



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Rutulinių guolių nuspėjamos priežiūros matematinis modelis

Baigiamasis magistro studijų projektas

Vitas Vainauskas

Projekto autorius

doc. dr. Mantas Landauskas

Vadovas

prof. dr. Daiva Dumčiuvienė

Vadovė

Kaunas, 2024



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Rutulinių guolių nuspėjamos priežiūros matematinis modelis

Baigiamasis magistro studijų projektas

Didžiųjų verslo duomenų analitika (6213AX001)

Vitas Vainauskas

Projekto autorius

doc. dr. Mantas Landauskas

Vadovas

prof. dr. Daiva Dumčiuvienė

Vadovė

prof. dr. Evaldas Vaičiukynas

Recenzentas

prof. dr. Vytautas Snieška

Recenzentas

Kaunas, 2024



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Vitas Vainauskas

Rutulinių guolių nuspėjamos priežiūros matematinis modelis

Akademinių 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.

Vitas Vainauskas

Patvirtinta elektroniniu būdu

Vainauskas Vitas. Rutulinių guolių nuspėjamos priežiūros matematinis modelis. Magistro studijų baigiamasis projektas / vadovas doc. dr. Mantas Landauskas, vadovė prof. dr. Daiva Dumčiuvienė; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

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

Reikšminiai žodžiai: nuspėjamoji techninė priežiūra, guolių defektai, sąnaudų ir naudos analizė, regresija, klasifikacija.

Kaunas, 2024. 68 p.

Santrauka

Rutulinių guolių nuspėjamoji priežiūra yra laikoma kritiškai svarbia norint užtikrinti ilgalaikį ir efektyvų mašinų bei įrenginių veikimą. Dėl aukšto efektyvumo ir patikimumo reikalavimų, šie komponentai plačiai naudojami įvairiose pramonės šakose, įskaitant automobilių gamybą, aviaciją ir laivų statybą. Guolių diagnostika ir būklės stebėseną Lietuvoje, palyginti su pasauliniais mastais, tebėra nauja ir mažai ištirta sritis, todėl šis tyrimas prisideda prie svarbaus mokslo ir pramonės plėtros. Nuspėjamosios priežiūros matematiniai modeliai, kurie suteikia galimybę optimizuoti priežiūros procesus, mažinti netikėtus gedimus ir pratęsti įrangos tarnavimo laiką, buvo ištirti šiame darbe. Pagrindinis šio tyrimo tikslas buvo nustatyti mašininio mokymo metodų taikymo efektyvumą guolių defektų aptikimui, defekto vietos ir dydžio nustatymui, naudojant realius duomenis. Tyrimo metu buvo naudojami atsitiktinių miškų, LightGBM ir dirbtinių neuroninių tinklų algoritmai. Rezultatai parodė, kad šie modeliai idealiai identifiko gedimus, pasižymėdami 100 % tikslumu. Nustatant defektų vietas, geriausi rezultatai buvo pasiekti naudojant LightGBM modelį: tikslumas – 96,39 %, makrovidutinio ploto po kreive reikšmė – 0,9951. O įvertinant defektų dydžius, LightGBM modelis užfiksavo mažiausią paklaidą ir determinacijos koeficientą – 0,9153. Tai pat atlikus sąnaudų ir naudos analizę, nustatyta, kad pasiūlytas nuspėjamosios techninės priežiūros sprendimas vienam traukinio vagonui yra finansiškai naudingas. Numatyta, kad nuspėjamos priežiūros sistema atsipirktų per 4,95 metų.

Vainauskas Vitas. A Mathematical Model for Predictive Maintenance of Ball Bearings. Master's Final Degree Project / supervisor Assoc. Prof. Dr. Mantas Landauskas, supervisor Prof. Dr. Daiva Dumčiuvienė; Faculty of Mathematics and Natural Sciences, Kaunas University of Technology.

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

Keywords: predictive maintenance, bearing defects, cost-benefit analysis, regression, classification.

Kaunas, 2024. 68.

Summary

Predictive maintenance of ball bearings is considered critical for the long-term and efficient operation of machinery and equipment. Due to their high performance and reliability requirements, these components are widely used in a wide range of industries including automotive, aviation and shipbuilding. Bearing diagnostics and condition monitoring in Lithuania is still a new and under-researched area compared to the global scale, and this research contributes to an important scientific and industrial development. Mathematical models for predictive maintenance, which enable optimisation of maintenance processes, reduction of unexpected failures and prolongation of equipment lifetime, have been investigated in this work. The main objective of this study was to determine the effectiveness of the application of machine learning methods for bearing defect detection, defect location and defect size using real data. Random forest, LightGBM and artificial neural network algorithms were used in the study. The results showed that these models were ideal for fault identification with 100% accuracy. The LightGBM model gave the best results in identifying defect locations, with an accuracy of 96.39% and a macro-average area under the curve of 0.9951. As for the defect sizes, the LightGBM model recorded the lowest errors and a coefficient of determination of 0.9153. The cost-benefit analysis also showed that the proposed predictive maintenance solution is financially viable per train car. The predictive maintenance system is estimated to pay for itself in 4.95 years.

Turinys

Lentelių sąrašas	8
Paveikslų sąrašas	9
Santrumpų ir terminų sąrašas	11
Įvadas.....	12
1. Literatūros apžvalga	13
1.1. Techninė priežiūra	13
1.2. Nuspėjamos techninės priežiūros svarba.....	14
1.3. Duomenų modalumai nuspėjamos priežiūros kontekste	15
1.3.1. Jutiklių duomenys.....	15
1.3.2. Nuotraukų ir vaizdo įrašų duomenys.....	15
1.3.3. Garso duomenys	16
1.3.4. Tekstiniai duomenys.....	16
1.4. Rutulinių guolių defektai ir defektų priežastys	16
1.4.1. Nuovargis	18
1.4.2. Nusidėvėjimas	18
1.4.3. Korozija	18
1.4.4. Elektrinė erozija.....	19
1.4.5. Plastinė deformacija	19
1.4.6. Įtrūkimai ir lūžiai.....	20
1.5. Nuspėjamos priežiūros ekonominiai aspektai	20
1.5.1. Sprendimo priėmimo metodai tyrimuose	21
1.5.2. Nuspėjamos priežiūros keitimas į korekcinę ir prastovas	22
1.5.3. Pinigų srutai, susiję su investicijomis į nuspėjamą techninę priežiūrą.....	22
1.6. Matematiniai metodai nuspėjamai priežiūrai	23
1.6.1. Dirbtiniai neuroniniai tinklai tyrimuose	25
1.6.2. Atraminiai vektorių klasifikatoriai tyrimuose	25
1.6.3. K artimiausi kaimynai tyrimuose	25
1.6.4. Atsitiktiniai miškai tyrimuose	26
1.6.5. Gilusis mokymas tyrimuose	26
1.7. Literatūros apibendrinimas.....	27
2. Tyrimo metodai	28
2.1. Tyrimo eiga	28
2.2. Naudojama programinė įranga	29
2.3. Tyrimo duomenys.....	29
2.4. Požymių išgavimas.....	30
2.5. Projekte naudoti metodai.....	31
2.5.1. Atsitiktinis miškai.....	31
2.5.2. LightGBM	32
2.5.3. Dirbtinis neuroninis tinklas	34
2.5.4. Sąnaudų ir naudos analizė	36
2.6. Hiperparametrų derinimas	37
2.7. Modelių vertinimas.....	38
2.7.1. Klasifikavimo modelių vertinimas	38
2.7.2. Regresijos modelių vertinimas	39

3. Tyrimo rezultatai.....	41
3.1. Žvalgomoji analizė ir požymių išskyrimas.....	41
3.2. Guolių defektų identifikavimas	46
3.3. Guolių defektų vietos nustatymas	49
3.4. Guolių defektų dydžio nustatymas	52
3.4.1. Guolių defektų dydžio nustatymas regresijos būdu.....	52
3.4.2. Alternatyvus guolių defektų dydžio nustatymas klasifikavimo būdu	55
3.5. Sąnaudų ir naudos analizė	58
3.6. Rekomendacijos	60
Išvados	62
Literatūros sąrašas	64
Informacijos šaltinių sąrašas	68
Priedai.....	69
1 priedas. OvR ROC kreivės kiekvienai defekto vietai: a) atsitiktiniai miškai, b) neuroniniai tinklai, c) LightGBM.....	69
2 priedas. OvR ROC kreivės kiekvienam defekto dydžiui: a) atsitiktiniai miškai, b) neuroniniai tinklai, c) LightGBM.....	69

Lentelių sąrašas

1 lentelė. Atlikti tyrimai guolių gedimų srityje	23
2 lentelė. Neuronų aktyvavimo funkcijos	35
3 lentelė. Modeliuose derinami hiperparametrai	37
4 lentelė. Sumaišymo matrica projekto atžvilgiu	38
5 lentelė. Klasifikavimo modelių vertinimo rodikliai	38
6 lentelė. Duomenų pasiskirstymas būklės, defekto vietos ir dydžio atžvilgiu	43
7 lentelė. Defekto identifikavimo modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)	46
8 lentelė. Guolių defekto identifikavimo modelių tikslumo metrikos	48
9 lentelė. Defekto vietos nustatymo modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)	49
10 lentelė. Guolių defekto vietos nustatymo modelių tikslumo metrikos	50
11 lentelė. Defekto vietos nustatymo regresijos modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)	52
12 lentelė. Guolių defekto dydžio nustatymo regresijos modelių paklaidos	54
13 lentelė. Defekto dydžio nustatymo klasifikavimo modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)	55
14 lentelė. Guolių defekto dydžio nustatymo modelių tikslumo metrikos	56
15 lentelė. Guolių nuspėjamos techninės priežiūros sistemai reikalingos investicijos	58
16 lentelė. Guolių nuspėjamos techninės priežiūros sistemos einamosios išlaidos	59
17 lentelė. Guolių nuspėjamos techninės priežiūros sistemos naudos	59
18 lentelė. Guolių nuspėjamos techninės kasmetinės sąnaudos ir nauda	59

Paveikslų sąrašas

1 pav. Techninės priežiūros strategijų klasių diagrama [1].....	13
2 pav. Guolio gyvavimo trukmė ir gedimai	16
3 pav. Guolių defektų kategorijos [13]	17
4 pav. Vidinis guolio žiedas: a) po paviršiumi inicijuotas atskilimas, b) bendras guolio vaizdas [14]	18
5 pav. Guolio narvelis: a) abrazyvinis narvelio kišenių nusidėvėjimas, b) bendras guolio vaizdas [14]	18
6 pav. Guolio išorinis žiedas: a) trinties korozijos defektai, b) bendras guolio vaizdas [14]	19
7 pav. Išorinis guolio žiedas ir rutuliukas: a) elektros erozija defektai, b) bendras guolio vaizdas [14]	19
8 pav. Vidinis guolio žiedas: a) plastinės deformacijos nuo perkrovos, b) bendras guolio vaizdas [14]	20
9 pav. Vidinis guolio žiedas: a) įtrūkimas, atsiradęs dėl per didelės trinties su nejudančiu guolio surinkimo agregatu, b) bendras guolio vaizdas [14].....	20
10 pav. Sąsaja tarp išlaidų ir prevencinės techninės priežiūros apimties [15]	22
11 pav. Pinigų srautai, susiję su investicijomis į nuspėjamą priežiūrą [15].....	23
12 pav. Tyrimo eigos procesas.....	28
13 pav. Case Western Reserve universiteto guolių bandymų stendas [34].....	30
14 pav. Atsitiktinių miškų supaprastintas algoritmas	31
15 pav. Histogramos algoritmas [40]	33
16 pav. Sprendimų medžių augimas pagal <i>leaf-wise</i> strategiją.....	34
17 pav. Bendras neuroninio tinklo modelis [41].....	34
18 pav. Neuroninis tinklas su dvejais paslėptais sluoksniais [41]	35
19 pav. Guolio be defektų vibracija, esant 1797 aps/min motoro greičiui	41
20 pav. Guolių vibracija, kai defektai yra 0,007, 0,014 ir 0,021 colio dydžio, esant 1797 aps/min motoro greičiui: a)–c) guolių su defektu ant vidinio žiedo vibracija, d)–f) guolių su defektu ant rutuliuko vibracija, g)–i) guolių su defektu ant išorinio žiedo vibracija	42
21 pav. Guolių su defektu ir be defekto statistinių charakteristikų tankio diagramos kiekvienai duomenų rinkmenai: a) vidurkis, b) asimetrijos koeficientas, c) ekscesas, d) standartinis nuokrypis, e) Shannon'o entropija, f) piko faktorius	44
22 pav. Statistinių charakteristikų vertės pagal defekto vietą: a) vidurkis, b) asimetrijos koeficientas, c) ekscesas, d) standartinis nuokrypis, e) Shannon'o entropija, f) piko faktorius.....	45
23 pav. Statistinių charakteristikų vertės pagal defekto dydį: a) vidurkis, b) asimetrijos koeficientas, c) ekscesas, d) standartinis nuokrypis, e) Shannon'o entropija, f) piko faktorius.....	46
24 pav. Defekto identifikavimo modelių sumaišymo matricos: a) atsitiktinių miškai, b) neuroniniai tinklai, c) LightGBM	47
25 pav. Defekto identifikavimo modelių ROC kreivės.....	48
26 pav. Defekto identifikavimo modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM....	49
27 pav. Defekto vietos nustatymo modelių sumaišymo matricos: a) atsitiktinių miškai, b) neuroniniai tinklai, c) LightGBM.....	50
28 pav. Defekto vietos nustatymo modelių makrovidutinės OvR ROC kreivės	51
29 pav. Defekto vietos nustatymo modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM	52
30 pav. Tikrų ir prognozuotų reikšmių (a–c grafikai) ir prognozės liekanų reikšmių (d–f grafikai) diagramos: a) ir d) atsitiktiniai miškai, b) ir e) neuroniniai tinklai, c) ir f) LightGBM	53

31 pav. Modelių prognozuotų reikšmių stačiakampės diagramos: a) atsitiktiniai miškai, b) neuroniniai tinklai, c) LightGBM.....	54
32 pav. Defekto dydžio nustatymo regresijos modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM	55
33 pav. Defekto dydžio nustatymo klasifikavimo modelių sumaišymo matricos: a) atsitiktinių miškai, b) neuroniniai tinklai, c) LightGBM.....	56
34 pav. Defekto dydžio nustatymo modelių makrovidutinės OvR ROC kreivės	57
35 pav. Defekto dydžio nustatymo klasifikavimo modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM	58

Santrumpų ir terminų sąrašas

Santrumpos:

NLP (angl. *Natural Language Processing*) – natūralios kalbos apdorojimas;

ISO (angl. *International Organization for Standardization*) – Tarptautinė standartizacijos organizacija;

PRR (angl. *Percentage Rate of Return*) – procentinė grąžos norma;

NPV (angl. *Net Present Value*) – grynoji dabartinė vertė;

ROI (angl. *Return on Investment*) – investicijų grąža;

IRR (angl. *Internal Rate of Return*) – vidinė grąžos norma;

BCR (angl. *Benefit-Cost Ratio*) – naudos ir sąnaudų indeksas;

ML (angl. *Machine Learning*) – mašininis mokymas;

DA (angl. *Dimensional Analysis*) – dimensinė analizė;

DNT (angl. *Artificial Neural Networks*) – dirbtiniai neuroniniai tinklai;

CBA (angl. *Cost-Benefit Analysis*) – sąnaudų ir naudos analizė;

AUC (angl. *Area Under the Curve*) – plotas po kreive;

OvO (angl. *One-vs-One*) – vienas prieš vieną ploto po kreive skaičiavimo strategija;

OvR (angl. *One-vs-Rest*) – vienas prieš visus ploto po kreive skaičiavimo strategija;

MAE (angl. *Mean Absolute Error*) – vidutinė absoliuti paklaida;

MSE (angl. *Mean Squared Error*) – vidutinė kvadratinė paklaida;

MAPE (angl. *Mean Absolute Percentage Error*) – vidutinė absoliutinė procentinė paklaida.

Ivadas

Darbo problematika ir aktualumas

Rutulinių guolių, kaip esminės mašinų ir įrenginių dalies, priežiūra yra kritiškai svarbi užtikrinant ilgalaikį ir efektyvų jų veikimą. Šie komponentai plačiai naudojami įvairiose pramonės šakose, įskaitant automobilių gamybą, aviaciją, laivų statybą, taip pat įvairias gamybos linijas, kur aukštas efektyvumas ir patikimumas yra būtini. Rutulinių guolių nuspėjamos priežiūros matematinis modelis yra svarbus tyrimų objektas, nes jis leidžia optimizuoti guolių priežiūros procesus, mažinti netikėtus gedimus ir pratęsti įrangos veikimo laiką. Naudodami šiuos modelius, inžinieriai gali nustatyti optimalius priežiūros intervalus, sumažinti priežiūros sąnaudas ir užtikrinti sklandų gamybos procesų tęstinumą. Be to, šių modelių plėtra ir taikymas prisideda prie tvarumo principų įgyvendinimo, mažinant atliekų kiekį ir energijos suvartojimą. Šių technologijų svarba dar labiau išauga šiandienos kontekste, kai pramonės sektoriai siekia maksimalaus efektyvumo ir minimalių eksploatacinių išlaidų, o technologijų pažanga suteikia vis naujų galimybių guolių būklės stebėsenai ir analizei.

Guolių diagnostika Lietuvoje, palyginti su pasauliniais mastais, tebėra santykinai nauja ir mažai ištirta sritis, suteikianti didelį potencialą moksliniams tyrimams ir inovacijoms. Nepaisant to, kad pasaulio mastu guolių būklės stebėsenos ir diagnostikos technologijos sparčiai vystomos, taikant pažangius analitinius metodus, tokius kaip dirbtinis intelektas, mašininis mokymasis ir duomenų analizė, Lietuvoje ši tema dar nėra sulaukusi pakankamo dėmesio. Tai suteikia mokslininkams ir inžinieriams unikalią galimybę prisidėti prie šios srities plėtos, kuriant pritaikytus modelius ir metodikas, atsižvelgiant į vietos pramonės ypatumus ir poreikius.

Realizuojant nuspėjamos priežiūros matematinį modelį trimis etapais, tyrimas išsiskiria iš kitų panašių darbų savo išsamumu ir struktūrizuotu požiūriu į problemos sprendimą. Pirmasis etapas orientuotas į defekto aptikimą, antrasis – į defekto vietos nustatymą ir trečiasis – į defekto dydžio įvertinimą. Toks metodologinis požiūris leidžia atskleisti ir išanalizuoti kiekvieną guolio gedimo aspektą atskirai, užtikrinant gilesnį ir tikslesnį gedimų analizės procesą. Skirtingai nuo daugumos studijų, kurios sutelkia dėmesį į siauresnę problemų aibę, šis tris etapus apimantis modelis suteikia galimybę ne tik nustatyti defekto buvimą, bet ir išsamiai įvertinti jo poveikį guolio veikimui.

Darbo tikslas ir uždaviniai

Pagrindinis darbo **tikslas** – naudojant viešai prieinamus realius duomenis ištirti mašininio mokymo metodų taikymą guolių defektams aptikti, klasifikuoti, įvertinti. Tikslui pasiekti buvo iškelti šie **uždaviniai**:

1. išanalizuoti tiriamą problemą, remiantis pasaulyje atliktų mokslinių tyrimų rezultatais;
2. apžvelgti ekonominius aspektus, lemiančius nuspėjamosios priežiūros sistemų įdiegimą ir taikymą;
3. parinkti tinkamus mašininio mokymo algoritmus;
4. apjungti panaudotus metodus į bendrą struktūrą, kuri leistų vykdyti nuspėjamą guolių priežiūrą;
5. įvertinti ekonomines sistemos sąnaudas ir naudą;
6. parengti rekomendacijas dėl tolimesnio sukurtos metodikos tobulinimo ir tolimesnių tyrimų atlikimo ir/arba tokios sistemos diegimo.

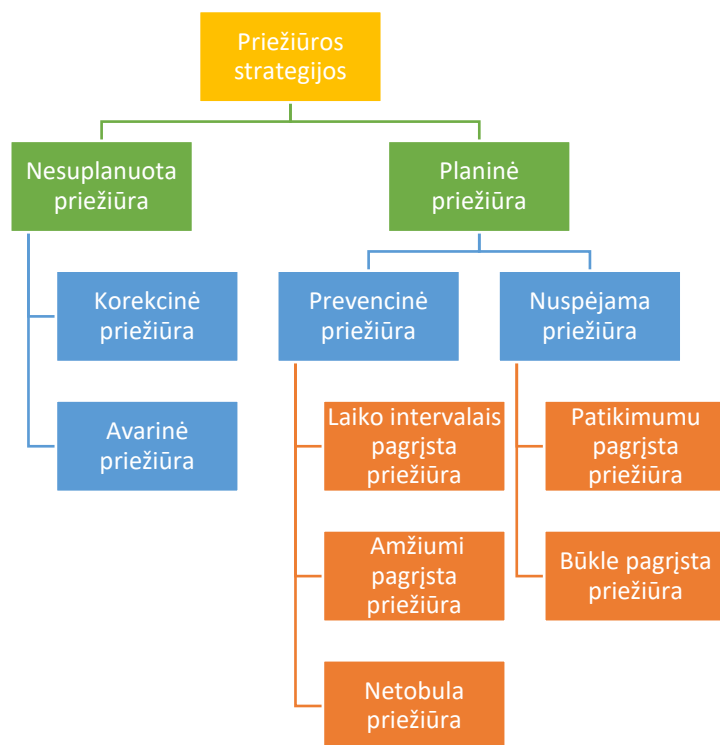
1. Literatūros apžvalga

Šiame skyriuje bus supažindinama su technine priežiūra ir jos strategijomis bei svarba. Skyriuje taip pat apžvelgiami duomenų modalumai, kurie svarbūs nuspėjamajai priežiūrai. Aptariami rutulinių guolių defektai, išskiriant pagrindines defektų kategorijas ir jų atsiradimo priežastis. Taip pat išsamiai analizuojami nuspėjamos priežiūros ekonominiai aspektai, įskaitant įvairių investicijų vertinimo metodikų naudojimą praktikoje. Skyriaus pabaigoje pristatomos įvairios matematinės metodikos, taikomos nuspėjamajai priežiūrai.

1.1. Techninė priežiūra

Techninė priežiūra apima veiksmus, skirtus išvengti arba šalinti įrangos gedimus, kurie gali atsirasti naudojant prietaisą. Pastaruoju metu techninė priežiūra tampa vis svarbesnė siekiant padidinti įrenginių efektyvumą ir tęstinumą. Įvairių techninės priežiūros strategijų apibendrinantį grafiką galima pamatyti **1 pav.** Techninės priežiūros strategijos gali būti skirstomos į planinę ir neplaninę priežiūrą. Neplaninė techninė priežiūra atliekama tik po to, kai išryškėja defektas, gedimas. Ji tinka objektuose, kuriuose įrengtų mašinų skaičius yra minimalus ir gamykla nėra visiškai priklausoma nuo atskiro įrenginio patikimumo. Ši techninė priežiūra taip pat gali būti tinkama, kai gedimų skaičius yra minimalus, gedimas nesukelia rimtų sąnaudų ar saugos pasekmių. Nesuplanuota techninė priežiūra sudaryta iš dviejų kategorijų [1]:

- korekcinės techninės priežiūros. Veikla atliekama po to, kai įvyko gedimas, ir skirta atkurti įrenginio būklę;
- avarinės techninės priežiūros. Veikla atliekama nedelsiant, kad būtų išvengta rimtų pasekmių. Techninės priežiūros dažnumui taikomi apribojimai, siekiant optimizuoti sąnaudas.



1 pav. Techninės priežiūros strategijų klasių diagrama [1]

Planuojamą techninę priežiūrą dar galima skirstyti į prevencinę ir nuspėjamą techninę priežiūrą. Daugeliu atvejų pastebimas geresnis išteklių panaudojimas, palyginti su nesusplanuotomis strategijomis. Prevencinė techninė priežiūra – tai strategija, organizuojama siekiant atlikti techninę priežiūrą iš anksto nustatytais intervalais, kad būtų sumažinta gedimo arba veikimo pablogėjimo tikimybė. Ją galima klasifikuoti į šias kategorijas [1]:

- laiko intervalais pagrįstą priežiūrą. Priežiūra atliekama nustatytais intervalais. Intervalai parenkami atsižvelgiant į tai, kad ilgesnių intervalų atveju sumažinama aukšta gedimo tikimybė, tuo tarpu trumpesnių intervalų atveju siekiama sumažinti didelės preventyvinės techninės priežiūros sąnaudas;
- amžiumi pagrįstą priežiūrą. Pagal šią strategiją prevencinė techninė priežiūra nustatytais intervalais atliekama tik tada, kai sistema pasiekia tam tikrą amžių;
- netobulą priežiūrą. Minėtose strategijose daroma prielaida, kad po prevencinės techninės priežiūros sistema bus atstatyta į pradinę būklę. Netobulos techninės priežiūros metu sistema paliekama su tam tikru degradacijos lygiu [1, 2].

Nuspėjamoji ir prevencinė techninė priežiūra skiriasi priežiūros planavimu. Prevencinė atliekama pagal fiksuotą grafiką, o nuspėjamoji – pagal adaptyvųjį. Nuspėjamoji priežiūra skirstoma į dvi kategorijas [1]:

- būkle pagrįstą priežiūrą. Sprendimas atlikti techninę priežiūrą priimamas stebint sistemos ir (arba) jos komponentų parametrus. Pavyzdžiui, rotacinių sistemų atveju tinka pasirinkti vibracijos charakteristiką;
- patikimumu pagrįstą priežiūrą. Taikant šią strategiją, sistemos patikimumo įverčiai naudojami ekonomiškai efektyviam techninės priežiūros tvarkaraščiui sudaryti. Orlaiviuose ir kitose su sauga susijusiose programose ekonomiskumas derinamas su sauga ir prieinamumu, siekiant kuo labiau sumažinti išlaidas ir prastovas, tačiau pašalinti gedimo tikimybę.

1.2. Nuspėjamos techninės priežiūros svarba

Nuspėjamos techninės priežiūros programų pritaikymas pramonėje teikia išpūdingą naudą, sumažindamas įrangos gedimų riziką ir optimizuodamas išteklių panaudojimą. Šių programų įgyvendinimas leidžia efektyviai planuoti techninės priežiūros procedūras, mažinant ar netgi eliminuojant neprojektuotų viršvalandžių darbą. Taip pat, dėl nuspėjamos techninės priežiūros, įmanoma sumažinti atsargų lygį, kartu užtikrinant, kad būtinos detalės būtų užsakomos ir pristatomos laiku, atitinkant būsimus techninės priežiūros poreikius. Be to, šios programos padeda optimizuoti įrangos veikimo efektyvumą, sumažinant energijos sąnaudas ir didinant gamybos patikimumą [3, 4, 5].

Ankstesnių studijų metu nustatyta, kad efektyviai valdoma nuspėjamos techninės priežiūros sistema gali lemti 8 iki 12 procentų lėšų sutaupymą, lyginant su sistemomis, kurios remiasi vien prevencine priežiūra. Atsižvelgiant į įrangos priklausomybę nuo neplaninės priežiūros ir būklės, galima identifikuoti net 30–40 procentų didesnę taupymo potencialą. Nepriklausomų apklausų duomenimis, diegiant veiksmingą nuspėjamos techninės priežiūros sistemą, vidutiniškai pasiekiamas šių pramonės lėšų taupymas [3]:

- investicijų grąžos koeficientas siekia 10 kartų;

- techninės priežiūros išlaidos mažėja 25–30 %;
- gedimų skaičius sumažėja 70–75 %;
- prastovų laikas mažėja 35–45 %;
- gamybos apimtys didėja 20–25 %.

Šie statistikos duomenys atspindi nuspėjamos techninės priežiūros svarbą ir efektyvumą pramoninėje aplinkoje, pabrėžiant jos teikiamą ekonominę naudą ir operacinio stabilumo užtikrinimą.

1.3. Duomenų modalumai nuspėjamos priežiūros kontekste

Duomenų modalumas nuspėjamojoje techninėje priežiūroje apibrėžia įvairių tipų duomenis, kuriuos galima rinkti ir analizuoti siekiant numatyti įrangos gedimus prieš jiems įvykstant. Pagrindiniai nuspėjamojoje priežiūroje sekami duomenų tipai:

- jutiklių duomenys;
- nuotraukų ir vaizdo įrašų duomenys;
- garso duomenys;
- tekstiniai duomenys.

1.3.1. Jutiklių duomenys

Jutiklių duomenys atlieka svarbų vaidmenį techninės priežiūros strategijose, ypač nuspėjamosios techninės priežiūros kontekste. Šio tipo duomenys, kuriems būdingi nuoseklūs matavimai, surinkti per tam tikrą laiką, suteikia dinamišką ir įžvalgų vaizdą apie įrangos veikimą, būklę ir veiklos efektyvumą. Jutiklių duomenys, naudojami nuspėjamojoje techninėje priežiūroje, paprastai apima jutiklių, stebinčių mašinų būklę, pavyzdžiui, temperatūrą, vibraciją, slėgį ir kitus svarbius parametrus, rodmenis. Šie duomenų taškai renkami reguliariais laiko intervalais, taip sukuriant istorinį įrašą apie įrangos veikimą laikui bėgant. Dėl nuoseklaus laiko eilučių duomenų pobūdžio jie yra unikaliai vertingi sudarant modelius, nustatant tendencijas ir galimas anomalijas, rodančias būsimums gedimus [6].

