parinktas pagal jo gebėjimą spręsti specifinius iššūkius. Giluminio mokymosi algoritmai išsiskyrė savo lankstumu ir gebėjimu modeliuoti sudėtingas priklausomybes, tikimybiniai modeliai suteikė galimybę įvertinti neapibrėžtumus, o tradiciniai statistiniai modeliai užtikrino stabilumą ir prognozių patikimumą. Šių metodų integracija sudarė

pagrindą aukštos kokybės prognozėms generuoti įvairiose srityse, nuo energetikos ir finansų iki socialinių tinklų ir sveikatos priežiūros sistemų analizės.

1.3.1. Hierarchinės prognozavimo strategijos: *Bottom-up*, *Top-down* ir *Middle-out*

Hierarchinio prognozavimo strategijos yra svarbus instrumentas verslo, ekonomikos, tiekimo grandinių valdymo ir kitose srityse, kai reikia integruoti skirtingų agregavimo lygių (pvz., regionų, produktų kategorijų, laikotarpių) prognozes. Pagrindinės hierarchinio prognozavimo strategijos yra *Bottom-up*, *Top-down* ir *Middle-out*, kurių kiekviena turi privalumų ir trūkumų, priklausomai nuo konteksto. Toliau analizuojamos šios strategijos, jų taikymo sritys, pranašumai ir ribotumai remiantis moksline literatūra.

Hierarchinės prognozės susideda iš kelių lygių, pavyzdžiui:

- Geografinis lygmuo (šalis → regionas → miestas)
- Produkto kategorijos (bendra pardavimų suma → produktų grupės → atskiri produktai)
- Laiko skalė (metai → ketvirčiai → mėnesiai)

Pagrindinis uždavinys – užtikrinti prognozių suderinamumą tarp skirtingų hierarchijos lygių, kad būtų išvengta neatitikimų [11].

Bottom-up strategija: prognozės formuojamos apatiniuose hierarchijos lygiuose (pvz., atskiruose pardavimo taškuose), o vėliau sumuojamos į aukštesnius lygius [12].

Taikymo sritys:

- Mažmeninė prekyba (pardavimų prognozės pagal parduotuves)
- Tiekimo grandinės valdymas (atskirų sandėlių paklausos prognozė)
- Finansų sektorius (akcijų rinkos tendencijų vertinimas pagal individualias įmones)

Privalumai:

- Gera atspindi lokalią paklausą, nes remiasi detaliais duomenimis.
- Mažesnė informacijos praradimo rizika dėl duomenų agregavimo.
- Tinkama, kai žemesni lygiai turi didesnę prognozės tikslumą.

Trūkumai:

- Reikalauja daug skaičiavimo resursų, ypač didelėms hierarchijoms.
- Gali būti mažai efektyvi, jei žemesni lygiai turi didelį triukšmingumą.
- Sunku valdyti sisteminės klaidas (pvz., visose parduotuvėse pasikartojančius nukrypimus).

Top-down strategija: prognozė pirmiausia sudaroma aukščiausiam hierarchijos lygyje (pvz., bendri šalies pardavimai), o vėliau perskirstoma į žemesnius lygius naudojant proporcijas (pvz., istorinius pardavimų santykius) [13].

Taikymo sritys:

- Makroekonominis planavimas (BVP pasiskirstymas pagal regionus)
- Energijos vartojimo prognozės (šalies paklausos paskirstymas pagal miestus)
- Finansų biudžetavimas (bendros įmonės išlaidos skirstomos padaliniais)

Privalumai:

- Paprasta įgyvendinti, reikalauja mažiau skaičiavimo resursų.
- Gera veikia, kai aukščiausio lygio prognozė yra patikima.
- Sumažina triukšmo įtaką žemuosiuose lygiuose.

Trūkumai:

- Nepaiso lokalių skirtumų (pvz., regioninių ekonominių sąlygų).
- Klaida aukščiausiam lygiui iškreipia visas žemesnes prognozes.
- Istoriniai pasiskirstymo koeficientai gali būti netikslūs ateities tendencijoms.

Middle-out strategija: prognozė pradedama viduriniame hierarchijos lygyje (pvz., regioniniuose pardavimuose), o vėliau ekstrapoliuojama į aukštesnius ir žemesnius lygius [14].

Taikymo sritys:

- Logistika (regioninių sandėlių paklausos prognozė)
- Marketingo planavimas (kategorijų lygio prognozės, skirstomos į produktus)
- Viešasis sektorius (sveikatos išlaidų prognozė pagal regionus)

Privalumai:

- Subalansuoja detalių duomenų ir bendro vaizdo naudą.
- Tinkama, kai viduriniai lygiai turi didžiausią prognozės tikslumą.
- Sumažina per didelio agregavimo arba triukšmo problemas.

Trūkumai:

- Reikalauja gerai apibrėžto vidurinio lygio pasirinkimo.
- Gali būti mažiau lanksti už *Bottom-up* metodą.
- Sudėtingesnė įgyvendinimo logistika.

2 lentelė. Hierarchinių prognozavimo strategijų palyginimas

Kriterijus	Bottom-up	Top-down	Middle-out
Tikslumas	Aukštas žemuose lygiuose	Aukštas aukščiausiam lygiui	Subalansuotas
Sudėtingumas	Didelis (daug skaičiavimų)	Žemas (paprasta agregacija)	Vidutinis
Atsparumas triukšmui	Jautrus	Atsparesnis	Vidutiniškai atsparus
Lankstumas	Didelis (lokalūs pokyčiai)	Mažas (standartizuota)	Vidutinis

Kadangi nei viena strategija nėra universali, mokslinėje literatūroje siūlomi hibridiniai metodai:

- Lyginamieji deriniai (*reconciliation methods*), pvz., *MinT* (Wickramasuriya et al., 2019) [15], kurie derina prognozes iš skirtingų lygių.
- Mašininis mokymas – neuroniniai tinklai gali optimizuoti hierarchinį prognozavimą.

Išvados:

- *Bottom-up* tinkamas, kai svarbūs detalūs duomenys, bet reikalauja resursų.

- *Top-down* efektyvus, kai aukščiausio lygio prognozė patikima, bet nepaiso lokalios dinamikos.
- *Middle-out* subalansuoja abu požiūrius, ypač naudingas vidutinio masto sistemose.
- Ateityje dominuos hibridiniai ir mašininio mokymo pagrindu sukurti metodai, užtikrinantys didesnę tikslumą ir adaptyvumą.

1.3.2. *TsBench*: Standartizuota laiko eilučių prognozavimo metodų vertinimo aplinka

Laiko eilučių prognozavimo tyrimuose vienas iš svarbiausių iššūkių yra objektyviai palyginti skirtingus metodus, kad būtų galima įvertinti jų tinkamumą įvairiems praktiniams uždaviniams. Šiam tikslui pasiekti buvo sukurta *TsBench* – standartizuota testavimo aplinka, kuri leidžia sistemingai vertinti prognozavimo modelių tikslumą, efektyvumą ir pritaikomumą. Ši sistema tampa neatsiejama akademinių tyrimų dalimi ir yra pritaikoma įvairiose srityse, tokiose kaip energetika, mažmeninė prekyba, logistika ir klimato mokslai.

Tyrimai, naudojantys *TsBench*, dažnai pabrėžia, kad tai yra galingas įrankis skirtingų metodų palyginimui įvairiuose kontekstuose. Pavyzdžiui, hierarchinių laiko eilučių prognozavimo srityje *TsBench* padeda analizuoti, kaip modeliai veikia skirtinguose lygmenyse ir kaip užtikrinamas prognozių suderinamumas. Be to, testuojant tiek paprastus statistinius modelius, tiek sudėtingus giluminio mokymosi algoritmus, *TsBench* pateikia objektyvias išvadas apie jų privalumus ir trūkumus.

Pagrindinės *TsBench* savybės:

- **Įvairių duomenų rinkinių naudojimas:** *TsBench* apima realius ir sintetinius duomenų rinkinius, kurie leidžia įvertinti modelių elgseną esant skirtingoms sąlygoms. Pavyzdžiui, realūs duomenys apima tokius rinkinius kaip energijos suvartojimo, transporto srautų ar mažmeninės prekybos duomenys, tuo tarpu sintetiniai duomenys yra generuojami kontroliuojamomis sąlygomis, siekiant ištirti modelių gebėjimą atpažinti tendencijas, sezoniskumą ar triukšmą.
- **Standartizuotos metrikos:** *TsBench* naudoja tokias vertinimo metrikas kaip *MAPE* (*Mean Absolute Percentage Error*), *RMSE* (*Root Mean Square Error*), *CRPS* (*Continuous Ranked Probability Score*) ir *NLL* (neigiamos logaritminės tikimybės funkcija). Šios metrikos užtikrina, kad modelių rezultatai būtų vertinami objektyviai ir tarpusavyje palyginami.
- **Skirtingų prognozavimo metodų integracija:** *TsBench* palaiko platų metodų spektrą, įskaitant statistinius modelius (pvz., *SARIMAX* ir *VAR*), tradicinio mašininio mokymosi algoritmus (pvz., *GBM* ir *XGBoost*) bei giluminio mokymosi modelius (pvz., *DeepAR*, *N-BEATS*, *LSTNet*). Tai leidžia tyrėjams sistemingai įvertinti kiekvieno metodo tinkamumą priklausomai nuo duomenų pobūdžio.
- **Hierarchinių duomenų prognozavimas:** *TsBench* ypač pritaikytas hierarchinių laiko eilučių prognozavimui, leidžiant užtikrinti suderinamumą tarp skirtingų hierarchijos lygių. Pavyzdžiui, mažmeninės prekybos atveju prognozės gali būti generuojamos atskiriems produktams, jų kategorijoms ar regioninėms grupėms, užtikrinant, kad visi lygmenys būtų suderinti.
- **Skaičiavimo kaštų įvertinimas:** Be tikslumo, *TsBench* taip pat leidžia įvertinti modelių skaičiavimo efektyvumą, kas yra itin svarbu dirbant su dideliais duomenų rinkiniais ar realaus laiko prognozavimo uždaviniais.

Įtraukdami *TsBench* vertinimo metodus į šį projektą, galima sistemingai palyginti pasirinktų modelių rezultatus ir įvertinti jų tinkamumą daugiamačių laiko eilučių prognozavimo uždaviniams.

1.3.3. Prognozavimo metodų palyginimo aspektai

Daugiamačių laiko eilučių prognozavimo metodų vertinimas reikalauja įvairiapusio požiūrio, kadangi skirtingi metodai gali skirtis ne tik prognozavimo tikslumu, bet ir skaičiavimo efektyvumu, interpretavimu, prisitaikymu prie hierarchinių struktūrų ar galimybe įtraukti papildomą informaciją (pvz., išorinius regresorius).

Vienas iš pagrindinių vertinimo aspektų – prognozavimo tikslumas. Tyrimuose dažniausiai naudojamos tokios klaidų metrikos kaip *RMSE* (šakninis vidutinio kvadratinio nuokrypio įvertis), *MASE* (normalizuota vidutinė absoliutinė paklaida), *SMAPE* (simetrinė vidutinė absoliutinė procentinė paklaida) ar *MAAPE* (vidutinė absoliuti procentinė paklaida). Šios metrikos leidžia objektyviai įvertinti prognozių artumą tikrosioms reikšmėms bei lyginti skirtingus modelius tarpusavyje.

Kitas svarbus aspektas – skaičiavimo efektyvumas. Daugumos giluminio mokymosi ar mašininio mokymosi modelių treniravimas reikalauja didelių skaičiavimo resursų, ypač kai taikomi didelio masto duomenims. Tai ypač svarbu realaus laiko sprendimuose ar prognozuojant tūkstančius laiko eilučių (kaip *M5* konkurse). Todėl praktiniuose scenarijuose neretai pasirenkami paprastesni, bet stabilesni modeliai, pavyzdžiui, *Exponential Smoothing* ar *ARIMA*.

Hierarchinės struktūros atitikimas – dar vienas esminis kriterijus. Kai kurie metodai yra natūraliai pritaikomi tik žemiausio lygio prognozėms (pvz., *Bottom-up*), tuo tarpu kiti leidžia išlaikyti nuoseklumą visoje hierarchijoje. *M5* konkurse naudota *WRMSSE* metrika (svertinis šakninis normalizuotas kvadratinis nuokrypis) kaip pagrindinis vertinimo kriterijus, nes ji leidžia įvertinti prognozių tikslumą visais 12 hierarchijos lygmenimis ir kartu atsižvelgia į skirtingą produktų pardavimo apimtį [7].

Dar vienas aspektas – metodo universalumas ir interpretacija. Pavyzdžiui, giluminio mokymosi modeliai gali gerai veikti su labai nestruktūruotais ar neturinčiais aiškių sezoniškumo signalų duomenimis, tačiau jų interpretacija dažnai būna sudėtinga. Tuo tarpu statistiniai modeliai, tokie kaip *SARIMA* ar *ETS*, gali būti lengviau suprantami ir paaiškinami verslo kontekste.

Atsižvelgiant į šiuos aspektus, tampa aišku, kad nėra vieno universaliai geriausio metodo – konkretus pasirinkimas priklauso nuo duomenų struktūros, verslo poreikių, prognozės horizonto bei turimų resursų. Būtent todėl dauguma naujausių tyrimų pabrėžia metodų derinimo arba kompozicinių modelių (*ensemble*) taikymo svarbą (Makridakis et al., 2020) [16].

1.3.4. Išoriniai kintamieji laiko eilučių prognozavimo modeliuose

Laiko eilučių prognozavimo modeliuose išoriniai kintamieji (angl. *exogenous variables*) – tai papildomi veiksniai, kurie nėra prognozuojamos laiko eilutės dalis, tačiau gali turėti įtakos jos elgsenai. Įtraukiant šiuos kintamuosius į modelius, siekiama pagerinti prognozavimo tikslumą, ypač kai prognozuojama sistema yra veikiamą išorinių veiksnių, tokių kaip oro sąlygos, ekonominiai rodikliai ar socialiniai įvykiai.

Klasikiniai modeliai, tokie kaip *ARIMAX* ir *SARIMAX*, leidžia įtraukti išorinius kintamuosius kaip papildomus regresorius. Pavyzdžiui, Eshragh'o ir kt. (2022) [17] tyrime, prognozuojant Australijos elektros energijos paklausą, į modelį buvo įtraukti aplinkos veiksniai, tokie kaip maksimali ir minimali temperatūra bei saulės spinduliuotė. Tai leido sumažinti prognozavimo klaidą (*MAPE*) nuo 6.35% iki 3.41%, palyginti su modeliu be išorinių kintamųjų.

Giluminio mokymosi modeliai taip pat pradėjo integruoti išorinius kintamuosius. Pavyzdžiui, *NBEATSx* modelis, išplėstas iš *N-BEATS* architektūros, leidžia įtraukti išorinius veiksnius į prognozavimo procesą. Olivares'as ir kt. (2023) [18] parodė, kad įtraukus išorinius kintamuosius į *NBEATSx* modelį, prognozavimo tikslumas pagerėjo iki 20% elektros kainų prognozavimo užduotyje.

Transformeriai, tokie kaip *TimeXer*, taip pat buvo pritaikyti darbui su išoriniais kintamaisiais. Wang'as ir kt. (2024) [19] pasiūlė *TimeXer* modelį, kuris naudoja specialias įterpimo sluoksnius ir dėmesio mechanizmus, kad efektyviai integruotų išorinius kintamuosius į prognozavimo procesą. Šis modelis pasiekė pažangiausius rezultatus dvylikoje realaus pasaulio prognozavimo užduočių.

Svarbu pažymėti, kad išorinių kintamųjų įtraukimas į modelius reikalauja atsargumo. Reikia užtikrinti, kad šie kintamieji būtų prieinami prognozavimo metu ir būtų tinkamai apdoroti, kad būtų išvengta duomenų nutekėjimo (angl. data leakage). Be to, būtina įvertinti šių kintamųjų stacionarumą ir galimą įtaką modelio stabilumui.

1.4. Pagrindiniai rezultatai ir išvalgos

Bendrai apžvelkime metodus bei modelių tipus.