1.3.2. Nuotraukų ir vaizdo įrašų duomenys

Nuotraukų ir vaizdo duomenys tampa vis vertingesni prognozuojamosios techninės priežiūros srityje, kurioje naudojami pažangūs kompiuterinės regos metodai, leidžiantys nustatyti, numatyti ir užkirsti kelią įrangos gedimams. Toks vaizdinių duomenų panaudojimas ypač veiksmingas tais atvejais, kai tradicinių jutiklių duomenys gali neužfiksuoti visų galimų problemų spektro [7].

Nuotraukos ir vaizdo įrašai taikomi nuspėjamojoje techninėje priežiūroje labai įvairiai. Šiluminio vaizdavimo metu infraraudonųjų spindulių kameros fiksuoja įrangos paviršių temperatūros pokyčius, nustatydamos karštus taškus, kurie rodo per didelę trintį, perkaitusius komponentus arba izoliacijos gedimus [8]. Didelės raiškos kameros naudojamos vizualiniams patikrinimams, jomis galima aptikti paviršiaus nusidėvėjimą, koroziją, įtrūkimus ir kitas struktūrines anomalijas, kurios gali būti ne iš karto matomos. Taip dronai ir robotai su kameromis gali pasiekti sunkiai pasiekiamas vietas ir išsamiai apžvelgti įrangos ir infrastruktūros būklę [9].

1.3.3. Garso duomenys

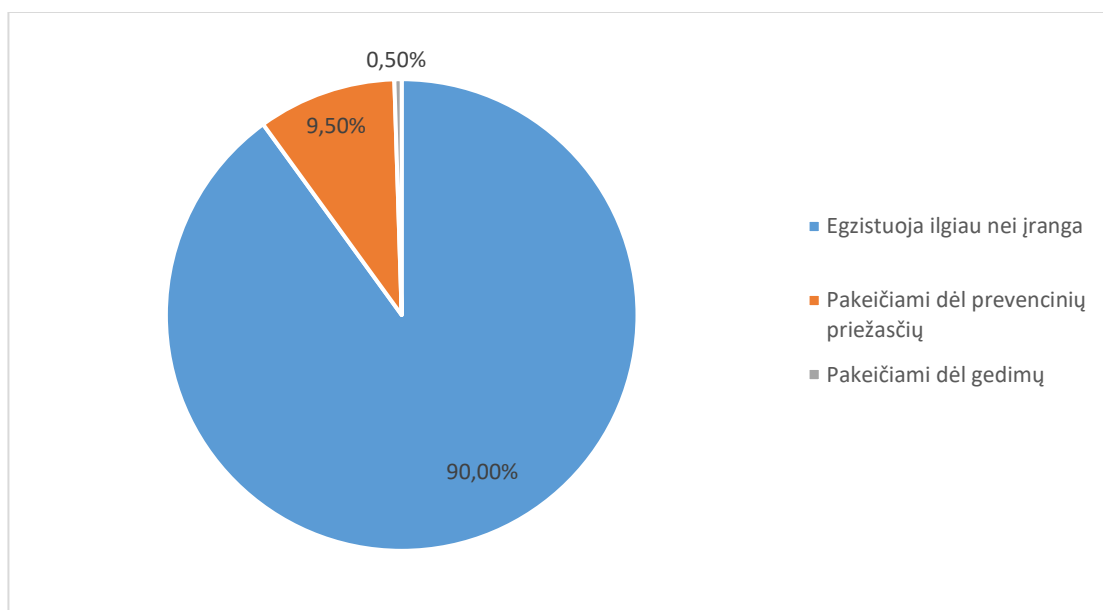
Garso duomenys yra nenutrūkstamas akustinių signalų, kuriuos generuoja įranga jos veikimo metu, srautas. Šių signalų charakteristikose, pavyzdžiui, dažnyje, amplitudėje ir trukmėje, yra vertingos informacijos, kurią tinkamai išanalizavus galima nustatyti mechaninių problemų atsiradimą gerokai anksčiau. Garso duomenys atspindi sudėtingą mechaninių komponentų sąveiką. Įprastomis eksploataavimo sąlygomis yra nustatytas bazinis akustinis profilis, kuris laikui bėgant yra gana stabilus. Tačiau bet koks nukrypimas nuo šio bazinio lygio, pavyzdžiui, neįprastas triukšmas, garso intensyvumo pokytis arba naujų akustinių modelių atsiradimas, gali rodyti besivystančias problemas [10].

1.3.4. Tekstiniai duomenys

Tekstinė informacija apima techninės priežiūros žurnalus, operatoriaus pastabas, pranešimus apie klaidas. Skirtingai nuo jutiklių duomenų, kurie yra struktūruoti ir skaitiniai, tekstiniai duomenys yra nestrukūruoti, todėl jiems analizuoti reikia specialių metodų, kurių pagrindą sudaro natūralios kalbos apdorojimo (angl. *Natural Language Processing*, NLP) sritis. Taikant NLP, iš tekstinių duomenų galima išgauti ir interpretuoti vertingą informaciją, kuri palengvina pasikartojančių problemų identifikavimą, įrangos gedimo būdų išskyrimą ir techninės priežiūros veiklos veiksmingumo supratimą [11]. Be to, tekstiniai duomenys gali atskleisti subtilias užuominas apie įrangos būklę, kurių tiesiogiai nefiksuoja jutikliai, pavyzdžiui, neįprastų garsų, kvapų aprašymus ar vizualinius nusidėvėjimo pastebėjimus, todėl galima susidaryti visapusišką vaizdą apie įrangos situaciją.

1.4. Rutulinių guolių defektai ir defektų priežastys

Kiekvienais metais visame pasaulyje pagaminama apie 10 milijardų guolių. Tik maža dalis visų naudojamų guolių tikrai sugenda (žr. **2 pav.**). Dauguma jų (apie 90 %) egzistuoja ilgiau nei pats įrenginys, kuriame jie yra įdiegti. Dalis guolių (9,5 %) pakeičiami prieš sugedimą dėl saugumo (prevencinės) priežasčių. Maždaug 0,5 % guolių pakeičiami, nes jie yra sugadinti arba sugenda. Tai reiškia, kad kasmet dėl pažeidimų ir gedimų pakeičiama apie 50 000 000 guolių [12].



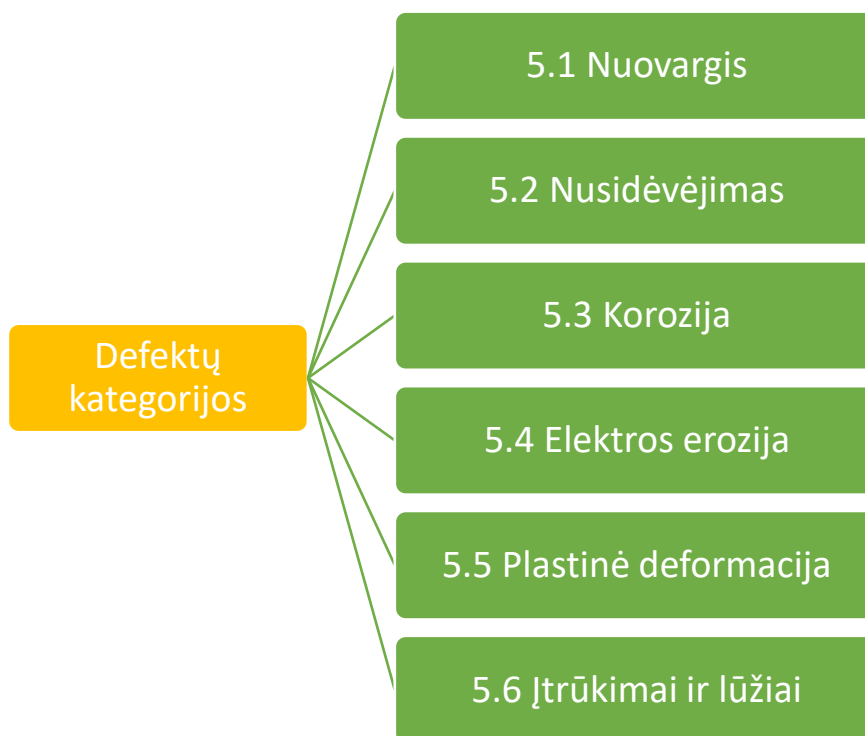
2 pav. Guolio gyvavimo trukmė ir gedimai

Yra keletas priežasčių, kodėl guoliai gali būti pažeisti arba sugesti [12]:

- 1/3 guolių sugenda dėl pastovaus darbo;
- 1/3 guolių sugenda dėl tepimo problemų (netinkamas tepalas, netinkamas kiekis, netinkami tepimo intervalai);
- 1/6 guolių sugenda dėl teršalų (neveiksmingi sandarikliai);
- 1/6 guolių sugenda dėl kitų priežasčių (netinkamas tvarkymas ir montavimas, sunkesnė arba kitokia nei tikėtasi apkrova, netinkami arba nepakankami tarpai).

Skaiciai gali skirtis priklausomai nuo pramonės ar taikymo srities. Pavyzdžiui, medienos ir popieriaus pramonėje pagrindinė guolių gedimo priežastis yra teršalai ir nepakankamas tepimas, o ne pastovus darbas. Laiko tarpas nuo pirmos (pradinės) žalos iki momento, kai guolis tampa netinkamas naudoti, gali labai skirtis. Esant didesniai greičiui, tai gali užtrukti keletą sekundžių. Didelėse, lėtai besisukančiose mašinose tai gali užtrukti mėnesius. Jei pažeistas guolis nepastebimas ir nepakeičiamas prieš jį katastrofiškai sugadinant, gali atsirasti antrinė žala mašinai ir jos komponentams. Be to, kai guolis sugenda katastrofiškai, gali būti sunku, nustatyti gedimo priežastį [12].

Įvairiuose leidiniuose guolių pažeidimai ir gedimai gali būti skirtingai klasifikuojami bei naudojama skirtinga terminologija. Siekiant užtikrinti, kad visame pasaulyje guolių gedimai ir pažeidimai būtų analizuojami ir aprašomi taip pat, buvo sudarytas ISO standartas 15243:2004, atnaujintas 2017 m. Standartas apibrėžia bendras gaires, kaip identifikuoti ir klasifikuoti įvairius guolių pažeidimus ir gedimus, remiantis padariniais ir vizualiniais požymiais. Pagal ISO 15243:2017 guolių defektai yra skirstomi į 6 pagrindines kategorijas (žr. **3 pav.**). Numeracija atitinka ISO poskyrių numeraciją [13].



3 pav. Guolių defektų kategorijos [13]

1.4.1. Nuovargis

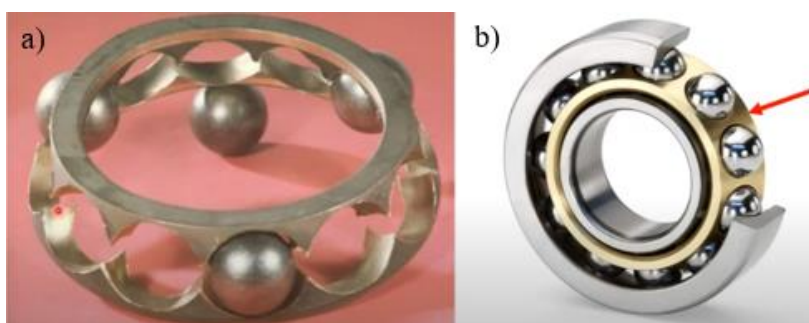
Riedėjimo kontaktų nuovargį sukelia pasikartojantys įtempiai, atsirandantys riedėjimo elementų ir riedėjimo takų kontaktuose. Nuovargis akivaizdžiai pasireiškia kaip struktūros (mikrostruktūros) pokyčiai ir kaip medžiagos atplaišos nuo paviršiaus (makrostruktūros), kurios daugeliu atveju gali būti mikrostruktūros pokyčių pasekmė (žr. **4 pav.**). Guolių defektai dėl nuovargio gali būti susiję su įtrūkimais, atsirandančiais paviršiuje arba po juo [13].



4 pav. Vidinis guolio žiedas: a) po paviršiumi inicijuotas atskilimas, b) bendras guolio vaizdas [14]

1.4.2. Nusidėvėjimas

Nusidėvėjimas tai eksploatacijos metu atsirandantis laipsniškas medžiagos dilimas nuo paviršiaus, dėl dviejų slystančių arba riedančių / slystančių besiliečiančių paviršių sąveikos. Nusidėvėjimas gali būti abrazyvinis arba adhezinis. Abrazyvinis dėvėjimasis atsiranda, kai vienas iš dviejų besiliečiančių paviršių yra kietesnis arba šiurkštesnis už kitą ir medžiaga pašalinama nuo minkštesnio paviršiaus [12]. Abrazyvinis nusidėvėjimas taip pat atsiranda, kai kietesnis trečiasis kūnas, įstrigęs tarp dviejų besiliečiančių paviršių, santykinai judėdamas nudilina medžiagą nuo minkštesnio paviršiaus (žr. **5 pav.**). Tuo tarpu adhezinis nusidėvėjimas kyla dėl vietinio sukibimo tarp besiliečiančių kietųjų paviršių, dėl kurio vyksta medžiagos pernešimas tarp šių dviejų paviršių arba praradimas nuo bet kurio iš jų [12].

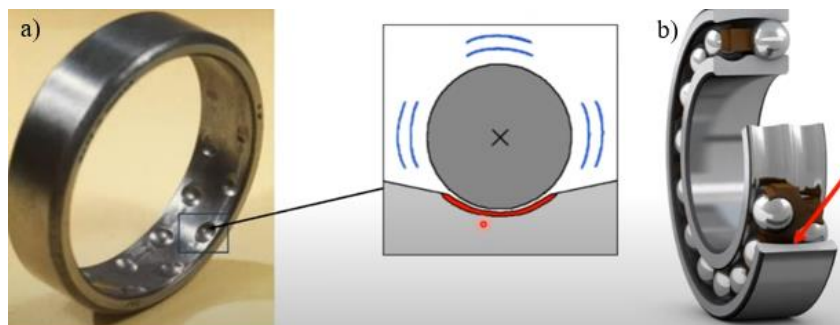


5 pav. Guolio narvelis: a) abrazyvinis narvelio kišenių nusidėvėjimas, b) bendras guolio vaizdas [14]

1.4.3. Korozija

Yra du pagrindiniai korozijos tipai: drėgmės korozija ir trinties korozija. Drėgmės korozija atsiranda, kai metalo paviršius veikiamas drėgmės ar agresyvios terpės, pavyzdžiui, vandens ar rūgščių. Ši sąveika sukelia paviršiaus oksidaciją, vadinamą korozija ar rūdijimu, dėl kurios atsiranda paviršiaus įdubimai ir, ilgainiui, paviršiaus suirimas. Trinties korozija yra cheminė reakcija, kurią suaktyvina santykiniai mikrojūduliai tarp besiliečiančių paviršių esant tam tikroms

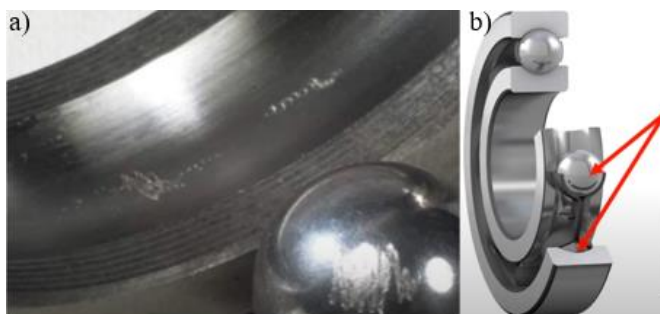
trinties ir apkrovos sąlygoms. Dėl šių mikrojudesių paviršiai oksiduojasi (žr. **6 pav.**), o išsiskyrusi medžiaga tampa matoma kaip miltelių pavidalo rūdys ir (arba) nuo vieno ar abiejų besiliečiančių paviršių prarandama medžiaga [13].



6 pav. Guolio išorinis žiedas: a) trinties korozijos defektai, b) bendras guolio vaizdas [14]

1.4.4. Elektrinė erozija

Elektrinė erozija – tai lokalizuotas mikrostruktūrinis pokytis ir medžiagos pašalinimas kontaktiniuose paviršiuose, kurį sukelia tekanti žalinga elektros srovė. Kai elektros įtampa tarp guolio žiedų ir riedėjimo elementų viršija izoliacijos pramušimo slenkstinę vertę, elektros srovė teka iš vieno guolio žiedo į kitą per riedėjimo elementus ir jų tepalo plėveles. Kontrasto zonose tarp riedėjimo takelių ir riedėjimo elementų įvyksta koncentruota iškrova, dėl kurios per labai trumpą laiko tarpą lokalizuotai įkaista, todėl kontakto zonos išsilydo, susijungia ir susidaro krateriai 500 μm skersmens. Kai žalinga elektros srovė nuolat veikia, erozija įgauna kitokią išvaizdą. Iš pradžių paviršiaus pažeidimai gali būti negilių, glaudžiai vienas prie kito išdėstytų ir labai mažų, maždaug mikrometro dydžio, kraterių pavidalo (žr. **7 pav.**). Vėliau susidaro grioveliai [13].



7 pav. Išorinis guolio žiedas ir rutuliukas: a) elektros erozija defektai, b) bendras guolio vaizdas [14]

1.4.5. Plastinė deformacija

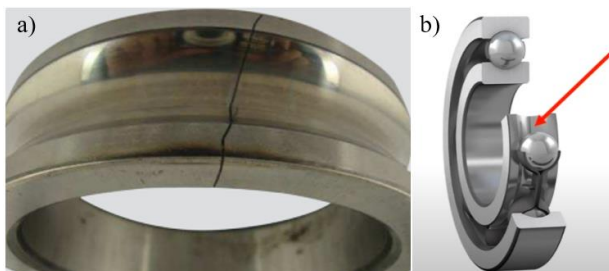
Dėl pernelyg didelės statinės arba smūginės apkrovos gali susidaryti plastiškai deformuotos įdubos, sutampančios su riedėjimo elemento ir bėgių takelio sąlyčio vietomis (žr. **8 pav.**). Plastinės deformacijos taip pat gali atsirasti dėl šiukšlių, patekusių tarp riedėjimo elemento ir riedėjimo takelio [12].



8 pav. Vidinis guolio žiedas: a) plastinės deformacijos nuo perkrovos, b) bendras guolio vaizdas [14]

1.4.6. Įtrūkimai ir lūžiai

Guolių sudedamųjų dalių lūžiai gali atsirasti, kai įtempių visuma viršija medžiagos tempimo stiprį, kai viršijamas medžiagos lenkimo nuovargio stipris arba kai susidaro didelis trinties įkaitimas. Įtempių koncentracija, kylanti dėl didelių smūginių jėgų, pavyzdžiui, atsirandančių dėl plaktukų ir kaltų šaltojo montavimo metu, sukuria smulkius įtrūkimus, kurie eksploatuojant guolį gali virsti įtrūkimais grūdinto plieno žieduose (žr. **9 pav.**). Įtempiai, atsirandantys kaitinant ir montuojant žiedus ant netinkamo dydžio velenų ir per mažo dydžio korpusų, taip pat gali sukelti ašinius įtrūkimus eksploatuojant guolį [12].



9 pav. Vidinis guolio žiedas: a) įtrūkimas, atsiradęs dėl per didelės trinties su nejudančiu guolio surinkimo agregatu, b) bendras guolio vaizdas [14]

1.5. Nuspėjamos priežiūros ekonominiai aspektai

Nuspėjama techninė priežiūra yra strateginė investicija; kaip ir su bet kuria kita su finansiniais įsipareigojimais ir ištekliais susijusia veikla, tikimasi, kad jos grąža viršys pradines išlaidas. Nuspėjamos techninės priežiūros sprendimai dažnai būna taip / ne arba reikia rinktis iš kelių alternatyvų. Priimant sprendimus dėl investicijų, reikia atsižvelgti į pinigų kainą laike. Nesvarbu, ar organizacija yra pelno siekianti, ne pelno siekianti, privati, valstybinė ar vyriausybinių, visi išteklių kainuoja. Išskiriami pagrindiniai sprendimų priėmimo metodai [15, 16]:

- atsipirkimo laikotarpis;
- procentinė grąžos norma (angl. *Percentage Rate of Return*, PRR);
- investicijų grąža (angl. *Return on Investment*, ROI);
- grynoji dabartinė vertė (angl. *Net Present Value*, NPV);
- vidinė grąžos norma (angl. *Internal Rate of Return*, IRR);
- naudos ir sąnaudų indeksas (angl. *Benefit-Cost Ratio*, BCR).

1.5.1. Sprendimo priėmimo metodai tyrimuose

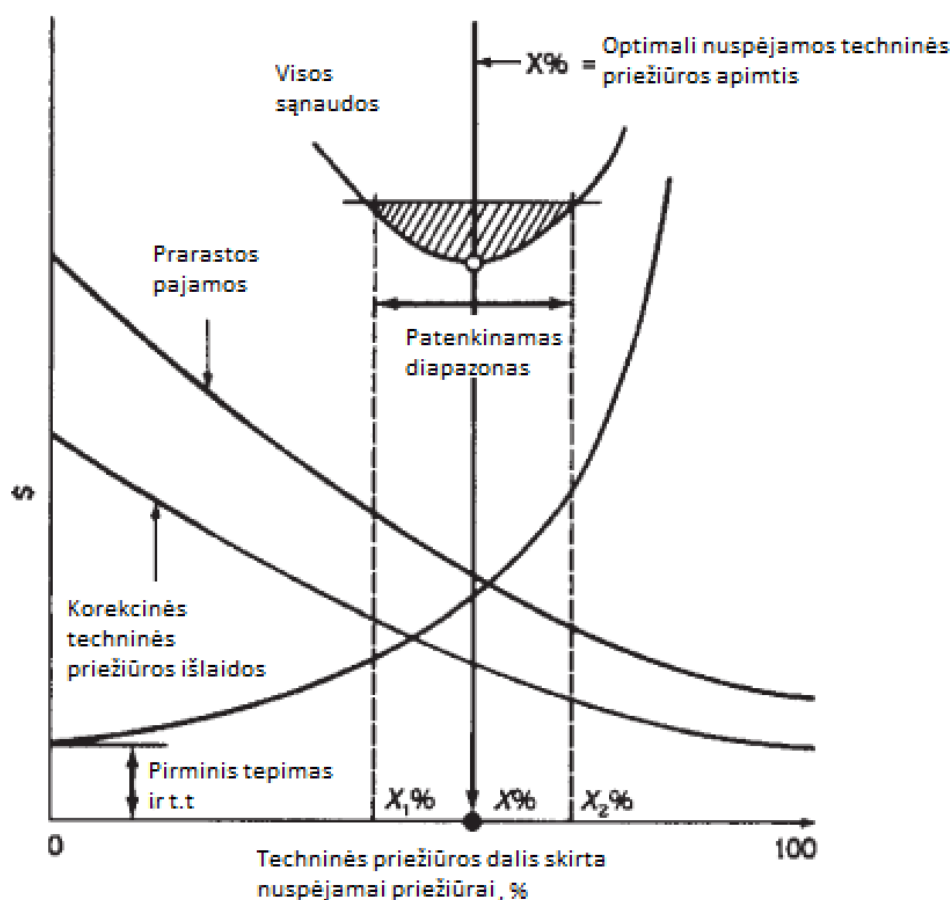
Shamayleh'io ir kt. tyrime [17] pristatytas ligoninės laboratorinės įrangos (imunoanalizatorių) nuspėjamos techninės priežiūros sistemos įgyvendinimas. Sistemoje naudojami belaidžiai akselerometro jutikliai ir duomenų rinkimo plokštė, skirta įrangos būklei stebėti ir vibracijos duomenims perduoti į debesiją apdorojimui. Finansinis įgyvendinimo aspektas išsamiai aprašytas sąnaudų ir naudos analizėje. Į pradinės išlaidas įeina įranga, darbo užmokestis, ryšiai, debesijos paslaugos ir programinės įrangos licencijos, iš viso 32 749 \$. Nauda, įvertinta 33 595 \$, gaunama dėl sumažėjusių prastovų, sumažėjusių atsargų sąnaudų dėl laiku užsakomų dalių, sumažėjusių gedimų ir alternatyvių sąnaudų. Atliekant analizę taip pat apskaičiuota grynoji dabartinė vertė (NPV) – 85 177 \$ per 11 metų, atsižvelgiant į 2,2 % infliaciją, 8 % minimalią patrauklią grąžos normą ir 5 % pridėtinės vertės mokestį (PVM).

Nuspėjamos techninės priežiūros taikymas JAV kariuomenės sunkiesiems išplėstinio mobilumo taktiniams sunkvežimiams išnagrinėtas Fisher'io [18]. Ekonominiu požiūriu tyrime analizuojama galima nuspėjamos priežiūros sistemos diegimo finansinė grąža, atsižvelgiant į tokius veiksnius, kaip sistemos techninės priežiūros išlaidos, naujų technologijų diegimo išlaidos ir galimos sutaupytos lėšos dėl sumažėjusių techninės priežiūros poreikių ir padidėjusio veiklos efektyvumo. Ataskaitoje pateiktose ekonominėse išvadose nurodoma, kad bendra apskaičiuota 230 sunkvežimių nuspėjamos techninės priežiūros technologijų diegimo kaina būtų apie 3 450 000 JAV dolerių, o kiekvienas vienetas kainuotų apie 15 000 JAV dolerių. Prognozuota, kad per penkerius metus nuspėjamos techninės priežiūros diegimas leistų sutaupyti apie 82 508 JAV dolerių, palyginti su baziniu modeliu, į kurį nuspėjamos priežiūros technologijos neįtrauktos. Dėl to apskaičiuotas atsipirkimo laikotarpis yra labai ilgas – maždaug 209 metai, esant dabartinei vieneto kainai investicija laikoma nepraktiška.

Wolf'o ir kt. ekonominis nuspėjamos techninės priežiūros modelio automobilių pramonėje vertinimas buvo sutelktas į vidaus ir iš išorės įsigytų sprendimų finansinio pagrįstumo palyginimą naudojant grynąją dabartinę vertę [19]. Buvo daromos prielaidos, kad nuspėjamai techninei priežiūrai bus taikomas 70 proc. atpažinimo ir 50 proc. tikslumo lygis, o pramoniniai robotai turėtų veikti 7 metus, jų vidutinis laikas tarp gedimų būtų 250 000 valandų, o remonto trukmė – 2 valandos, kiekvieno jų kaina – 6 000 eurų. Neplanuotų prastovų sąnaudos buvo pastebimai didesnės, palyginti su planuotų prastovų sąnaudomis, kurių kaina kėbulų ceche siekė 50 eurų už minutę, o surinkimo linijoje – 20 000 eurų už minutę. Pasirinkus išorinį sprendimą, buvo numatytas 150 eurų mokestis už vieną robotą per metus, o vidinės kūrimo sąnaudos buvo numatytos 100 000 eurų, pridėjus nuolatinės personalo išlaidas. Taikant šias prielaidas, atlikus grynosios dabartinės vertės skaičiavimus paaiškėjo, kad išorinio sprendimo rezultatas kėbulų gamykloje yra nuostolingas (277 500 eurų), tuo tarpu vidaus sprendimas davė teigiamą grąžą (128 000 eurų). Taikant surinkimo linijos scenarijų, išorės sprendimas davė gerokai didesnę teigiamą NPV (307 500 eurų), palyginti su vidaus alternatyva (46 730 eurų). Šie rezultatai rodo, kad didelę įtaką priimant sprendimą dėl nuspėjamos techninės priežiūros sprendimų kūrimo įmonėje ar įsigijimo iš išorės daro mastas ir techninės priežiūros sąnaudos: didesnėms operacijoms paprastai naudingesni vidaus sprendimai, nes pastoviosios sąnaudos paskirstomos didesniam turto kiekiui, o aplinkoje, kurioje yra mažiau mašinų ir didesnės vieneto techninės priežiūros sąnaudos, ekonomiškai naudingesni gali būti išorės sprendimai.

1.5.2. Nuspėjamos priežiūros keitimas į korekcinę ir prastovas

Nuspėjamos techninės priežiūros, korekcinės techninės priežiūros ir prarastų gamybos pajamų sąsajos pavaizduotos **10 pav.** Vertikaliąją skalę yra doleriai. Horizontalioji skalė – prevencinei techninei priežiūrai skirta visos techninės priežiūros procentinė dalis. Nuspėjamos techninės priežiūros procentinė dalis svyruoja nuo nulio iki beveik 100 proc. Kreivė nesiekia 100 proc., nes visada yra tam tikrų gedimų, dėl kurių reikalinga atlikti korekcinę techninę priežiūrą. Kuo daugiau bet kokios rūšies techninės priežiūros atliekama, tuo daugiau kainuoja šios veiklos atlikimas. Tačiau atliekant daugiau nuspėjamos techninės priežiūros turėtų mažėti ir korekcinės techninės priežiūros, ir prastovų sąnaudos. Prastovų sąnaudos šioje iliustracijoje yra didesnės nei nuspėjamos ar korekcinės techninės priežiūros sąnaudos [15]. Daugelio gamybos linijų prastovos sąnaudos viršija 10 000 JAV dolerių per valandą. Tokios prastovos išlaidos gerokai viršija bet kokias techninės priežiūros, darbo jėgos ar net medžiagų, kurias galima panaudoti darbui, sąnaudas [20].



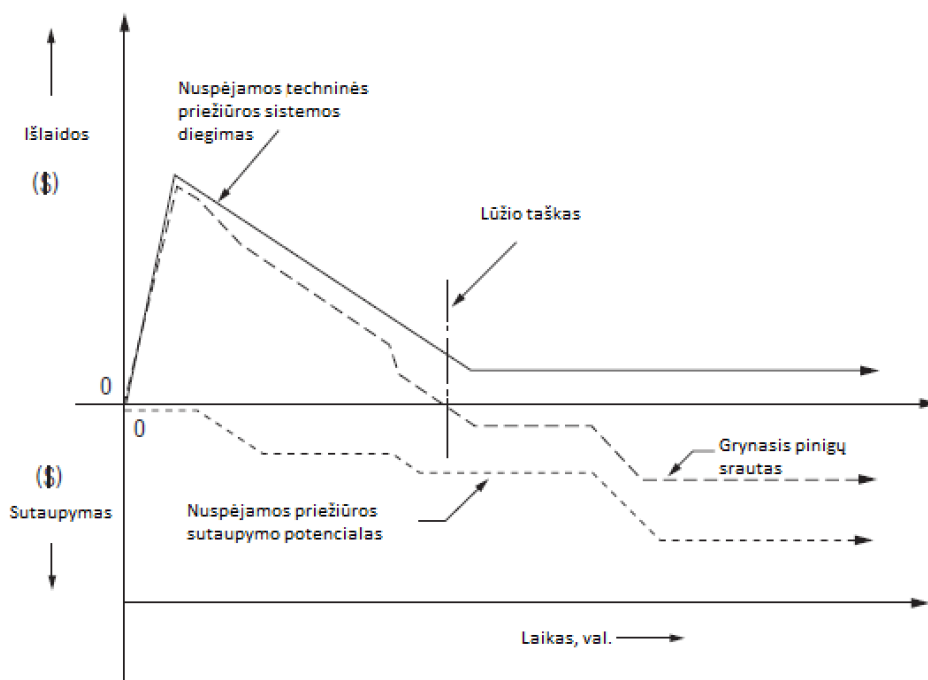
10 pav. Sąsaja tarp išlaidų ir prevencinės techninės priežiūros apimtys [15]

Kaip parodyta **10 pav.**, siekiamybė yra patenkinamas diapazonas, o ne optimumo taškas. Grafiškai kiekvienas bendrųjų sąnaudų kreivės taškas yra nuspėjamos ir korekcinės priežiūros sąnaudų bei prarastų pajamų sąnaudų suma.

1.5.3. Pinigų srautai, susiję su investicijomis į nuspėjamą techninę priežiūrą

Iš pradžių nuspėjamos priežiūros sistemos sąnaudos yra didelės, nes investuojama į įrangą ir darbuotojų mokymus. Tačiau pasiekus lūžio tašką pradedamas gauti pastovus pelnas. Tai yra todėl,

kad sumažėjusios techninės priežiūros išlaidos ir pagerėjusios eksploatacinės savybės tampa bendra finansine nauda [15]. **11 pav.** parodyta, kaip grynasis pinigų srautas pereina per lūžio tašką į pastovaus teigiamo finansinio pelno regioną.



11 pav. Pinigų srautai, susiję su investicijomis į nuspėjamą priežiūrą [15]

1.6. Matematiniai metodai nuspėjamai priežiūrai

Dirbtinio intelekto srityje mašininis mokymas (angl. *Machine Learning*, ML) tapo galinga priemone, leidžiančia kurti prognozavimo algoritmus. ML metodais galima apdoroti didelės dimensijos ir daugiamatėje erdvėje esančius duomenis ir išgauti paslėptus ryšius sudėtingoje ir dinamiškoje aplinkoje [21]. Apačioje pateikta **1 lentelė**, apibendrinanti tyrimus guolių gedimų srityje. Lentelę sudaro penki stulpeliai:

1. autoriai. Šiame stulpelyje pateiktos tyrimo atlikėjų pavardės;
2. modelis. Nurodytos, koks mašininio mokymo modelis buvo naudotas tyrimo metu;
3. duomenų rinkinys. Stulpelyje aprašyta, iš kur buvo gauti duomenys, naudoti modeliui mokyti arba testuoti;
4. defekto tipas. Pateikti specifiniai defektų ar būklių, kuriuos modelis bandė nustatyti, tipai;
5. tikslumas. Pateikta, kaip tiksliai modelis galėjo nustatyti įvairias būkles ar defektus.

1 lentelė. Atlikti tyrimai guolių gedimų srityje

Autoriai	Modelis	Duomenų rinkinys	Defekto tipas	Tikslumas
Kumbhar ir kt. [22]	Dirbtiniai neuroniniai tinklai	Autorių eksperimentiniai duomenys	Išorinio žiedo defektai (trys dydžiai)	97,79 %
Li ir kt. [23]	Dirbtiniai neuroniniai	Duomenys gauti iš <i>Matlab Simulink</i>	Be defektų būklė, defektai ant besisukančio objekto	97,9 %

Autoriai	Modelis	Duomenų rinkinys	Defekto tipas	Tikslumas
	tinklai	programoje sukurtas asinchroninio variklio matematinio modelio	Be defektų būklė, defektai ant vidinio žiedo	88,2 %
			Be defektų būklė, guolio laisvumas	96,5 %
Zarei ir kt. [24]	Dirbtiniai neuroniniai tinklai	Autorių eksperimentiniai duomenys	Be defektų būklė, defektai ant išorinio žiedo, defektai ant vidinio žiedo	100 %
Saidi ir kt. [25]	Atraminių vektorių klasifikatorius	Case Western Reserve universiteto guolių defektų duomenų bazė	Be defektų būklė, defektai ant išorinio žiedo, defektai ant vidinio žiedo, defektai ant besisukančio objekto	99,99 %
Rojas ir Nandi [26]	Atraminių vektorių klasifikatorius	Eksperimentiniai duomenys iš kitų autorių	Be defektų būklė, būklė su defektais	99,1 %
			Be defektų būklė, defektai ant išorinio žiedo, defektai ant vidinio žiedo, defektai ant besisukančio objekto, defektai ant narvo, naudotas guolis	Daugiau nei 96 %
Sharma ir kt. [27]	K artimiausi kaimynai	Autorių eksperimentiniai duomenys	Be defektų būklė, defektai ant vidinio žiedo, defektai ant išorinio žiedo	100 %
Tian ir kt. [28]	K artimiausi kaimynai	Autorių eksperimentiniai duomenys	-	Ekspertinis / vizualus vertinimas
Han ir Jiang [29]	Atsitiktiniai miškai	Case Western Reserve universiteto guolių defektų duomenų bazė	Be defektų būklė, defektai ant išorinio žiedo, defektai ant vidinio žiedo, defektai ant rutuliuko	100 %
			0,007, 0,014, 0,021 defekto dydžiai	99,7 %
Sanchez ir kt. [30]	K artimiausi kaimynai	Case Western Reserve universiteto guolių defektų duomenų bazė	Be defektų būklė, būklė su defektais	97,5 %
	Atsitiktiniai miškai			97,5 %
Janssens ir kt. [31]	Sąsukų neuroniniai tinklai	Autorių eksperimentiniai duomenys	Be defektų būklė, išorinio žiedo defektai, lengvas netinkamas sutepimas, itin prastas sutepimas	93,61 %
Shao ir kt. [32]	Autošifatoriai (angl. <i>Autoencoders</i>)	Autorių eksperimentiniai duomenys	Be defektų būklė, defektai ant išorinio žiedo, defektai ant vidinio žiedo, defektai ant besisukančio objekto, defektai ant išorinio ir vidinio žiedo, defektai ant besisukančio objekto ir vidinio žiedo, defektai ant besisukančio objekto ir išorinio žiedo, defektai ant besisukančio objekto, išorinio žiedo ir vidinio žiedo	87,8 %

Šiems algoritmams taikyti reikia daug srities žinių bei išmanyti požymių inžineriją. Pirmiausia duomenų rinkiniui atliekama išsami tiriamoji duomenų analizė, o po to taikomi matmenų mažinimo metodai, pavyzdžiui, pagrindinių komponentų analizė (angl. *Principal Component Analysis, PCA*) ir kt. Galiausiai reprezentatyviausi požymiai perduodami ML algoritmui. Tačiau dėl darbo srities žinių trūkumo gali būti sudėtinga atrinkti tinkamas savybes, iš kurių modeliai galėtų mokytis. Be to, modelis, kuris buvo išmokytas atlikti užduotį vienoje srityje, be tam tikrų korekcijų gali blogai atlikti užduotį kitoje srityje.

1.6.1. Dirbtiniai neuroniniai tinklai tyrimuose

Naujoviška metodika guolių defektų dydžio diagnostikai, integruojant dimensinę analizę (angl. *Dimensional Analysis*, DA) ir dirbtinius neuronų tinklus, pristatyta Kumbhar'o ir kt. [22]. Eksperimentams buvo naudojami dirbtiniai pažeisti guoliai, o defektų dydžiai buvo atidžiai kontroliuojami per elektroerozinį apdirbimą. Vibracijų duomenys buvo renkami naudojant specialiai sukurtą bandymų įrangą. DA buvo taikyta kaip preliminarus metodas defektų dydžio įvertinimui, o DNT – kaip pagrindinis įrankis defektų klasifikacijai ir tikslesniam dydžio nustatymui. Šis integruotas požiūris leido pasiekti aukštą diagnostikos tikslumą, demonstruojant metodikos efektyvumą ir praktinį pritaikomumą guolių defektų dydžių nustatyme.

DNT taikymas indukcinio variklio guolių gedimų diagnostikai aptartas Li ir kt. moksliniame darbe [23]. Darbe buvo išgautos guolių vibracijos dažnių savybės, kurios yra svarbios nustatant guolio būklę, ir pateikta metodika, kaip šios savybės gali būti naudojamos kartu su DNT variklio guolių gedimams diagnozuoti. Tyrimo metu buvo naudojamas vibracijos simuliacijos metodas, kad būtų galima projektuoti įvairias diagnostikos strategijas. Gauti rezultatai parodė, kad DNT gali būti efektyvus įrankis įvairių guolio gedimų diagnozei atlikti, interpretuojant guolio vibracijos signalus.

Dirbtinių neuroninių tinklų taikymas vibracijos signalų analizei, siekiant nustatyti ir klasifikuoti guolių gedimus indukcinuose varikliuose, aprašytas Zarei ir kt. [24]. Tyrimo autorių sukurtas filtras iš vibracijos signalo sugebėjo pašalinti nesusijusias su guolio gedimais komponentes. Šis filtras, pavadintas RNFC (angl. *Removing Non-bearing Fault Component*), leidžia geriau atskirti gedimo signalus. Po filtravimo proceso, išfiltruotas signalas naudojamas kaip įvestis antrajame neuroniniame tinkle, kuris, remiantis laiko srities signalo požymiais, atpažįsta ir klasifikuoja skirtingus guolių gedimo tipus.

1.6.2. Atraminiai vektorių klasifikatoriai tyrimuose

Ritininių guolių būklės stebėjimo ir gedimų diagnostikos metodika, derinant aukštesnės eilės spektrų analizės ypatybes ir atraminių vektorių klasifikatorių kaip modeliavimo įrankį, pristatyta Saidi ir kt. [25]. Tyrime atskleista, kad pasiūlytas metodas patikimai atpažįsta skirtingus ritininių guolių gedimų tipus remiantis vibracijos signalais, taip patikrinant aukštesniosios eilės statistikos savybių ir atraminių vektorių klasifikatorių kombinacijos potencialą efektyviai rotacinės įrangos gedimų diagnostikai.

Naujas būdas naudoti atraminių vektorių klasifikatorių su branduolio funkcija ir nuosekliuoju minimaliu optimizavimo algoritmu, siekiant nustatyti ir klasifikuoti guolių gedimus, buvo pristatytas Rojas'o ir Nandi [26]. Tyrimas atliktas naudojant du vibracijos duomenų rinkinius, apimančius įvairias guolių būkles. Taikant atraminių vektorių klasifikatorių ir naudojant spektrines ir statistines vibracijos signalo savybes, buvo pasiektas 99,1 % tikslumas. Tyrime pabrėžiamas pasiūlyto klasifikavimo schemos praktiškumas ir veiksmingumas riedėjimo elementų guolių gedimų nustatymo srityje.

1.6.3. K artimiausi kaimynai tyrimuose

Svorinis artimiausių kaimynų metodas guolių gedimų nustatymui buvo naudojamas Sharma'os ir kt. [27]. Tyrimo metu buvo naudojama vibracijos signalų analizė, siekiant klasifikuoti guolių būklę.

Tyrime Mahalanob'io atstumas buvo taikomas artimiausių kaimynų modeliui, o pagrindinių komponentų metodas – duomenų dimensijai sumažinti. Toks technologinis pasirinkimas leido efektyviai nustatyti įvairius guolių gedimus, demonstruojant tyrimo inovatyvumą ir praktinę reikšmę pramonės mašinų priežiūros srityje.

Metodas elektromotorių guolių gedimų nustatymui, naudojant spektrinį ekscesą ir k artimiausių kaimynų atstumų analizę, buvo sukurtas Tian'o ir kt. [28]. Metodas išsiskiria gebėjimu aptikti įvairius guolio gedimus, tokius kaip įdubimai, įtrūkimai ar dėvėjimasis, taikant spektrinio eksceso pagrindu sukuriama kaltės požymių išgavimą ir k artimiausių kaimynų atstumo matavimą. Derinant šiuos metodus su pagrindinių komponentų analize buvo formuotas sveikatos indeksas, kuris leido nustatyti guolio būklę.

1.6.4. Atsitiktiniai miškai tyrimuose

Han'o ir Jiang'o tyrime buvo pristatytas guolių gedimų diagnostikos metodas, naudojant variacinę modų dekompoziciją signalo skaidymui, autoregresinį modelį gedimo požymių išskyrimui ir atsitiktinių miškų klasifikatorių modelių atpažinimui [29]. Metodas pasižymi aukštu tikslumu gedimų lokalizavimo ir sunkumo lygio nustatymo užduotyse, demonstruodamas žymiai geresnius rezultatus palyginti su standartiniais atraminių vektorių klasifikatoriais, tuo pačiu pabrėžiant atsitiktinių miškų klasifikatorių pranašumą dėl jo mažo skaičiavimo laiko.

Tyrime, atliktame Sanchez'o ir kt., buvo pristatytas metodas sukamųjų mašinų, įskaitant guolius, gedimams nustatyti, naudojant atsitiktinius miškus ir k artimiausių kaimynų algoritmus bei savybių reitingavimo metodus, tokius kaip *ReliefF*, *Chi* kvadratų ir informacijos prieaugio [30]. Tyrimas apėmė penkias duomenų bazines, iš kurių keturios (DB1–DB4) buvo skirtos įvairių pavarų dėžių konfigūracijų gedimams nustatyti, o penktoji (DB5), paimta iš *Case Western Reserve* universiteto guolių defektų duomenų bazės, buvo skirta guolių gedimų diagnostikai. Analizės metu buvo atskleista, kad laiko savybių, įskaitant neįprastas elektromiografijos srities savybių, naudojimas gali pagerinti gedimų nustatymo tikslumą, suteikiant naujų įžvalgų.

1.6.5. Gilusis mokymas tyrimuose

Tyrimas, kuriame buvo naudojamas sąsūkų neuroninis tinklas įvairių guolių defektų klasifikavimui, pristatytas Janssens'o ir kt. [31]. Šis metodas leido automatiškai išrinkti naudingas savybes guolių gedimų nustatymui tiesiogiai iš duomenų, mažinant poreikį rankiniams savybių inžinerijos procesams. Sąsūkų neuroninio tinklo modelis parodė 93,61 % tikslumą, pranokdamas klasikinių požymių inžinerijos metodą, kuris naudojo rankiniu būdu sukurtas savybes ir atsitiktinių miškų klasifikatorių. Tyrimas parodė, kaip šiuolaikinės mašininio mokymo technikos gali pagerinti gedimų nustatymo procesus sukantiesiems mechanizmams.

Gilusis autošifravimo mokymo metodas sukamųjų mašinų gedimų diagnozei pristatytas Shao ir kt. [32]. Metodas pagrįstas naujos autošifravimo mokymo funkcijos, kuriai taikomas maksimalus korentropijos principas, sukūrimu siekiant pagerinti gedimo požymių išgavimą iš matuojamų vibracijos signalų. Be to, naudojamas dirbtinių žuvų pulko algoritmas (angl. *Artificial Fish Swarm Algorithm*, AFSA) svarbiausių autošifravimo mokymo parametrų optimizavimui, kad būtų prisitaikyta prie signalo ypatybių. Taikant šį metodą, autoriai pasiekė aukštą tikslumą diagnozuojant įvairius gedimus tiek pavarų dėžės, tiek elektrinių lokomotyvų ritininiuose guoliuose, parodant metodikos efektyvumą ir patikimumą sudėtingų gedimų atpažinimo užduotyse.

1.7. Literatūros apibendrinimas

Išnagrinėjus mokslinius šaltinius, pastebėta, kad rutulinių guolių nuspėjamos priežiūros sistemos yra neatsiejamas pramonės sektoriaus elementas, lemiantis tiek įrangos patikimumą, tiek jos efektyvumą. Tyrimai rodo, kad naudojant pažangius matematinius modelius ir technologijas, įmanoma ne tik ženkliai sumažinti gedimų skaičių, bet ir minimizuoti prastovas, optimizuojant gamybos procesų veiklą ir išteklių panaudojimą. Šiuolaikinės nuspėjamosios priežiūros sistemos sugeba efektyviai prognozuoti įrangos gedimus, remiantis iš anksto renkamais duomenimis ir jų analize. Tai leidžia organizacijoms laiku atlikti būtinas priežiūros procedūras, išvengiant brangių ir laiko reikalaujančių avarinių remonto darbų.

Tyrimai dirbtinio intelekto srityje, ypač mašininio mokymo metodų taikymas guolių gedimų diagnozei, atskleidžia įvairių metodikų potencialą ir pritaikomumą pramonėje. Dirbtinių neuroninių tinklų, atraminių vektorių klasifikatorių, k artimiausių kaimynų metodų ir atsitiktinių miškų naudojimas leidžia efektyviai nustatyti ir klasifikuoti gedimus remiantis vibracijos signalais. Gilieji mokymosi metodai, tokie kaip sąsukų neuroniniai tinklai ir gilieji autošifravimo modeliai, taip pat pasižymi aukštu tikslumu diagnozuojant įvairius guolių gedimus, mažinant poreikį rankiniams savybių inžinerijos procesams. Be to, šie metodai puikiai dirba ir su kitais duomenų modalumais: nuotraukomis ir vaizdo įrašais, garso duomenimis bei tekstiniais duomenimis, kurių analizė suteikia papildomų įžvalgų įrangos būklei vertinti. Tokios inovatyvios metodikos ne tik demonstruoja aukštą diagnostikos tikslumą, bet ir pabrėžia šiuolaikinių technologijų potencialą sudėtingose pramoninės priežiūros užduotyse.

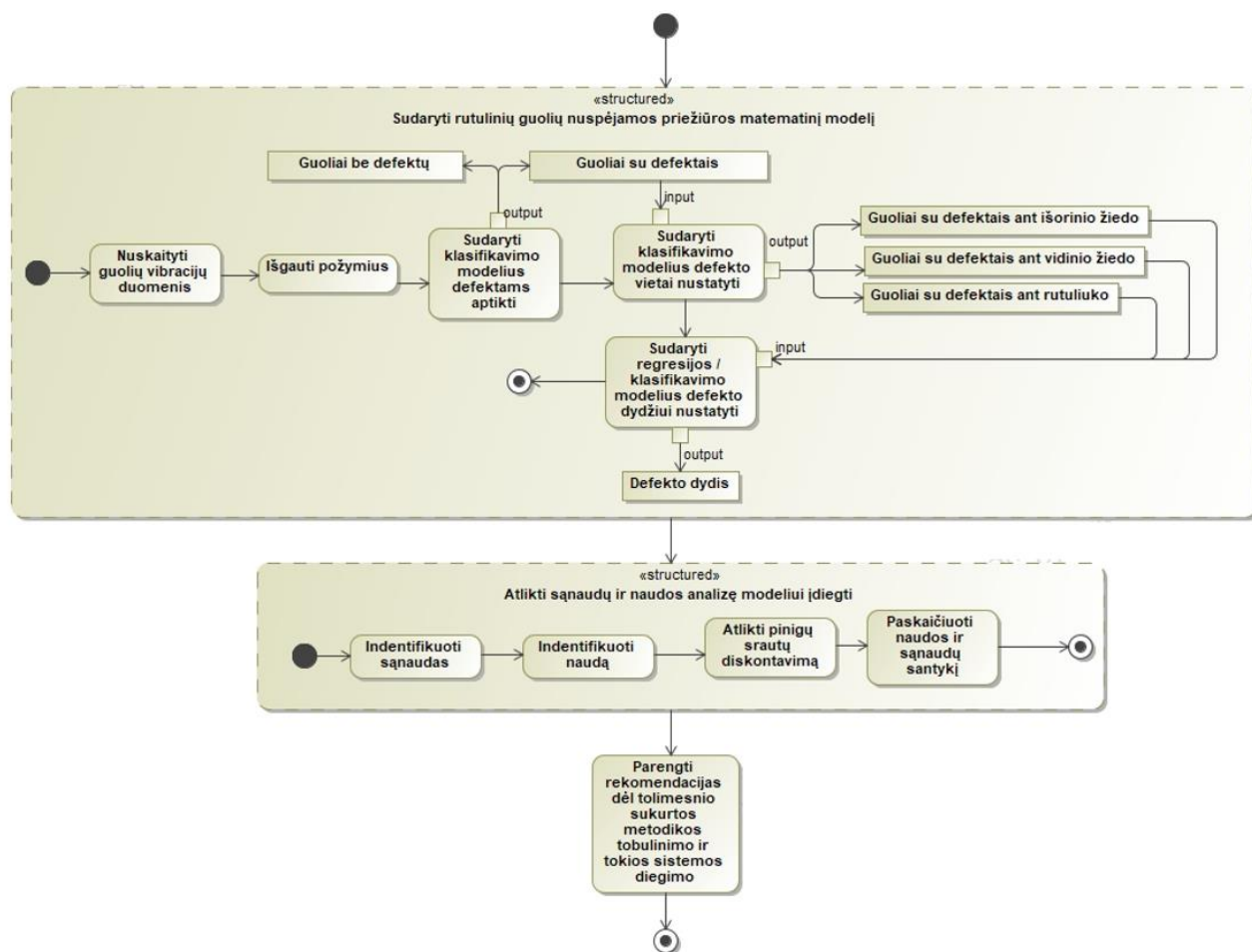
Nors šių sistemų diegimas susijęs su pradinių investicijų poreikiu, įskaitant mokymus, įrangos įsigijimą ir duomenų analizės sistemas, jų nauda yra ilgalaikė ir matoma dėl sumažėjusių eksploatacinių išlaidų ir didesnio įrangos naudojimo efektyvumo. Šiandienos pramonė vis labiau vertina tokių technologijų taikymą, ypač atsižvelgiant į augančius efektyvumo ir tvarumo reikalavimus. Vystantis technologijoms, nuspėjamoji priežiūra tampa vis labiau integruota į įvairias pramonės šakas, pradedant nuo automobilių gamybos ir baigiant aeronautika, kur kiekvienas efektyvumo pagerėjimas gali reikšti didelius finansinius sutaupymus. Todėl svarbu atkreipti dėmesį į naujausias inovacijas nuspėjamosios priežiūros srityje, nes jos leidžia ne tik sekti įrangos būklę realiu laiku, bet ir prognozuoti gedimus, užtikrinant sklandų ir pertrūkių neturintį gamybos procesą.

2. Tyrimo metodai

Skyriuje detaliai apžvelgiama metodika, kuri bus taikoma rutulinių guolių nuspėjamosios priežiūros modeliavimui. Tyrimo eigos aprašyme nurodomi pagrindiniai etapai, kuriais remiantis buvo vykdomas tyrimas. Toliau apibūdinama tyrimui atlikti naudojama įranga, tyrimo duomenys, jų kilmė bei pristatoma, kokie požymiai bus išgauti. Taip pat pristatomi pasirinkti prognozavimo metodai. Sąnaudų ir naudos analizės skyrelyje išdėstoma metodika, pagal kurią bus įvertinamas ekonominis sistemos diegimas. Galiausiai pristatomas hiperparametrų derinimo metodas ir pateikiamos modelių tinkamumo vertinimo metrikos.

2.1. Tyrimo eiga

Tyrimo eigos proceso diagrama pateikta **12 pav.** Guolių vibracijų duomenų rinkimas yra pirmasis žingsnis, leidžiantis suformuoti pagrindą tolesniam guolių būklės stebėjimui. Šie duomenys apdorojami siekiant išgauti svarbiausius požymius, kurie apibūdina guolio veikimą ir galimus nesklandumus. Išgauti požymiai tampa pagrindu klasifikavimo modelių kūrimui, kurie leidžia atskirti guolius be defektų nuo guolių su defektais. Kai defektai yra aptikti, kitas žingsnis yra defektų vietos nustatymas. Modeliai yra mokomi atpažinti defektų buvimą išoriniame žiede, vidiniame žiede arba ant rutuliuko. Toliau sukuriama regresijos / klasifikavimo modelių sudarymo procesai apima modelių treniravimą, vertinimą, derinimą bei reikšmingiausių kintamųjų nustatymą.



12 pav. Tyrimo eigos procesas

Atliekant sąnaudų ir naudos analizę yra identifikuojamos ir apskaičiuojamos galimos sąnaudos. Taip pat yra nustatoma galima ekonominė nauda rutulinių guolių nuspėjamos priežiūros sistemos. Diskontavus pinigų srautus, apskaičiuojamas naudos ir sąnaudų santykis, pagal kurį galima įvertinti rutulinių guolių nuspėjamos priežiūros sistemos diegimo patrauklumą. Galiausiai, remiantis visais atliktais tyrimais ir analize, parengiamos rekomendacijos dėl metodikos tobulinimo ir sistemos diegimo. Šios rekomendacijos apima technologinius atnaujinimus, procesų integraciją, siekiant maksimaliai išnaudoti naujai sukurto modelio galimybes.

2.2. Naudojama programinė įranga

Tyrime buvo naudojama *Python* programavimo kalba. Tai yra aukšto lygio, plačiai naudojama programavimo kalba, pasižyminti savo aiškia sintakse ir kodų skaitymo lengvumu. *Python* programavimo kalba yra populiari dėl savo universalumo ir palaikymo įvairių programavimo paradigmų, kas leidžia lengvai pritaikyti ją įvairiems programavimo uždaviniams ir duomenų analizės projektams [33]. Tyrimui buvo panaudotos šios *Python* bibliotekos:

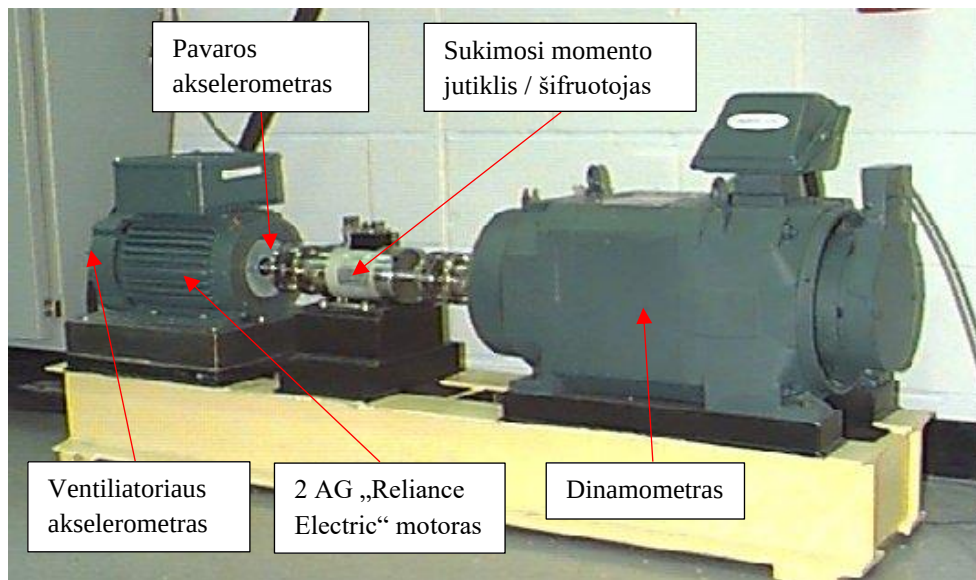
- *matplotlib* – biblioteka, kuri leidžia kurti aukštos kokybės grafikus ir diagramas;
- *numpy* – biblioteka, skirta efektyviai atlikti skaičiavimus su daugiamačiais masyvais ir matricomis;
- *pandas* – duomenų analizės ir manipuliavimo biblioteka *Python* programavimo kalboje;
- *keras* – biblioteka, skirta lengvai ir greitai projektuoti, mokyti, vertinti DTN modelius;
- *lightgbm* – biblioteka, skirta LightGBM modeliams kurti;
- *scikit-learn* – viena populiariausių *Python* bibliotekų mašiniam mokymui. Ji suteikia įrankius duomenų apdorojimui, modelių kūrimui ir vertinimui, įskaitant klasifikaciją, regresiją.