3 lentelė. Prognozavimo metodų palyginimas

Metodas	Modelio tipas	Duomenų rinkiniai	Hierarchinių duomenų analizė	Pagrindinės išvalgos
<i>DeepVAR</i> (<i>VEC-LSTM</i>)	Giliojo mokymosi (<i>LSTM</i> pagrindu)	<i>M5, Traffic, Wikipedia</i>	Taip	Veiksmingas hierarchinėms laiko eilutėms, tačiau reikalauja daug resursų.
<i>LSTNet 1</i>	<i>CNN + RNN</i> kombinacija	<i>Electricity, Traffic</i>	Ne	Naudingas trumpalaikėms prognozėms, tačiau gali prastai veikti su ilgalaikiais ryšiais.
<i>LSTNet 2</i>	<i>CNN + RNN</i> kombinacija	Energetikos duomenys	Ne	Sezoniškumą užfiksuoja geriau nei tradiciniai <i>LSTM</i> modeliai.
<i>LSTNet 3</i>	<i>CNN + RNN</i> kombinacija	Finansų ir ekonomikos duomenys	Ne	Patobulinta versija, geriau valdanti duomenų triukšmą.
<i>GRAformer</i>	Grafų transformerių modelis	Transporto srautai, socialiniai tinklai	Taip	Efektyvus modeliuojant erdvinis ir laiko duomenis kartu.
<i>iTransformer</i>	<i>Transformer</i> architektūra	Sintetiniai ir realūs laiko eilučių duomenys	Ne	Pagerintas interpretacijos lygis, tačiau didelis skaičiavimo kaštas.
<i>TimeXer</i>	<i>Transformer</i> variacija	Multimodalūs duomenys (finansai, klimatas, energetika)	Ne	Optimizuota struktūra leidžia greitesnį mokymąsi nei klasikiniai transformeriai.

<i>MTGNN</i>	<i>GNN</i> (grafų neuroniniai tinklai)	Transporto ir tinklų duomenys	Taip	Ypač tinkamas prognozuoti laiko eilutes, turinčias erdvinis ryšius.
<i>STFGNN</i>	Erdvėlaikio grafų neuroninis tinklas	Eismo ir energetikos duomenys	Taip	Užtikrina didesnę tikslumą modeliuojant erdvinis priklausomumus.
<i>StemGNN</i>	Spektrinis grafų modelis	Skirtingi realūs duomenys	Taip	Tinka labai dinamiškiems duomenims; stabilus hierarchinėse prognozėse.
<i>DSCM</i>	Kombinuotas modelis (tradiciniai + giliojo mokymosi metodai)	Prekybos ir ekonomikos duomenys	Ne	Lankstus metodas, tačiau priklauso nuo duomenų kokybės.
<i>DeepAR</i>	Tikimybinis <i>RNN</i>	<i>M5</i> , <i>Traffic</i> , <i>Wikipedia</i>	Taip	Pateikia patikimumo intervalus, tačiau gali per daug prisitaikyti prie duomenų.
<i>N-BEATS</i>	Giliojo mokymosi (<i>MLP</i> pagrindu)	Finansai, klimatas, mažmeninė prekyba	Ne	Gera prognozuoja be papildomų duomenų, bet prastai modeliuoja sudėtingas priklausomybes.
<i>SARIMAX</i>	Statistinis modelis	Ekonomikos, finansų, energetikos duomenys	Ne	Stiprus sezonizmo modeliavime, bet ribotas sudėtingose daugiamatėse sistemose.
<i>XGBoost</i>	Tradicinis mašininis mokymasis	<i>M5</i> , mažmeninės prekybos duomenys	Taip	Efektyvus su dideliais duomenų rinkiniais, bet reikalauja daug hiperparametrų derinimo.
<i>DeepGLO</i>	<i>CNN</i> + <i>RNN</i> kombinacija	Transportas, energija, socialiniai tinklai	Taip	Apjungia globalų ir lokalų mokymą, tinka didelėms daugiamatėms laiko eilučių sistemoms.

Daugiamatėjų laiko eilučių prognozavimo tyrimuose pateikti rezultatai rodo, kad taikyti metodai yra itin veiksmingi sprendžiant sudėtingas prognozavimo užduotis, tačiau kiekvieno metodo naudingumas priklauso nuo duomenų pobūdžio ir tyrimų tikslų. Galima toliau pažvelgti kiekvieno metodo veiksmingumą bei naudotas tikslumo metrikas.

1.4.1. Giluminio mokymosi metodai

Giluminiai neuroniniai tinklai pasirodė itin efektyvūs modeliuojant sudėtingas daugiamatėjų laiko eilučių priklausomybes. Pavyzdžiui, modeliai, derinantys globalius ir lokalius duomenų ypatumus, parodė reikšmingą prognozavimo tikslumo padidėjimą, ypač didelio masto duomenų rinkiniuose (tokiuose kaip energijos suvartojimas ar finansiniai rodikliai) [9].

Dažniausiai naudotos tikslumo metrikos buvo *MAPE*, *RMSE* ir *MAE* (*Mean Absolute Error*). Šios metrikos leido objektyviai įvertinti prognozių tikslumą įvairiuose kontekstuose.

DeepGLO modelis elektros suvartojimo prognozėse pasiekė *MAPE* reikšmę 0.341 ir *RMSE* reikšmę 6.49 su nereguliuotais duomenimis. Šis modelis lenkė tradicinius metodus, tokius kaip *SARIMAX* ir *Prophet*, demonstruodamas didesnę tikslumą net esant sudėtingoms duomenų struktūroms.

DeepAR modelis *Wikipedia* duomenų rinkinyje pasiekė *MAPE* reikšmę 3.335, o *TCN* metodas pasiekė 1.322, kai buvo naudojami nereguliuoti duomenys. Šie rezultatai rodo, kad giluminio mokymosi metodai ypač veiksmingi modeliuojant trumpalaikius ir ilgalaikius ryšius.

Giluminio mokymosi metodai gali būti naudingi projektuojant magistrinį darbą, jei numatoma naudoti didelius, nevienalyčius duomenų rinkinius, kuriuose yra svarbios tiek globalios, tiek lokalios struktūros.

1.4.2. Tikimybiniai modeliai

Tyrimuose, kuriuose buvo taikomi tikimybiniai metodai, pastebėta, kad galimybė pateikti patikimumo intervalus ir rizikos vertinimus padidino prognozių naudingumą realiaame pasaulyje. Šie metodai buvo ypač veiksmingi klimato prognozėse ir finansinių rinkų analizėje, kur prognozavimo tikslumas neatsiejamas nuo neapibrėžtumo valdymo [6].

Naudota *CRPS* ir logaritminio tikėtinumo funkcija (*Log-Likelihood*), kurios įvertino modelių gebėjimą tiksliai prognozuoti tikimybinius pasiskirstymus.

Tikimybiniai metodai, tokie kaip *LogNormMix* ir *DSFlow*, sintetiniuose duomenyse pasiekė reikšmingai mažas neigiamos logaritminės tikėtinumo funkcijos (*NLL*) reikšmes. Pavyzdžiui, *LogNormMix* modelis sintetiniuose *Poisson* duomenyse užfiksavo *NLL* reikšmę 0.25, o *DSFlow* - 0.254. Šie rezultatai patvirtina šių metodų gebėjimą tiksliai modeliuoti tikimybinius pasiskirstymus, ypač esant netipiniams duomenų pasiskirstymams.

Tikimybiniai modeliai tinka, jei magistriniame projekte reikėtų prognozuoti rezultatus su neapibrėžtumu, pavyzdžiui, atliekant rizikos vertinimą.

1.4.3. Autoregresiniai modeliai

SARIMAX ir *VAR* modeliai parodė didelį efektyvumą prognozuojant sezonines ir struktūrines laiko eilučių ypatybes. *SARIMAX* buvo ypač efektyvus analizuojant laikotarpius, kai išoriniai veiksniai darė reikšmingą įtaką, o *VAR* leido modeliuoti kelių kintamųjų tarpusavio priklausomybes [5, 20].

Naudotos *AIC* (*Akaike Information Criterion*) ir *BIC* (*Bayesian Information Criterion*) modelio tinkamumui įvertinti, taip pat *MAE* ir *MAPE* prognozių tikslumui matuoti.

Šie modeliai galėtų būti taikomi magistriniame projekte, jei analizuojami duomenys pasižymi aiškiu sezoniškumu arba jei reikėtų modeliuoti tarpusavio priklausomybes tarp kintamųjų.

1.4.4. Hibridiniai metodai

Hibridiniai modeliai, kurie derino giluminio mokymosi algoritmus su tradiciniais statistiniais metodais, buvo efektyvūs sprendžiant sudėtingas prognozavimo užduotis, pavyzdžiui, hierarchinių duomenų prognozes. Šie metodai padėjo suderinti stabilumą ir lankstumą, užtikrinant aukštą tikslumą skirtinguose hierarchijos lygiuose.

Be tradicinių tikslumo matų, tokie modeliai dažnai buvo vertinami naudojant *HITS* (*Hierarchical Time Series Score*), kuris matuoja prognozių suderinamumą tarp hierarchijos lygių [4].

Magistriniame projekte šie metodai galėtų būti naudingi, jei numatoma dirbti su hierarchiniais ar sudėtingos struktūros duomenimis.

1.4.5. Srautų normalizavimo metodai

Srautų normalizavimo metodai (*normalized flows*) buvo naudojami taškiniais procesams modeliuoti, kai duomenys buvo nereguliariai paskirstyti. Rezultatai parodė, kad šis metodas leidžia tiksliai modeliuoti įvykių dažnumą ir laiką be intensyvumo funkcijų parametrizavimo [2, 10].

Kullback-Leibler divergencija (*KL divergence*) buvo naudojama įvertinti, kiek tiksliai modelis atkartoja realų duomenų pasiskirstymą.

Šis metodas gali būti naudingas analizuojant duomenis su nereguliariais intervalais, pavyzdžiui, socialinių tinklų vartotojų aktyvumą.

1.5. Baigiamojo magistro projekto temos ir uždavinių pagrindimas

Daugiamačių laiko eilučių prognozavimo tyrimai atskleidžia didelį metodologinį potencialą, tačiau kartu išryškėja ir reikšmingi iššūkiai. Literatūros apžvalga parodė, kad giluminio mokymosi metodai, tokie kaip *DeepGLO*, *DeepAR* ir *N-BEATS*, dažnai pateikia didelio tikslumo prognozes sudėtingiems ir dideliems duomenų rinkiniams. Visgi, paprastesni statistiniai ar tradicinio mašininio mokymosi metodai, tokie kaip *SARIMAX* ar *GBM*, kai kuriais atvejais pasižymi panašiu arba net didesniu efektyvumu.

Šio projekto tyrimų svarbą pagrindžia hierarchinių laiko eilučių prognozavimo iššūkiai. Nėra iki galo aišku, ar geriausia prognozes generuoti atsižvelgiant į hierarchijos žemiausius lygius, o vėliau rekonstruoti aukštesnio lygmens duomenis, ar atvirkščiai – pradėti nuo bendresnių duomenų ir juos detalizuoti. „M5 varžybos“ atskleidė, kad nei viena komanda negalėjo pasiūlyti universalaus sprendimo, užtikrinančio aukščiausią tikslumą visais hierarchijos lygmenimis. Tai patvirtina būtinybę gilintis į metodų suderinimo strategijas.

Projekto metu planuojama tirti, kurie iš esamų metodų tinkamiausi įvairioms prognozavimo užduotims, kaip prognozių tikslumą veikia skirtingos duomenų struktūros ir kokius kompromisus reikia priimti tarp tikslumo ir skaičiavimo resursų poreikio. Šie aspektai yra itin aktualūs praktiškai pritaikant tyrimų rezultatus tokiose srityse kaip logistikos planavimas, energijos paskirstymas ar prekybos tinklų optimizavimas. Projekte bus siekiama įvertinti, ar sudėtingesni giluminio mokymosi metodai pateisina investicijas į jų diegimą, ar paprastesni statistiniai modeliai vis dar gali būti tinkamas pasirinkimas.

2. Tyrimo metodai

Daugiamačių laiko eilučių prognozavimui būtina pasirinkti tinkamus modelius ir taikyti atitinkamas duomenų paruošimo bei vertinimo strategijas, siekiant užtikrinti prognozių tikslumą ir patikimumą. Šiame skyriuje aprašomi tyrime taikyti prognozavimo modeliai, pateikiamos duomenų apdorojimo procedūros, išoriniai kintamieji ir prognozavimo kriterijai. Kad būtų užtikrintas tyrimo pakartojamumas ir rezultatų patikimumas, detalai pristatomi modelių parinkimo motyvai, duomenų transformacijos žingsniai bei vertinimo metrikos. Tyrimo metu naudotas vienas duomenų rinkinys – *M5* konkurso duomenys, kurio struktūra išsamiau aprašyta tyrimų rezultatų skyriuje.

2.1. Naudoti prognozavimo metodai

Siekiant įvertinti daugiamačių laiko eilučių prognozavimo metodų veiksmingumą, tyrimo metu buvo taikomos dvi skirtingos modelių grupės: standartiniai *M5* konkurso organizatorių pateikti etaloniniai modeliai ir individualiai sukurtas autoregresinis modelis su išoriniais regresoriais (*ARIMAX*). Etaloninių modelių taikymas leido objektyviai įvertinti darbo rezultatus, palyginant juos su plačiai priimtomis prognozavimo praktikomis. Sukurto modelio tikslas buvo pritaikyti papildomą informaciją apie išorinius kintamuosius, siekiant pagerinti prognozių tikslumą hierarchinėse struktūrose. Šiame poskyryje detaliau pristatomi naudoti modeliai, jų taikymo logika bei pagrindiniai skirtumai tarp metodų.

2.1.1. Etaloniniai varžybų (*benchmark*) modeliai

Tyrimo metu buvo taikomi keli baziniai prognozavimo modeliai, kurie sudarė atskaitos tašką (angl. *benchmark*) sukurtų metodų palyginimui. Šie modeliai parinkti atsižvelgiant į *M5* konkurso organizatorių rekomendacijas bei literatūroje pripažintą taikymo praktiką. Žemiau pateikiami taikytų modelių pagrindiniai principai ir matematiniai aprašymai.

- ***Naive***. *Naive* metodas prognozuoja, kad ateities reikšmės bus lygios paskutinei žinomai reikšmei:

$$\hat{y}_{t+h} = y_t,$$

kur \hat{y}_{t+h} - prognozuojama reikšmė po h žingsnių, y_t – paskutinė stebėta reikšmė.

- ***Seasonal Naive (sNaive)***. *Seasonal Naive* modelis taiko sezoniškumo principą, prognozuodamas reikšmę pagal praėjusio sezono duomenis:

$$\hat{y}_{t+h} = y_{t+h-m(k+1)},$$

kur m – sezono ilgis (pvz., 365 dienos metiniam sezoniškumui), k – sveikasis skaičius, tenkinantis $h \leq m(k+1)$.

- ***Simple Exponential Smoothing (SES)***. *SES* modelis prognozuoja kaip svertinį visų ankstesnių stebėjimų vidurkį, kur svoriai mažėja eksponentiškai:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t,$$

kur $\alpha \in (0, 1)$ – glotninimo parametras, \hat{y}_t – ankstesnė prognozė.

- **Moving Average (MA).** MA modelis prognozuoja pagal paskutinių q reikšmių aritmetinį vidurkį:

$$\hat{y}_{t+1} = \frac{1}{q} \sum_{i=0}^{q-1} y_{t-i},$$

kur q – lango dydis (kiek ankstesnių stebėjimų imti).

- **Croston.** Croston metodas naudojamas retų paklausų prognozei. Jis atskirai prognozuoja paklausos dydį (z) bei tarpų tarp paklausų trukmę (p).

Paklausos dydžio prognozė:

$$z_t = z_{t-1} + \alpha(y_t - z_{t-1}).$$

Intervalo prognozė:

$$p_t = p_{t-1} + \alpha(\tau_t - p_{t-1}).$$

Prognozė:

$$\hat{y}_{t+h} = \frac{z_t}{p_t},$$

kur τ_t – tarpas tarp dviejų paklausų.