Be to, analizės procesas buvo vykdytas naudojant *Jupyter Notebook*, interaktyvią aplinką, kuri leidžia apjungti kodą, tekstą ir vizualizacijas vienoje vietoje. O sąnaudų ir naudos analizė buvo atliekama su *Microsoft Excel* skaičiuokle.

2.3. Tyrimo duomenys

Case Western Reserve universiteto guolių duomenų centro svetainėje pateikiama išsami įprastų ir su defektais rutulinių guolių duomenų bazė [34]. Šis duomenų rinkinys buvo pasirinktas matematiniams modeliams sudaryti. Duomenys gauti atlikus daugybę kruopščiai dokumentuotų eksperimentų, kuriuose daugiausia dėmesio skirta 2 AG *Reliance Electric* varikliui (žr. **13 pav.**). Šių eksperimentų rutuliniai guoliai buvo specialiai pažeisti, naudojant elektroerozinį mechaninį apdirbimą, kad būtų sukurti įvairaus dydžio defektai – nuo 0,007 colio iki 0,021 colio skersmens. Pažeidimai buvo atlikti skirtingose guolių vietose, įskaitant vidinį žiedą, patį rutuliuką ir išorinį žiedą.

Po gedimo indukcijos guoliai buvo vėl sumontuoti variklyje, o vibracijos duomenys buvo registruojami esant 0–3 arklio galių variklio apkrovai ir 1720–1797 aps/min. greičiui. Duomenis fiksuojančią aparatūrą sudarė akselerometrai, strategiškai išdėstyti svarbiausiuose variklio korpuso taškuose, konkrečiai 12 valandų padėtyje tiek pavaros, tiek ventiliatoriaus galuose, o kartais ir ant variklio pagrindo plokštės. Tokia sąranka leido tiksliai išmatuoti vibracijas, kurios yra svarbios vertinant guolių būklę ir veikimą.



13 pav. Case Western Reserve universiteto guolių bandymų stendas [34]

Tyrime naudoti vibracijos duomenys buvo renkami 12 000 įrašų per sekundę. Šis aukšto dažnio duomenų rinkimas buvo labai svarbus siekiant nustatyti subtilius vibracijos niuansus, rodančius skirtingas gedimų būsenas. Akselerometrai buvo prijungti prie 16 kanalų DAT įrašymo įrenginio, o gauti duomenys buvo apdorojami ir sutvarkyti *Matlab* aplinkoje [34]. Svetainėje pateikiami duomenų failai su ventiliatoriaus ir pavaros guolių vibracijų duomenimis, taip pat variklio sukimosi greičiu. Šis duomenų rinkinys saugomas *Matlab* (*.mat) formatu. Duomenų rinkinį sudaro daugiau kaip 13 200 000 įrašų.

Sąnaudų ir naudos analizei atlikti, duomenys buvo paimti iš viešai prieinamų šaltinių internete. Duomenų šaltiniai yra pateikti informacijos šaltinių sąrašė [1–6].

2.4. Požymių išgavimas

Efektyvią diagnozę ir prognozę galima atlikti išskiriant reikšmingiausius ir jautriausius požymius. Netikslūs ir netinkami požymiai mažina bendrą gedimų diagnozavimo ir prognozavimo patikimumą [35]. Šiame tyrime kiekviena nuskaityta laiko eilutė buvo padalijama į segmentus po 3000 įrašų. Šis dydis buvo pasirinktas taip, kad kiekvienas segmentas turėtų pakankamai duomenų statistinėms charakteristikoms apskaičiuoti, bet tuo pačiu būtų pakankamai mažas, kad atspindėtų lokalizuotas vibracijos savybes. Kiekvienam segmentui apskaičiuoti šie požymiai:

- vidurkis – charakteristika, apskaičiuojama kaip visų duomenų stebinių suma padalinta iš jų skaičiaus;
- asimetrijos koeficientas – charakteristika, apibūdinanti duomenų pasiskirstymo simetriškumą apie vidurkį;
- ekscesas – charakteristika, apibūdinanti duomenų viršūnės aštrumą;
- standartinis nuokrypis – charakteristika, matuojanti duomenų išsibarstymą apie vidurkį;
- piko faktorius (angl. *Crest Factor*, CF) – santykis tarp didžiausios vibracijos amplitudės ir kvadratų vidurkio;
- Shannon'o entropija – vidutinis informacijos, gaunamos iš stochastinio duomenų šaltinio, greitis [36]. Aukšta entropijos reikšmė rodo didelį atsitiktinumą duomenyse.

Signalas $X(t)$ Shannon'o entropija apibrėžiama pagal (1) formulę:

$$E(X) = - \sum_{i=1}^N p(x_i) \log_2(p(x_i)); \quad (1)$$

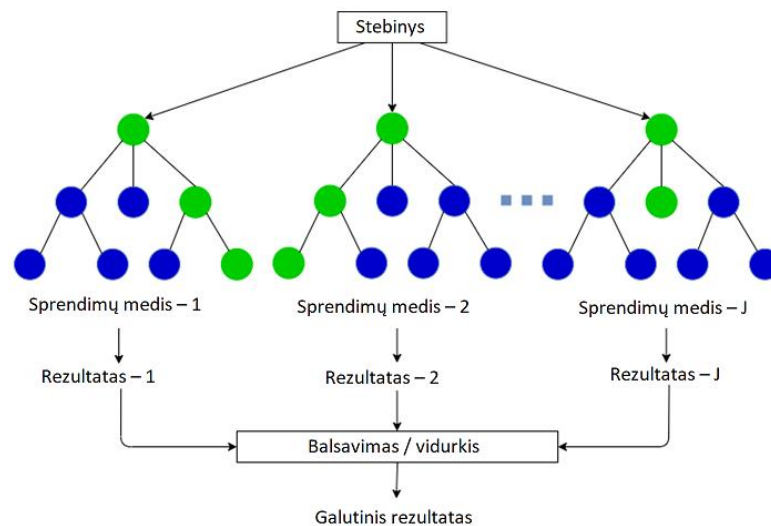
čia $p(x_i)$ – tikimybė, kad įvyks x_i .

2.5. Projekte naudoti metodai

Šiame poskyryje aprašyti pasirinkti prognozavimo metodai: atsitiktiniai miškai, LightGBM, dirbtiniai neuroniniai tinklai. Taip pat pristatyta sąnaudų ir naudos analizės metodika.

2.5.1. Atsitiktinis miškai

Atsitiktinių miškų algoritmas, kurį pasiūlė Breiman'as [37], yra kolektyvinio mokymo metodas, naudojamas tiek regresijos, tiek klasifikacijos problemoms spręsti. Algoritmas prasideda nuo duomenų atrankos *bootstrap* būdu. Iš duomenų rinkinio $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$, kur x_i yra požymių vektorius, o y_i – prognozuojamas kintamasis, atsitiktiniu būdu su pakartojimais traukiami stebiniai, kiekvienas iš jų naudojamas atskiram medžiui auginti (žr. **14 pav.**).



14 pav. Atsitiktinių miškų supaprastintas algoritmas

Medžių augimo proceso metu kiekvieno mazgų skaidymui iš p turimų kintamųjų atsitiktinai pasirenkama m kintamųjų, kur $m \leq p$. Paprastai m yra fiksuotas dydis: $m = \sqrt{p}$ klasifikavimui ir $m = N/3$ regresijai. Kiekvieno medžio mazgo skaidymo metu surandamas geriausias dvejetainis padalinimas tarp visų įmanomų dvejetainių padalinių m kintamųjų atžvilgiu. Regresijos atveju tai atliekama minimizuojant kvadratinių paklaidų sumą. Klasifikavimo atveju mazgų skaidymas vykdomas maksimizuojant informacijos prieaugį, minimizuojant entropiją arba minimizuojant Gini indeksą [38]. Gini indeksas apskaičiuojamas pagal (2) formulę:

$$Q = \sum_{k \neq k'}^K \hat{p}_k \hat{p}_{k'}; \quad (2)$$

čia $\hat{p}_k$ yra k klasės stebėjimų dalis mazge. $\hat{p}_k$ apskaičiuojamas pagal (3) formulę:

$$\hat{p}_k = \frac{1}{n} \sum_{i=1}^n (y_i = k). \quad (3)$$

Medžio augimas tęsiamas tol, kol nėra pasiekta viena iš stabdymo sąlygų: mazge esama mažiau duomenų nei minimalus leistinas duomenų kiekis, pasiektas maksimalus leistinas medžio gylis, arba papildomas mazgų skaidymas nebeturi reikšmingo poveikio klaidos mažinimui. Galutinės prognozės formavimas skiriasi priklausomai nuo uždavinio tipo. Regresijos atveju galutinė prognozė yra visų medžių prognozių aritmetinis vidurkis, kuris apskaičiuojamas pagal (4) formulę:

$$\hat{f}(x) = \frac{1}{J} \sum_{j=1}^J \hat{h}_j(x); \quad (4)$$

čia $\hat{h}_j(x)$ – j -ojo medžio prognozė taškui x . Klasifikavimo atveju, galutinė klasė yra ta, už kurią daugiausiai medžių „nubalsavo“, tai yra, kuri buvo dažniausiai prognozuojama. Galutinė klasė nustatoma pagal (5) formulę:

$$\hat{f}(x) = \arg \max_y \sum_{j=1}^J I(\hat{h}_j(x) = y); \quad (5)$$

čia I – indikatorinė funkcija, kuri lygi 1, kai prognozė $\hat{h}_j$ sutampa su y , ir 0 kitu atveju [38].

Kai *bootstrap* būdu sudaroma pradinė imtis, kai kurie stebėjimai į pradinę imtį nepatenka. Šie duomenys vadinami *out-of-bag* ir yra naudingi apibendrinimo paklaidai ir kintamųjų svarbai įvertinti.

2.5.2. LightGBM

LightGBM metodą pasiūlė Ke ir kt. [39], siekdami išspręsti problemą, kad esant dideliame duomenų kiekiui gerokai sumažėja gradientinių sprendimų medžių (angl. *Gradient Boosting Decision Trees*, GBDT) tikslumas ir efektyvumas. Šis metodas apjungia gradientu pagrįstą vienpusę atranką (angl. *Gradient-based One Side Sampling*, GOSS) ir išskirtinių požymių susiejimo (angl. *Exclusive Feature Bundling*, EFB) priemones su GBDT. GBDT modelis apibūdinamas pagal (6) formulę:

$$F(x; w) = \sum_{t=0}^T \alpha_t h_t(x; w); \quad (6)$$

čia $h_t(x; w)$ – t sprendimų medžio modelis; $F(x; w)$ – GBDT modelio prognozuojamosios vertės; x – įvesties imtis, w – sprendimų medžio parametrai, α – kiekvieno medžio svoris. Optimalus modelis išsprendžiamas pagal (7) formulę:

$$\hat{F} = \arg \min_F \sum_{i=0}^N L(y, F(x; w)); \quad (7)$$

čia $L(y, F(x; w))$ – minimizuojama nuostolių funkcija, skirta erdvei x atvaizduoti į erdvę y .

LightGBM naudoja GOSS kaip atrankos algoritmą. Didelius gradientus turinčios imtys išsaugomos, o mažus gradientus turinčios imtys atrenkamos atsitiktinai ir joms suteikiamas pastovus svoris. Tokiu būdu GOSS daugiau dėmesio skiria nepakankamai treniruotoms imtims, nekeisdamas pirminių duomenų pasiskirstymo. Dispersijos prieaugis skirstant atvejus pagal A ir B požymių poaibius yra apibrėžtas (8) lygtimi:

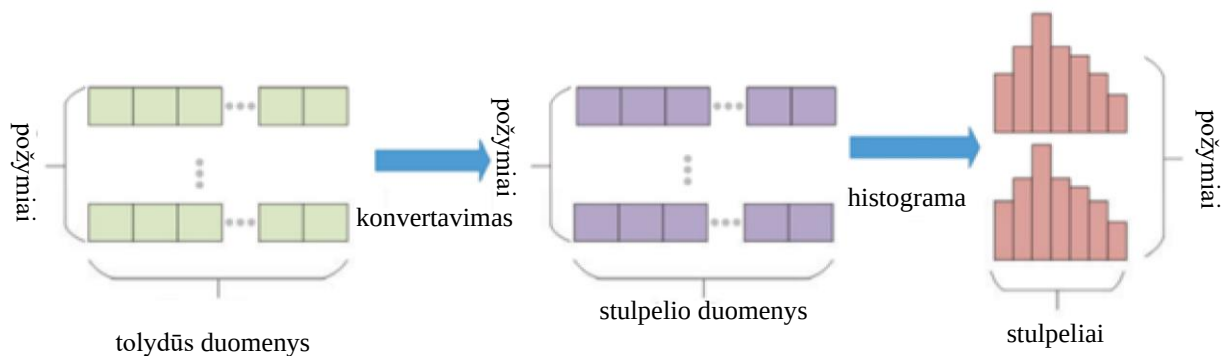
$$\tilde{V}_j(d) = \frac{1}{n}(L_1 + L_2), \quad (8)$$

$$L_1 = \frac{(\sum_{x_i \in A_l} g_i + \frac{1-a}{b} \sum_{x_i \in B_l} g_i)^2}{n_l^j(d)}, \quad (9)$$

$$L_2 = \frac{(\sum_{x_i \in A_r} g_i + \frac{1-a}{b} \sum_{x_i \in B_r} g_i)^2}{n_r^j(d)}; \quad (10)$$

čia kintamieji L_1 ir L_2 apskaičiuojami atitinkamai pagal (9) ir (10) lygtis. A poaibį sudaro didžiausi $a \times 100\%$ stebinių su dideliais gradientais. A_c yra $(1-a) \times 100\%$ stebinių, o B poaibis yra diskretiškai atrinktas $b \times |A_c|$ dydžio. Kiekvienoje gradientinio didinimo iteracijoje neigiami nuostolių funkcijos gradientai GBDT modelio išvesties atžvilgiu žymimi kaip g_i . Be to, $A_l = \{x_i \in A: x_{ij} \leq d\}$, $A_r = \{x_i \in A: x_{ij} > d\}$, $B_l = \{x_i \in B: x_{ij} \leq d\}$, $B_r = \{x_i \in B: x_{ij} > d\}$, n yra charakteristikos x dydis, a , b ir d yra konstantos [39].

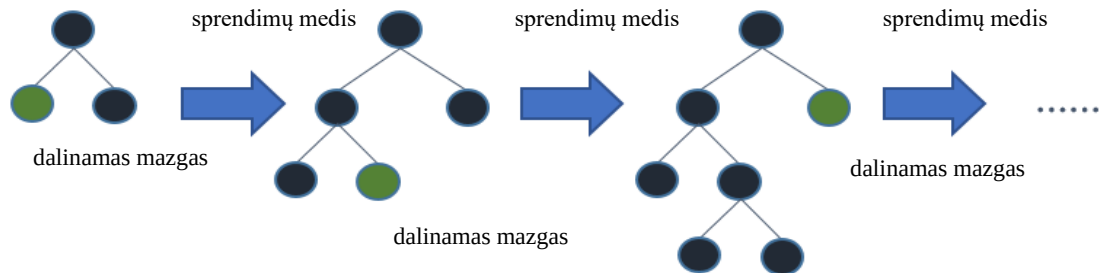
LightGBM algoritmas ne tik optimizuoja mokymo stebinius pagal GOSS, bet ir naudoja EFB požymiams išgauti, kad dar labiau padidintų tinklo mokymo greitį. Kadangi didelės dimensijos duomenys paprastai yra reti, tarpusavyje nesusiję reti požymiai EFB sujungiami į naujus požymius, o tada pagal naujus požymius sudaroma histograma (žr. **15 pav.**). Histogramos algoritmas taikomas diskretizuojant tikrines reikšmes, kad būtų galima sudaryti histogramą. Peržiūrint mokymo duomenis, skaičiuojama kiekvienos diskrečiosios histogramos reikšmės kaupiamoji statistika, kad būtų nustatytas optimalus segmentavimo taškas, o tai veiksmingai sumažina atminties sąnaudas ir laiko sudėtingumą [40].



15 pav. Histogramos algoritmas [40]

LightGBM naudoja augimo strategiją *leaf-wise* (žr. **16 pav.**). Ją galima aiškinti kaip naudingiausių lapų mazgų pasirinkimą augimui kiekviename dalijamajame mazge, kad būtų išvengta nepelningo mazgų skaidymo ir būtų sutaupomi skaičiavimo ištekliai. Tuo pat metu didžiausias gylis

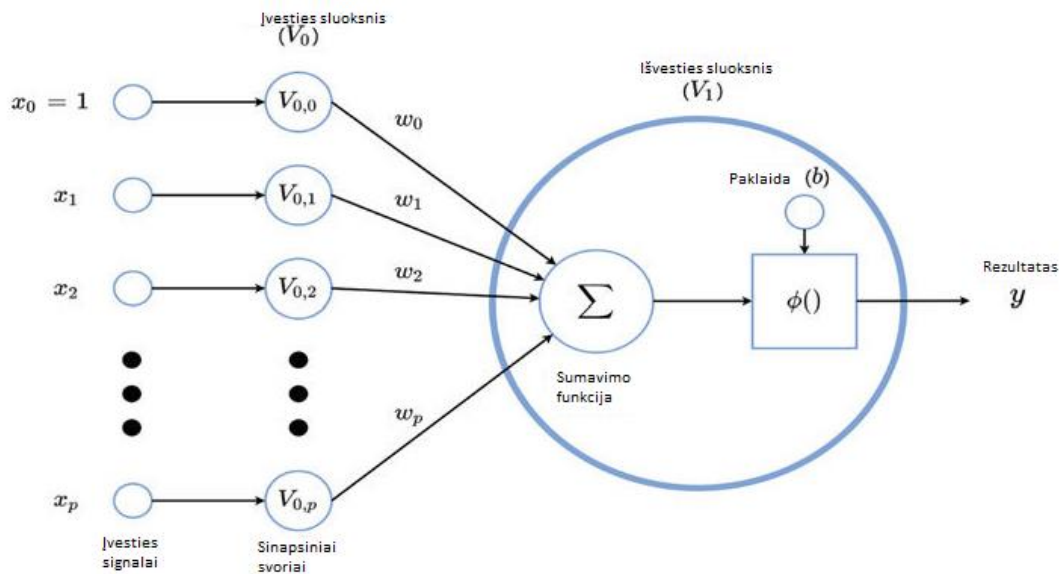
naudojamas siekiant apriboti medžio augimą, taip kontroliuojamas tinklo sudėtingumas ir išvengiama persimokymo problemos. Didinant tinklo mokymo greitį, taip pat užtikrinamas LightGBM modelio bendrinimo gebėjimas [40].



16 pav. Sprendimų medžių augimas pagal *leaf-wise* strategiją

2.5.3. Dirbtinis neuroninis tinklas

Dirbtinis neuroninis tinklas yra kompiuterių mokslo technologija, įkvėpta žmogaus smegenų struktūros ir funkcijos. Ji sudarytas iš tarpusavyje sujungtų elementų, vadinamų dirbtiniais neuronais, kurie imituoja natūralių neuronų veikimą. Bendras neuroninio tinklo modelis pateiktas **17 pav.**



17 pav. Bendras neuroninio tinklo modelis [41]

$x_1, \dots, x_p$ – tai informacija (įvestis), kurią neuronas gauna iš išorinės sistemos arba iš kitų neuronų, su kuriais jis turi ryšį. $w = (w_1, \dots, w_p)$ – tai sinapsinių svorių vektorius, kuris modifikuoja gautą informaciją, imituodamas sinapsę tarp biologinių neuronų. Juos galima interpretuoti kaip koeficientus, galinčius susilpninti arba sustiprinti vertes, kurias norima perduoti neuronui. Parametras b_j vadinamas neurono paklaida. DNT atveju mokymasis reiškia ryšio tarp nurodyto tinklo mazgų (neuronų) svorių keitimo metodą [41]. Skirtingas neuronų gaunamas vertes keičia sinapsiniai svoriai, kurie sudedami ir gaunamas grynasis įėjimas (angl. *Net Input*) apskaičiuojamas pagal (11) formulę:

$$v_j = \sum_{i=1}^p w_{ij}x_i + b. \quad (11)$$

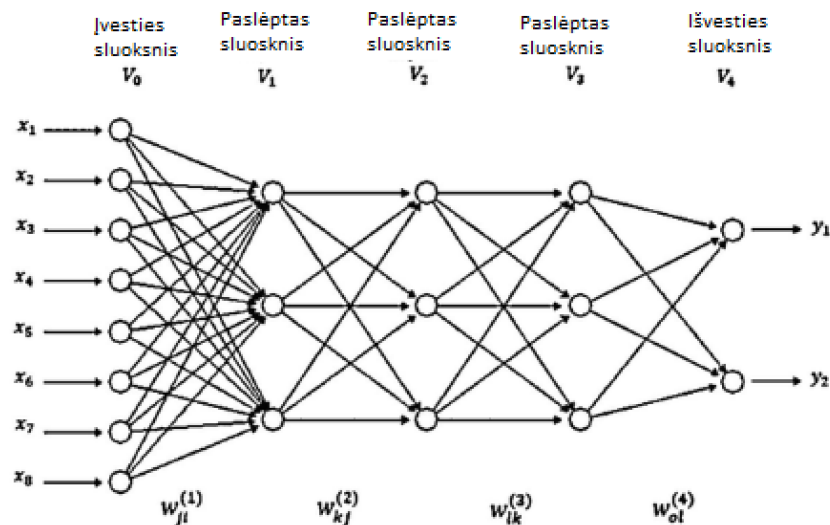
Šis grynasis įėjimas v_j lemia, ar neuronas bus aktyvuotas, ar ne. Neurono aktyvavimas priklauso nuo aktyvavimo funkcijos. Tinklo įvestis įvertinama pagal šią funkciją ir gaunama tinklo išvestis, kaip parodyta (12) formulėje:

$$y_j = g(v_j); \quad (12)$$

čia g – aktyvavimo funkcija.

Neuronai tinkle būna suskirstyti į šiuos sluoksnius [41, 42]:

- įvesties sluoksnį, tai neuronų rinkinys, kuris tiesiogiai gauna informaciją iš išorinių tinklo šaltinių. Neuronų skaičius įvesties sluoksnyje dažniausiai sutampa su tinklui pateiktų įvesties aiškinamųjų kintamųjų skaičiumi;
- paslėptus sluoksnius, kurie sudaryti iš vidinių tinklo neuronų rinkinio. Paslėptų sluoksnių skaičius gali būti 0, 1 arba daugiau (žr. **18 pav.**). Paprastai kiekvieno paslėpto sluoksnio neuronai dalijasi tos pačios rūšies informacija, dėl šios priežasties jie vadinami paslėptaisiais sluoksniais;
- išvesties sluoksnį, tai neuronų rinkinys, kuris perduoda tinklo apdorotą informaciją į išorę. Išėjimo neuronai atitinka išėjimo kintamuosius.



18 pav. Neuroninis tinklas su dvejais paslėptais sluoksniais [41]

DNT įėjimo ir paslėptų sluoksnių susiejimą lemia aktyvavimo funkcijos. Aktyvinimo funkcijos vieno sluoksnio mazgų išvestį perduoda kitam sluoksniui (iki išvesties sluoksnio imtinai). Aktyvinimo funkcijos yra skaliarinės funkcijos, suteikiančios konkrečią išvestį iš neurono. Apačioje pateikta **2 lentelė**, apibendrinanti pagrindines neuronų aktyvavimo funkcijas [41]:

2 lentelė. Neuronų aktyvavimo funkcijos

Aktyvavimo funkcija	Aprašymas	Formulė
Tiesinė	Tiesinėje funkcijose priklausomas kintamasis yra	$g(z) = Wz$

Aktyvavimo funkcija	Aprašymas	Formulė
	tiesiogiai proporcingai susijęs su nepriklausomu kintamuoju.	
ReLU	Kai įvestis mažesnė už nulį, išvestis lygi nuliui, tačiau kai įvestis pakyla virš tam tikros ribos, atsiranda tiesinis ryšys su priklausomu kintamuoju. Ši aktyvavimo funkcija dažniausiai naudojama paslėptuosiuose sluoksniuose ir išvesties sluoksniuose, kai išvesties kintamasis yra tolydus ir didesnis už nulį.	$g(z) = \max(0, z)$
Nesandarus ReLU	Tai yra modifikacija standartinės ReLU funkcijos, kuri grąžina 0 visiems neigiamiems įvesties signalams. Skirtingai nuo ReLU, nesandarus ReLU leidžia mažą teigiamą gradientą, kai įvestis yra neigiama, siekiant išvengti problemos, kai neuronai tampa neaktyvūs ir nebegali mokytis efektyviai.	$g(z) = \begin{cases} z, & \text{jeigu } z > 0 \\ \alpha z, & \text{kitu atveju} \end{cases}$
Sigmoidė	Transformuoja įvesties signalą į reikšmę, kuri yra tarp 0 ir 1. Ši aktyvavimo funkcija yra viena iš labiausiai paplitusių aktyvavimo funkcijų tipų, naudojamų kuriant DNT, kai rezultatas yra tikimybė arba dvejetainis rezultatas.	$g(z) = (1 + e^{-z})^{-1}$
Softmax	Softmax yra sigmoidės aktyvavimo funkcijos apibendrinimas, skirtas kategoriniams rezultatams. Softmax yra funkcija, kurią dažnai randama klasifikatoriaus, turinčio daugiau nei dvi kategorijas, išvesties sluoksnyje.	$g(z_j) = \frac{e^{z_j}}{1 + \sum_{c=1}^C e^{z_c}}, j = 1, \dots, C$
Tanh	Tanh funkcija paverčia bet kokią įvesties signalą į išvestį, kuri yra tarp -1 ir 1. Tai suteikia funkcijai galimybę normalizuoti išvestį taip, kad ji būtų simetriška apie nulį, o tai yra naudinga daugeliui uždavinių, kur balansuoti duomenys yra svarbūs.	$\tanh(z) = \frac{\sinh(z)}{\cosh(z)} = \frac{e^z - e^{(-z)}}{e^z + e^{(-z)}}$

2.5.4. Sąnaudų ir naudos analizė

Tyrimo objektas – vienas traukinio vagonas. Sąnaudų ir naudos analizė bus atliekama 5 metams.

Sąnaudų ir naudos analizė (angl. *Cost-Benefit Analysis*, CBA) yra metodas, naudojamas vertinant ekonominius sprendimus. Tai finansinis įrankis, padedantis palyginti projekto teikiamą naudą su projekto susijusiomis sąnaudomis bei įvertinti, ar konkretus projektas yra ekonomiškai pagrįstas [43]. Sąnaudų ir naudos analizė yra sudaryta iš šių etapų:

1. sąnaudų identifikavimo. Įvertinamos visos tiesioginės ir netiesioginės sąnaudos, susijusios su projektu. Tai apima pradines išlaidas dėl įrangos pirkimo, įrengimo, mokymų bei einamąsias išlaidas dėl techninės priežiūros, prenumeratų, personalo.
2. naudos identifikavimo. Tai apima sutaupytas lėšas dėl sumažėjusių prastovų, išlaidų sutaupymą dėl ilgesnio įrangos tarnavimo laiko, naudą dėl didesnio efektyvumo, didesnes pajamas dėl geresnės gaminių kokybės ir mažesnio defektų skaičiaus;
3. pinigų srautų diskontavimo. Būsimos naudos ir sąnaudos yra diskontuojamos į dabartinę vertę, leidžiant jas objektyviai įvertinti ir palyginti laiko atžvilgiu.
4. naudos ir sąnaudų santykio (angl. *Benefit-Cost Ratio*, BCR) apskaičiavimo ir įvertinimo. Jeigu BCR yra didesnis nei 1, projektas yra laikomas finansiškai prasmingu, kadangi jo teikiamos naudos viršija sąnaudas. Jei BCR yra lygus 1, projektas yra ekonomiškai neutralus, o jei

mažesnis nei 1, projektas laikomas ekonomiškai neefektyviu, nes sąnaudos viršija gaunamas naudas. BCR apskaičiuojamas pagal (13) formulę:

$$BCR = \frac{\sum_{t=1}^n \frac{B_t}{(1+r)^t}}{\sum_{t=0}^n \frac{C_t}{(1+r)^t}}; \quad (13)$$

čia n – analizuojamų metų skaičius; t – pinigų srauto atsiradimo laikotarpis; r – diskonto norma; B_t – veiklos naudos t metais; C_t – veiklos sąnaudos t metais.

2.6. Hiperparametrų derinimas

Hiperparametrų derinimas yra procesas mašininio mokymosi modelių kūrimo, kurio tikslas yra atrasti optimalius hiperparametrų reikšmių derinius [44]. Tyrimui buvo pasirinktas atsitiktinės paieškos (angl. *Random Search*) metodas. Šio metodo esmė yra atsitiktinai parinkti hiperparametrų reikšmes iš nustatytų intervalų ir įvertinti, kaip pasirinktos reikšmės veikia modelio našumą. Atsitiktinė paieška dažnai naudojama dėl savo paprastumo ir gebėjimo efektyviai naršyti didelėse parametrų erdvėse, ypač kai keli hiperparametrai turi nevienodą poveikį modelio efektyvumui [44]. Visiems analizuojamiems modeliams parametrų derinimas buvo atliekamas 200 iteracijų, taikant 5 dalių kryžminę validaciją. **3 lentelė** apibendrinanti tyrimo modeliuose derinamus hiperparametrus pateikt apačioje.