- **Optimized Croston (optCroston), SBA, TSB.** Šie modeliai yra Croston metodo modifikacijos:

- *optCroston* optimizuoja α reikšmę automatiškai.
- *SBA* (Syntetos-Boylan Approximation) pritaiko korekciją Croston prognozei:

$$\hat{y}_{SBA} = \frac{2z_t}{p_t+1}.$$

- *TSB* (Teunter-Syntetos-Babai) papildomai modeliuoja tikimybę, kad paklausa egzistuoja. Kitaip sakant, atskirai modeliuojama paklausos tikimybė ir dydis.
- **ADIDA ir iMAPA.** ADIDA metodas suagreguoja duomenis į žemesnio dažnio eilutę, jiems pritaikomas prognozavimas, o po to duomenys deagreguojami. iMAPA yra pagerinta ADIDA versija, naudojanti kelis agregavimo lygius vienu metu ir jungiant prognozes.
- **Bottom-up metodai: ES_bu ir ARIMA_bu.** Šie metodai prognozuoja žemiausio lygio laiko eilutes atskirai, o po to agreguoja rezultatus į aukštesnius hierarchijos lygius. Pavyzdžiui, taikant *Exponential Smoothing* arba *ARIMA* kiekvienai prekei atskirai.
- **MLP_l ir RF_l.** MLP_l (*Multi-Layer Perceptron local*) – dirbtinis neuroninis tinklas taikomas kiekvienai laiko eilutei atskirai. RF_l (*Random Forest local*) – sprendimų medžių ansamblis treniruojamas kiekvienai laiko eilutei individualiai. Šie abu metodai leidžia modeliuoti sudėtingus netiesinius ryšius tarp laiko eilučių reikšmių.

Taikyti etaloniniai modeliai leido įvertinti prognozavimo užduoties sudėtingumą bei sudaryti pagrindą tolimesniam sukurtų metodų rezultatų palyginimui. Visų modelių tikslumas buvo

vertinamas *WRMSSE* metrika, apimančia prognozavimo tikslumą skirtingais hierarchijos lygmenimis. Toliau pateikiama modelių palyginimo lentelė (žr. 4 lentelę), įvertinant treniravimo laiką, resursų poreikį bei praktinį taikomumą.

4 lentelė. Modelių palyginimo lentelė

Modelis	Treniravimo trukmė	Resursų poreikis	Pritaikomumas dideliui kiekiui laiko eilučių	Praktinis taikymas
<i>Naive / sNaive</i>	Labai greitai	Minimalus (<i>CPU</i>)	Tobulas	Labai paprasta
<i>Exponential Smoothing</i>	Lėta	Mažas	Geras	Tinka mažoms sistemoms
<i>ARIMA</i>	Vidutinė	Vidutinis	Prastėja didėjant eilutėms	Ribotas praktikoje
<i>Random Forest</i>	Vidutinė-didelė	Vidutinis-didelis	Gali būti lėtas be paralelizavimo	Galimas pritaikymas
<i>MLP (neural net)</i>	Lėta	Didelis (<i>RAM, CPU/GPU</i>)	Ribotas	Netinka masinei prognozei
<i>LSTM / GRU</i>	Labai lėta	Reikia <i>GPU</i>	Geras, tačiau reikalauja didelės inžinerijos	Tinka tik su stipria infrastruktūra
<i>Bottom-up Exponential Smoothing</i>	Lėta	Mažas	Gera pritaikomas	Vienas geriausių variantų
<i>Bottom-up ARIMA</i>	Vidutinė	Vidutinis	Vidutinis	Gana universalus

2.1.2. *ARIMAX* su išoriniais regresoriais

Tyrimo metu, siekiant pagerinti daugiamatį laiko eilučių prognozavimo tikslumą, buvo sukurtas modelis, pagrįstas sezoniniu autoregresiniu integruotuoju slankiuoju vidurkiu su išoriniais regresoriais (*ARIMAX*). Šis modelis leidžia prognozuoti tikslinį kintamąjį atsižvelgiant ne tik į jo pačio ankstesnes reikšmes, bet ir į papildomą informaciją apie išorinius veiksnus, galinčius turėti reikšmingos įtakos pardavimų dinamikai.

Bendroji *ARIMAX* modelio formulė yra išreiškiama taip:

$$\Phi_p(B)(1-B)^d(1-B^s)^D y_t = \Theta_q(B)\Theta_Q(B^s)\epsilon_t + \beta X_t,$$

kur y_t – prognozuojama reikšmė laiko momentu t , B – poslinkio operatorius ($B^k y_t = y_{t-k}$), d – ne sezoninių skirtumų laipsnis, D – sezoninių skirtumų laipsnis su periodiškumu s , $\Phi_p(B)$ ir $\Theta_q(B)$ – atitinkamai autoregresinių (*AR*) ir slankiųjų vidurkių (*MA*) komponentų polinamai, $\Theta_Q(B^s)$ – sezoninio slankiojo vidurkio komponentas, ϵ_t – atsitiktinės paklaidos komponentas (baltasis triukšmas), X_t – išorinių regresorių matrica, β – regresorių koeficientų vektorius.

Šiame tyrime kaip išoriniai regresoriai X_t buvo naudojami septyni komponentai:

- Atitinkamos produktų kategorijos pardavimų suma.
- Atitinkamo departamento pardavimų suma.
- Šventinės dienos (binarinis kintamasis: 1 – šventė, 0 – įprasta diena).

- Atitinkamos valstijos pardavimų suma.
- Atitinkamos parduotuvės pardavimų suma.
- Atitinkamos valstijos-kategorijos pardavimų suma.
- Kainų dinamikos regresorius (7 dienų slankusis vidurkis).

Kiekvienas regresorius buvo kuriamas remiantis produkto priklausomybe konkrečiai parduotuvei, departamentui, kategorijai ar regionui. Regresorių laiko eilutės buvo prognozuojamos į ateitį naudojant geriausiai kiekvienam kintamajam pritaikytą vienmatį modelį (atrinktą pagal kelias prognozės tikslumo metrikas ir skaičiavimo efektyvumą). Šventinių dienų stulpelis buvo laikomas žinomiu iš anksto.

Modelių atranka:

1. Modelių rinkinys: kiekvieno išorinio regresoriaus (išskyrus šventinių dienų kintamojo) laiko eilutei taikyti 15 vienmačių prognozavimo metodų (kai kurie jau buvo minėti anksčiau tarp etaloninių modelių).

- **Seasonal Trend** - paprastas modelis, kuris derina pastovią tendenciją ir sezoniškumą, dažnai naudojamas kaip pradinė bazinė versija kompleksiniams metodams.
- **Seasonal Naive**.
- **Croston**.
- **Holt-Winters** – metodas su multiplikatyviu sezoniškumu, leidžianti modeliuoti tiek tendenciją, tiek sezoniškumą:

$$\begin{cases} l_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \\ s_t = \gamma \left(\frac{y_t}{l_t} \right) + (1 - \gamma)s_{t-m} \\ \hat{y}_{t+h} = (l_t + hb_t)s_{t-m+h} \end{cases}.$$

- **Theta** – metodas, sėkmingai naudotas M3 prognozavimo konkurse, kombinuoja tiesinį tendencijos modelį su sezoniniais svyravimais.
- **Exponential Smoothing**.
- **Exponential Smoothing with damped trend** – eksponentinis glotninimas su mažėjančia tendencija, kuri ilgai „gesinasi“ (*damped trend*).
- **ARIMA**.
- **ARIMA with damped trend** – standartinis ARIMA modelis, papildytas slopstančia tendencijos komponento versija, siekiant išvengti pernelyg didelio ilgalaikio augimo.
- **TBATS** – metodas, tinkamas labai sudėtingoms sezoninėms struktūroms ir ilgoms laiko eilutėms. Jis jungia Box-Cox transformaciją (B), ARMA klaidų komponentus (A), tendrą (T), sezonines komponentes su trigonometrija (S).

- **PROPHET** – „Facebook“ sukurtas modelis, skirtas verslo duomenų prognozėms. Apjungia adityvinį sezoniškumą, švenčių efektus ir tendencijų pertraukas:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t,$$

kur $g(t)$ – tendencija, $s(t)$ – sezoniškumas, $h(t)$ – šventės, ϵ_t – paklaida.

- **THETAh** – *Theta* modelio hierarchinė versija, taikoma su papildomu duomenų padalijimu ir prognozių sujungimu.
- **Exponential Smoothing hierarchical** – eksponentinis glodinimas, taikomas hierarchiniu būdu panašiai kaip „bottom-up“ strategijoje.
- **ARIMA hierarchical** – analogiškai kaip Exponential Smoothing hierarchical, tik naudojant ARIMA kaip bazinį metodą.
- **Exponential Smoothing with Box-Cox transformation** - *Exponential Smoothing* versija, kuri naudoja *Box-Cox* transformaciją duomenims stabilizuoti ir pagerinti modelio prisitaikymą.

Visi minėti metodai buvo įtraukti į automatinę atrankos procedūrą, užtikrinančią, kad kiekvienam regresoriui būtų pritaikytas tinkamiausias metodas pagal bendrą efektyvumo ir tikslumo įvertį. Ši strategija leido padidinti ARIMAX modelio patikimumą ir dinamiškumą, nepriklausomai nuo regresorių pobūdžio.

2. Slankiojo lango metodas:

- Sudaryti trys persidengiantys mokymo langai, kiekvienas po 12 žingsnių.
- Langai persidengė 50%.

3. Modelių vertinimas:

- Prognozės vertintos pagal penkias klaidų metrikas (*RMSE*, *MDAE*, *MASE*, *SMAPE*, *MAAPE*).
- Kiekvienai metrikai suteiktas reitingas (1 – geriausias modelis, 15 - prasčiausias).

4. Skaičiavimo trukmė:

- Modelių vykdymo laikas matuotas sekundėmis, greitesni modeliai gavo aukštesnį reitingą.

5. Galutinis sprendimas:

- Apskaičiuota kiekvieno modelio vidutinė reitingų vertė.
- Geriausiu laikomas modelis, turintis mažiausią bendrą reitingą, užtikrinantį balansą tarp tikslumo ir efektyvumo.

Taigi kiekvienam regresoriui buvo pritaikytas atskirai atrinktas geriausias modelis. Šie modeliai buvo naudojami ateities reikšmėms prognozuoti. Šventinės dienos buvo integruojamos kaip egzogeninis, bet žinomas kintamasis.

Darbe taikyta *Middle-out* strategija apžvelgta praeitame skyriuje. *ARIMAX* modelis buvo treniruojamas tik 9 lygyje, o prognozės agreguojamos į aukštesnius lygius bei proporcingai išskaidomos į žemesnius.

Kiekvienai 9-ojo hierarchijos lygio laiko eilutei (t. y. kiekvienai parduotuvės-departamento kombinacijai) prognozuojant naudojami šie išoriniai regresoriai:

- Atitinkamo departamento prognozuojamos pardavimų reikšmės – kiekvienai eilutei priskiriama to departamento prognozė;
- Atitinkamos kategorijos prognozuojamos pardavimų reikšmės – kategorija nustatoma pagal produkto priklausomybę;
- Šventinių dienų indikatorius – bendra visoms eilutėms binarinė reikšmė (1 – šventė, 0 – įprasta diena);
- Atitinkamos valstijos prognozuojamos pardavimų reikšmės – priskiriamos pagal parduotuvės buvimo vietą;
- Atitinkamos parduotuvės prognozuojamos pardavimų reikšmės – kiekvienai eilutei priskiriamas atitinkamos parduotuvės laiko eilutės prognozės stulpelis;
- Valstijos-kategorijos pardavimų prognozės – apibendrina regioninį ir tematinį aspektą vienu kintamuoju;
- Kainų dinamikos regresorius – visoms eilutėms bendras laiko kintamasis, apskaičiuotas kaip visų prekių vidutinės kainos 7 dienų slankusis vidurkis.

Šie regresoriai buvo parenkami individualiai kiekvienai eilutei pagal jos unikalią kombinaciją: parduotuvė-departamentas-kategorija-valstija.

2.2. Duomenų paruošimo strategijos

Prognozavimo modeliams, norint efektyviai panaudoti tiek tikslinius kintamuosius, tiek išorinius regresorius, būtina tinkamai paruošti pirminius duomenis. Nors tyrime naudoti *M5* konkurso duomenys buvo gana išsamūs ir gerai struktūruoti, vis dėlto reikėjo atlikti keletą papildomų transformacijų ir kintamųjų konstravimo veiksmų, siekiant užtikrinti sklandų modelių taikymą.

2.2.1. Šventinių dienų identifikavimas

Naudojant „*calendar.csv*“ duomenų failą, buvo suformuotas naujas binarinis kintamasis „*is_holiday*“, žymintis, ar konkreti diena buvo šventinė. Šventinė diena buvo laikoma tokia, jei bent viename iš stulpelių „*event_name_1*“, „*event_type_1*“, „*event_name_2*“ arba „*event_type_2*“ buvo nurodytas įvykis. Visos šventės yra JAV šventės, todėl jos atspindi amerikietišką mažmeninės prekybos ritmą, įskaitant tokias dienas kaip Padėkos diena, Kalėdos, Darbininkų diena (pirmasis rugsėjo pirmadienis) ir kt.

Binarinis kintamasis buvo priskirtas visoms dienoms:

- „*is_holiday*“ = 1, jei yra bent vienas įvykis,
- „*is_holiday*“ = 0 – visais kitais atvejais.

2.2.2. Laiko eilučių transformacijos

Originalus „*sales_train_validation.csv*“ duomenų failas pateikia informaciją „*wide*“ formatu, kur kiekviena produkto-parduotuvės kombinacija yra pateikiama kaip viena eilutė, o kiekviena diena –

atskiras stulpelis ($d_1, d_2, \dots, d_{1913}$). Kad būtų galima atlikti laiko eilučių analizę, duomenys buvo transformuoti į „long“ formatą, kur kiekviena eilutė atspindi vieną produkto pardavimą konkrečią dieną. Šis žingsnis būtinas tiek apibendrintų regresorių apskaičiavimui, tiek duomenų vizualizacijai.

2.2.3. Išorinių regresorių formavimas

Kiekvienam prognozuojamam objektui buvo priskirti septyni išoriniai regresoriai, atspindintys tiek pardavimų struktūrą, tiek kontekstinius veiksnius:

1. Atitinkamo departamento bendri pardavimai kiekvieną dieną.
2. Atitinkamos produktų kategorijos bendri pardavimai kiekvieną dieną.
3. Binarinis šventinių dienų indikatorius (1 – šventė, 0 – ne šventė).
4. Atitinkamos valstijos bendri pardavimai kiekvieną dieną.
5. Atitinkamos parduotuvės bendri pardavimai kiekvieną dieną.
6. Atitinkamos valstijos–kategorijos kombinacijos bendri pardavimai.
7. Kainų dinamikos rodiklis – visų prekių vidutinės kainos 7 dienų slankusis vidurkis.

Departamentų, kategorijų, valstijų, parduotuvių bei valstijos-kategorijos pardavimai buvo gauti transformuojant duomenis į ilgąjį formatą, priskiriant kalendorines datas (naudojant *left_join* su kalendoriaus lentele), ir tuomet juos agreguojant su *group_by()* ir *summarise()* funkcijomis. Pvz., departamentų pardavimai buvo apskaičiuoti su *group_by(date, dept_id)*, o valstijos–kategorijos – su *group_by(date, state_id, category)*.

Kad būtų išvengta trūkstančių duomenų, naudota *values_fill = 0* strategija, kuri visoms dienoms užtikrina reikšmes, net jei tą dieną nevyko joks pardavimas.

Tuo tarpu kainų dinamikos regresorius buvo apskaičiuotas iš *sell_prices.csv* duomenų, agreguojant kainas pagal dieną ir skaičiuojant 7 dienų slankųjį vidurkį. Šis kintamasis buvo bendras visoms prognozuojamoms eilutėms, atspindintis bendrą rinkos kainų tendenciją.

Šių regresorių reikšmės buvo naudojamos tiek treniravimo laikotarpiu, tiek ateities prognozėse – todėl visos jų laiko eilutės buvo iš anksto prognozuojamos vienmačiais modeliais.

2.2.4. Laiko eilučių stacionarumo įvertinimas

Kadangi *ARIMAX* modeliai remiasi prielaida, jog bent dalis laiko eilučių komponentų yra stacionarūs, tyrimo metu buvo atliktas kelių tipinių prognozuojamų laiko eilučių bei išorinių regresorių stacionarumo įvertinimas naudojant *ADF* (*Augmented Dickey–Fuller*) testą. *ADF* testas tikrina nulinę hipotezę, kad laiko eilutėje egzistuoja vieneto šaknis, t. y. kad ji nėra stacionari. Jei *p* reikšmė yra mažesnė už pasirinktą reikšmingumo lygį (pvz., 0.05), nulinė hipotezė atmetama, ir laiko eilutė laikoma stacionaria.

ADF testas buvo pritaikytas kelioms reprezentatyvioms prognozuojamoms eilutėms (pvz., *CA_1_FOODS_1*, *TX_2_HOUSEHOLD_2*) bei kai kuriems pagrindiniams išoriniams kintamiesiems (*FOODS*, *avg_price_price*). Visais atvejais *p* reikšmės buvo mažesnės nei 0.01, o tai rodo, jog eilutės gali būti laikomos stacionariomis. Todėl buvo nuspręsta netaikyti papildomų transformacijų (pvz., diferencijavimo), o tai palengvino prognozavimo proceso interpretaciją ir sumažino modelio sudėtingumą.