3 lentelė. Modeliuose derinami hiperparametrai

Modelis	Hiperparametras	Paaiškinimas	Parametrų erdvė
Atsitiktiniai miškai	<i>max_depth</i>	Maksimalus medžio gylis.	Atsitiktinis sveikas skaičius nuo 4 iki 10 (tolygiai pasiskirstęs).
	<i>min_samples_split</i>	Minimalus stebinių skaičius skaidant mazgą.	Atsitiktinis sveikas skaičius nuo 2 iki 10 (tolygiai pasiskirstęs).
	<i>min_samples_leaf</i>	Minimalus skaičius stebinių kiekvienam lapui.	Atsitiktinis sveikas skaičius nuo 1 iki 10 (tolygiai pasiskirstęs).
Neuroniniai tinklai	<i>neurons_1</i>	Pirmojo paslėptojo sluoksnio neuronų skaičius.	64, 128, 256.
	<i>neurons_2</i>	Antrojo paslėptojo sluoksnio neuronų skaičius.	16, 32.
	<i>activation_1</i>	Pirmojo paslėpto sluoksnio aktyvacijos funkcija.	<i>relu, tanh</i> .
	<i>activation_2</i>	Antrojo paslėpto sluoksnio aktyvacijos funkcija.	<i>relu, tanh</i> .
	<i>optimizer</i>	Optimizavimo algoritmas, naudojamas modelio mokymui.	<i>adam, SGD</i> .
	<i>epochs</i>	Mokymo epochų skaičius.	50, 100.
	<i>batch_size</i>	Mokymo partijos dydis. Nurodo, kiek mokymo pavyzdžių bus perduota tinklui vienu metu.	16, 32, 64.
LightGBM	<i>num_leaves</i>	Lapų skaičius viename medyje.	Atsitiktinis sveikas skaičius nuo 20 iki 60 (tolygiai pasiskirstęs).
	<i>learning_rate</i>	Mokymosi greitis, kuris nurodo žingsnio dydį kiekvienoje	Atsitiktinis realus skaičius nuo 0,05 iki 0,1 (tolygiai

Modelis	Hiperparametras	Paaškinimas	Parametrų erdvė
		iteracijoje, kad pasiektų globalų minimumą.	pasiskirstęs).
	<i>n_estimators</i>	Medžių skaičius.	Atsitiktinis sveikas skaičius nuo 50 iki 300 (tolygiai pasiskirstęs).
	<i>max_depth</i>	Maksimalus medžio gylis.	Atsitiktinis sveikas skaičius nuo 3 iki 10 (tolygiai pasiskirstęs).
	<i>min_child_weight</i>	Minimalus vaikų svoris, reikalingas padalijimui.	Atsitiktinis sveikas skaičius nuo 1 iki 11 (tolygiai pasiskirstęs).
	<i>subsample</i>	Duomenų poaibio dalis, naudojama kiekviename medyje.	Atsitiktinis realus skaičius nuo 0,6 iki 1 (tolygiai pasiskirstęs).
	<i>colsample_bytree</i>	Stulpelių dalis, naudojama kiekviename medyje.	Atsitiktinis realus skaičius nuo 0,6 iki 1 (tolygiai pasiskirstęs).

2.7. Modelių vertinimas

Šiame poskyryje pateikiama informacija apie klasifikavimo ir regresijos modelių vertinimo metrikas ir metodus.

2.7.1. Klasifikavimo modelių vertinimas

Sumaišymo matrica (angl. *Confusion Matrix*) – vienas iš pagrindinių metodų klasifikavimo problemoms vertinti. Tai yra lentelė, kuri rodo santykius tarp tikrųjų ir prognozuojamų klasių, suteikdama vizualinį būdą suprasti, kiek gerai modelis atpažįsta kiekvieną klasę [45]. Apačioje pateikta **4 lentelė**, kurioje atvaizduota dvejetainio klasifikavimo sumaišymo matrica.

4 lentelė. Sumaišymo matrica projekto atžvilgiu

		Tikroji klasė	
		Teigiama	Neigiama
Prognozuojama klasė	Teigiama	TP	FN
	Neigiama	FP	TN

Sumaišymo matrica sudaryta iš šių pagrindinių elementų:

- TP (angl. *True Positive*) – modelio teisingai nuspėtos teigiamos klasės atvejų skaičius;
- FN (angl. *False Negative*) – modelio klaidingai nuspėtos neigiamos klasės atvejų skaičius;
- FP (angl. *False Positive*) – modelio klaidingai nuspėtos teigiamos klasės atvejų skaičius;
- TN (angl. *True Negative*) – modelio teisingai nuspėtos neigiamos klasės atvejų skaičius.

Nors naudojant sumaišymo matricą galima gauti svarbios informacijos apie klasifikavimo algoritmų veikimą, ši matrica ne visada yra patogiausias įrankis algoritmų efektyvumo vertinimui ir palyginimui. Dažnai naudojami įvairūs rodikliai, leidžiantys palyginti klasifikavimo algoritmų efektyvumą. Žemiau pateikta **5 lentelė**, apibendrinanti įvairius klasifikavimo rodiklius.

5 lentelė. Klasifikavimo modelių vertinimo rodikliai

Rodiklis	Aprašymas	Formulė
Tikslumas (angl. <i>Accuracy</i>)	Apibrėžia, kokia dalis visų prognozuotų atvejų buvo teisingai identifikuota.	$\frac{TP + TN}{TP + TN + FN + FP}$

Rodiklis	Aprašymas	Formulė
Preciziškumas (angl. <i>Precision</i>)	Apibrėžia, kokia dalis nuspėtos teigiamos klasės atvejų buvo iš tikrųjų teisingos.	$\frac{TP}{TP + FP}$
Jautrumas (angl. <i>Sensitivity</i>)	Apibrėžia, kokia dalis tikrųjų teigiamų atvejų buvo teisingai identifikuoti kaip teigiami.	$\frac{TP}{TP + FN}$
Specifiškumas (angl. <i>Specificity</i>)	Apibrėžia, kokia dalis tikrųjų neigiamų atvejų buvo teisingai identifikuoti kaip neigiami.	$\frac{TN}{TN + FP}$
F ₁ įvertis (angl. <i>F₁ Score</i>)	Harmoninis vidurkis tarp preciziškumo ir jautrumo, matuojantis klasifikatoriaus tikslumą esant klasių disbalansui.	$2 \times \frac{\text{preciziškumas} \times \text{jautrumas}}{\text{preciziškumas} + \text{jautrumas}}$

Taip pat, vertinant klasifikavimo modelius, naudojama ROC (angl. *Receiver Operating Characteristics*) kreivė ir AUC (angl. *Area Under the Curve*) rodiklis, kuris apibūdina plotą po ROC kreive. ROC kreivės abscisių ašyje vaizduojamas *jautrumas*, o ordinačių ašyje – *1 – specifiškumas*. Didelis AUC reiškia, kad modelis gerai atskiria klases, o mažas AUC rodo, kad modelio gebėjimai yra artimi atsitiktiniam spėjimui. AUC yra ypač naudingas modelių lyginimui, nes suteikia vieną matavimą, apibendrinantį modelio našumą esant įvairioms ribinėms vertėms [45].

Esant daugiau nei dviem klasėms, klasifikavimo modelių vertinimo rodikliai, gali būti apskaičiuojami kiekvienai klasei atskirai, o galutinis rezultatas dažnai yra apibendrinamas, taikant svertinį vidurkį arba makrovidurkį. Plotas po ROC dažniausiai naudojamas dvejų klasių problemoms, kadangi matuojama, kaip gerai atskiriamos dvi klasės. Vis dėlto, AUC taip pat gali būti pritaikytas daugiau nei dviem klasėms klasifikuoti remiantis šiomis strategijomis:

- vienas prieš visus (angl. *One-vs-Rest*, OvR). Ši strategija apima AUC skaičiavimą kiekvienai klasei, laikant ją kaip teigiamą, o visas kitas klases – kaip neigiamas.
- vienas prieš vieną (angl. *One-vs-One*, OvO). Ši metodika reikalauja skaičiuoti AUC kiekvienai porai klasių atskirai.

2.7.2. Regresijos modelių vertinimas

Regresijos modelių vertinimas yra būtinas žingsnis siekiant įvertinti modelio tikslumą ir patikimumą duomenų prognozavimo kontekste. Tai atliekama įvertinant modelio sukurtų prognozių ir faktinių duomenų skirtumus, naudojant įvairius matavimo rodiklius [46].

Vidutinė absoliuti paklaida (angl. *Mean Absolute Error*, MAE) – matematinis rodiklis, naudojamas prognozių tikslumui įvertinti. MAE apskaičiuojamas pagal (14) formulę kaip vidutinė absoliutinių skirtumų suma tarp prognozuojamų ir tikrųjų reikšmių:

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i|; \quad (14)$$

čia $\hat{y}_i$ – imties prognozuojama reikšmė; y_i – imties tikroji reikšmė; n – stebinių kiekis imtyje.

Vidutinė kvadratinė paklaida (angl. *Mean Squared Error*, MSE) – matematinis rodiklis, matuojantis vidutinį kvadratų skirtumą tarp modelio prognozuojamų reikšmių ir tikrųjų duomenų reikšmių. MSE apskaičiuojamas pagal (15) formulę:

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2; \quad (15)$$

čia $\hat{y}_i$ – imties prognozuojama reikšmė; y_i – imties tikroji reikšmė; n – stebinių kiekis imtyje.

Determinacijos koeficientas (R^2) yra statistikos rodiklis, naudojamas modeliavime ir regresijos analizėje, siekiant įvertinti, kokia dalis duomenų variacijos yra paaiškinama naudojant regresijos modelį. R^2 apskaičiuojamas pagal (16) formulę:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}; \quad (16)$$

čia $\hat{y}_i$ – imties prognozuojama reikšmė; y_i – imties tikroji reikšmė; n – stebinių kiekis imtyje; $\bar{y}_i$ – visų imties reikšmių vidurkis.

Vidutinė absoliutinė procentinė paklaida (angl. *Mean Absolute Percentage Error*, MAPE) – matematinis rodiklis, matuojantis vidutinės absoliučias procentines paklaidas tarp modelio prognozuojamų reikšmių ir tikrųjų reikšmių. MAPE apskaičiuojamas pagal (17) formulę:

$$MAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} \frac{|y_i - \hat{y}_i|}{|y_i|} \times 100; \quad (17)$$

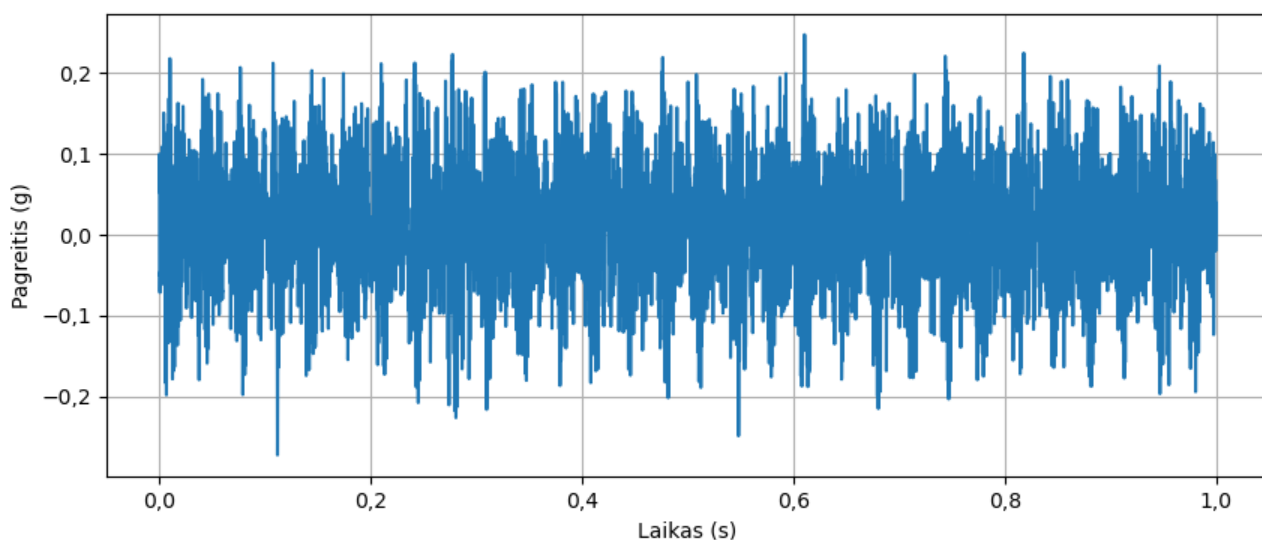
čia $\hat{y}_i$ – imties prognozuojama reikšmė; y_i – imties tikroji reikšmė; n – stebinių kiekis imtyje.

3. Tyrimo rezultatai

Skyrius pradedamas nuo žvalgomosios analizės ir požymių išskyrimo, kurioje apžvelgiamos duomenų savybės, svarbios guolių būklei įvertinti. Toliau skyriuje aptariamas defektų identifikavimas, lokalizavimas bei jų dydžio nustatymas, naudojant įvairius mašininio mokymo algoritmus. Skyrius baigiamas sąnaudų ir naudos analize, siekiant išnagrinėti, ar finansiškai naudingas būtų sistemos diegimas. Pabaigoje pateikiamos rekomendacijos dėl tolimesnio sukurtos metodikos tobulinimo, diegimo.

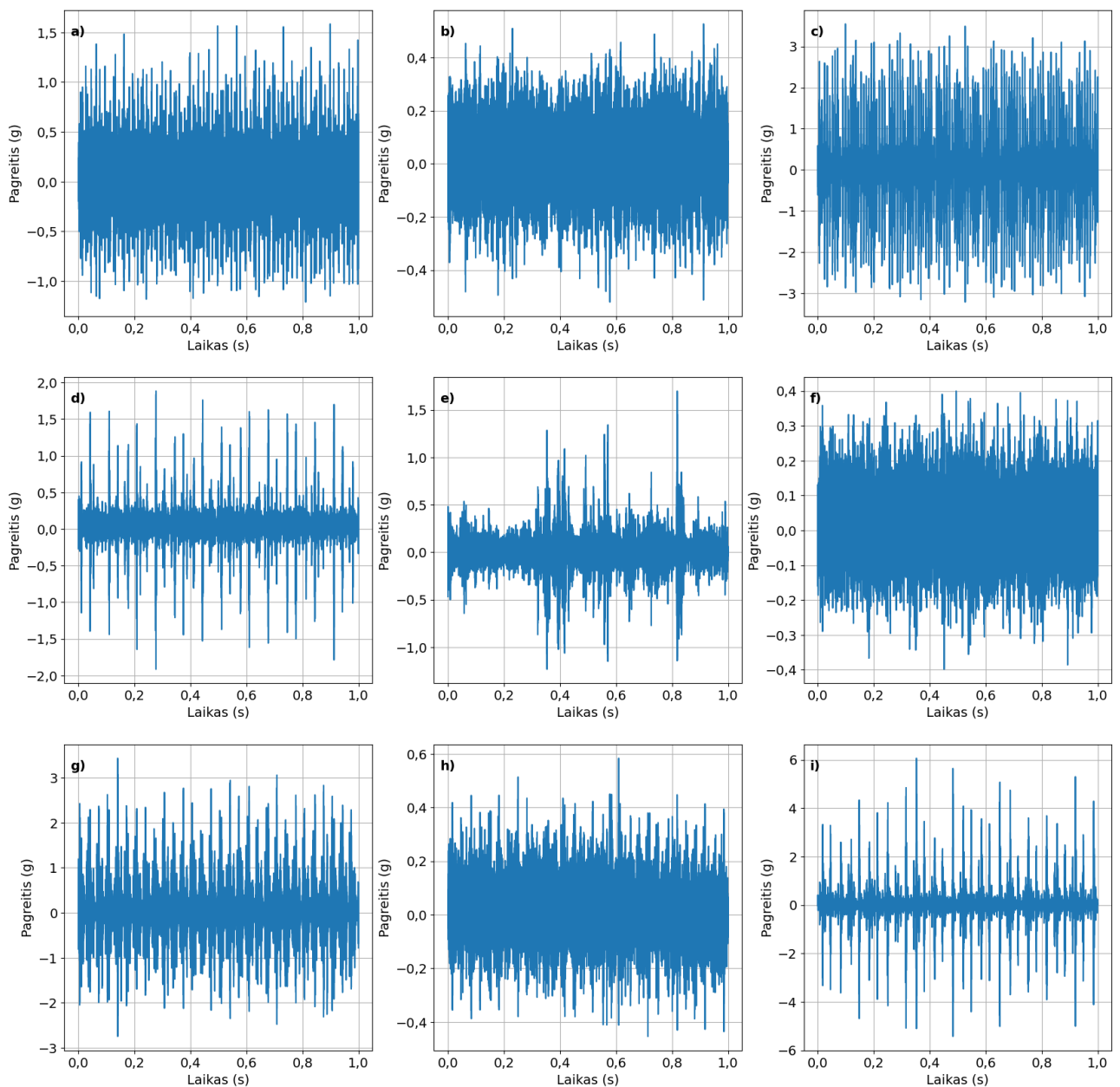
3.1. Žvalgomoji analizė ir požymių išskyrimas

Guolio be defektų vibracijos signalą, esant 1797 aps/min motoro greičiui, galima pamatyti **19 pav.** Signalas yra ties nulių, o jo amplitudė 1 sekundės trukmės metu svyruoja nuo $-0,2728$ iki $0,2474$ g; čia g – laisvojo kritimo pagreitis, kuris lygus $9,8 \frac{m}{s^2}$.



19 pav. Guolio be defektų vibracija, esant 1797 aps/min motoro greičiui

20 pav. pateikiami devyni grafikai, pažymėti nuo a) iki i), kuriuose pavaizduoti guolių su skirtingo dydžio defektais 1 sekundės trukmės vibracijos duomenys, išmatuoti varikliui veikiant 1797 aps/min greičiu. Pirmoje eilutėje a)–c) grafikai atspindi guolių su defektu ant vidinio žiedo vibraciją. Antroje eilutėje d)–f) grafikuose pavaizduotos vibracijos guolių, turinčių defektą ant rutuliuko. Trečioje eilutėje g)–i) grafikuose vaizduojama guolių su defektu ant išorinio žiedo vibraciją. Grafikų stulpeliuose pateikiami vibracijų duomenys esant skirtingiems defektų dydžiams: pirmame stulpelyje – 0,007 colio defektams, antrame – 0,014 colio defektams, ir trečiame – 0,021 colio defektams. Iš grafikų matyti, kad visų tipų defektai (vidinio žiedo, rutuliuko, išorinio žiedo) sukelia reikšmingus vibracijų pokyčius. Guolio be defektų vibracijos amplitudė yra žymiai mažesnė už bet kurios kategorijos guolių su defektais amplitudę.



20 pav. Guolių vibracija, kai defektai yra 0,007, 0,014 ir 0,021 colio dydžio, esant 1797 aps/min motoro greičiui: a)–c) guolių su defektu ant vidinio žiedo vibracija, d)–f) guolių su defektu ant rutuliuko vibracija, g)–i) guolių su defektu ant išorinio žiedo vibracija

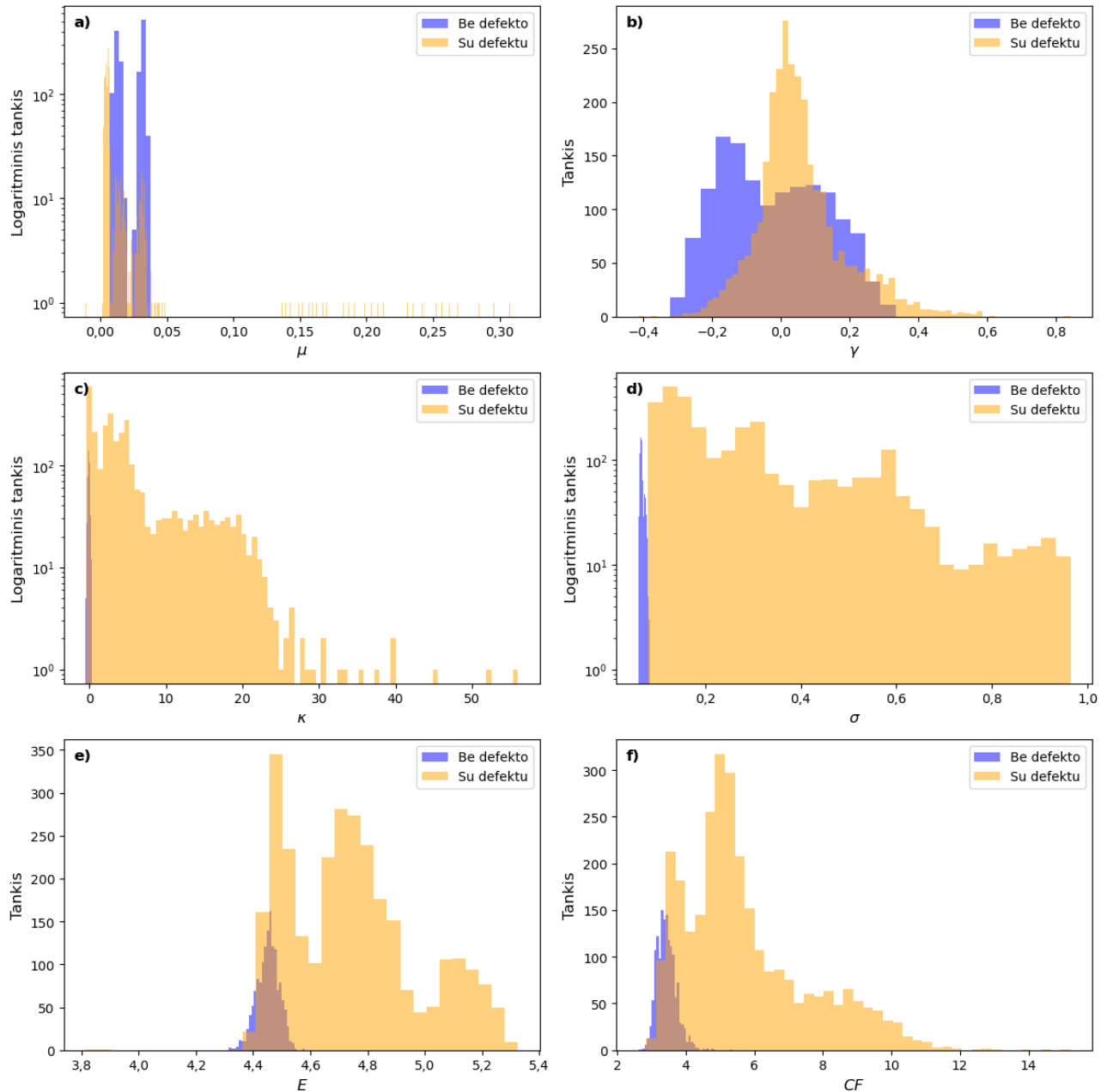
Vibracijos duomenų analizės metu buvo išgauti požymiai. Padalinus kiekvieną laiko eilutę į segmentus po 3000 įrašų, kiekvienam segmentui apskaičiuotos duomenų statistinės charakteristikos: vidurkis, asimetrijos koeficientas, ekscesas, standartinis nuokrypis, Shannon'o entropija, piko faktorius. Požymiai analizei atlikti įrašyti į *pandas* tipo duomenų lentelę.

6 lentelė, apibendrinanti gautų duomenų pasiskirstymą po požymių išgavimo pateikta apačioje. Duomenų rinkinys yra sudarytas iš 4 412 įrašų, iš kurių 2 952 (67 %) įrašai yra guolių su defektu ir 1 460 (33 %) – guolių be defekto. Defektų vieta yra skirstoma į tris kategorijas: ant rutuliuko, ant vidinio žiedo ir ant išorinio žiedo, kiekvienoje kategorijoje esant po 984 stebinius. Nurodyti trys skirtingi defektų dydžiai: 0,007, 0,014 ir 0,021, kiekvienas sudarantis po 33,(3) % visų defektų.

6 lentelė. Duomenų pasiskirstymas būklės, defekto vietos ir dydžio atžvilgiu

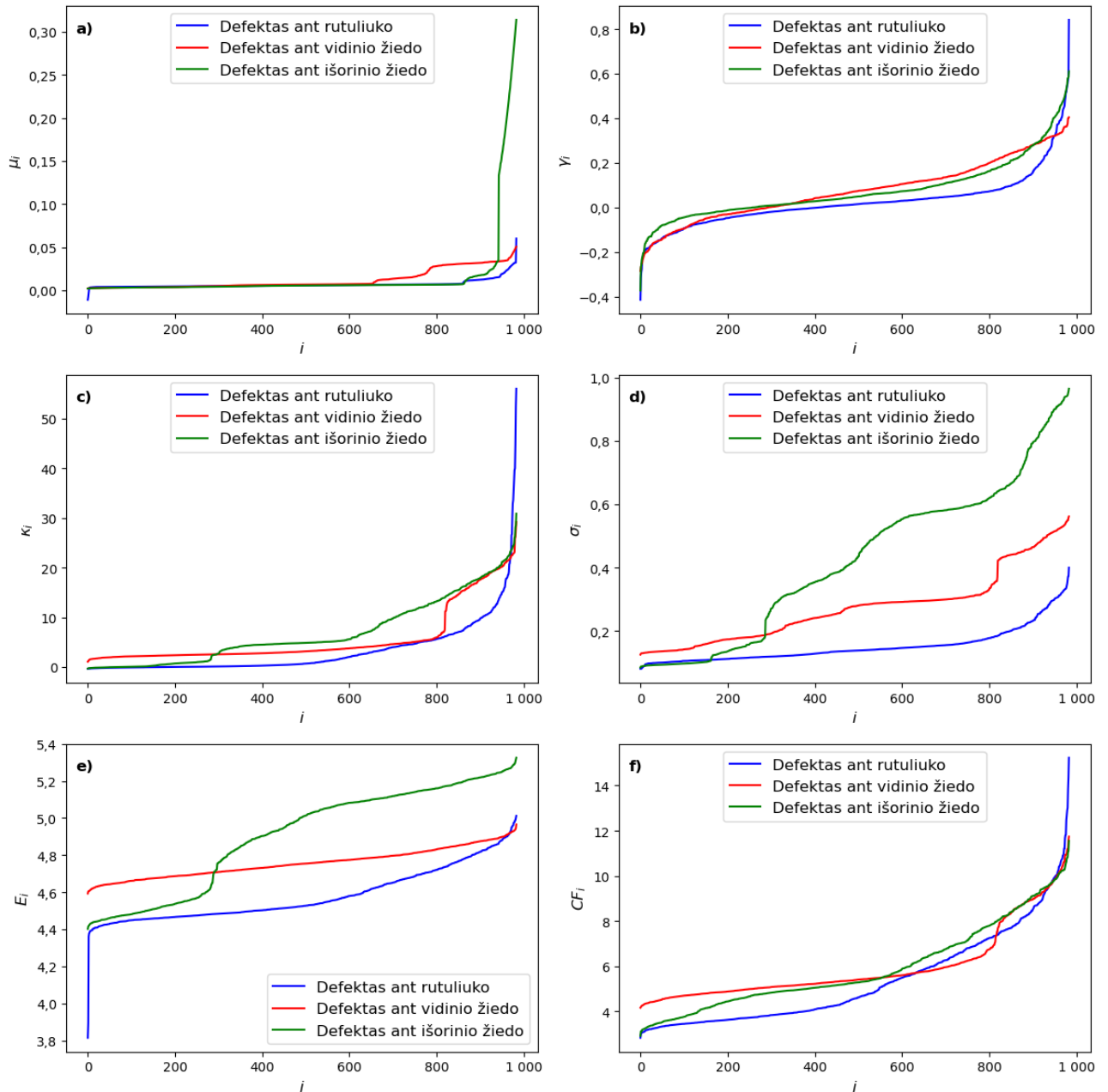
Atributas	Reikšmė	Kiekis	Procentinė dalis
Būklė	Su defektu	2952	67
	Be defekto	1460	33
Defekto vieta	Ant rutuliuko (R)	984	33,(3)
	Ant vidinio žiedo (VŽ)	984	33,(3)
	Ant išorinio žiedo (IŽ)	984	33,(3)
Defekto dydis	0,007	984	33,(3)
	0,014	984	33,(3)
	0,021	984	33,(3)

Palyginus guolius be defektų ir guolius su defektais, pastebėti dideli skirtumai tarp jų statistinių savybių (žr. **21 pav.**). Mažesnės vidutinės vertės buvo nustatytos tokioms statistinėms charakteristikoms kaip vidurkis, asimetrijos koeficientas, ekscesas, standartinis nuokrypis, Shannon'o entropija ir piko faktorius guoliuose be defektų. Mažesnė taip pat buvo ir guolių be defektų statistinių charakteristikų sklaida. Priešingai, guolių su defektais atveju didesni buvo tiek statistinių charakteristikų vidurkiai, tiek ir jų sklaida. Šis skirtumas ypač pastebimas c) ir d) grafikuose, kur buvo didesnis ekstremalių nuokrypių skaičius, pabrėžiantis nereguliarias eksploatacines charakteristikas.