2.2.5. Papildomos pastabos

Atliekant duomenų paruošimą bei prognozavimo modelių konstravimą buvo atsižvelgta į kelis svarbius aspektus:

- Nebuvo taikyta papildoma normalizacija ar rankinis stacionarumo užtikrinimas, nes pasirinkti modeliai (pvz., *ARIMA*, *ETS*, *PROPHET*) turi integruotas stacionarumo valdymo priemones. Taip pat atlikus *Augmented Dickey–Fuller* testą, paaiškėjo, kad dauguma pagrindinių prognozuojamų laiko eilučių ir regresorių yra stacionarūs, todėl papildomos transformacijos buvo nereikalingos.
- Visi išoriniai regresoriai buvo tiksliai susieti su prognozuojamomis laiko eilutėmis pagal datą, departamentą, kategoriją, valstiją ir parduotuvę. Regresoriai buvo formuojami atsižvelgiant į realią prekių priklausomybę hierarchijoje, užtikrinant semantinį suderinamumą.
- Modeliavimas buvo suplanuotas taip, kad išliktų galimybė tiek agreguoti prognozes į aukštesnius hierarchijos lygius, tiek proporcingai jas paskirstyti į žemesnius lygius. Tokiu būdu užtikrintas suderinamumas su *M5* konkurso duomenų hierarchine struktūra bei palaikomas *Middle-out* prognozavimo strategijos įgyvendinimas.

2.3. Modelių vertinimo metrikos

Modelių veikimo vertinimas yra esminė laiko eilučių prognozavimo tyrimo dalis. Tinkamai parinktos klaidų metrikos leidžia ne tik įvertinti modelių tikslumą, bet ir palyginti skirtingų metodų pranašumus tiek kiekybiškai, tiek kokybiškai. Šiame tyrime buvo naudojamas kelių metrikų rinkinys, siekiant išvengti priklausomybės nuo vienos klaidos funkcijos savybių ir užtikrinti objektyvesnį modelių vertinimą.

2.3.1. Root Mean Squared Error (RMSE)

RMSE yra viena populiariausių metrikų regresijos ir prognozavimo užduotyse. Ji skiria didesnę svorį didelėms klaidoms, todėl yra jautri išskirtinėms reikšmėms:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2},$$

kur y_t – tikroji reikšmė laiko momentu t , \hat{y}_t – prognozuota reikšmė, n – prognozuojamų taškų skaičius.

Mažesnė RMSE reikšmė indikuoja didesnę prognozės tikslumą. Tačiau ši metrika yra priklausoma nuo skalės, todėl netinka lyginti tarp skirtingų laiko eilučių ar produktų.

2.3.2. Median Absolute Error (MDAE)

MDAE skaičiuoja medianą vietoje vidurkio, todėl yra atspari išskirtinėms reikšmėms:

$$MDAE = \text{median}(|y_t - \hat{y}_t|).$$

MDAE yra naudinga vertinant modelio stabilumą bendroje prognozės kokybėje, nes neleidžia pavienėms klaidoms iškraipyti bendro vaizdo.

2.3.3. Mean Absolute Scaled Error (MASE)

MASE yra mastelio atžvilgiu nepriklausoma metrika, leidžianti palyginti prognozės kokybę tarp skirtingų eilučių ar net duomenų rinkinių. Ji lygina modelio absoliučią klaidą su bazinio naiviojo modelio klaida:

$$MASE = \frac{\frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|}{\frac{1}{n-1} \sum_{t=2}^n |y_t - y_{t-1}|}.$$

$MASE < 1$ reiškia, kad modelis geresnis nei naivus prognozavimas. $MASE > 1$ – blogesnis.

2.3.4. Symmetric Mean Absolute Percentage Error (SMAPE)

SMAPE vertina prognozės tikslumą procentine forma, tačiau naudoja simetrišką formulę, kad sumažintų šališkumą tarp mažų ir didelių reikšmių:

$$SMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{(|y_t| + |\hat{y}_t|)/2}.$$

SMAPE visada išreiškiama procentais, ir vertybės svyruoja tarp 0% (idealu) ir 200%. Tai naudinga, kai reikia vertinti klaidas, nepriklausomai nuo originalių reikšmių mastelio.

2.3.5. Mean Arctangent Absolute Percentage Error (MAAPE)

MAAPE – tai naujesnė metrika, siūloma kaip alternatyva *MAPE* ir *SMAPE*, nes ji geriau valdo mažas reikšmes bei išvengia padalijimo iš nulio:

$$MAAPE = \frac{1}{n} \sum_{t=1}^n \arctan \left(\left| \frac{y_t - \hat{y}_t}{y_t} \right| \right).$$

Šios metrikos reikšmės yra tarp 0 ir $\frac{\pi}{2}$ (maždaug 1.57). Kadangi funkcija *arctangent* yra ribota, metrika yra stabili ir lengviau interpretuojama, ypač kai $y_t \approx 0$.

2.3.6. Weighted Root Mean Squared Scaled Error (WRMSSE)

WRMSSE – tai specifinė metrika, sukurta ir taikyta *M5* prognozavimo konkurse, skirta objektyviai vertinti prognozių kokybę esant hierarchiniams duomenims. Skirtingai nei standartinės klaidų metrikos, *WRMSSE* ne tik normalizuoja duomenų mastelius, bet ir suteikia didesnę svarbą prognozėms, kurios turi didesnę verslo reikšmę (pvz., generuoja daugiau pardavimų).

Bendra *WRMSSE* formulė:

$$WRMSSE = \sum_{i=1}^N w_i \cdot RMSSE_i,$$

kur N – prognozuojamų eilučių skaičius, w_i – kiekvienos eilutės svoris, susijęs su jos reikšmingumu (pvz., pardavimų verte), $RMSSE_i$ – šakninis vidutinis kvadratinis skalės paklaidos įvertis konkrečiai eilutei, apskaičiuojamas kaip:

$$RMSSE_i = \sqrt{\frac{\sum_{t=n+1}^{n+h} (y_{t,i} - \hat{y}_{t,i})^2}{\frac{1}{n-1} \sum_{t=2}^n (y_{t,i} - y_{t-1,i})^2}},$$

kur $y_{t,i}$ – tikroji reikšmė, $\hat{y}_{t,i}$ – prognozuota reikšmė, n – mokymo laikotarpio ilgis, h – prognozuojamų dienų skaičius.

M5 konkurse kiekviena laiko eilutė turėjo svorį, apskaičiuotą pagal:

- pardavimų vertę paskutinėmis 28 dienomis (matuojama pinigais: „sell_price“ \times „units_sold“),
- hierarchinę poziciją (kai kurios eilutės yra produktai, kitos – visos kategorijos ar regionai).

Tai reiškia, kad netiksli prognozė svarbiai eilutei (pvz., „FOODS“ kategorijai Kalifornijoje) turėjo daug didesnę poveikį galutiniam rezultatui nei mažareikšmės prekės nišinėje parduotuvėje.

WRMSSE leidžia įvertinti ne tik atskirų prognozių tikslumą, bet ir kaip modelis elgiasi skirtingais hierarchijos lygmenimis (pvz., prekė \rightarrow departamentas \rightarrow kategorija \rightarrow regionas \rightarrow visa šalis). Be to, ši metrika yra jautri tiek prognozės kokybei, tiek bazių reikšmių pokyčiams, todėl labai tinkama daugiakopėms hierarchijoms vertinti.

Kadangi RMSSE daliklis susijęs su vidutiniu pokyčiu, skirtingos pardavimų apimtys nesukelia metrikos iškraipymo. Tai ypač svarbu, kai vienos prekės parduodamos tūkstančiais, o kitos – vos keli vienetai.

Naudodama svorius, WRMSSE pabrėžia, kad verslui svarbios prognozės (pvz., populiarios prekės) turi būti tikslios, o mažareikšmės prognozės nėra tokios kritinės.

Tyrimo metu WRMSSE buvo naudojama:

- kiekvienam taikytam „benchmark“ modeliui įvertinti,
- ARIMAX modelio rezultatams įvertinti atskirai kiekviename hierarchijos lygyje,
- lyginamajai analizei tarp modelių atlikti,
- pagrįsti, kaip tikslesnės prognozės gali prisidėti prie geresnio atsargų valdymo verslo kontekste.

Taigi WRMSSE yra ne tik techninė, bet ir strateginė metrika, puikiai atspindinti tiek prognozavimo kokybę, tiek jos praktinį poveikį verslui. Dėl šios priežasties ji yra esminis vertinimo įrankis šiame tyrime.

2.3.7. Modelių reitingavimo strategija

Vertinant regresorių prognozavimo modelius, kiekviena iš aukščiau aprašytų penkių metrikų (RMSE, MDAE, MASE, SMAPE, MAAPE) buvo naudojama reitingavimui:

- Kiekvienam regresoriui visi taikyti modeliai buvo įvertinti pagal visas metrikas.
- Modeliams buvo priskirtas reitingas (1 – geriausias modelis pagal konkrečią metriką).
- Be klaidų metrikų, papildomai buvo fiksuojamas vykdymo laikas (sekundėmis), ir taikomas toks pat reitingavimas.

Galutinė vidutinė reitingų vertė buvo naudojama atrinkti geriausią modelį, užtikrinant balansą tarp tikslumo ir skaičiavimo efektyvumo. Tokia strategija leido adaptuotis prie regresorių skirtingų savybių bei automatizuoti optimalaus modelio pasirinkimą kiekvienam išoriniam kintamajam.

Tolimesniuose skyriuose pateikiama prognozių rezultatų analizė pagal šias metrikas, įskaitant *WRMSSE* įverčius kiekvienam hierarchijos lygiui bei lyginamąją analizę su alternatyviais sprendimais.

2.4. Skaičiavimų aplinka ir techniniai sprendimai

Visi duomenų apdorojimo, prognozavimo modelių treniravimo ir rezultatų analizės žingsniai buvo atlikti naudojant R programavimo aplinką. Darbe naudotos šios pagrindinės bibliotekos:

- *forecast, fable, prophet* – laiko eilučių modeliavimui,
- *dplyr, tidyr, lubridate* – duomenų transformacijai ir filtravimui,
- *tictoc, microbenchmark* – skaičiavimo laikui matuoti.

3. Tyrimo rezultatai

Šiame skyriuje pristatomi tyrimo metu gauti rezultatai, pradedant naudoto duomenų rinkinio charakteristikomis ir baigiant taikytų modelių prognozių analizėmis. Pateikiami tiek aprašomieji statistiniai duomenų tyrimai, tiek kiekybiniai prognozavimo rezultatų įverčiai, gauti taikant pasirinktus metodus. Skiriamas dėmesys rezultatų interpretacijai bei palyginimui su alternatyviais modeliais, taip pat įvertinama praktinė prognozių nauda mažmeninės prekybos kontekste.

Pirmiausia apžvelgiamas naudotas duomenų rinkinys, jo struktūra ir ypatybės, kurios turėjo įtakos tolesniems sprendimams ir modelių pasirinkimui.

3.1. M5 duomenų rinkinio aprašymas

Tyrime buvo naudojamas M5 prognozavimo konkurso duomenų rinkinys, paskelbtas „Kaggle“ platformoje 2020 metais. Duomenys atspindi penkerius metus trukusius „Walmart“ tinklo kasdienes pardavimus trijose JAV valstijose – Kalifornijoje (CA), Teksase (TX) ir Viskonsine (WI). Rinkinys skirtas daugiamačių ir hierarchinių laiko eilučių prognozavimui, todėl pasižymi išskirtinai turtinga struktūra tiek laiko, tiek prekinės, tiek geografinės hierarchijos požiūriu.

3.1.1. Duomenų failų struktūra

M5 rinkinį sudaro trys pagrindiniai CSV failai, kurių struktūra pateikta 5 lentelėje.

5 lentelė. Trijų pagrindinių M5 rinkinio CSV failų struktūra

CSV failo pavadinimas	CSV failo aprašymas	Stulpeliai CSV faile
<i>sales_train_validation.csv</i>	Šiame faile pateikti dienos pardavimai kiekvienai prekei ir parduotuvei nuo pirmosios dienos (d_1) iki paskutinės mokymo laikotarpio dienos (d_{1913}).	<ul style="list-style-type: none">• <i>item_id</i> – prekės identifikatorius,• <i>dept_id</i> – departamento ID (pvz., <i>FOODS_1</i>),• <i>cat_id</i> – kategorijos ID (pvz., <i>FOODS</i>),• <i>store_id</i>, <i>state_id</i> – parduotuvės ir valstijos identifikatoriai,• <i>d_1</i> – <i>d_1913</i> – kiekvienos dienos pardavimų duomenys.
<i>calendar.csv</i>	Šiame faile pateikiama informacija apie kalendorines datas.	<ul style="list-style-type: none">• <i>date</i> – faktinė data (yyyy-mm-dd),• <i>wm_yr_wk</i> – savaitės kodas (naudojamas susiejant su kainų failu).• <i>event_name_1</i>, <i>event_type_1</i>, <i>event_name_2</i>, <i>event_type_2</i> – šventės ar įvykiai.
<i>sell_prices.csv</i>	Šiame faile pateikiamos produktų kainos pagal laiką ir parduotuves.	<ul style="list-style-type: none">• <i>store_id</i>, <i>item_id</i> – nurodo kainos objektą,• <i>wm_yr_wk</i> – savaitės kodas (naudojamas susiejimui),• <i>sell_price</i> – konkretaus produkto kaina konkrečioje parduotuvėje konkrečią savaitę.

3.1.2. Duomenų hierarchija

M5 duomenys pasižymi daugiapakopiu hierarchiniu pobūdžiu. Viso yra 12 hierarchijos lygių, apimančių 30490 laiko eilučių. Kiekviena laiko eilutė priklauso vienam ar keliems aukštesniems lygmenims. Apibendrintai hierarchija gali būti schematizuojama taip:

Visas tinklas (Level 1) → **Valstijos** (CA, TX, WI) → **Parduotuvės** (10 skirtingų) → **Kategorijos** (*FOODS, HOBBIES, HOUSEHOLD*) → **Departamentai** (*FOODS_1, FOODS_2* ir t.t.) → **Prekės** (3049)

Toks hierarchinis išdėstymas leidžia taikyti agregavimo bei prognozių suderinimo metodus (pvz., *bottom-up, top-down, middle-out*).

3.1.3. Tyrime naudotų duomenų specifikacija

Tyrime buvo naudojami visi *sales_train_validation.csv* failo įrašai, aprėpiantys visų prekių pardavimus visose parduotuvėse per visą mokymo laikotarpį (1913 dienų). Nors pradiniai duomenys yra smulkiausiame (12-ajame) hierarchijos lygyje, prognozės buvo orientuotos į 9-ąjį lygį, t. y. kiekvienos parduotuvės–departamento kombinacijos dieninius pardavimus. Vėliau šios prognozės buvo naudojamos visų kitų hierarchijos lygių rezultatams išgauti taikant *Middle-out* strategiją.

Laiko eilučių papildymui ir kontekstualizavimui buvo suformuoti keli išoriniai regresoriai, sukurti remiantis kelių duomenų šaltinių informacija:

- *calendar.csv* naudotas identifikuoti šventinėms dienoms ir sudaryti binarinį *is_holiday* kintamąjį.
- *sell_prices.csv* panaudotas apskaičiuoti kainų dinamikos regresorių – 7 dienų slankųjį vidurkį, agreguotą kategorijų lygiu, imituoju bendras kainų tendencijas rinkoje.
- Iš *sales_train_validation.csv* duomenų suformuoti papildomi agreguoti pardavimų regresoriai:
 - Departamento pardavimai (L5),
 - Kategorijos pardavimai (L4),
 - Valstijos pardavimai (L2),
 - Parduotuvės pardavimai (L3),
 - Valstijos–kategorijos pardavimai (L8).

Šie regresoriai buvo priskirti kiekvienai prognozuojamai laiko eilutei atsižvelgiant į jos priklausymą tam tikram departamentui, kategorijai, parduotuvei ir valstijai.

Be to, buvo naudojamas ir *sales_test_validation.csv* failas, kuriame pateikti tikrieji testavimo laikotarpio pardavimai. Šis failas pasitarnavo kaip pagrindas *WRMSSE* tikslumo metrikai apskaičiuoti, įvertinant prognozių kokybę visos hierarchijos mastu.

Visos laiko eilutės buvo suderintos pagal kalendoriaus datas, transformuotos į ilguosius ir plačiuosius formatus, užpildytos trūkstamos reikšmės (naudojant *values_fill = 0*), o duomenų jungimas atliktas su *dplyr*, *tidyr* ir kitomis R aplinkos priemonėmis.

3.1.4. Duomenų savybės

- Laiko aprėptis: ~6 metai, 1913 dienų mokymo laikotarpis + 28 dienos testavimui.
- Duomenų dažnis: dieninis.
- Pardavimų pasiskirstymas: kai kurios prekės rodo sezoniškumą (pvz., šventinių prekių), kai kurios pasižymi nereguliaria paklausa.
- Kintamųjų tipai: skaitiniai (pardavimai, kaina), nominalūs (ID), binariniai (*is_holiday*).

3.1.5. Duomenų išskirtinumai

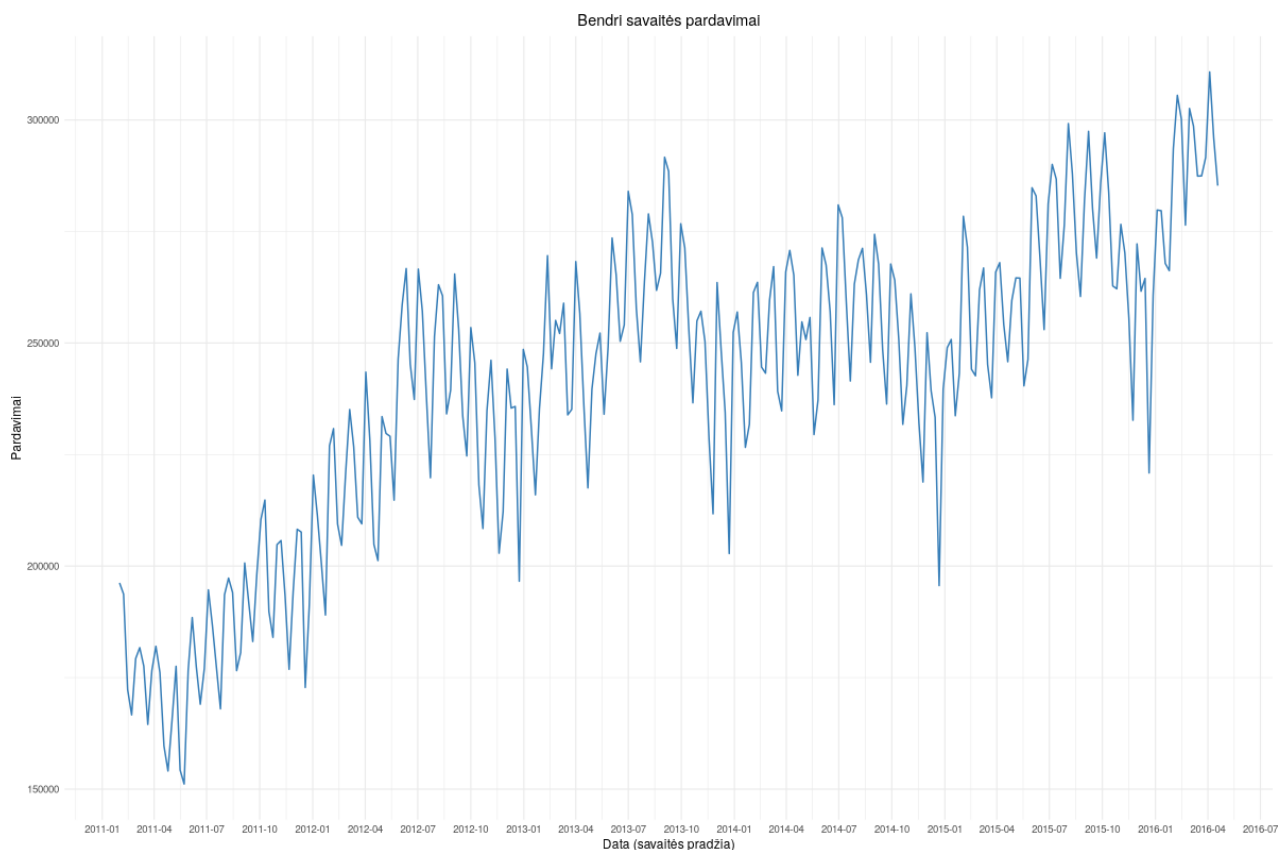
- Kai kurių prekių paklausa yra intermituojanti – ilgą laiką nulis, po to staigus pardavimas.
- Dalis prekių rodo stiprų sezoniškumą (pvz., gruodžio mėn. pakilimai).
- Duomenys iš esmės atspindi realios mažmeninės prekybos operacijas, todėl gali būti interpretuojami kontekstiškai.

3.2. Pirminė duomenų analizė

Šiame poskyryje analizuojami *M5* konkurso duomenys, siekiant geriau suprasti jų struktūrą bei išryškinti dėsningumus, galinčius turėti įtakos prognozavimo tikslumui.

3.2.1. Pardavimų dinamika laike

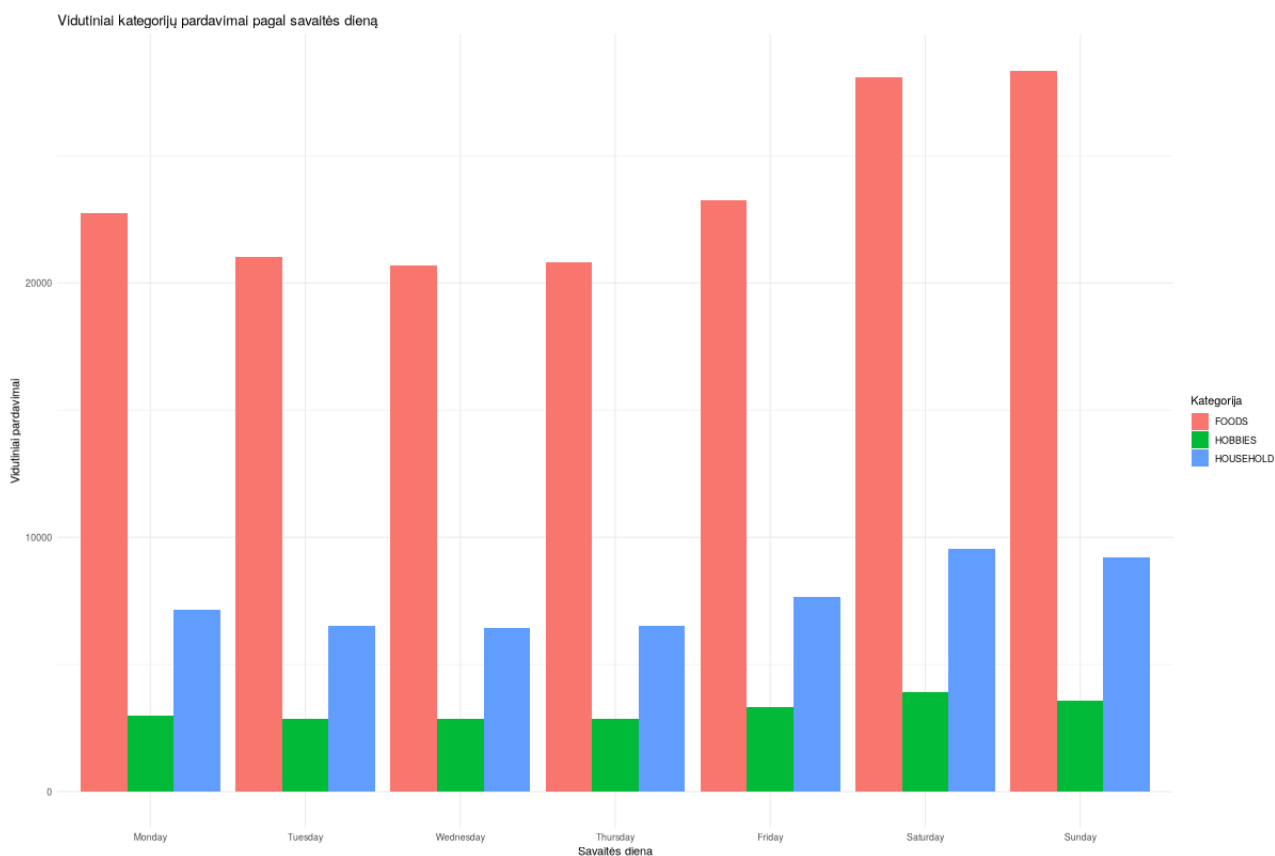
Bendri visų prekių pardavimai laikui bėgant pateikti 1 pav. Pastebima aiški sezoniškumo įtaka, su reikšmingais nuosmukiais metų pabaigoje. Taip pat pastebima, jog kiekvienais metais prekių pardavimai tolygiai didėja.



1 pav. Kiekvienos savaitės pardavimų grafikas

3.2.2. Pardavimų pasiskirstymas pagal savaitės dieną

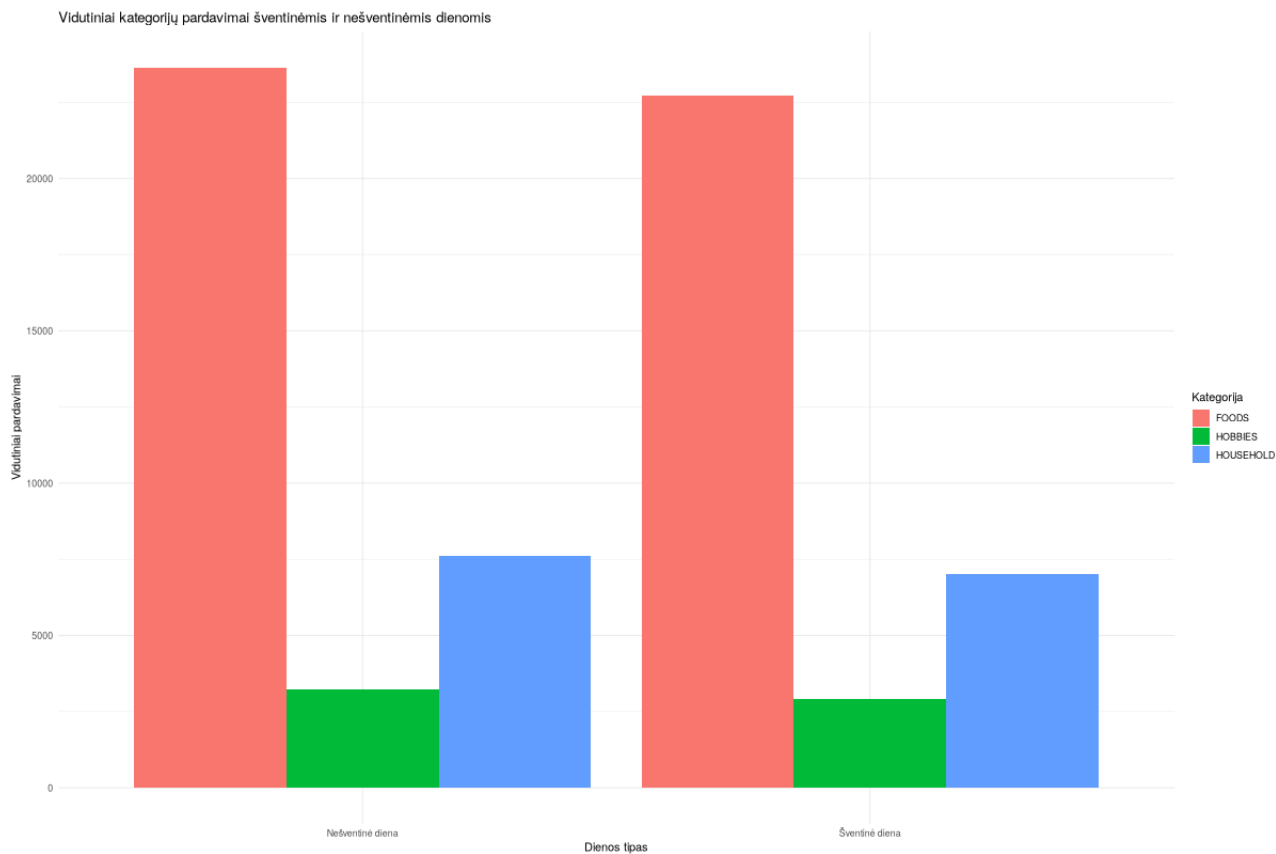
Apskaičiuoti vidutiniai pardavimai pagal savaitės dieną rodo, kad vartotojų aktyvumas nestipriai priklauso nuo savaitės eigos. Daugiausiai pardavimų fiksuojama sekmadieniais ir šeštadieniais, o mažiausiai – trečiadieniais ir ketvirtadieniais. Tai svarbu modelių treniravimui, ypač trumpalaikėms prognozėms. Be to *FOODS* kategorija fiksuoja žymiai daugiau pardavimų nei kitos kategorijos ir tarp pardavimų pastebima didesnė dinamika. Kategorijų pardavimų pagal savaitės dieną grafikas pateiktas 2 pav.



2 pav. Vidutinių kategorijų pardavimų pagal savaitės dieną grafikas

3.2.3. Šventinių dienų įtaka

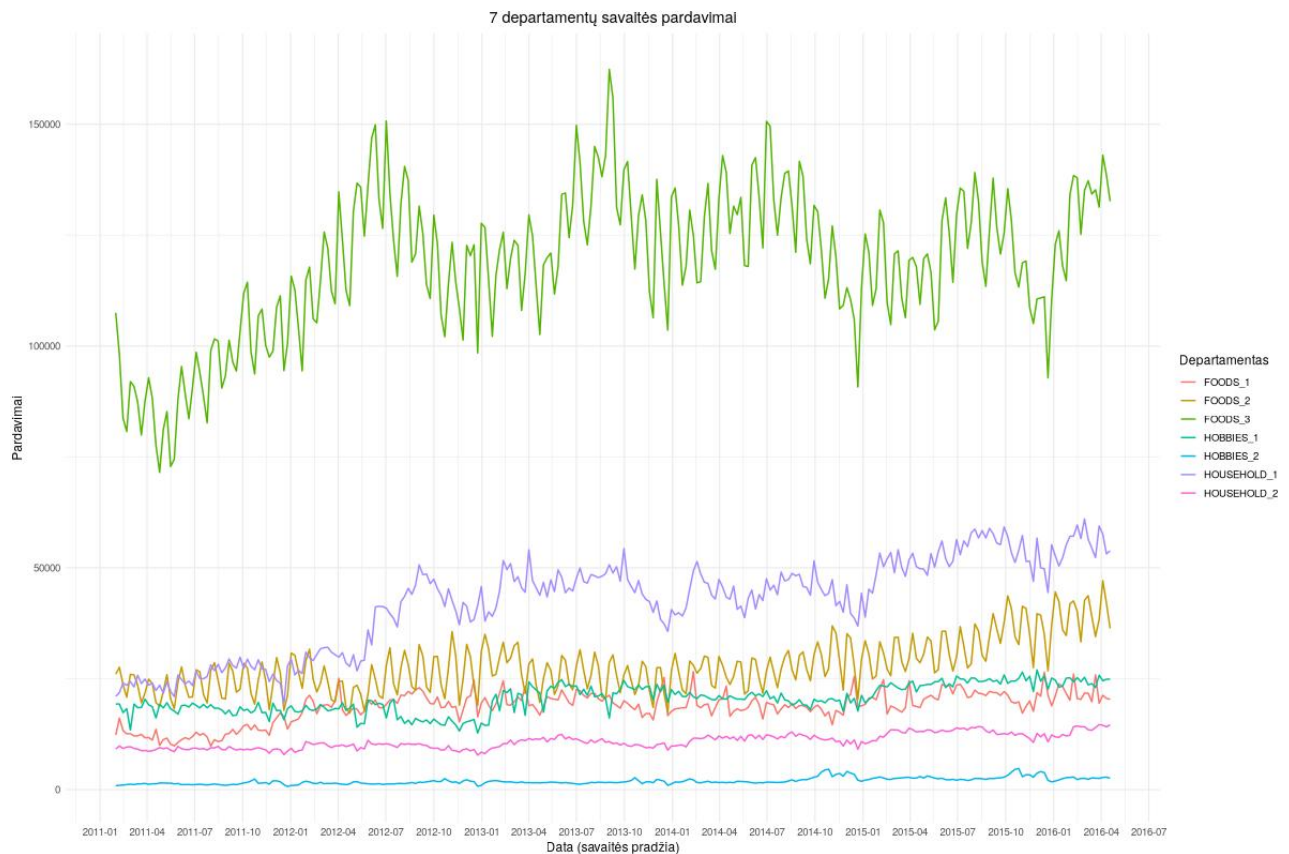
Lyginant šventinių ir nešventinių dienų vidutinius pardavimus, didelis skirtumas nebuvo pastebėtas. Pardavimai nešventinėmis dienomis buvo šiek tiek didesni. Kategorijų pardavimų pasiskirstymas pateiktas 3 pav.