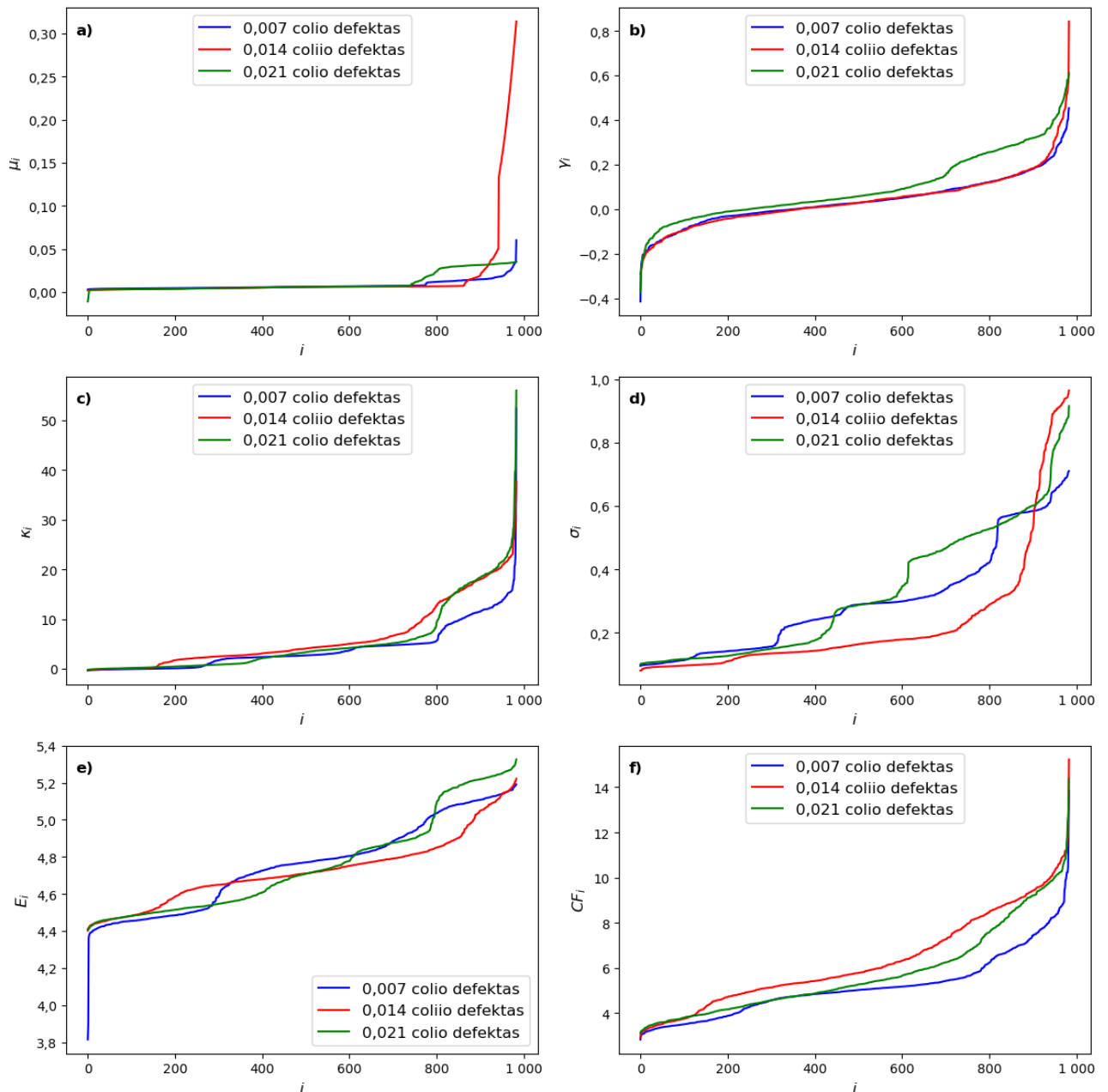
21 pav. Guolių su defektu ir be defekto statistinių charakteristikų tankio diagramos kiekvienai duomenų rinkmeni: a) vidurkis, b) asimetrijos koeficientas, c) ekscesas, d) standartinis nuokrypis, e) Shannon'o entropija, f) piko faktorius

Išnagrinėjus statistines charakteristikas guolių su defektais ant rutuliukų, vidinių žiedų ir išorinių žiedų, galime pastebėti keletą dėsningumų (žr. **22 pav.**). Iš b) ir c) grafikų matoma, kad asimetrijos koeficiento ir eksceso maksimalios reikšmės yra didesnės esant defektams ant rutuliuko. Taip pat vidutinės standartinio nuokrypio, Shannon'o entropijos, vidurkio reikšmės yra didesnės guolių su defektais ant išorinio žiedo.



22 pav. Statistinių charakteristikų vertės pagal defekto vietą: a) vidurkis, b) asimetrijos koeficientas, c) ekscesas, d) standartinis nuokrypis, e) Shannon'o entropija, f) piko faktorius

Buvo lyginamos ir guolių su 0,007, 0,014, 0,021 colio defektais statistinės charakteristikos (žr. 23 pav.). Pastebėta, kad visų charakteristikų vidutinės vertės trimis kategorijoms (0,007, 0,014 ir 0,021) yra panašios, su nedideliais svyravimais. Smarkiai išsiskiria vidurkio charakteristikos 0,014 colio defekto maksimali reikšmė nuo kitų defekto dydžių maksimalių vidurkio charakteristikos reikšmių.



23 pav. Statistinių charakteristikų vertės pagal defekto dydį: a) vidurkis, b) asimetrijos koeficientas, c) ekscesas, d) standartinis nuokrypis, e) Shannon'o entropija, f) piko faktorius

3.2. Guolių defektų identifikavimas

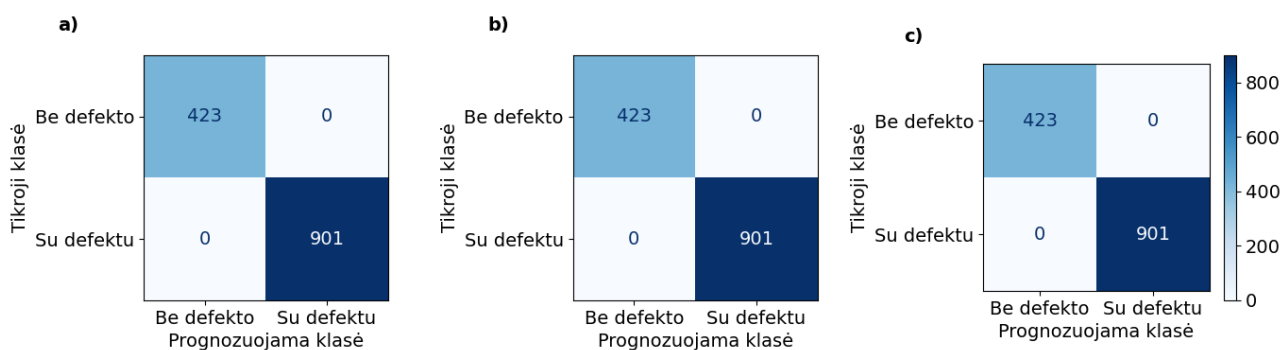
Duomenų rinkinys buvo padalintas į mokymo ir testavimo imtis. 70 % (3088) duomenų buvo skirta mokymui ir 30 % (1324) – duomenų testavimui. Hiperparametrų derinimas buvo atliktas siekiant maksimizuoti F_1 įvertį. Toliau pateikta **7 lentelė** su geriausiomis hiperparametrų derinimo metu gautomis reikšmėmis.

7 lentelė. Defekto identifikavimo modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)

Modelis	Hiperparametras	Reikšmė
Atsitiktiniai miškai	<i>max_depth</i>	9
	<i>min_samples_split</i>	5

Modelis	Hiperparametras	Reikšmė
	<i>min_samples_leaf</i>	1
Neuroniniai tinklai	<i>neurons_1</i>	256
	<i>neurons_2</i>	16
	<i>activation_1</i>	<i>relu</i>
	<i>activation_2</i>	<i>relu</i>
	<i>optimizer</i>	<i>adam</i>
	<i>epochs</i>	50
	<i>batch_size</i>	32
LightGBM	<i>num_leaves</i>	46
	<i>learning_rate</i>	0,0456
	<i>n_estimators</i>	178
	<i>max_depth</i>	8
	<i>min_child_weight</i>	1
	<i>subsample</i>	0,7708
	<i>colsample_bytree</i>	0,6488

Klasių disbalanso problema buvo spręsta kiekvienai klasei automatiškai priskiriant svorius, atvirkščiai proporcingus klasių dažnumui duomenų imtyje. Su atsitiktiniais miškais, neuroniniais tinklais ir LightGBM buvo nustatyti 423 atvejai be defekto ir 901 atvejis su defektu. Visų trijų modelių sumaišymo matricos yra identiškos, nėra nei vienos klaidingos prognozės (žr. **24 pav.**).

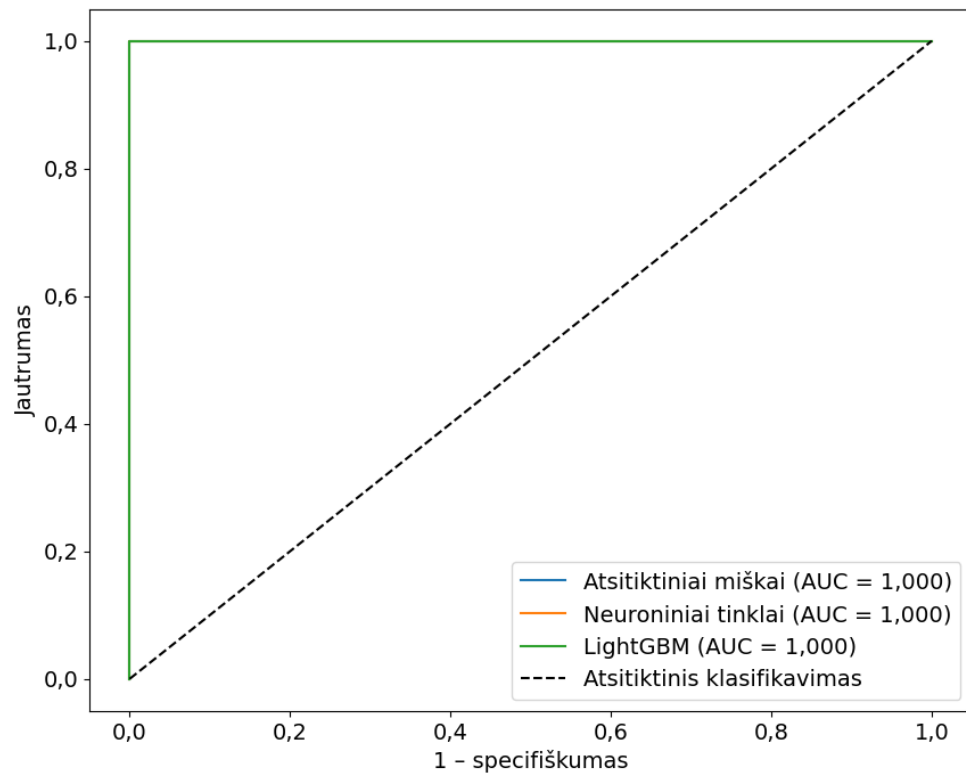


24 pav. Defekto identifikavimo modelių sumaišymo matricos: a) atsitiktinių miškai, b) neuroniniai tinklai, c) LightGBM

8 lentelė apibendrina mašininio mokymosi modelių vertinimo rezultatus. Naudojant atsitiktinius miškus, neuroninius tinklus ir LightGBM buvo pasiekti tobuli rezultatai. **25 pav.** pateiktos defekto identifikavimo modelių ROC kreivės. Visų modelių plotas po kreive yra 1. Pagrindinis skirtumas tarp šių modelių yra mokymo laikas: atsitiktiniai miškai buvo išmokyti per 227 sekundes, neuroniniai tinklai per 985 sekundes, o LightGBM per 210 sekundes.

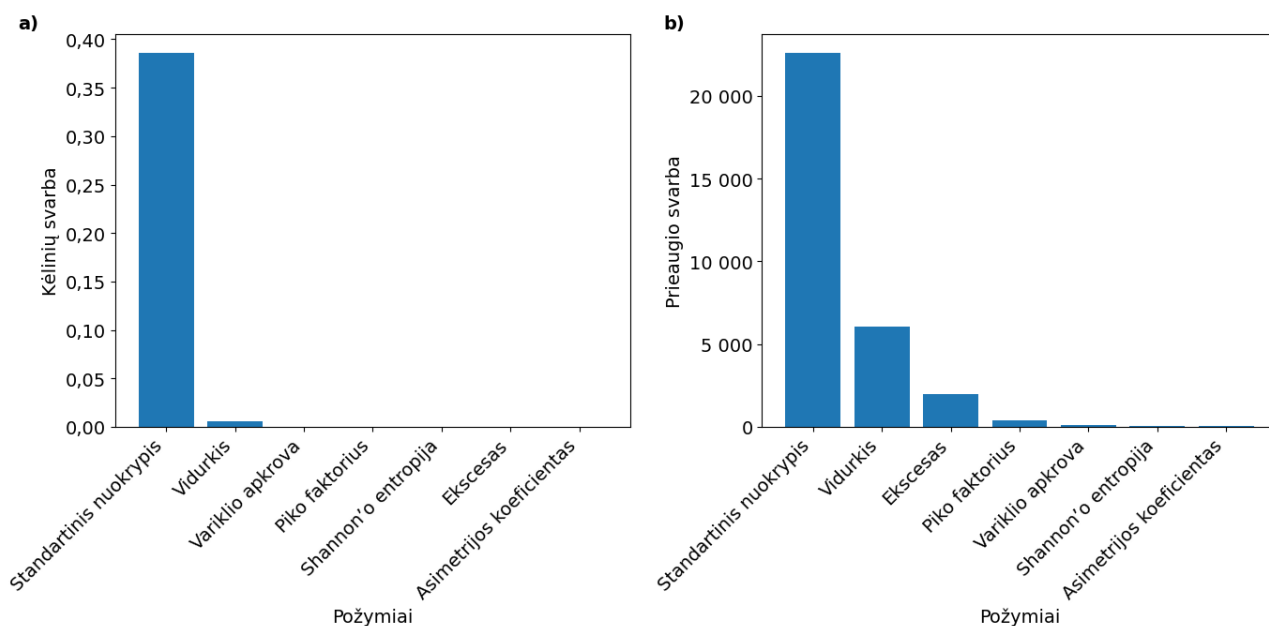
8 lentelė. Guolių defekto identifikavimo modelių tikslumo metrikos

Modelis	Tikslumas, %	Preciziškumas	Jautrumas	Specifiškumas	F ₁ įvertis	AUC	Mokymo laikas, s
Atsitiktiniai miškai	100	1	1	1	1	1	227
Neuroniniai tinklai	100	1	1	1	1	1	985
LightGBM	100	1	1	1	1	1	210



25 pav. Defekto identifikavimo modelių ROC kreivės

Kintamųjų svarbai apskaičiuoti atsitiktiniams miškams hiperparametro $n_estimators$ reikšmė buvo padidinta iki 1000. Lyginant atsitiktinių miškų ir LightGBM modelius, matoma, kad abiem atvejais svarbiausios savybės yra standartinis nuokrypis ir vidurkis (žr. **26 pav.**). Tokie požymiai kaip asimetrijos koeficientas, variklio apkrova ir Shannon'o entropija abiejų modelių laikomi nesvarbiais.



26 pav. Defekto identifikavimo modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM

Kitų autorių atliktų tyrimų rezultatai identifikuojant gedimą buvo truputį prastesni (žr. **1 lentelę**). Rojas'as ir Nandi [26] taikant atraminių vektorių klasifikatoriaus metodą pasiekė virš 98 % modelio tikslumą. O Sanchez'as ir kt. [30] su k artimiausių kaimynų ir atsitiktinių miškų modeliais pasiekė 97,5 % tikslumą.

Defekto identifikavimą praktikoje galima atlikti naudojant logistinę regresiją, kuri reikalautų mažiau resursų. Nepaisant to, dabartiniai modeliai (neuroniniai tinklai, LightGBM, atsitiktiniai miškai) yra sąlyginai greitai apmokomi dėl pažangių skaičiavimo technologijų. Be to, naudojant sudėtingesnius modelius dažniausiai gaunami geresni rezultatai, nes jie gali geriau įvertinti sudėtingesnius duomenų struktūros ir sąveikų aspektus.

3.3. Guolių defektų vietos nustatymas

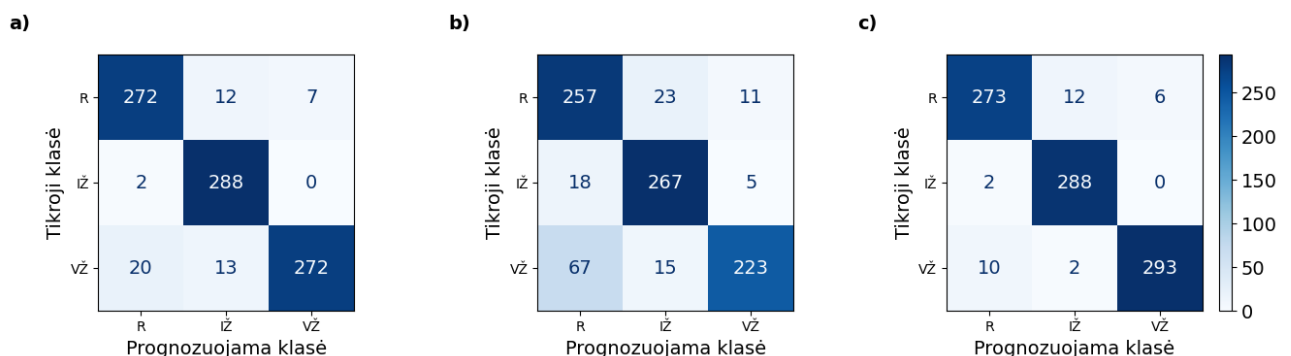
Duomenų rinkinys buvo padalintas į mokymo ir testavimo imtis. 70 % (2066) duomenų buvo skirta mokymui ir 30 % (886) – duomenų testavimui. Hiperparametrų derinimas buvo atliktas siekiant maksimizuoti tikslumą. Toliau pateikta **9 lentelė** su geriausiomis hiperparametrų derinimo metu gautomis reikšmėmis.

9 lentelė. Defekto vietos nustatymo modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)

Modelis	Hiperparametras	Reikšmė
Atsitiktiniai miškai	<i>max_depth</i>	9
	<i>min_samples_split</i>	2
	<i>min_samples_leaf</i>	5
Neuroniniai tinklai	<i>neurons_1</i>	256
	<i>neurons_2</i>	32
	<i>activation_1</i>	<i>relu</i>
	<i>activation_2</i>	<i>relu</i>

Modelis	Hiperparametras	Reikšmė
LightGBM	<i>optimizer</i>	<i>adam</i>
	<i>epochs</i>	100
	<i>batch_size</i>	32
	<i>num_leaves</i>	23
	<i>learning_rate</i>	0,0703
	<i>n_estimators</i>	183
	<i>max_depth</i>	6
	<i>min_child_weight</i>	1
	<i>subsample</i>	0,8668
	<i>colsample_bytree</i>	0,6597

Guolių defektų vietos nustatymas buvo atliktas naudojant duomenų rinkinį, kuriame nebuvo klasių disbalanso. Pateikti sumaišymo matricų rezultatai rodo klasifikavimo modelių našumą trijų skirtingų klasifikatorių: atsitiktinių miškų, neuroninių tinklų ir LightGBM (žr. **27 pav.**). Visi trys modeliai pasižymi geru rezultatu. Atsitiktinių miškų ir LightGBM modelių rezultatai buvo panašūs, tačiau neuroninių tinklų atpažinimas defektų ant rutuliuko (R) ir ant išorinio žiedo (IŽ) buvo prastesnis.



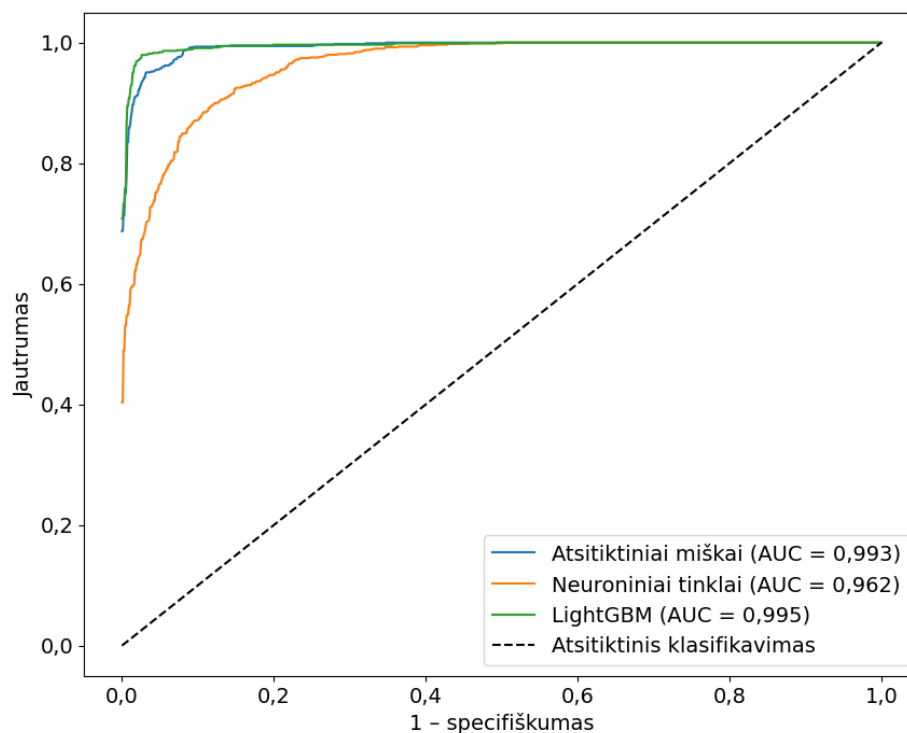
27 pav. Defekto vietos nustatymo modelių sumaišymo matricos: a) atsitiktinių miškų, b) neuroniniai tinklai, c) LightGBM

Toliau pateikta **10 lentelė**, apibendrinanti mašininio mokymosi modelių vertinimo rezultatus. Iš šių duomenų matyti, kad visi trys modeliai labai tiksliai nustatė defekto vietą. Tačiau, palyginus juos tarpusavyje, galima pastebėti, kad LightGBM buvo efektyviausias pagal visas metrikas. LightGBM modelis pasižymėjo aukščiausiu tikslumu (96,84 %), makrovidutiniu specifiškumu, makrovidutine OvR AUC reikšme (žr. **28 pav.**) ir kitomis svarbiomis vertinimo metrikomis. OvR ROC kreivės kiekvienai klasei yra pateiktos **1 priede**. Be to, LightGBM modelio mokymo trukmė buvo trumpiausia, kas yra taip pat svarbu praktiniam naudojimui.

10 lentelė. Guolių defekto vietos nustatymo modelių tikslumo metrikos

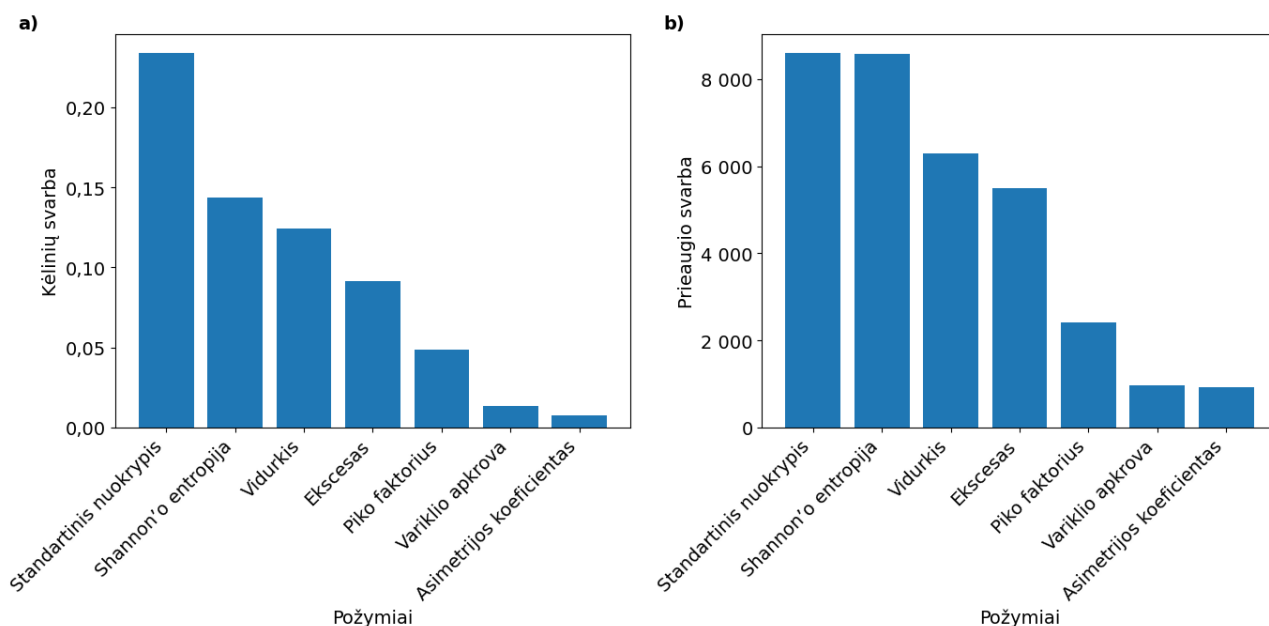
Modelis	Defekto vietos ir vidurkio tipas	Tikslumas	Preciziškumas	Jautrumas	Specifiškumas	F1 įvertis	OvR AUC	Mokymo laikas, s
Atsitiktiniai	R	93,91 %	0,9252	0,9347	0,9679	0,9299	0,9894	197

Modelis	Defekto vietos ir vidurkio tipas	Tikslumas	Preciziškumas	Jautrumas	Specifiškumas	F ₁ įvertis	OvR AUC	Mokymo laikas, s
miškai	IŽ		0,9201	0,9931	0,9965	0,9552	0,9957	
	VŽ		0,9749	0,8918	0,9456	0,9315	0,9929	
	Makro		0,9401	0,9342	0,9700	0,9389	0,9927	
Neuroniniai tinklai	R	84,31 %	0,7515	0,8832	0,9375	0,8120	0,9436	815
	IŽ		0,8754	0,9207	0,9604	0,8975	0,9839	
	VŽ		0,9331	0,7311	0,8733	0,8199	0,9568	
	Makro		0,8545	0,8431	0,9237	0,8427	0,9615	
LightGBM	R	96,39 %	0,9579	0,9381	0,9700	0,9479	0,9915	185
	IŽ		0,9536	0,9931	0,9966	0,9730	0,9956	
	VŽ		0,9799	0,9607	0,9796	0,9702	0,9979	
	Makro		0,9638	0,9640	0,9821	0,9637	0,9951	



28 pav. Defekto vietos nustatymo modelių makrovidutinės OvR ROC kreivės

Požymių svarba buvo analizuojama dviem skirtingiems metodams: atsitiktiniams miškams ir LightGBM algoritmui (žr. **29 pav.**) Kintamųjų svarbai apskaičiuoti atsitiktiniams miškams hiperparametro $n_estimators$ reikšmė buvo padidinta iki 1000. Atsitiktinių miškų ir LightGBM algoritmui svarbiausi požymiai buvo standartinis nuokrypis ir Shannon'o entropija. Abiejų algoritmų buvo nustatyta, kad mažiausiai svarbūs požymiai – variklio apkrova, asimetrijos koeficientas.



29 pav. Defekto vietos nustatymo modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM

Kitų autorių rezultatai nustatant gedimo vietą buvo geresni (žr. **1 lentelę**), dažnai siekiant ar net viršijant 97 % tikslumą. Zarei ir kt. [24] su dirbtiniais neuroniniais tinklais pasiekė 100 % tikslumą, o Saidi ir kt. [25] su atraminių vektorių modeliu – 99,99 %, abu naudojant specializuotus duomenų rinkinius.

3.4. Guolių defektų dydžio nustatymas

Šiame poskyryje aprašyti regresijos ir klasifikavimo modelių guolių defekto dydžio nustatymo rezultatai. Praktikoje dažniausiai būtų taikomas regresijos būdas, nes defekto dydis gali būti įvairus ir tikslus defekto dydžio įvertinimas yra svarbus siekiant tinkamai prižiūrėti įrenginį. Tačiau šiuo atveju klasifikavimo būdas taip pat yra tinkamas, nes defektų dydžiai buvo suskirstyti į tris kategorijas: 0,007, 0,014 ir 0,021.