3 pav. Vidutinių kategorijų pardavimų pasiskirstymo pagal šventes grafikas

3.2.4. Pardavimai pagal departamentus

Stebint departamentų pardavimus per laiką, matoma, kad kai kurie departamentai (pvz., *FOODS_3* ar *HOUSEHOLD_1*) ne tik fiksuoja daugiau pardavimų, bet ir turi labiau išreikštas sezoniškumo tendencijas nei kiti. Taip pat departamentų pardavimų amplitudė labai skiriasi – tai patvirtina, kad hierarchinė struktūra (pvz., *Bottom-up* ar *Top-down* modeliavimas) yra svarbi prognozavimo tikslumui. Departamentų pardavimų grafikas pateiktas 4 pav.



4 pav. Departamentų pardavimų grafikas

Atlikta duomenų analizė parodė, jog *M5* duomenų rinkinio pardavimai pasižymi tiek sezoniskumu, tiek struktūriniais skirtumais tarp kategorijų ir departamentų. Nors savaitės dienos ir šventinės dienos turi tam tikrą įtaką pardavimų lygiui, didžiausi skirtumai pastebimi tarp skirtingų prekių grupių. Tai rodo, jog prognozuojant svarbu atsižvelgti ne tik į laikines tendencijas, bet ir į duomenų hierarchinę struktūrą bei skirtingų segmentų elgseną. Toks kontekstas yra būtinas tinkamam modelių treniravimui ir prognozių interpretavimui.

3.3. Etaloninių modelių rezultatų analizė

Modelių palyginimui buvo naudota *WRMSSE* (*Weighted Root Mean Squared Scaled Error*) metrika, kuri vertina prognozių tikslumą atsižvelgiant į hierarchinę struktūrą, sezoninį skaidymą ir kiekvieno elemento svarbą bendrame pardavimų kontekste. Skirtingai nei klasikinės *RMSE* ar *MAPE* metrikos, *WRMSSE* yra jautresnė struktūriniam duomenų pasiskirstymui bei suteikia daugiau svorio dažniau ir daugiau parduodamiems elementams. Tai padeda tiksliau įvertinti modelių kokybę realiame verslo scenarijuje.

Lyginami modeliai apima tiek bazinius, tiek pažangesnius metodus – nuo paprasto „naive“ iki klasikinių laiko eilučių modelių (pvz., *ARIMA*), bei mašininio mokymosi algoritmų (*Random Forest*, *MLP*). Vertinimas atliktas visose 12 hierarchijos lygmenyse, kurie apima skirtingą informacijos detalumą – nuo visų prekių pardavimų iki konkretaus produkto vienoje parduotuvėje.

Buvo palyginti šie modeliai:

- Naivieji modeliai (*Naive*, *Seasonal Naive*)

- Statistiniai modeliai (*Moving Average, Exponential Smoothing, ARIMA*)
- Giluminio mokymosi modeliai (*MLP, LSTM* ir kt.)
- Hierarchiniai modeliai (pvz., *Bottom-Up* versijos)

6 lentelė. Etaloninių modelių *WRMSSE* rezultatai

Lygmuo	Naive	sNaive	SES	MA	Croston	optCroston	SBA	TSB	ADIDA	iMAPA	ES_bu	ARIMA_bu	MLP_1	RF_1
L1	1.488	0.751	1.137	1.186	1.120	1.120	1.182	1.134	1.146	1.142	0.685	1.103	1.160	1.123
L2	1.517	0.754	1.116	1.155	1.095	1.096	1.149	1.113	1.120	1.118	0.697	1.077	1.146	1.116
L3	1.535	0.799	1.116	1.147	1.090	1.094	1.140	1.113	1.117	1.116	0.733	1.079	1.169	1.141
L4	1.520	0.762	1.136	1.175	1.118	1.118	1.177	1.132	1.142	1.139	0.700	1.105	1.156	1.121
L5	1.606	0.817	1.200	1.239	1.167	1.168	1.216	1.192	1.201	1.200	0.762	1.152	1.221	1.184
L6	1.540	0.783	1.115	1.143	1.094	1.096	1.143	1.111	1.116	1.115	0.718	1.079	1.139	1.111
L7	1.614	0.830	1.155	1.185	1.127	1.129	1.165	1.149	1.153	1.153	0.764	1.107	1.181	1.155
L8	1.539	0.839	1.093	1.114	1.073	1.077	1.111	1.091	1.092	1.092	0.748	1.060	1.135	1.113
L9	1.562	0.892	1.111	1.130	1.090	1.095	1.115	1.108	1.107	1.108	0.793	1.070	1.151	1.132
L10	1.271	1.032	0.926	0.931	0.943	0.938	0.945	0.922	0.924	0.924	0.833	0.915	0.946	0.946
L11	1.211	1.075	0.893	0.899	0.898	0.898	0.897	0.891	0.890	0.890	0.842	0.880	0.907	0.909
L12	1.165	1.100	0.868	0.874	0.865	0.867	0.861	0.864	0.864	0.865	0.845	0.856	0.879	0.879
Average	1.464	0.870	1.072	1.098	1.057	1.058	1.092	1.068	1.073	1.072	0.760	1.040	1.099	1.078

Pagal 6 lentelėje pateiktus rezultatus:

- Geriausią vidutinį *WRMSSE* rezultatą visuose hierarchijos lygmenyse pasiekė *Bottom-up Exponential Smoothing* modelis, kurio vidurkis siekia 0.76. Tai rodo, jog šis metodas efektyviai prisitaiko prie skirtingų produktų ir parduojuvių kombinacijų ir yra itin naudingas hierarchinėje prognozėje.
- Antrasis pagal tikslumą – *Seasonal Naive* modelis (vidutinis *WRMSSE* – 0.87), kurio rezultatai buvo netikėtai geri, ypač žemesniuose lygmenyse. Tai leidžia daryti prielaidą, kad sezoniniai pasikartojimai duomenyse yra gana stiprūs ir nuspėjami.
- Visų kitų modelių vidutinis *WRMSSE* viršijo vieneta, kas rodo prastesnį bendrą tikslumą lyginant su dviem aukščiau minėtais metodais.
- *Naive* modelis pasirodė prasčiausiai, visų lygių *WRMSSE* reikšmės viršijo 1.1, todėl šis modelis galėtų būti laikomas žemiausia atskaitos riba.
- Antras pagal prastumą buvo *MLP_1* modelis, kuris galimai nėra tinkamas šio tipo laiko eilučių prognozei. Jo rezultatai buvo šiek tiek geresni aukščiausiuose lygmenyse (10-12), kur sumažėja individualios eilučių dispersijos.
- Trečiasis pagal gerumą – *Bottom-up ARIMA* modelis (vidutinis *WRMSSE* – 1.04), kuris pranoko ne tik klasikinį *ARIMA*, bet ir kai kuriuos pažangesnius giluminio mokymosi metodus. Tai leidžia daryti prielaidą, kad *Bottom-up* prognozavimo metodai yra gana veiksmingi, ypač kai kiekviena laiko eilutė turi pakankamai istorinių duomenų ir atitinkamą sezoniskumą.

Dvyliktame hierarchijos lygyje, kuris atitinka mažiausią agregavimo lygį (produktų agreguotų pagal parduotuvę pardavimai), visų modelių *WRMSSE* reikšmės buvo labai artimos viena kitai. Tai gali reikšti, kad:

- Arba šis lygis iš esmės yra prognoziškai sunkiausias ir modeliai konverguoja į panašius ribinius sprendimus,
- Arba duomenų dispersija yra mažesnė, todėl net ir paprastesni metodai sugeba pasiekti priimtina tikslumą.

Vis dėlto, daugelio modelių *WRMSSE* šiame lygyje išliko gana mažas (0.8-0.9), tad galutinė išvada lieka neapibrėžta ir reikalauja detalesnės analizės.

Taigi atlikta analizė parodė, kad:

- Hierarchiniai modeliai, ypač *Bottom-up Exponential Smoothing*, yra itin tinkami šio tipo uždaviniui. Jų gebėjimas prisitaikyti prie kiekvienos laiko eilutės dinamikos, išsaugant struktūrinę hierarchiją, leidžia pasiekti geriausius rezultatus.
- Sezoninis naivus modelis, nepaisant savo paprastumo, pasirodė netikėtai gerai, pabrėždamas aiškią sezoniskumo svarbą duomenyse.
- *Naive* ir *MLP_1* modeliai nepasiteisino – pirmasis dėl savo paprastumo, antrasis – tikriausiai dėl architektūros, kuri nėra tinkama prognozuoti šio tipo hierarchines eiles.
- *Bottom-up* metodų efektyvumas (pvz., naudojant *ARIMA*) rodo, kad lokalios prognozės gali būti naudingos, jei tinkamai agreguojamos.

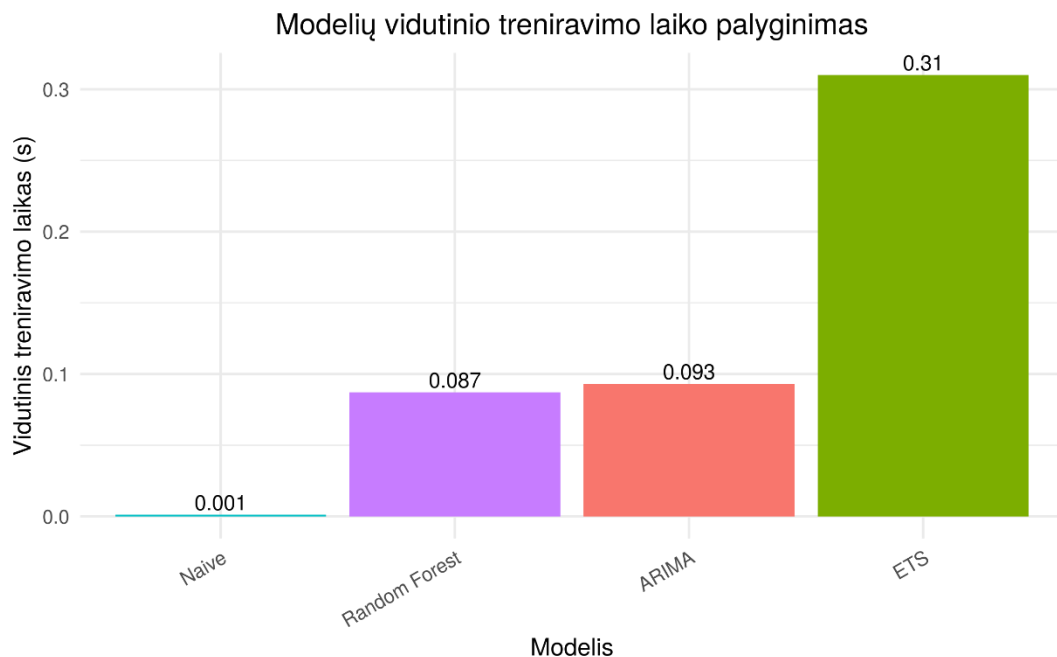
Bendrai paėmus, sudėtingesni modeliai nebūtinai garantuoja geresnį rezultatą. Ši analizė atskleidė, kad modelio paprastumas ir gebėjimas išnaudoti sezoniskumą gali būti stipresni veiksniai nei gilios architektūros sudėtingumas.

3.4. Modelių taikymo praktikoje ir resursų poreikio analizė

Tyrimo metu išbandyti prognozavimo modeliai pasižymi ne tik skirtingu tikslumu, bet ir labai skirtingais skaičiavimo bei pritaikymo aspektais. Tam, kad prognozavimo metodai būtų naudingi praktinėse situacijose, svarbu įvertinti ne tik rezultatų kokybę, bet ir resursų poreikius, modeliavimo trukmę bei galimybę automatizuoti procesus didelio masto sistemose.

3.4.1. Modelių palyginimas pagal treniravimo sudėtingumą ir resursus

Eksperimentiškai palyginus modelius, kurių resursų poreikis yra mažiausias pastebimas aiškus treniravimo laiko skirtumas tarp modelių. Rezultatai pateikti 5 pav.



5 pav. Modelių vidutinio treniravimo laiko vienai laiko eilutei palyginimo grafikas

Galima pastebėti, jog visi šie modeliai veikia gan greitai. Gal tik *Exponential Smoothing* metodas nebūtų pats praktiškiausias (apie tris kartus lėtesnis nei *Random Forest* bei *ARIMA*) jei laiko eilučių būtų daug. Pavyzdžiui, pagal atliktą eksperimentą 100 tūkstančių laiko eilučių prognozavimas užtruktų apytiksliai daugiau nei 8 valandas.

Apibendrinant – kai reikia sparčių, patikimų ir lengvai palaikomų prognozių, *Bottom-up Exponential Smoothing* ir *Seasonal Naive* tampa itin patrauklūs pasirinkimai. Tuo tarpu giluminio mokymosi modeliai reikalauja žymiai daugiau resursų ir sudėtingesnės infrastruktūros.

3.5. Sukurto *ARIMAX* modelio aprašymas ir taikymo principas

ARIMAX (*AutoRegressive Integrated Moving Average with eXogenous variables*) modelis leidžia atlikti laiko eilučių prognozavimą, pasitelkiant papildomą informaciją apie su laiko eilute susijusius išorinius veiksnius (regresorius). Šio darbo tikslui buvo sukurtas *ARIMAX* modelis, kuriam pritaikyti net 7 išoriniai regresoriai, siekiant pagerinti prognozavimo tikslumą sudėtingame *M5* hierarchinių duomenų rinkinyje.

Naudoti išoriniai regresoriai:

1. Šventinės dienos indikatorius (*is_holiday*)
2. Atitinkamo departamento pardavimai
3. Atitinkamos kategorijos pardavimai
4. Atitinkamos valstijos pardavimai
5. Atitinkamos parduotuvės pardavimai
6. Valstijos-kategorijos pardavimai
7. Kainų dinamikos rodiklis (*avg_price_price*), apibrėžtas kaip 7 dienų slankusis vidurkis

Prognozavimo pradžioje modelis buvo išbandytas žemiausiame (12) lygyje, tikintis tiksliausių rezultatų dėl maksimalios informacijos detalumo. Vis dėlto, gauti rezultatai parodė, kad prognozių

kokybė tokiam lygyje buvo nepakankama ($WRMSSE = 1,1387$). Analoginė situacija pastebėta ir vienuoliktame ($WRMSSE = 0,9574$) bei dešimtam ($WRMSSE = 1,0125$) lygiuose. Šiuose lygiuose modelis aplenkė tik *Naive* bei *Seasonal Naive* metodus. Šių lygių prognozės dažnai kenčia nuo didelio duomenų triukšmo bei nepastovumo. Tokiuose lygiuose geriau veikia etaloniniai modeliai arba diskretūs modeliai, tokie kaip *Poisson*.

Atsižvelgus į tai, buvo nuspręsta pritaikyti *ARIMAX* modelį devintam lygiui (parduotuvių-departamentų lygmuo), t.y. prognozes atlikti pagal kiekvienos prekių kategorijos departamento pardavimus kiekvienoje parduotuvėje. Šiame lygyje duomenų agregacija leidžia sumažinti triukšmą, tačiau išlaikyti pakankamą specifiką, o mažesnis laiko eilučių kiekis (70) padeda efektyviau valdyti skaičiavimo resursus, sumažinti prognozavimo laiką ir padidinti prognozės interpretavimo galimybes. Devintame lygyje gautas reikšmingai geresnis rezultatas – $WRMSSE = 0.5404$ lyginant ne tik su rezultatais žemesniuose lygiuose, bet ir su daugeliu etaloninių modelių. Dėl išorinės informacijos įtraukimo, modelis sugebėjo geriau įvertinti sezoninius šuolius bei specifinius šventinių dienų efektus.