3.4.1. Guolių defektų dydžio nustatymas regresijos būdu

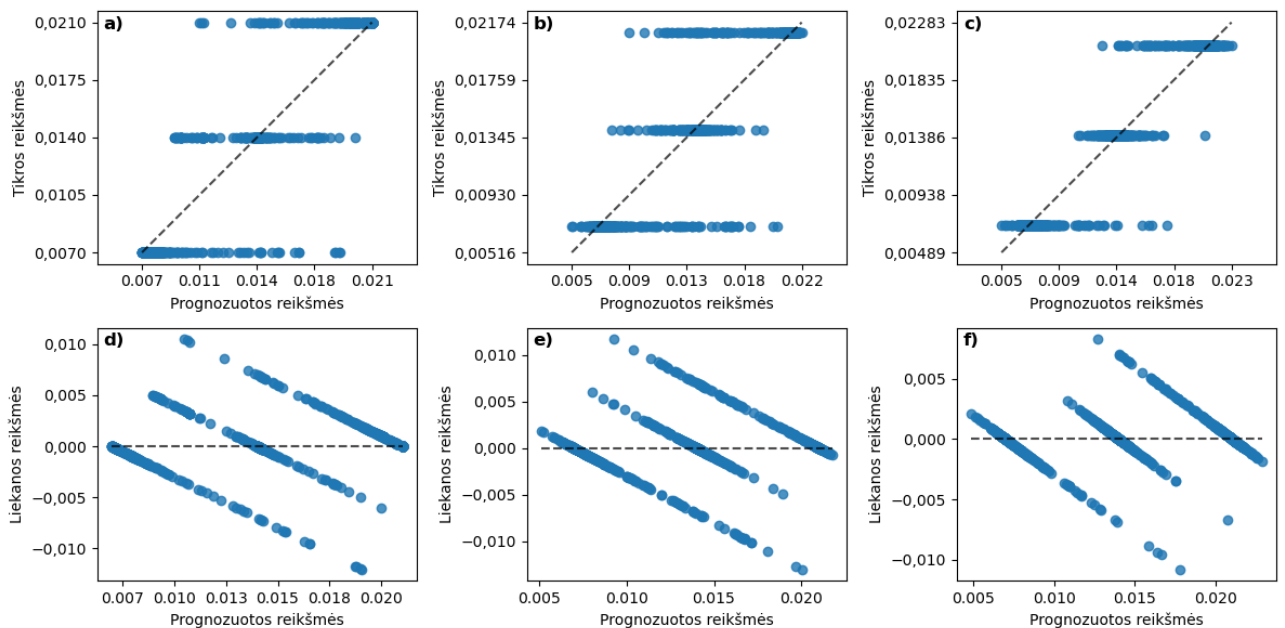
Duomenų rinkinys buvo padalintas į mokymo ir testavimo imtis. 70 % (2066) duomenų buvo skirta mokymui ir 30 % (886) – duomenų testavimui. Hiperparametrų derinimas buvo atliktas siekiant minimizuoti vidutinę absoliutinę paklaidą. Toliau pateikta **11 lentelė** su geriausiomis hiperparametrų derinimo metu gautomis reikšmėmis.

11 lentelė. Defekto vietos nustatymo regresijos modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)

Modelis	Hiperparametras	Reikšmė
Atsitiktiniai miškai	<i>max_depth</i>	9
	<i>min_samples_split</i>	1
	<i>min_samples_leaf</i>	2
Neuroniniai tinklai	<i>neurons_1</i>	256
	<i>neurons_2</i>	128
	<i>activation_1</i>	<i>relu</i>

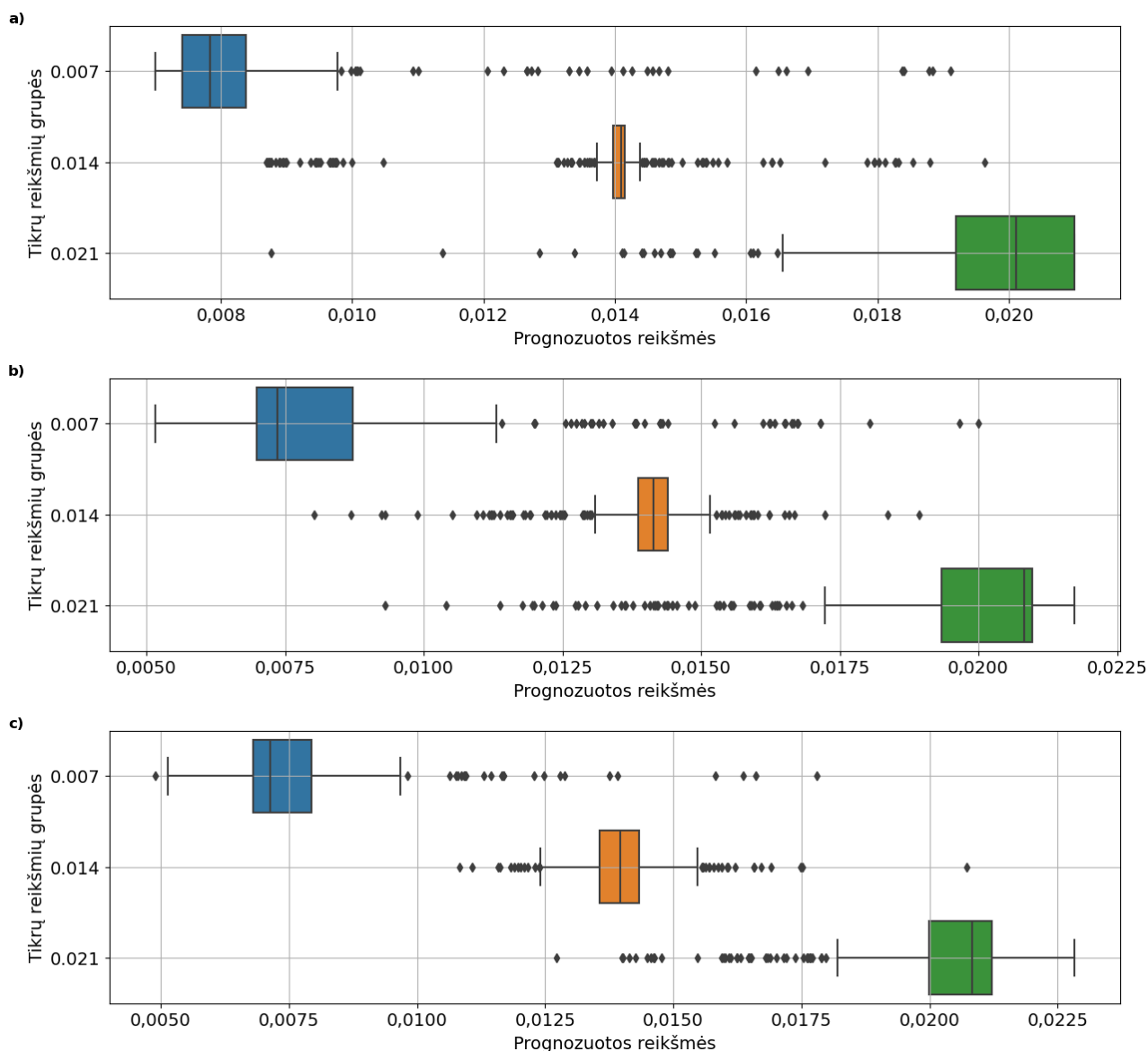
Modelis	Hiperparametras	Reikšmė
	<i>activation_2</i>	<i>relu</i>
	<i>optimizer</i>	<i>adam</i>
	<i>epochs</i>	100
	<i>batch_size</i>	16
LightGBM	<i>num_leaves</i>	56
	<i>learning_rate</i>	0,0689
	<i>n_estimators</i>	256
	<i>max_depth</i>	9
	<i>min_child_weight</i>	7
	<i>subsample</i>	0,8481
	<i>colsample_bytree</i>	0,9167

Guolių defektų dydžio nustatymas buvo atliktas naudojant duomenų rinkinį, kuriame nebuvo klasių disbalanso. Vizualiai atvaizduotas atsitiktinių miškų, neuroninių tinklų ir LightGBM regresijos modelių paklaidas galima pamatyti **30 pav.** Modelių tikrų ir prognozuotų reikšmių santykis atvaizduotas pirmoje eilutėje, liekanų ir prognozuotų reikšmių – antroje eilutėje. Galima pastebėti, kad LightGBM modelio liekanos yra mažiausios.



30 pav. Tikrų ir prognozuotų reikšmių (a–c grafikai) ir prognozės liekanų reikšmių (d–f grafikai) diagramos:
a) ir d) atsitiktiniai miškai, b) ir e) neuroniniai tinklai, c) ir f) LightGBM

Regresijos modelių prognozių stačiakampės diagramos pateiktos **31 pav.** Pirmos ir trečios grupės (0,007 ir 0,021) reikšmės yra gerai atspėtos su nedideliais nuokrypiais. Antroje grupėje (0,014) yra pastebėtas didesnis prognozių išsibastymas apie imties plotį, ypač a) ir b) grafikuose, kur yra daugiau išskirčių. Vidutinės prognozės visų modelių yra arti tikrųjų verčių, todėl visi modeliai yra gana tikslūs. Taip pat galima pastebėti, kad LightGBM modelio prognozės turi mažiausiai išskirčių.



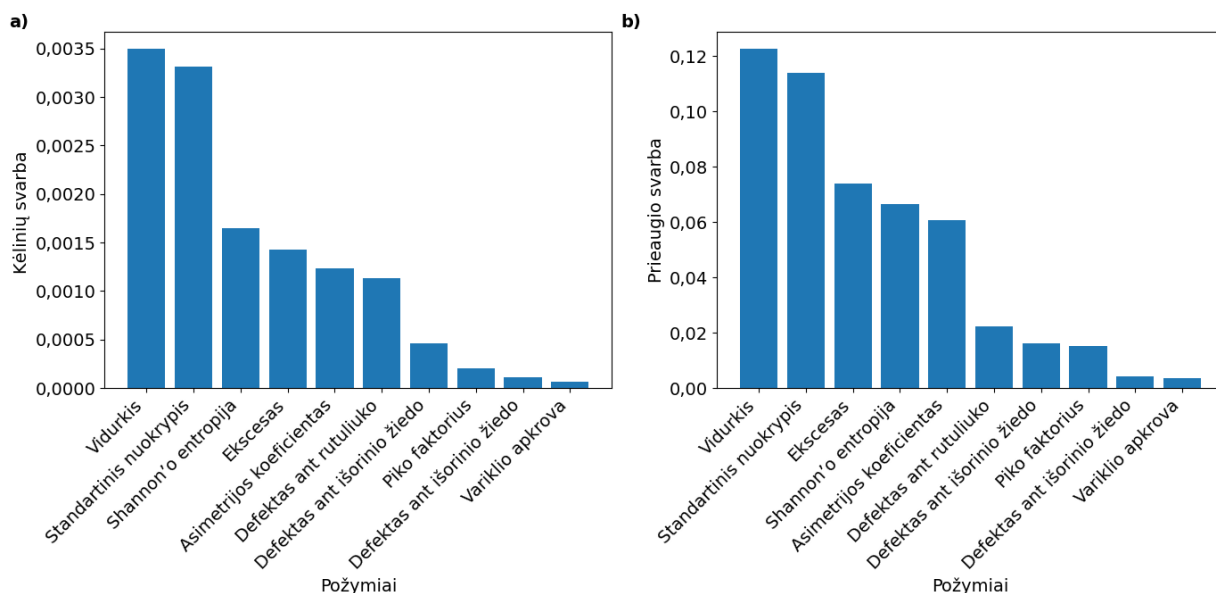
31 pav. Modelių prognozuotų reikšmių stačiakampės diagramos: a) atsitiktiniai miškai, b) neuroniniai tinklai, c) LightGBM

Toliau pateikta **12 lentelė**, apibendrinanti mašininio mokymosi modelių vertinimo rezultatus. Atsitiktinių miškų modelis buvo išmokintas per 239 sekundes, o MAE rodiklis buvo 0,0011, MSE – 4,9537, MAPE – 0,1074 ir R^2 – 0,8484. Neuroninių tinklų modeliui, kuris buvo išmokintas per ilgiausią laiką – 812 sekundžių, užfiksuoti šie rodikliai: MAE – 0,0013, MSE – 6,8757, MAPE – 0,1231 ir R^2 – 0,7896. Tuo tarpu LightGBM modelis, išmokintas per 196 sekundes, parodė geriausius rezultatus: MAE – 0,0009, MSE – 2,7673, MAPE – 0,0806 ir R^2 – 0,9153.

12 lentelė. Guolių defekto dydžio nustatymo regresijos modelių paklaidos

Modelis	MAE	MSE	MAPE	R^2	Mokymo laikas, s
Atsitiktiniai miškai	0,0011	4,9537	0,1074	0,8484	239
Neuroniniai tinklai	0,0013	6,8757	0,1231	0,7896	812
LightGBM	0,0009	2,7673	0,0806	0,9153	196

Analizuojant du skirtingus modelius – atsitiktinius miškus ir LightGBM – pastebima, kad kintamųjų svarba abiem modeliams išdėstyta panašiai (žr. **32 pav.**). Kintamųjų svarbai apskaičiuoti atsitiktiniams miškams hiperparametro *n_estimators* reikšmė buvo padidinta iki 1000. Atsitiktinių miškų modeliui didžiausią svarbą turintys kintamieji yra vidurkis, standartinis nuokrypis. LightGBM modeliui šie kintamieji taip pat yra svarbiausi. Kintamųjų, tokių kaip ekscesas ir Shannon'o entropija, pozicijos šiek tiek skiriasi, tačiau abiem atvejais jie laikomi svarbiais. Mažiausiai svarbių kintamųjų grupę sudaro požymiai defektas ant išorinio žiedo ir variklio apkrova.



32 pav. Defekto dydžio nustatymo regresijos modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM

3.4.2. Alternatyvus guolių defektų dydžio nustatymas klasifikavimo būdu

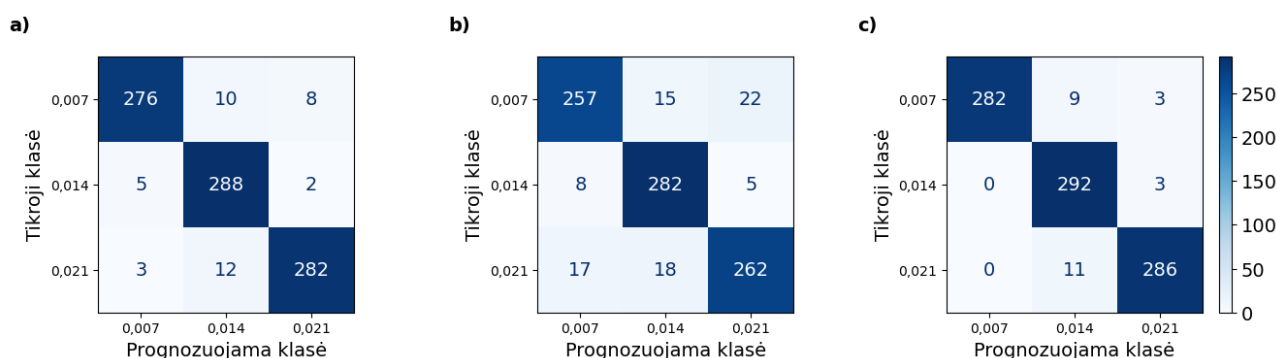
Duomenų rinkinys buvo padalintas į mokymo ir testavimo imtis. 70 % (2066) duomenų buvo skirta mokymui ir 30 % (886) – duomenų testavimui. Hiperparametrų derinimas buvo atliktas siekiant maksimizuoti tikslumą. Toliau pateikta **13 lentelė** su geriausiomis hiperparametrų derinimo metu gautomis reikšmėmis.

13 lentelė. Defekto dydžio nustatymo klasifikavimo modelių hiperparametrų derinimo rezultatai (nustatytos optimalios vertės)

Modelis	Hiperparametras	Reikšmė
Atsitiktiniai miškai	<i>max_depth</i>	9
	<i>min_samples_split</i>	1
	<i>min_samples_leaf</i>	2
Neuroniniai tinklai	<i>neurons_1</i>	256
	<i>neurons_2</i>	32
	<i>activation_1</i>	<i>relu</i>
	<i>activation_2</i>	<i>relu</i>
	<i>optimizer</i>	<i>adam</i>
	<i>epochs</i>	100
	<i>batch_size</i>	32

Modelis	Hiperparametras	Reikšmė
LightGBM	<i>num_leaves</i>	41
	<i>learning_rate</i>	0,0852
	<i>n_estimators</i>	206
	<i>max_depth</i>	7
	<i>min_child_weight</i>	1
	<i>subsample</i>	0,7129
	<i>colsample_bytree</i>	0,7246

Guolių defektų dydžio nustatymas buvo atliktas naudojant duomenų rinkinį, kuriame nebuvo klasių disbalanso. Pateikti sumaišymo matricų rezultatai rodo klasifikavimo modelių našumą trijų skirtingų klasifikatorių: atsitiktiniai miškų, neuroninių tinklų ir LightGBM (žr. 27 pav.). Prasčiausiai klasifikavimas buvo atliktas neuroninių tinklų modelio su daugiausiai klaidingų prognozių. Aukščiausias tikslumas buvo pademonstruotas LightGBM modelio.



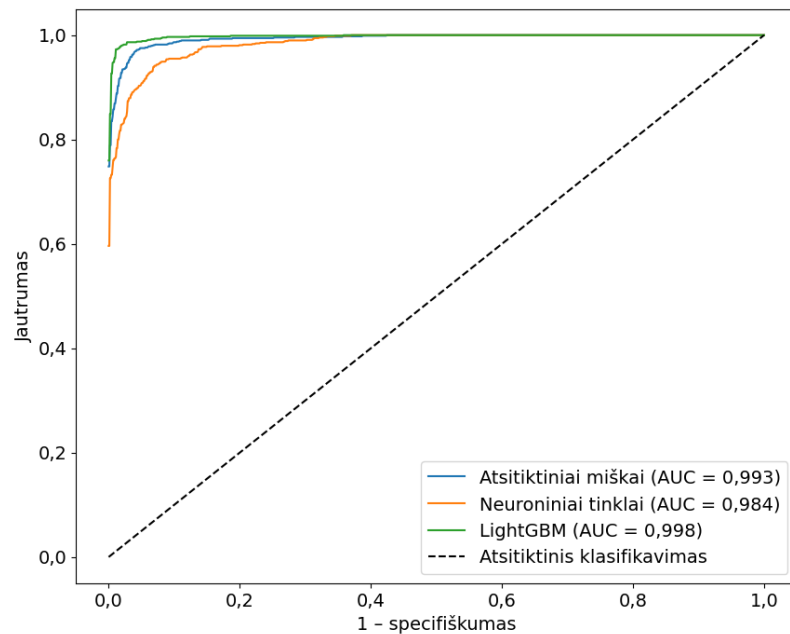
33 pav. Defekto dydžio nustatymo klasifikavimo modelių sumaišymo matricos: a) atsitiktinių miškų, b) neuroniniai tinklai, c) LightGBM

Toliau pateikta **14 lentelė**, apibendrinanti mašininio mokymosi modelių vertinimo rezultatus. Iš šių duomenų matyti, kad visi trys modeliai labai tiksliai nustatė defekto dydį. Galima pastebėti, kad LightGBM buvo efektyviausias pagal visas metrikas. LightGBM modelis pasižymėjo aukščiausiu tikslumu (97,07 %), makrovidutiniu specifiškumu, makrovidutine OvR AUC reikšme (žr. 34 pav.) ir kitomis svarbiomis tikslumo metrikomis. OvR ROC kreivės kiekvienai klasei yra pateiktos **2 priede**. Be to, LightGBM modelio mokymo trukmė buvo užfiksuota trumpiausia (210 sekundžių).

14 lentelė. Guolių defekto dydžio nustatymo modelių tikslumo metrikos

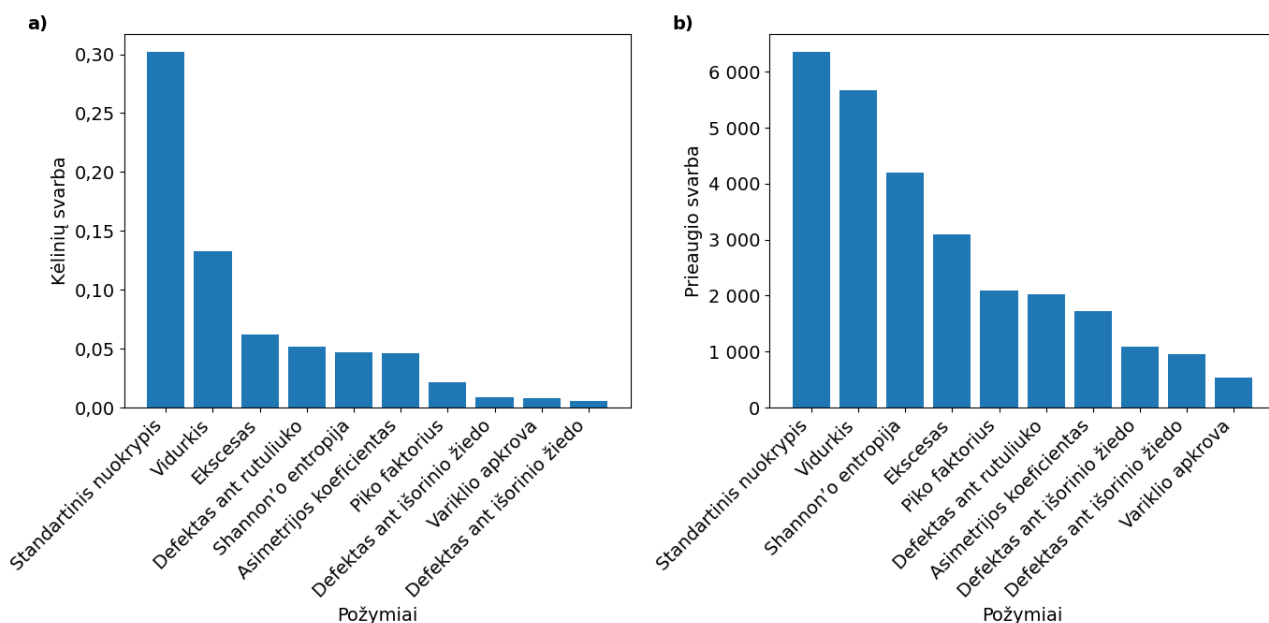
Modelis	Defekto vietos ir vidurkio tipas	Tikslumas	Preciziškumas	Jautrumas	Specifiškumas	F ₁ įvertis	OvR AUC	Mokymo laikas, s
Atsitiktiniai miškai	0,007	95,49 %	0,9718	0,9388	0,9701	0,9550	0,9924	222
	0,014		0,9290	0,9763	0,9878	0,9521	0,9937	
	0,021		0,9658	0,9495	0,9747	0,9576	0,9938	
	Makro		0,9555	0,9548	0,9776	0,9549	0,9934	
Neuroniniai	0,007	90,41 %	0,9113	0,8741	0,9387	0,8924	0,9795	901

Modelis	Defekto vietos ir vidurkio tipas	Tikslumas	Preciziškumas	Jautrumas	Specifiškumas	F ₁ įvertis	OvR AUC	Mokymo laikas, s
tinklai	0,014		0,8952	0,9559	0,9772	0,9246	0,9912	
	0,021		0,9066	0,8822	0,9414	0,8942	0,9810	
	Makro		0,9044	0,9041	0,9524	0,9037	0,9839	
LightGBM	0,007	97,07 %	1,0000	0,9592	0,9801	0,9792	0,9984	210
	0,014		0,9359	0,9898	0,9948	0,9621	0,9973	
	0,021		0,9795	0,9630	0,9815	0,9711	0,9968	
	Makro		0,9718	0,9707	0,9855	0,9708	0,9975	



34 pav. Defekto dydžio nustatymo modelių makrovidutinės OvR ROC kreivės

Požymių svarba buvo analizuojama dviem skirtingiems metodams: atsitiktiniams miškams ir LightGBM algoritmui (žr. 35 pav.). Kintamųjų svarbai apskaičiuoti atsitiktiniams miškams hiperparametro $n_estimators$ reikšmė buvo padidinta iki 1000. Abiejų modelių atveju, svarbiausi požymiai yra standartinis nuokrypis ir vidurkis, kurie turi didžiausią įtaką modelių tikslumui. Shannon'o entropija ir ekscesas taip pat yra reikšmingi požymiai, nors jų svarba skiriasi tarp modelių. Mažiausiai svarbūs abiem modeliams yra defektas ant išorinio žiedo ir variklio apkrova, kurie turi mažesnę įtaką modelio tikslumui.



35 pav. Defekto dydžio nustatymo klasifikavimo modelių požymių svarba: a) atsitiktiniai miškai, b) LightGBM

1 lentelėje aptarti autoriai, atlikę defekto dydžio nustatymo tyrimus, problemą sprendę klasifikavimo būdu. Han'o ir Jiang'o [29] bei Kumbhar'io ir kt. [22] sukurti modeliai pasiekė tikslumus didesnius nei 97 %. Panašų tikslumą 97,07 % pavyko pasiekti ir šiame tyrime taikant LightGBM modelį.

3.5. Sąnaudų ir naudos analizė

Guolių nuspėjamos priežiūros sistemos diegimui vienam traukinio vagonui, turinčiam 4 aširačius, numatytos šios investicijos: keturi akselerometrai *PCB Piezotronics 352C68* (kaina – 1630,48 Eur), vienas mikrokompiuteris *Raspberry Pi SC0194(9)* (kaina – 51,19 Eur), viena SIM kortelė iš *Telia* (kaina – 3 Eur), 50 metrų *Pro Base* laidų (kaina – 34,5 Eur), įrengimo darbai, kurių sąnaudos apskaičiuotos 8 valandoms, esant 17 Eur už valandą tarifui (kaina – 136 Eur) bei viena *SanDisk Extreme Pro UHS-I 256 GB* atminties kortelė (kaina – 38,53 Eur). Iš viso investicijų 1 893,7 Eur. Žemiau pateikta **15 lentelė**, apibendrinanti investicijas, reikalingas guolių nuspėjamos techninės priežiūros sistemai.

15 lentelė. Guolių nuspėjamos techninės priežiūros sistemai reikalingos investicijos

Įranga / medžiagos / darbo sąnaudos	Tiekėjas / modelis	Kiekis	Kaina, Eur	Suma, Eur
Akselerometras (vnt.)	<i>PCB Piezotronics 352C68</i>	4	407,62	1 630,48
Mikrokompiuteris (vnt.)	<i>Raspberry Pi SC0194(9)</i>	1	51,19	51,19
SIM kortelė (vnt.)	<i>Telia</i>	1	3,00	3,00
Laidai (m)	<i>PRO BASE</i>	50	0,69	34,50
Įrengimo darbai (val.)	-	8	17,00	136,00
Atminties kortelė	<i>SanDisk Extreme Pro UHS-I (256 GB)</i>	1	38,53	38,53
INVESTICIJOS IŠ VISO, Eur:				1 893,7

Toliau pateikta **16 lentelė** apibendrinanti einamąsias išlaidas, reikalingas guolių nuspėjamos techninės priežiūros sistemai. Einamosios išlaidos, susijusios su guolių vibracijos stebėjimo sistema, yra apskaičiuotos metų laikotarpiui. Ryšio paslaugos, teikiamos *Telia*, kainuotų 358,8 Eur vieniems metams. Debesų duomenų saugyklos paslaugos sudarytų 648 Eur per tą patį laikotarpį. Darbo jėga, atstovaujama 1/10 inžinieriaus etato, būtų apie 300 Eur. Bendros metinės einamosios išlaidos, apimančios būtinas paslaugas ir darbo jėgą, reikalingą sistemos palaikymui ir jos veikimui, siektų 1 306,8 Eur.

16 lentelė. Guolių nuspėjamos techninės priežiūros sistemos einamosios išlaidos

Prenumeratos / darbo sąnaudos	Tiekėjas / modelis	Kiekis	Kaina, Eur	Suma, Eur
Ryšio išlaidos (mėn.)	<i>Telia</i>	12	29,90	358,8
Duomenų saugykla (mėn.)	AWS S3	12	54,00	648
1/10 inžinierius etato (mėn.)	-	12	25,00	300
EINAMOSIOS IŠLAIDOS IŠ VISO, Eur:				1 306,8

Nauda, susijusi su guolių vibracijos sistemos diegimu, yra galimybė išvengti neplanuotų remontų, leidžiančių sutaupyti lėšas, kurios būtų išleistos darbo užmokesčiui. Toliau pateikta **17 lentelė**, apibendrinanti sutaupytas išlaidas, suskaičiuotas pagal darbo valandas, kurios būtų reikalingos įvairiems specialistams. Paminos meistrui už 8 valandas būtų mokama 136 Eur, o šešiams technikams – po 112 Eur. Kokybės ekspertui būtų mokama 120 Eur už 8 valandas. Bendra nauda, išvengiant vieno neplaninio gedimo, sudarytų 928 Eur. Ši suma atspindi darbo užmokesčio išlaidas, kurios būtų patirtos, jei reiktų skubiai spręsti gedimo pasekmes.

17 lentelė. Guolių nuspėjamos techninės priežiūros sistemos naudos

Nauda	Kiekis	Kaina, Eur	Suma, Eur
Paminos meistras (val.)	8	17,00	136
Technikas (val.)	6 × 8	14,00	672
Kokybės ekspertas (val.)	8	15,00	120
Nauda neplaninio gedimo atveju, Eur:			928

18 lentelė su apskaičiuotomis tiek paprastosiomis, tiek diskontuotomis išlaidomis ir naudomis pateikta apačioje. Buvo pasirinkta 8 % diskonto norma. Taip pat laikysime, kad sistema padėtų išvengti dviejų neplaninių remontų. Nuo pirmųjų iki penktųjų metų kiekvienais metais fiksuojamos vienodos metinės sąnaudos, kurios siekia 1 306,8 eurų, tačiau jų diskontuotos vertės mažėja kiekvienais metais dėl diskontavimo. Diskontuotos suminės sąnaudos auga kasmet, prasidedant nuo 1 893,70 eurų ir pasiekiant 7 111,37 eurų penktaisiais metais. Naudos, skaičiuojamos nuo pirmųjų metų, taip pat nurodomos vienodos (1 856 eurų kasmet), bet jų diskontuotos vertės mažėja, pradedant nuo 1 718,52 eurų ir baigiant 1 263,16 eurų. Diskontuota suminė nauda taip pat auga, nuo 1 718,52 eurų iki 7 410,47 eurų.

18 lentelė. Guolių nuspėjamos techninės kasmetinės sąnaudos ir nauda

Metai	Sąnaudos, Eur	Diskontuotos sąnaudos, Eur	Diskontuotos suminės sąnaudos, Eur	Nauda, Eur	Diskontuota nauda, Eur	Diskontuota suminė nauda, Eur
0	1 893,70	1 893,7	1 893,70			

Metai	Sąnaudos, Eur	Diskontuotos sąnaudos, Eur	Diskontuotos suminės sąnaudos, Eur	Nauda, Eur	Diskontuota nauda, Eur	Diskontuota suminė nauda, Eur
1	1 306,80	1 210,00	3 103,70	1 856,00	1 718,52	1 718,52
2	1 306,80	1 120,37	4 224,07	1 856,00	1 591,22	3 309,74
3	1 306,80	1 037,38	5 261,45	1 856,00	1 473,35	4 783,09
4	1 306,80	960,54	6 221,99	1 856,00	1 364,22	6 147,31
5	1 306,80	889,39	7 111,37	1 856,00	1 263,16	7 410,47

Atsipirkimo laikas, apskaičiuotas remiantis pateiktais duomenimis, yra **4,95 metų**. Tai reiškia, kad investicija bus pilnai atsipirkusi per maždaug penkerius metus, atsižvelgiant į sąnaudas ir gautą naudą, diskontuotus į dabartinę vertę. O naudos ir sąnaudų indeksas, skaičiuotas pagal pateiktus duomenis, yra **1,05**. Projektas yra laikomas finansiškai prasmingu.

3.6. Rekomendacijos

Remiantis sukurtu guolių nuspėjamos priežiūros matematiniu modeliu, galima pateikti keletą rekomendacijų, kaip tobulinti sukurtą metodiką ir toliau plėtoti tokios sistemos diegimą.

Pirma, siūloma plėsti duomenų rinkinį įtraukiant daugiau duomenų iš skirtingų eksploatacinių sąlygų ir kitokio tipo guolių defektų. Šiuo metu duomenų rinkinys yra gana ribotas ir fokusuotas tik į tam tikras defektų rūšis ir dydžius. Duomenų įvairovės didinimas padėtų modeliams geriau apibendrinti informaciją ir padidintų jų pritaikomumą realiame pasaulyje, esant įvairioms eksploatacinėms sąlygoms.

Antra, aukštiesiems rezultatams pasiekti nebūtinai reikalingi sudėtingi modeliai. Rekomenduojama išbandyti paprasčiausią logistinę regresiją ar tiesinę diskriminantų analizę, kurių nereikėtų derinti. Šie paprastesni modeliai gali suteikti patikimus rezultatus su mažesnėmis skaičiavimo sąnaudomis ir paprastesniu diegimu bei palaikymu.

Trečia, įdiegimo ir eksploataavimo sąnaudų optimizavimui, būtų tikslinga analizuoti alternatyvius jutiklių ir aparatūros sprendimus. Nors dabartiniai komponentai yra veiksmingi, galbūt egzistuoja ekonomiškesni arba efektyvesni sprendimai, kurie galėtų sumažinti bendras sistemos diegimo ir palaikymo išlaidas. Be to, verta apsvarstyti galimybę naudoti pažangesnius mikrokompiuterius ar debesų kompiuterijos paslaugas, kurios galėtų pagerinti sistemos duomenų apdorojimo greitį ir efektyvumą.

Ketvirtas aspektas, kurį verta tobulinti, yra nuolatinė modelių efektyvumo stebėsena ir atnaujinimas. Kadangi vibracijos signalų duomenys gali keistis laikui bėgant, svarbu reguliariai atnaujinti modelius ir jų hiperparametrus, siekiant išlaikyti aukštą identifikavimo tikslumą. Tai gali būti pasiekta įdiegus automatinio modelių atnaujinimo ir mokymo sistemą, kuri nuolat analizuos naujus duomenis ir atnaujins modelius pagal naujausias tendencijas.

Penktas aspektas, kuris būtų svarbus diegiant guolių vibracijos stebėjimo sistemą, yra jautrumo tyrimas, skirtas sąnaudų ir naudos analizei. Jautrumo analizė padėtų įvertinti, kaip skirtingi įvesties parametrai veikia galutinį projektų ekonominį efektyvumą ir finansinį atsipirkimą. Toks vertinimas yra gyvybiškai svarbus suprantant rizikas ir nustatant, kurie projektų aspektai yra labiausiai linkę keistis dėl įvairių neapibrėžtumų.

Galiausiai, diegiant tokią sistemą praktiškai, būtina užtikrinti, kad inžinieriai ir technikai būtų tinkamai apmokyti naudotis šia technologija. Tai padės efektyviai interpretuoti gaunamus rezultatus ir laiku imtis reikalingų veiksmų, siekiant užkirsti kelią galimiems guolių gedimams ir užtikrinti nenutrūkstamą įrangos veikimą. Be to, svarbu užtikrinti sistemos integraciją su esamomis priežiūros ir stebėjimo platformomis, siekiant maksimaliai padidinti naudingumą ir efektyvumą realiame pramoniniame kontekste.

Išvados

1. Atlikus mokslinės literatūros analizę, buvo nustatyta, kad:
 - techninės priežiūros strategijos, ypač nuspėjamoji, gali lemti didesnę įrenginių patikimumą ir operacinį efektyvumą, sumažinant nenumatytas prastovas ir gedimų skaičių;
 - į nuspėjamąją priežiūrą integruojant įvairius duomenų tipus, tokius kaip jutiklių, vaizdų, garso ir tekstinius duomenis, suteikiamas išsamus įrangos veiklos ir būklės vaizdas;
 - guoliai gali sugesti dėl įvairių priežasčių, įskaitant netinkamą tepimą, teršalus, per didelę apkrovą ar netinkamą montavimą, darančių įtaką guolio patikimumui ir ilgaamžiškumui;
 - dirbtinio intelekto taikymas, naudojant dirbtinius neuroninius tinklus, atraminių vektorių klasifikatorius ir atsitiktinius miškus, leidžia tiksliai nustatyti įvairius gedimų tipus.
2. Apžvelgus ekonominius aspektus, lemiančius nuspėjamosios priežiūros sistemų diegimą ir taikymą, buvo išsiaiškinta, kad:
 - nuspėjama techninė priežiūra yra svarbi investicija, kurios finansinė grąža turėtų viršyti pradines išlaidas;
 - sprendimai dėl nuspėjamos priežiūros dažnai susiję su pasirinkimu tarp kelių alternatyvų ir būtinybe vertinti pinigų kainą laike;
 - teigiamos ekonominės naudos, lyginant su korekcinės priežiūros sąnaudomis ir prastovomis, matomos šiuolaikinių nuspėjamos techninės priežiūros įgyvendinimo pavyzdžiuose;
 - atitinkamai parinkti investicijų vertinimo metodai gali padėti optimizuoti resursų paskirstymą ir maksimizuoti grąžą.
3. Buvo parinkti trys mašininio mokymo algoritmai: atsitiktiniai miškai, LightGBM, neuroniniai tinklai. Atsitiktinių miškų algoritmas yra puikus pasirinkimas tiek klasifikacijai, tiek regresijai dėl savo gebėjimo adaptuotis prie įvairių duomenų tipų. LightGBM algoritmas gali efektyviai spręsti didelių duomenų rinkinių uždavinius, užtikrindamas greitį ir tikslumą. Tuo tarpu dirbtinis neuroninis tinklas yra tinkamas sudėtingoms analitikos problemoms dėl savo giluminio mokymosi ir daugiasluoksnių struktūrų, kurios leidžia efektyviai modeliuoti kompleksines sistemas.
4. Įvertinus sudarytus mašininio mokymo modelius buvo nustatyta, kad:
 - atsitiktinių miškų, neuroninių tinklų ir LightGBM modeliais buvo tobulai identifikuoti gedimai, remiantis tikslumo, preciziškumo, jautrumo, specifiškumo ir F_1 įverčiais. Iš visų modelių, greičiausias mokymo laikas buvo pasiektas naudojant LightGBM, užtrukta vos 210 sekundžių. Šiam modeliui svarbiausios savybės buvo standartinis nuokrypis ir vidurkis;
 - defektų vietos, efektyviausiai nustatytos LightGBM modelio, buvo identifikuotos su 96,39 % tikslumu bei pasiekta 0,9951 makrovidutinė (vienas prieš visus) ploto po kreive reikšmė. Šio modelio mokymo laikas buvo 185 sekundės. LightGBM modeliui svarbiausios savybės buvo standartinis nuokrypis ir Shannon'o entropija;
 - defekto dydžio matavimai buvo atlikti efektyviausiai naudojant LightGBM regresijos modelį, kurio užfiksuotos mažiausios paklaidos: vidutinė absoliutinė paklaida – 0,0009, vidutinė kvadratinė paklaida – 2,7673, vidutinė procentinė absoliutinė paklaida – 0,0806. Buvo pasiektas aukščiausias determinacijos koeficientas – 0,9153, rodantis aukštą duomenų variacijos paaiškinimo lygį. LightGBM modelio mokymo laikas, siekiantis tik 196

sekundes, buvo trumpiausias iš visų modelių. Šiam modeliui svarbiausios savybės buvo vidurkis ir standartinis nuokrypis;

- defektų dydžiai, efektyviausiai nustatyti LightGBM klasifikavimo modelio, buvo identifikuoti su 97,07 % tikslumu bei pasiekta 0,9975 makrovidutinė (vienas prieš visus) ploto po kreive reikšmė. Šio modelio mokymo laikas buvo 210 sekundžių. LightGBM modeliui svarbiausios savybės buvo standartinis nuokrypis ir Shannon'o entropija;
- visų optimizuotų modelių integravimas į vieningą struktūrą užtikrino aukštos kokybės sprendimų priėmimą.

5. Atlikus sąnaudų ir naudos analizę buvo nustatyta, kad:

- pradinis guolių nuspėjamos techninės priežiūros sistemos įdiegimo investicijų dydis yra 1 893,7 eurų, o kasmetinės einamosios išlaidos siekia 1 306,8 eurų;
- sistema potencialiai padėtų išvengti dviejų neplaninių gedimų per metus, kasmet sutaupant 1 856 eurų;
- atsipirkimo laikotarpis, apskaičiuotas remiantis diskontuotomis sąnaudomis ir nauda, yra 4,95 metų;
- sistema yra finansiškai prasminga, remiantis 1,05 naudos ir sąnaudų indeksu.

6. Siekiant tobulinti guolių nuspėjamos priežiūros sistemą, rekomenduojama išplėsti duomenų rinkinį, įtraukiant įvairesnes eksploatacines sąlygas ir skirtingus defektų tipus, taip pat įvertinti papildomus vibracijos signalų apdorojimo metodus. Svarbu analizuoti alternatyvius jutiklių ir aparatūros sprendimus, siekiant optimizuoti įdiegimo ir eksploataavimo sąnaudas, bei apsvarstyti pažangesnių mikrokompiuterių ar debesų kompiuterijos paslaugų naudojimą. Nuolatinė modelių efektyvumo stebėsena ir atnaujinimas yra būtina, siekiant išlaikyti aukštą tikslumą, o jautrumo analizė padėtų vertinti ekonominį projekto efektyvumą. Be to, būtina užtikrinti, kad inžinieriai ir technikai būtų tinkamai apmokyti naudotis šia technologija ir integruoti ją su esamomis priežiūros sistemomis, taip užtikrinant nenutrūkstamą įrangos veikimą ir efektyvią defektų prevenciją.

Literatūros sąrašas

1. KOTHAMASU, Ranganath, Samuel H. HUANG, ir William H. VERDUIN. System health monitoring and prognostics — a review of current paradigms and practices. *The International Journal of Advanced Manufacturing Technology* [interaktyvus]. 2006, **28**(9-10), 1012–1024 [žiūrėta 2024-02-04]. ISSN 1433-3015. Prieiga per: doi:10.1007/s00170-004-2131-6
2. PHAM, Hoang, ir Hongzhou WANG. Imperfect maintenance. *European Journal of Operational Research* [interaktyvus]. 1996, **94**(3), 425–438 [žiūrėta 2024-02-04]. ISSN 0377-2217. Prieiga per: doi:10.1016/s0377-2217(96)00099-9
3. SULLIVAN, Greg, et al. *Operations & Maintenance Best Practices - A Guide to Achieving Operational Efficiency (Release 3)* [interaktyvus]. Office of Scientific and Technical Information (OSTI), rugpjūtis 2010 [žiūrėta 2024-02-04]. Prieiga per: doi:10.2172/1034595
4. UCAR, Aysegul, Mehmet KARAKOSE, ir Necim KIRIMÇA. Artificial intelligence for predictive maintenance applications: key components, trustworthiness, and future trends. *Applied Sciences* [interaktyvus]. 2024, **14**(2), 898 [žiūrėta 2024-02-15]. ISSN 2076-3417. Prieiga per: doi:10.3390/app14020898
5. MEDDAOUI, Anwar, Mustapha HAIN, ir Adil HACHMOUD. The benefits of predictive maintenance in manufacturing excellence: a case study to establish reliable methods for predicting failures. *The International Journal of Advanced Manufacturing Technology* [interaktyvus]. 2023 [žiūrėta 2024-02-06]. ISSN 1433-3015. Prieiga per: doi:10.1007/s00170-023-12086-6
6. KASHPRUK, Nataliia, Cezary PISKOR-IGNATOWICZ, ir Jerzy BARANOWSKI. Time series prediction in industry 4.0: a comprehensive review and prospects for future advancements. *Applied Sciences* [interaktyvus]. 2023, **13**(22), 12374 [žiūrėta 2024-02-10]. ISSN 2076-3417. Prieiga per: doi:10.3390/app132212374
7. XIE, Jiawei, et al. Systematic literature review on data-driven models for predictive maintenance of railway track: implications in geotechnical engineering. *Geosciences* [interaktyvus]. 2020, **10**(11), 425 [žiūrėta 2024-02-10]. ISSN 2076-3263. Prieiga per: doi:10.3390/geosciences10110425
8. ULLAH, Irfan, et al. Predictive maintenance of power substation equipment by infrared thermography using a machine-learning approach. *Energies* [interaktyvus]. 2017, **10**(12), 1987 [žiūrėta 2024-03-08]. ISSN 1996-1073. Prieiga per: doi:10.3390/en10121987
9. KUMAR, Uday, Diego GALAR, ir Dammika SENEVIRATNE. *Robots, drones, uavs and ugvs for operation and maintenance*. Taylor & Francis Group, 2020. ISBN 9780429452260.
10. ROUSOPOULOU, Vaia, et al. Data analytics towards predictive maintenance for industrial ovens. Iš: *Lecture notes in business information processing* [interaktyvus]. Cham: Springer International Publishing, 2019, pp. 83–94 [žiūrėta 2024-03-07]. ISBN 9783030209476. Prieiga per: doi:10.1007/978-3-030-20948-3_8
11. USUGA-CADAVID, Juan Pablo, et al. Using deep learning to value free-form text data for predictive maintenance. *International Journal of Production Research* [interaktyvus]. 2021, 1–28 [žiūrėta 2024-03-17]. ISSN 1366-588X. Prieiga per: doi:10.1080/00207543.2021.1951868
12. DOLL, Gary. *Rolling bearing tribology: tribology and failure modes of rolling element bearings*. Elsevier, 2022. ISBN 9780128221419

13. ISO. ISO 15243:2017, *Rolling bearings — Damage and failures — Terms, characteristics and causes*. Geneva: ISO, 2017 [žiūrėta 2024-03-29]. Prieiga per: <https://www.iso.org/standard/59619.html>
14. SKF. Bearing damage and failure analysis.[interaktyvus]. 2017. [žiūrėta 2024-03-30]. Prieiga per internetą: https://cdn.skfmediahub.skf.com/api/public/0901d1968064c148/pdf_preview_medium/0901d1968064c148_pdf_preview_medium.pdf
15. MOBLEY, R. Keith. *Introduction to predictive maintenance*. Elsevier Science & Technology Books, 2002. ISBN 9780080478692.
16. BERMAN, Karen. *Financial intelligence: a manager's guide to knowing what the numbers really mean*. Boston: Harvard Business School Press, 2006. ISBN 1591397642.
17. SHAMAYLEH, Abdulrahim, Mahmoud AWAD, ir Jumana FARHAT. IoT based predictive maintenance management of medical equipment. *Journal of Medical Systems* [interaktyvus]. 2020, **44**(4) [žiūrėta 2024-04-10]. ISSN 1573-689X. Prieiga per: doi:10.1007/s10916-020-1534-8
18. FISCHER, Ken. *ENM 590: case studies in engineering management. condition based maintenance plus return on investment analysis* [interaktyvus]. Fort Belvoir, VA: Defense Technical Information Center, January 2011 [žiūrėta 2024-04-13]. Prieiga per: doi:10.21236/ada537960
19. WOLF, Christian, et al. Model to assess the economic profitability of predictive maintenance projects. Iš: *2019 international conference on high performance computing & simulation (HPCS)* [interaktyvus]. IEEE, 2019 [žiūrėta 2024-04-12]. ISBN 9781728144849. Prieiga per: doi:10.1109/hpcs48598.2019.9188221
20. SIEMENS. The true cost of downtime 2022. *SIEMENS* [interaktyvus]. 2022 [žiūrėta 2024-04-04]. Prieiga per: <https://assets.new.siemens.com/siemens/assets/api/uuid:3d606495-dbe0-43e4-80b1-d04e27ada920/dics-b10153-00-7600truecostofdowntime2022-144.pdf>
21. WUEST, Thorsten, et al. Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research* [interaktyvus]. 2016, **4**(1), 23–45 [žiūrėta 2024-04-06]. ISSN 2169-3277. Prieiga per: doi:10.1080/21693277.2016.1192517
22. KUMBHAR, Surajkumar G., R. G. DESAVALE, ir Nagaraj V. DHARWADKAR. Fault size diagnosis of rolling element bearing using artificial neural network and dimension theory. *Neural Computing and Applications* [interaktyvus]. 2021 [žiūrėta 2024-04-06]. ISSN 1433-3058. Prieiga per: doi:10.1007/s00521-021-06228-8
23. LI, B., et al. Neural-network-based motor rolling bearing fault diagnosis. *IEEE Transactions on Industrial Electronics* [interaktyvus]. 2000, **47**(5), 1060–1069 [žiūrėta 2024-04-06]. ISSN 0278-0046. Prieiga per: doi:10.1109/41.873214
24. ZAREI, Jafar, Mohammad Amin TAJEDDINI, ir Hamid Reza KARIMI. Vibration analysis for bearing fault detection and classification using an intelligent filter. *Mechatronics* [interaktyvus]. 2014, **24**(2), 151–157 [žiūrėta 2024-04-05]. ISSN 0957-4158. Prieiga per: doi:10.1016/j.mechatronics.2014.01.003
25. SAIDI, Lotfi, Jaouher BEN ALI, ir Farhat FNAIECH. Application of higher order spectral features and support vector machines for bearing faults classification. *ISA Transactions* [interaktyvus]. 2015, **54**, 193–206 [žiūrėta 2024-04-06]. ISSN 0019-0578. Prieiga per: doi:10.1016/j.isatra.2014.08.007

26. ROJAS, A., ir A. K. NANDI. Detection and classification of rolling-element bearing faults using support vector machines. Iš: *2005 IEEE workshop on machine learning for signal processing* [interaktyvus]. IEEE, 2005, [žiūrėta 2024-04-06]. ISBN 0780395174. Prieiga per: doi:10.1109/mlsp.2005.1532891
27. SHARMA, Amandeep, et al. Bearing fault diagnosis using weighted k-nearest neighbor. Iš: *2018 2nd international conference on trends in electronics and informatics (ICOEI)* [interaktyvus]. IEEE, 2018 [žiūrėta 2024-04-07]. ISBN 9781538635704. Prieiga per: doi:10.1109/icoei.2018.8553800
28. TIAN, Jing, et al. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with k-nearest neighbor distance analysis. *IEEE Transactions on Industrial Electronics* [interaktyvus]. 2016, **63**(3), 1793–1803 [žiūrėta 2024-04-04]. ISSN 1557-9948. Prieiga per: doi:10.1109/tie.2015.2509913
29. HAN, Te, ir Dongxiang JIANG. Rolling bearing fault diagnostic method based on VMD-AR model and random forest classifier. *Shock and Vibration* [interaktyvus]. 2016, **2016**, 1–11 [žiūrėta 2024-04-06]. ISSN 1875-9203. Prieiga per: doi:10.1155/2016/5132046
30. SÁNCHEZ, René-Vinicio, et al. Feature ranking for multi-fault diagnosis of rotating machinery by using random forest and KNN. *Journal of Intelligent & Fuzzy Systems* [interaktyvus]. 2018, **34**(6), 3463–3473 [žiūrėta 2024-04-07]. ISSN 1875-8967. Prieiga per: doi:10.3233/jifs-169526
31. JANSSENS, Olivier, et al. Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration* [interaktyvus]. 2016, **377**, 331–345 [žiūrėta 2024-04-07]. ISSN 0022-460X. Prieiga per: doi:10.1016/j.jsv.2016.05.027
32. SHAO, Haidong, et al. A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mechanical Systems and Signal Processing* [interaktyvus]. 2017, **95**, 187–204 [žiūrėta 2024-04-08]. ISSN 0888-3270. Prieiga per: doi:10.1016/j.ymssp.2017.03.034
33. What is python? Executive summary. *Python.org* [interaktyvus]. [žiūrėta 2024-04-25]. Prieiga per: <https://www.python.org/doc/essays/blurb/>
34. Bearing data center | case school of engineering | case western reserve university. *Case School of Engineering* [interaktyvus]. [žiūrėta 2024-04-19]. Prieiga per: <https://engineering.case.edu/bearingdatacenter>
35. SHARMA, Aditya, M. AMARNATH, ir Pk KANKAR. Feature extraction and fault severity classification in ball bearings. *Journal of Vibration and Control* [interaktyvus]. 2014, **22**(1), 176–192 [žiūrėta 2024-04-25]. ISSN 1741-2986. Prieiga per: doi:10.1177/1077546314528021
36. TORRES-GARCÍA, Alejandro A., et al. Pre-processing and feature extraction. Iš: *Biosignal processing and classification using computational learning and intelligence* [interaktyvus]. Elsevier, 2022, pp. 59–91 [žiūrėta 2024-04-29]. Prieiga per: doi:10.1016/b978-0-12-820125-1.00014-2
37. BREIMAN, Leo. *Machine Learning* [interaktyvus]. 2001, **45**(1), 5–32 [žiūrėta 2024-04-18]. ISSN 0885-6125. Prieiga per: doi:10.1023/a:1010933404324
38. CUTLER, Adele, D. Richard CUTLER, ir John R. STEVENS. Random forests. Iš: *Ensemble machine learning* [interaktyvus]. Boston, MA: Springer US, 2012, pp. 157–175 [žiūrėta 2024-04-17]. ISBN 9781441993250. Prieiga per: doi:10.1007/978-1-4419-9326-7_5
39. KE, Guolin, et al. LightGBM: a highly efficient gradient boosting decision tree. Iš: I. GUYON, et al., eds. *Advances in neural information processing systems* [interaktyvus]. Curran

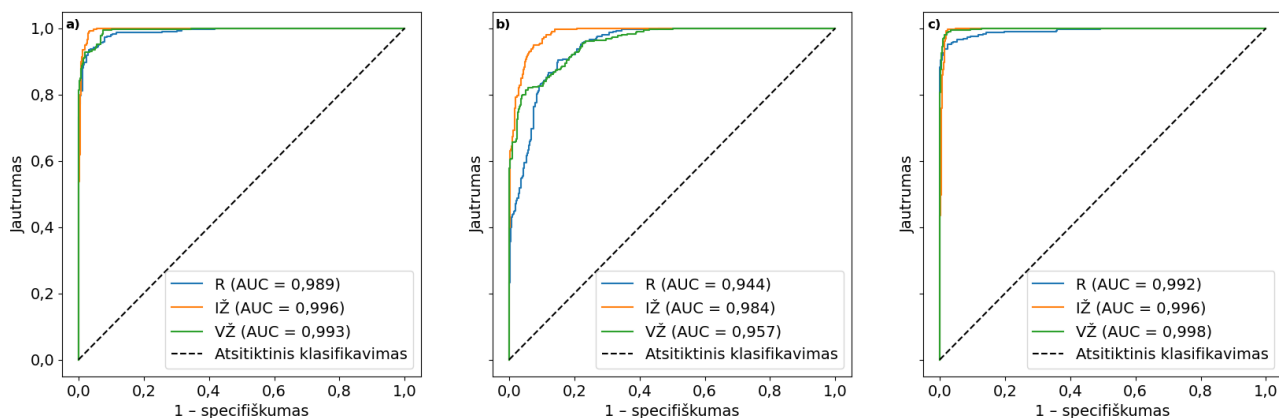
- Associates, Inc., 2017 [žiūrėta 2024-04-20]. Prieiga per: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
40. WANG, Ronghua, et al. Power system transient stability assessment based on bayesian optimized lightgbm. Iš: *2019 IEEE 3rd conference on energy internet and energy system integration (EI2)* [interaktyvus]. IEEE, 2019 [žiūrėta 2024-04-20]. ISBN 9781728131375. Prieiga per: doi:10.1109/ei247390.2019.9062027
 41. MONTESINOS LÓPEZ, Osva Antonio, Abelardo MONTESINOS LÓPEZ, and Jose CROSSA. Fundamentals of artificial neural networks and deep learning. Iš: *Multivariate statistical machine learning methods for genomic prediction* [interaktyvus]. Cham: Springer International Publishing, 2022, pp. 379–425 [žiūrėta 2024-04-20]. ISBN 9783030890094. Prieiga per: doi:10.1007/978-3-030-89010-0_10
 42. ABRAHAM, Ajith. Artificial neural networks. *Handbook of measuring system design*. 2005.
 43. SCHWAB, Bernhard, ir Peter LUSZTIG. A comparative analysis of the net present value and the benefit-cost ratio as measures of the economic desirability of investments. *The Journal of Finance* [interaktyvus]. 1969, **24**(3), 507–516 [žiūrėta 2024-04-25]. ISSN 0022-1082. Prieiga per: doi:10.1111/j.1540-6261.1969.tb00369.x
 44. HUTTER, Frank, Jörg LÜCKE, ir Lars SCHMIDT-THIEME. Beyond manual tuning of hyperparameters. *KI - Künstliche Intelligenz* [interaktyvus]. 2015, **29**(4), 329–337 [žiūrėta 2024-05-01]. ISSN 1610-1987. Prieiga per: doi:10.1007/s13218-015-0381-0
 45. VUJOVIC, Željko Đ. Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications* [interaktyvus]. 2021, **12**(6) [žiūrėta 2024-04-23]. ISSN 2156-5570. Prieiga per: doi:10.14569/ijacsa.2021.0120670
 46. PLEVRIS, V., et al. Investigation of performance metrics in regression analysis and machine learning-based prediction models. Iš: *8th european congress on computational methods in applied sciences and engineering* [interaktyvus]. CIMNE, 2022 [žiūrėta 2024-04-23]. Prieiga per: doi:10.23967/eccomas.2022.155

Informacijos šaltinių sąrašas

1. Model 352C68 | PCB Piezotronics. *PCB Piezotronics | Sensors to measure vibration, acoustics, force, pressure, load, strain, shock & torque* [interaktyvus]. [žiūrėta 2024-04-09]. Prieiga per: <https://www.pcb.com/products?m=352c68>
2. Raspberry pi 4 model B (4GB). *Electromaker.io - Unleash Your Creativity with DIY Electronics & Maker Projects* [interaktyvus]. [žiūrėta 2024-04-09]. Prieiga per: <https://www.electromaker.io/shop/product/raspberry-pi44gb>
3. Raspberry pi 4 model B (4GB). *Electromaker.io - Unleash Your Creativity with DIY Electronics & Maker Projects* [interaktyvus]. [žiūrėta 2024-04-11]. Prieiga per: <https://www.electromaker.io/shop/product/raspberry-pi44gb>
4. Mobiliojo ryšio pokalbių planai. *Telia* [interaktyvus]. [žiūrėta 2024-04-11]. Prieiga per: <https://www.telia.lt/privatiems/mobiliojo-rysio-planai?24-men-sutartis>
5. Koaksialinis kabelis - TV, SAT, CCTV, radijas | 75 Ω 50 Ω PB580C RG58 | žema kaina | pirkti | internetinė parduotuvė electrobase.lt. Electrobase ®  Noliktavā 2000+ Unikālas Preces  Ātra Piegāde [interaktyvus]. [žiūrėta 2024-04-11]. Prieiga per: <https://electrobase.lv/lt/katalogas/koaksialinis-kabelis-tv-sat-cctv-radijas-75-o-50-o/koaksialais-kabelis-pro-base-rg58-0>
6. Naujausi darbo skelbimai | cvbankas.lt. *Šiandienos darbo skelbimai* | *CVbankas.lt* [interaktyvus]. [žiūrėta 2024-04-10]. Prieiga per: <https://www.cvbankas.lt/>

Priedai

1 priedas. OvR ROC kreivės kiekvienai defekto vietai: a) atsitiktiniai miškai, b) neuroniniai tinklai, c) LightGBM



2 priedas. OvR ROC kreivės kiekvienam defekto dydžiui: a) atsitiktiniai miškai, b) neuroniniai tinklai, c) LightGBM