Be pačios $WRMSSE$ reikšmės, galima pastebėti ir šiuos rezultatų aspektus:

- Didžiausi pardavimai užfiksuoti *FOODS_3* ir *HOUSEHOLD_1* departamentuose, kurių modeliuotos reikšmės atitiko istorinių duomenų sezoniškumo struktūrą.
- *HOBBIES* kategorijos prognozės išsiskyrė kiek didesne paklaida, tačiau vis tiek išliko priimtino tikslumo ribose.
- Dienų su šventėmis metu *ARIMAX* modelis sėkmingai sumažino prognozės šuolius, kurie kitu atveju būtų nerealistiškai aukšti ar žemi.

Kadangi išorinių regresorių (išskyrus šventinių dienų indikatorius) ateities reikšmės yra nežinomos, jos buvo prognozuojamos taikant vienmačių laiko eilučių prognozavimo metodus (*ARIMA*, *ETS*, *TBATS*, *Prophet*, *THETA* ir kt.). Kiekvienam regresoriui buvo atrinktas optimalus modelis naudojant slankiojo lango metodiką (trys persidengiantys langai po 12 dienų), atsižvelgiant į penkias prognozės kokybės metrikas ($RMSE$, $MDAE$, $MASE$, $SMAPE$ ir $MAAPE$) bei prognozavimo laiką. Regresoriui buvo parenkamas modelis, kuris pasiekė mažiausią vidutinį reitingą tarp metrikų ir efektyvumo. Kategorijų ir departamentų atrinkti geriausi modeliai bei jų klaidos pateikti 7 lentelėje. Likusiųjų regresorių rezultatai pateikti pirmame priede.

7 lentelė. Išoriniams regresoriams atrinktų modelių klaidos

Regresorius	Modelis	RMSE	MDAE	MASE	SMAPE	MAAPE
<i>FOODS_1</i>	<i>Prophet</i>	646,65	507,67	1,54	15,74	15,28
<i>FOODS_2</i>	<i>ARIMAd</i>	981,62	559,73	1,29	11,54	10,87
<i>FOODS_3</i>	<i>ARIMAh</i>	1792,95	984,82	0,62	6,74	6,59
<i>HOBBIES_1</i>	<i>TBATS</i>	383,19	265,4	0,74	8,16	8,03
<i>HOBBIES_2</i>	<i>Prophet</i>	53,93	24,68	0,92	10,7	10,91
<i>HOUSEHOLD_1</i>	<i>ARIMAd</i>	992,57	666,03	0,9	9,71	9,78
<i>HOUSEHOLD_2</i>	<i>TBATS</i>	211,3	128,59	0,67	6,7	6,5
<i>FOODS</i>	<i>ARIMAh</i>	2717,23	1405,29	0,68	6,7	6,49
<i>HOBBIES</i>	<i>ARIMAh</i>	430,01	201,64	0,77	8,02	7,77
<i>HOUSEHOLD</i>	<i>ARIMAd</i>	1162,35	752,17	0,85	8,87	8,9

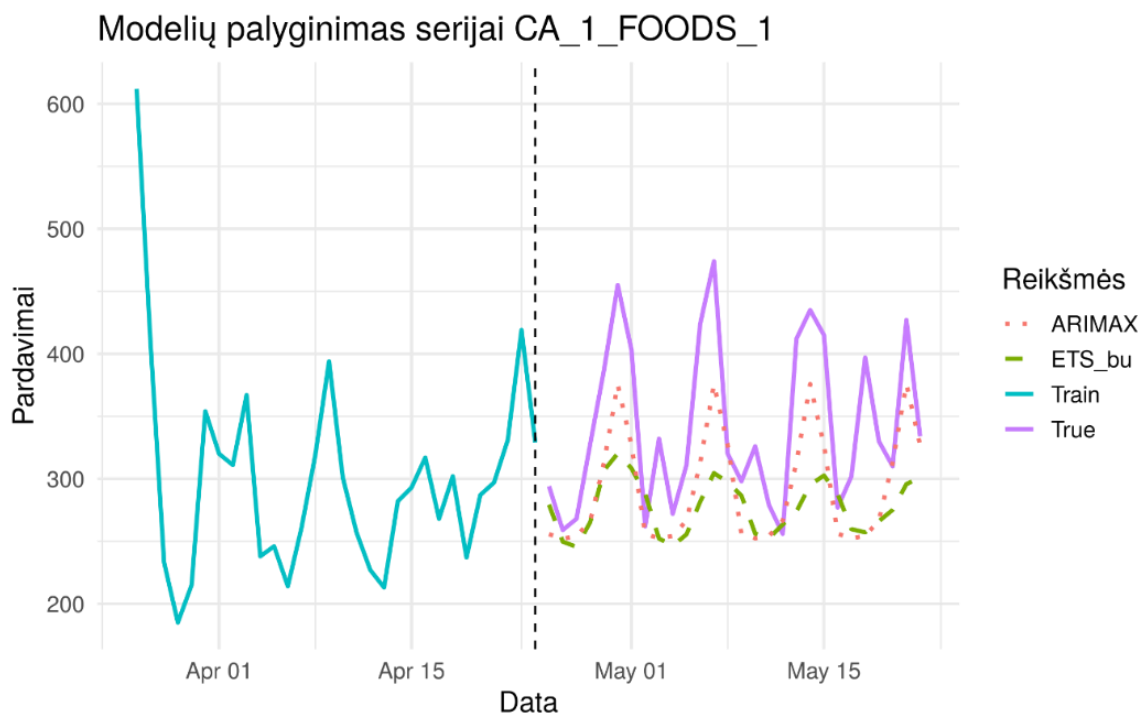
Po to, kai buvo sugeneruotos devinto lygio prognozės, tolesnė prognozavimo strategija buvo grindžiama *Middle-out* principu: aukštesni (1–8) hierarchijos lygiai buvo gauti agreguojant prognozes, o žemesni lygiai (10–12) buvo išskaidyti proporcingai pagal paskutinių treniravimo dienų faktinius pardavimus. Ši strategija leido išlaikyti gerą prognozių kokybę aukštesniuose lygiuose, tuo pačiu sumažinant modelio jautrumą duomenų triukšmui žemesniuose lygmenyse.

3.6. Etaloninių ir *ARIMAX* modelių palyginimas

Toliau atliekamas kiekybinis palyginimas tarp sukurto *ARIMAX* modelio ir keliolikos etaloninių prognozavimo modelių, naudotų *M5* konkurso kontekste. Modelių lyginimui naudota *WRMSSE* metrika.

8 lentelė. *ARIMAX* ir *Bottom-up Exponential Smoothing* modelių lyginimas

Hierarchijos lygis	<i>ARIMAX</i> modelio <i>WRMSSE</i>	<i>ETS_bu</i> modelio <i>WRMSSE</i>
L1 (visas)	0,2643	0,6855
L2 (valstija)	0,4394	0,6972
L3 (parduotuvė)	0,5282	0,7335
L4 (kategorija)	0,3916	0,6997
L5 (departamentas)	0,4464	0,7623
L6 (valstija-kategorija)	0,4895	0,7176
L7 (valstija-departamentas)	0,5327	0,7645
L8 (parduotuvė-kategorija)	0,5544	0,7475
L9 (parduotuvė-departamentas)	0,5404	0,7933
L10 (prekių pardavimai)	1,0901	0,8332
L11 (prekė-valstija)	1,0322	0,8424
L12 (prekė-parduotuvė)	0,9948	0,8455
Vidurkis	0,6087	0,7602



6 pav. *ARIMAX* ir *Bottom-up Exponential Smoothing* modelių prognozavimo palyginimo grafikas

ARIMAX modelis pasiekė geriausius rezultatus 1-9 lygiuose, ypač pirmame ($L1 = 0,2643$), antrame ($L2 = 0,4394$) ir ketvirtame ($L4 = 0,3916$). 10-12 lygiuose pastebimai išauga *WRMSSE*: daugiau nei 1,0, kas rodo, kad šiuose žemiausiuose lygiuose *ARIMAX* modelis praranda tikslumą. Tai tikėtina dėl duomenų triukšmo, mažesnio pardavimų kiekio, diskretumo ir didesnio atsitiktinumo. Šis rezultatas dar kartą patvirtina, kad *Middle-out* strategija buvo efektyviausia, nes sumažina prognozuojamų eilučių skaičių, leidžia modeliui geriau išnaudoti regresorius (kurių reikšmės dažnai nėra pasiekiamos žemiausiems lygiams), sumažina triukšmo ir nulinių reikšmių įtaką.

Nors *ARIMAX* modelis nenaudojo giluminio mokymosi ar sudėtingų ansamblių, jis pasiekė labai konkurencingus rezultatus prognozuodamas nuo 9 lygio. Naudojant 7 išorinius regresorius ir taikant hierarchinį išskleidimą (*Middle-out*), buvo pasiektas vidutinis *WRMSSE* 0,61, pranokstantis etaloninių modelių rezultatus. Tikslingai pasirinkus prognozavimo lygį ir struktūrą, statistinis modelis gali efektyviai konkuruoti su sudėtingesniais metodais.

Papildomai buvo išbandyta alternatyvi regresorių prognozavimo strategija – prognozuojamų kintamųjų reikšmės buvo gaunamos kaip trijų geriausių vienmačių modelių prognozių vidurkis, o ne tik vieno. Tikėtasi, kad šis metodas padidins įvesties duomenų stabilumą ir pagerins *ARIMAX* modelio rezultatus. Vis dėlto eksperimentai parodė, kad *WRMSSE* reikšmės daugeliu atvejų nepasikeitė arba net pablogėjo. Išsamesnis šio metodo aprašymas pateikiamas priede (žr. antrąjį priedą).

3.7. Hierarchinių prognozių nauda mažmeninėje prekyboje

Hierarchinės laiko eilučių prognozės yra itin aktualios mažmeninės prekybos sektoriuje, kur įmonės valdo tūkstančius prekių, išsidėsčiusių skirtingose parduotuvėse, departamentuose ir geografinėse vietovėse. Tokia struktūra sudaro kelių lygių hierarchiją, kurioje sprendimai turi būti priimami ne tik produktų, bet ir visos grandinės mastu.

Vienas pagrindinių tokio tipo prognozavimo privalumų – duomenų konsistencija tarp hierarchijos lygių. Tradiciniai metodai dažnai pateikia prieštaringas prognozes: pavyzdžiui, departamentų lygio prognozė gali nesutapti su visų jų sudarančių produktų sumine prognoze. Tokie neatitikimai gali lemti netikslius atsargų planavimo, logistikos ar marketingo sprendimus.

Atlikus eksperimentinį tyrimą su *M5* duomenų rinkiniu, buvo pastebėta, kad žemiausiuose lygiuose prognozavimo tikslumas dažnai sumažėja dėl didelio atsitiktinumo (pvz., prekė piršta tik kelias dienas), didelio nulinių reikšmių kiekio, sunkumų priskiriant tinkamus išorinius regresorius.

Kitavertus prognozuojant 9 lygio laiko eilutes (parduotuvė + departamentas), rezultatai buvo ne tik žymiai geresni ($WRMSSE = 0.5404$), bet ir stabilesni. Šis lygmuo leido pritaikyti išorinius regresorius, atspindinčius šventines dienas, departamentų bei kategorijų dinamiką, valstijų ir pardotuvių ypatumus, bei kainų pokyčius.

Šiame darbe išbandyti skirtingi hierarchinio prognozavimo metodai – tiek *Bottom-up* (prognozės pradedamos nuo žemiausio lygio ir agreguojamos), tiek *Middle-out* (prognozės generuojamos viduriniame lygyje ir išskaidomos į žemesnius bei agreguojamos į aukštesnius lygius). Būtent *Middle-out* metodas pasirodė optimalus, nes:

- sumažino prognozuojamų laiko eilučių skaičių nuo 30 tūkst. iki 70;
- leido panaudoti daugiau prasmingų išorinių regresorių (kurių nebūtų galima patikimai priskirti pavieniams produktams);
- sumažino triukšmo įtaką prognozėms, ypač žemesniuose lygiuose;
- užtikrino logišką rezultatų struktūrą – suma žemesniuose lygiuose atitiko aukštesnius lygius.

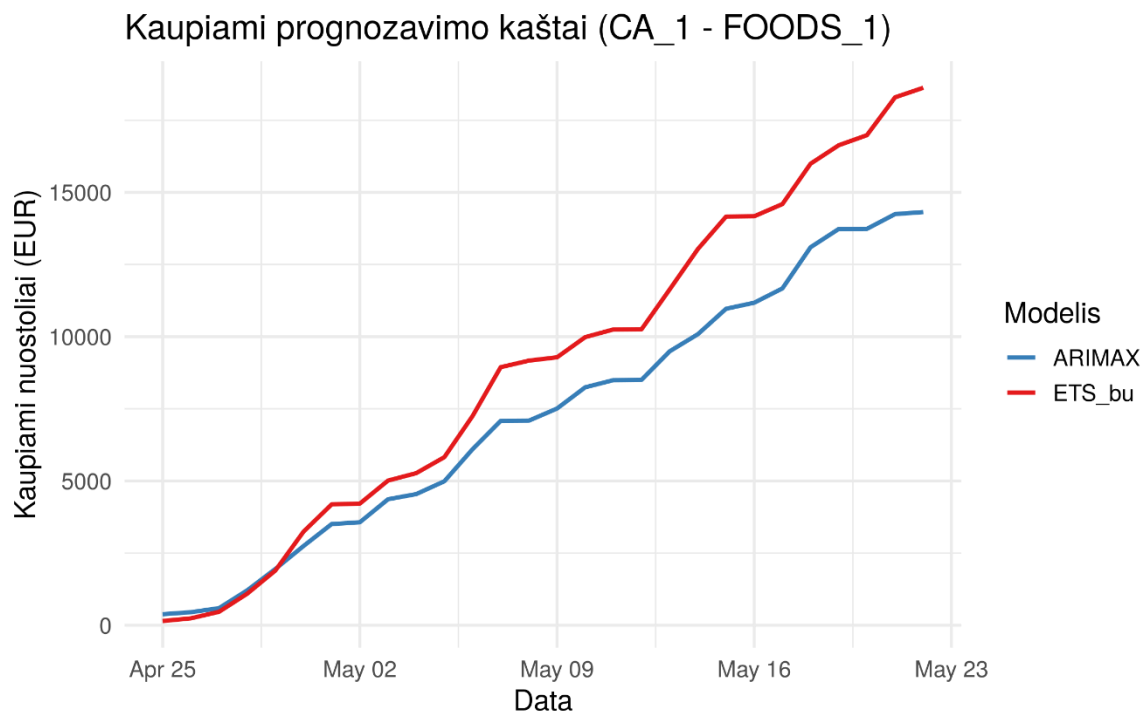
Mažmeninės prekybos kontekste tai reiškia, kad prognozės tampa tinkamos strateginiam planavimui, pvz.:

- užtikrinti tikslias prekių atsargas skirtingose pardotuovėse (parduotuvės-departamento lygmuo),
- prognozuoti paklausos šuolius šventinių laikotarpių metu (naudojant *is_holiday* regresorių),
- reaguoti į kainų pokyčius (naudojant *avg_price_price* regresorių),
- vertinti regioninius paklausos skirtumus (naudojant valstijos-kategorijos ar pardotuovės regresorius).

Be to, prognozių efektyvumas buvo įvertintas per praktinį piniginių kaštų analizės scenarijų, susijusį su atsargų planavimu. Šiame tyrime buvo įvertinti du pagrindiniai rizikos tipai:

- *Stockout* - situacija, kai prognozė buvo per maža ir nepakanka prekių. Šios situacijos kaštai modeliuoti kaip 10 EUR per vienetą (dėl prarasto pelno, klientų nepasitenkinimo ir kt.).
- *Overstock* - situacija, kai prognozė buvo per didelė ir liko nepanaudotų prekių. Kaštai – 1 EUR per vienetą (susiję su sandėliavimo ar nuolaidų išlaidomis).

Naudojant šį scenarijų, buvo apskaičiuoti kaupiamieji dienos nuostoliai pagal skirtingus modelius. Kaip parodyta 7 pav., *ARIMAX* modelis per visą testavimo laikotarpį sukaupė mažesnius prognozavimo kaštus nei tradicinis *Bottom-up Exponential Smoothing* modelis, pabrėžiant jo praktinį pranašumą realaus planavimo sprendimuose.



7 pav. Kaupiamieji prognozavimo kaštai pagal *ARIMAX* ir *ETS_bu* modelius (vienai parduotuvės–departamento serijai)

3.8. Rezultatų interpretacija ir tyrimo ribojimai

Atlikto tyrimo rezultatai atskleidė keletą svarbių įžvalgų apie daugiamačių laiko eilučių prognozavimo metodų taikymą didelio masto hierarchiniuose duomenyse. Pritaikius keletą skirtingų prognozavimo strategijų ir palyginus jų tikslumą su *M5* konkurso etaloniniais modeliais, paaiškėjo, kad tinkamai parinktas prognozavimo lygmuo bei išoriniai regresoriai gali reikšmingai pagerinti prognozės tikslumą.

Nors iš pradžių buvo planuota taikyti *ARIMAX* modelį tiesiogiai žemiausiam – dvyliktam – hierarchijos lygiui, gauti rezultatai ($WRMSSE \approx 1.14$) buvo prastesni nei daugelio paprastų etaloninių modelių. Panaši situacija pasikartojo ir vienuoliktame bei dešimtame lygiuose, kur modelis sunkiai prisitaikė prie itin diskrečių, triukšmingų ir dažnai nulinių laiko eilučių.

Geriausias rezultatas buvo pasiektas prognozuojant devintąjį lygį (parduotuvės ir departamento kombinacijas), kur *ARIMAX* modelio $WRMSSE$ siekė 0.5404. Šis lygmuo pasižymėjo pakankamu duomenų kiekio stabilumu, o tuo pačiu išlaikė ir reikšmingą detalumo lygį, leidžiantį pritaikyti regresorius, susijusius su šventinėmis dienomis, departamentais, kategorijomis, parduotuvėmis, kainų dinamika bei regioniniais skirtumais. Modelio rezultatai buvo išskaidyti į žemesnius lygius ir agreguoti į aukštesnius pasitelkiant *Middle-out* strategiją.

Pagrindiniai rezultatai:

- Dauguma etaloninių modelių pasiekė $WRMSSE > 1$, o tik *Bottom-up Exponential Smoothing* (*ES_bu*) (0,76) ir *Seasonal Naive* (0,87) turėjo konkurencingus rezultatus.
- *ARIMAX* modelis buvo žymiai tikslesnis už visus etaloninius modelius 1-9 lygiuose, tačiau prastesnis 10-12 lygiuose, kur duomenys buvo pernelyg smulkūs, kad regresoriai išlaikytų reikšmingą įtaką.

- Naudoti 7 išoriniai regresoriai leido sumažinti klaidas aukštesniuose lygmenyse, bet jų reikšmingumas smuko žemiausiuose lygiuose dėl padidėjusio duomenų atsitiktinumo.

Tyrimo ribojimai:

- *ARIMAX* modelis pateikia tolydžias prognozes, kurios nėra idealiai pritaikytos diskretiems pardavimų duomenims. Visgi dėl didelio prognozių kiekio ir jų agregavimo rezultatai lieka praktiškai vertingi.
- Modelio treniravimas su 7 regresoriais buvo apskaičiavimo požiūriu intensyvus – reikėjo kruopščiai optimizuoti modelį ir regresorių priskyrimą kiekvienai eilutei.
- Dalis regresorių (pvz., kainų dinamika ar regioninė kategorija) buvo identiški visoms eilutėms, todėl jų įtaka ne visada buvo reikšminga kiekvienam atvejui.
- *Middle-out* metodas reikalauja papildomų veiksmų žemyn skalaujant prognozes, o tai gali įnešti tam tikrą papildomą prognozės paklaidą, ypač kai duomenys žemiau yra itin triukšmingi.

Apibendrinant galima teigti, kad geriausi rezultatai buvo pasiekti ten, kur duomenys buvo pakankamai agreguoti, tačiau dar išlaikė struktūrinę reikšmę. Tai rodo, jog prognozavimo tikslumas ne visada priklauso nuo smulkiausios detalės, bet dažnai nuo gebėjimo pritaikyti struktūrinius ryšius tarp eilučių per išorinius kintamuosius.

Išvados

1. Apžvelgus daugiamačių laiko eilučių prognozavimo metodus, nustatyta, kad šiuo metu populiariausi sprendimai apima statistinius (pvz., *ARIMA*, *ETS*), tradicinio mašininio mokymosi (pvz., *Random Forest*, *MLP*), bei giluminio mokymosi (pvz., *LSTM*, *TCN*) metodus. Literatūros analizė atskleidė, kad giluminio mokymosi metodai dažnai reikalauja didelių duomenų ir skaičiavimo resursų, todėl praktikoje vis dar dažnai taikomi pažangesni statistiniai modeliai ar modelių kolektyvai.
2. Eksperimentinis palyginimas su *M5* konkurso duomenimis parodė, kad etaloninių modelių *WRMSSE* rezultatai dažniausiai viršijo 1, išskyrus kelis atvejus (pvz., *Bottom-up Exponential Smoothing*: 0.76, *Seasonal Naive*: 0.87). Tuo tarpu *ARIMAX* modelis su 7 išoriniais regresoriais pasiekė reikšmingai geresnius rezultatus daugelyje lygių – nuo 0.26 iki 0.55 *WRMSSE* reikšmėmis pirmuose devyniuose lygmenyse.
3. Tyrimo metu išnagrinėtos trys pagrindinės hierarchinės prognozavimo strategijos – *Bottom-up*, *Top-down* ir *Middle-out*. Pradžioje buvo siekiama prognozuoti detaliausiam hierarchijos lygyje (*L12*), taikant *Bottom-up* metodą, tačiau dėl didelio triukšmo ir diskretiškumo žemiausiose laiko eilutėse prognozių tikslumas buvo prastas ($WRMSSE \approx 1.14$). Įvertinus šiuos apribojimus, tyrime pasirinkta pereiti prie agreguotesnio prognozavimo lygmens (*L9*) ir taikyti *Middle-out* strategiją, kuri leido efektyviai balansuoti tarp tikslumo ir išorinių regresorių taikymo galimybių. Tai leido ne tik pagerinti rezultatų stabilumą žemuosiuose lygiuose, bet ir išlaikyti duomenų struktūros nuoseklumą visoje hierarchijoje. Šie rezultatai patvirtina, kad prognozavimo strategijos ir lygmens parinkimas yra esminis veiksnys siekiant maksimalaus prognozių tikslumo.
4. Išteklių analizė parodė, kad sudėtingesni prognozavimo metodai, tokie kaip *ARIMAX* su kelių tipų išoriniais regresoriais, yra žymiai skaičiavimų intensyvesni nei paprastesni statistiniai modeliai. Net naudojant paralelizaciją, prognozių generavimas viduriniame hierarchijos lygyje (apie 70 laiko eilučių) pareikalavo daugiau nei valandos skaičiavimų. Tai rodo, kad tokie metodai tinkami periodiniam planavimui ir analizei, tačiau mažiau pritaikomi realaus laiko prognozėms arba operatyviam reagavimui.
5. Remiantis atliktu darbu, rekomenduojama rinktis prognozavimo metodą ir hierarchijos lygį atsižvelgiant į duomenų struktūrą ir praktinį taikymą. Jei laiko eilutės pasižymi dideliu diskretiškumu ir atsiktinumu, tikslinga taikyti prognozavimą viduriniuose ar aukštesniuose lygmenyse. Tokiu būdu galima pasinaudoti išoriniais kintamaisiais, o rezultatus vėliau paskirstyti žemyn remiantis istoriniu pasiskirstymu.

Literatūros sąrašas

1. SEN, Rajat; YU, Hsiang-Fu; DHILLON, Inderjit ir Amazon. Think Globally, Act Locally: A Deep Neural Network Approach to High-Dimensional Time Series Forecasting (2019).
2. SHCHUR, Oleksandr; BILOŠ, Marin ir ÜNNEMANN, Stephan G. INTENSITY-FREE LEARNING OF TEMPORAL POINT PROCESSES. Prieiga per: <https://arxiv.org/abs/2402.19072..>
3. SALINAS, David; BOHLKE-SCHNEIDER, Michael; CALLOT, Laurent; MEDICO, Roberto ir GASTHAUS, Jan. High-Dimensional Multivariate Forecasting with Low-Rank Gaussian Copula Processes (2019). Prieiga per: <http://arxiv.org/abs/1910.03002..>
4. SUNDAR, Syama; WERNER, Lucien D.; BENIDIS, Konstantinos; MERCADO, Pedro; GASTHAUS, Jan ir kt. *End-to-End Learning of Coherent Probabilistic Forecasts for Hierarchical Time Series*. .
5. BALOCH, Mazhar; HONNURVALI, Mohamed Shaik; KABBANI, Adnan; JUMANI, Touqeer Ahmed ir CHAUHDARY, Sohaib Tahir. An Intelligent SARIMAX-Based Machine Learning Framework for Long-Term Solar Irradiance Forecasting at Muscat, Oman. *Energies (Basel)*, vol. 17 (2024), nr. 23, pp. 6118. Prieiga per: <https://www.proquest.com/docview/3144106791..>
6. GASTHAUS, Jan; BENIDIS, Konstantinos; WANG, Yuyang; RANGAPURAM, Syama S.; SALINAS, David ir kt. *Probabilistic Forecasting with Spline Quantile Function RNNs*. .
7. MAKRIDAKIS, Spyros; SPILIOTIS, Evangelos ir ASSIMAKOPOULOS, Vassilios. The M5 competition: Background, organization, and implementation. *International Journal of Forecasting*, vol. 38 (2022), nr. 4, pp. 1325–1336. Prieiga per: <https://www.sciencedirect.com/science/article/pii/S0169207021001187..>
8. MAKRIDAKIS, Spyros; SPILIOTIS, Evangelos ir ASSIMAKOPOULOS, Vassilios. M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, vol. 38 (2022), nr. 4, pp. 1346–1364. Prieiga per: <https://www.sciencedirect.com/science/article/pii/S0169207021001874..>
9. LAI, Guokun; CHANG, Wei-Cheng; YANG, Yiming ir LIU, Hanxiao. Modeling Long-and Short-Term Temporal Patterns with Deep Neural Networks. .
10. RIZVI, Syed Tahir Hussain; KANWAL, Neel; NAEEM, Muddasar; CUZZOCREA, Alfredo ir CORONATO, Antonio. Bridging Simplicity and Sophistication using GLinear: A Novel Architecture for Enhanced Time Series Prediction (2025). Prieiga per: <https://arxiv.org/abs/2501.01087..>
11. CHEN, Huijing ir BOYLAN, John E. *Top-down, bottom-up, and middle-out seasonal forecasting*. University of Salford-Management & Management Sciences Research Institute, 2009..
12. HYNDMAN, Rob J. ir ATHANASOPOULOS, George. *Forecasting*. Second print edition Lexington, Ky: Otexts, 2018. Prieiga per: http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=030400257&sequence=000001&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA..

13. ATHANASOPOULOS, George; AHMED, Roman A. ir HYNDMAN, Rob J. Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting*, vol. 25 (2009), nr. 1, pp. 146–166. Prieiga per: <https://www.sciencedirect.com/science/article/pii/S0169207008000691>..
14. KOURENTZES, Nikolaos ir ATHANASOPOULOS, George. Elucidate structure in intermittent demand series. *European Journal of Operational Research*, vol. 288 (2021), nr. 1, pp. 141–152. Prieiga per: <https://www.sciencedirect.com/science/article/pii/S0377221720304926>..
15. WICKRAMASURIYA, Shanika L.; ATHANASOPOULOS, George ir AND HYNDMAN, Rob J. Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization. *Journal of the American Statistical Association*, vol. 114 (2019), nr. 526, pp. 804–819. Prieiga per: <https://doi.org/10.1080/01621459.2018.1448825>..
16. MAKRIDAKIS, Spyros; SPILIOTIS, Evangelos ir ASSIMAKOPOULOS, Vassilios. The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, vol. 34 (2018), nr. 4, pp. 802.
17. ESHRAGH, Ali; GANIM, Benjamin; PERKINS, Terry ir BANDARA, Kasun. The Importance of Environmental Factors in Forecasting Australian Power Demand. *Environmental modeling & assessment*, vol. 27 (2022), nr. 1, pp. 1–11. Prieiga per: <https://link.springer.com/article/10.1007/s10666-021-09806-1>..
18. OLIVARES, Kin G.; CHALLU, Cristian; MARCJASZ, Grzegorz; WERON, Rafał ir DUBRAWISKI, Artur. Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with NBEATSx. *International Journal of Forecasting*, vol. 39 (2023), nr. 2, pp. 884–900. Prieiga per: <https://www.sciencedirect.com/science/article/pii/S0169207022000413>..
19. WANG, Yuxuan; WU, Haixu; DONG, Jiayang; QIN, Guo; ZHANG, Haoran ir kt. TimeXer: Empowering Transformers for Time Series Forecasting with Exogenous Variables (2024).
20. ADENAN, Moh; HAQ, Mujab Syaiful; NASIR, M. Abd ir SOSECO, Thomas. The Effects of Macroprudential Policies on the Performance of Conventional Banks in Indonesia. *ECONOMICS*, vol. 13 (2024), nr. 1, pp. 369.

Priedai

1 priedas. Likusiems išoriniams regresoriams atrinktų modelių klaidų lentelė

Čia pateikiama likusių išorinių regresorių geriausių atrinktų modelių bei jų klaidų lentelė.

9 lentelė. Likusiems išoriniams regresoriams atrinktų modelių klaidos

Regresorius	Modelis	RMSE	MDAE	MASE	SMAPE	MAAPE
CA	TBATS	1844,81	1340,94	0,78	8,27	8,44
TX	ARIMAh	1095,71	866,53	0,67	8,09	8,19
WI	ARIMAd	2143,79	1490,32	1,2	12,98	12,59
CA_1	TBATS	621,14	447,51	0,76	10,53	10,41
CA_2	ARIMA	872,47	415,2	1,15	14,16	14,37
CA_3	TBATS	703,13	550,02	0,83	9	9,29
CA_4	ARIMAh	261,65	138,97	0,85	7,39	7,15
TX_1	TBATS	398,44	220,17	0,69	9,28	9,24
TX_2	TBATS	456,52	310,55	0,67	9,25	9,5
TX_3	TBATS	359,77	272,55	0,71	8,01	8
WI_1	ARIMA	711,43	499,44	1,11	14,26	14,14
WI_2	HoltW	887,85	636,4	1,4	14	13,58
WI_3	ARIMAh	665,47	535,74	1,03	13,99	13,91
CA_FOODS	TBATS	1212,43	817,02	0,76	8,26	8,41
CA_HOBBIES	TBATS	244,38	171,95	0,87	10,35	10,36
CA_HOUSEHOLD	ARIMA	529,26	301,1	0,74	8,2	7,98
TX_FOODS	TBATS	841,57	595,15	0,69	9,42	9,31
TX_HOBBIES	ARIMAh	150,02	93,35	0,85	10,83	10,5
TX_HOUSEHOLD	TBATS	292,14	209,56	0,74	8,27	8,23
WI_FOODS	HoltW	1749,71	1182,53	1,32	14,99	14,99
WI_HOBBIES	ARIMAh	188,45	133,09	0,96	15,75	15,59
WI_HOUSEHOLD	ARIMAh	542,86	382,29	1,34	16,23	16,47

2 priedas. Alternatyvus regresorių prognozės agregavimo metodas

Šiame priede aprašomas papildomas metodas, išbandytas siekiant pagerinti *ARIMAX* modelio prognozių tikslumą. Vietoje įprastai taikomo metodo, kai kiekvienam išoriniam regresoriui parenkamas vienas geriausias vienmatis prognozavimo modelis (pagal kryžminio patikrinimo metrikas), šiame eksperimente buvo taikomas trijų geriausių modelių prognozių vidurkio metodas.

Metodika:

- Kiekvienai regresoriaus laiko eilutei buvo apskaičiuoti trijų geriausių modelių (pagal *SMAPE/MASE/MAAPE*) prognozės.
- Buvo imtas šių trijų modelių vidurkis kiekvienai dienai.

- Gautos prognozės naudotos kaip įvestis *ARIMAX* modeliui, analogiškai kaip pagrindiniame variante.

Rezultatai:

- Gauti *ARIMAX WRMSSE* rezultatai buvo labai panašūs į tuos, kurie gauti naudojant vieno geriausio modelio prognozes.
- Kai kuriais hierarchijos lygmenimis rezultatai net šiek tiek pablogėjo (pvz., devinto lygmens *WRMSSE* padidėjo nuo 0.5404 iki 0.5531).
- Pagrindinė galima priežastis – vidurkinimas išlygina sezoniškumus ir tendencijas, kurios kartais būna svarbios tikslioms prognozėms.

Kadangi šis metodas nepateikė apčiuopiamų rezultatų, jis nebuvo naudojamas galutiniuose eksperimentuose, tačiau vertas paminėjimo kaip potencialus tolesnių tyrimų objektas.