



**Vilnius
universitetas**

Duomenų Mokslo ir Skaitmeninių Technologijų Institutas






Informatikos inžinerijos krypties doktorantų konferencija
Veiklos ataskaita už 2022 m. spalio 1d. – 2023 m. kovo 21d.

Dalia BRESKUVIENĖ – Informatikos inžinerija T 007 doktorantė

Darbo vadovas – prof. habil. dr. Gintautas DZEMYDA

Doktorantūros pradžios ir pabaigos metai: 2021.12.01 – 2025.11.30



Disertacijos tema, tyrimo objektai ir tikslas

- Preliminari disertacijos tema:
- Klasifikatoriaus (nesubalansuotos) mokymo aibės optimizavimas, siekiant geresnės klasifikavimo kokybės.
- Tyrimo objektai:
- Finansinio sukčiavimo duomenų optimizavimas, siekiant tikslesnių mašininio mokymosi metodų rezultatų.
- Tikslas:
- Sukurti arba patobulinti nesubalansuotų duomenų optimizavimo metodą, skirtą finansinio sukčiavimo identifikavimui ir prevencijai.



Tyrimo uždaviniai

- Identifikuoti tinkamus metodus nesubalansuotos mokymo aibės optimizavimui;
- Identifikuoti aktualias mokslines problemas, kylančias uždaviniuose, susijusiuose su finansinio sukčiavimo aptikimu;
- Sukurti arba patobulinti algoritmą nesubalansuotos mokymo aibės optimizavimui atsižvelgiant į naujo taško klasifikavimą;
- Pritaikyti sukurtą arba patobulintą metodą nesubalansuotiems duomenims ir atlikti gautų duomenų analizę, rezultatų apibendrinimą, išvadų parengimą.

2021–2025m.

Vilniaus
universitetas

STUDIJŲ PLANAS IR JO VYKDYMO SUVESTINĖ

Studijų metai	Egzaminai ¹		Dalyvavimas konferencijose ²				Publikacijos ³					
			Tarptautinėse		Nacionalinėse		Su citav. rodikliu			Be citav. rodiklio		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė ⁴	Planas	Įvykdyta	Būklė ⁴
I (2021/2022)	3	3			1	1						
II (2022/2023)	1	1	1	1		1				1	1	Publikuota
III (2023/2024)			1				1					
IV (2024/2025)							1					
Iš viso:	4	4	2	1	1	2	2			1	1	

2022–2023m.

Vilniaus
universitetas

ATASKAITINIŲ METŲ DARBO PLANAS IR JO SUVESTINĖ

Egzaminai			
Fundamentalieji informatikos ir informatikos inžinerijos metodai		Išlaikyta	
Dalyvavimas konferencijose 2022/2023 (I pusmetis)			
Planas	Įvykdyta	Konferencijos tipas	
13th Conference: Data Analysis Methods for Software Systems, 2022.12.01/03 , Druskininkai	D. Breskuvienė, G. Dzemyda, „Autoencoder for Fraudulent Transactions Data Feature Engineering “, 13th Conference: Data Analysis Methods for Software Systems, 2022.12.01/03 , Druskininkai	Nacionalinė konferencija	
Publikacijos 2022/2023 (I pusmetis)			
Planas	Įvykdyta	Būklė	Publikacijos tipas
	D. Breskuvienė, G. Dzemyda (2022) „Autoencoder for Fraudulent Transactions Data Feature Engineering “ In: Bernatavičienė J. (2021) “Data Analysis Methods for Software Systems”, Vilnius University Proceedings, 17, pp. 1-82. doi: 10.15388/DAMSS.12.2021	Publikuota	Be cituojamumo rodiklio
Chapter in the book “Data Science in Application“	D. Breskuvienė, G. Dzemyda (2023). Imbalanced data classification approach based on clustered training set. In: Dzemyda G., Bernatavičienė J., Kacprzyk J. (Eds.), Data Science in Applications. Studies in Computational Intelligence. Springer	Publikuota	Be cituojamumo rodiklio

2021–2025m.

Vilniaus
universitetas

Dalyvavimas tarptautinėse konferencijose

Aprašas

D.Breskuvienė, G.Dzemyda „Clustering-based optimization in fraud detection classifier training“
EURO2022 2022.07.03 /06 ESPOO, FINLAND

Mokslinių tyrimų ir disertacijos rengimo planas:

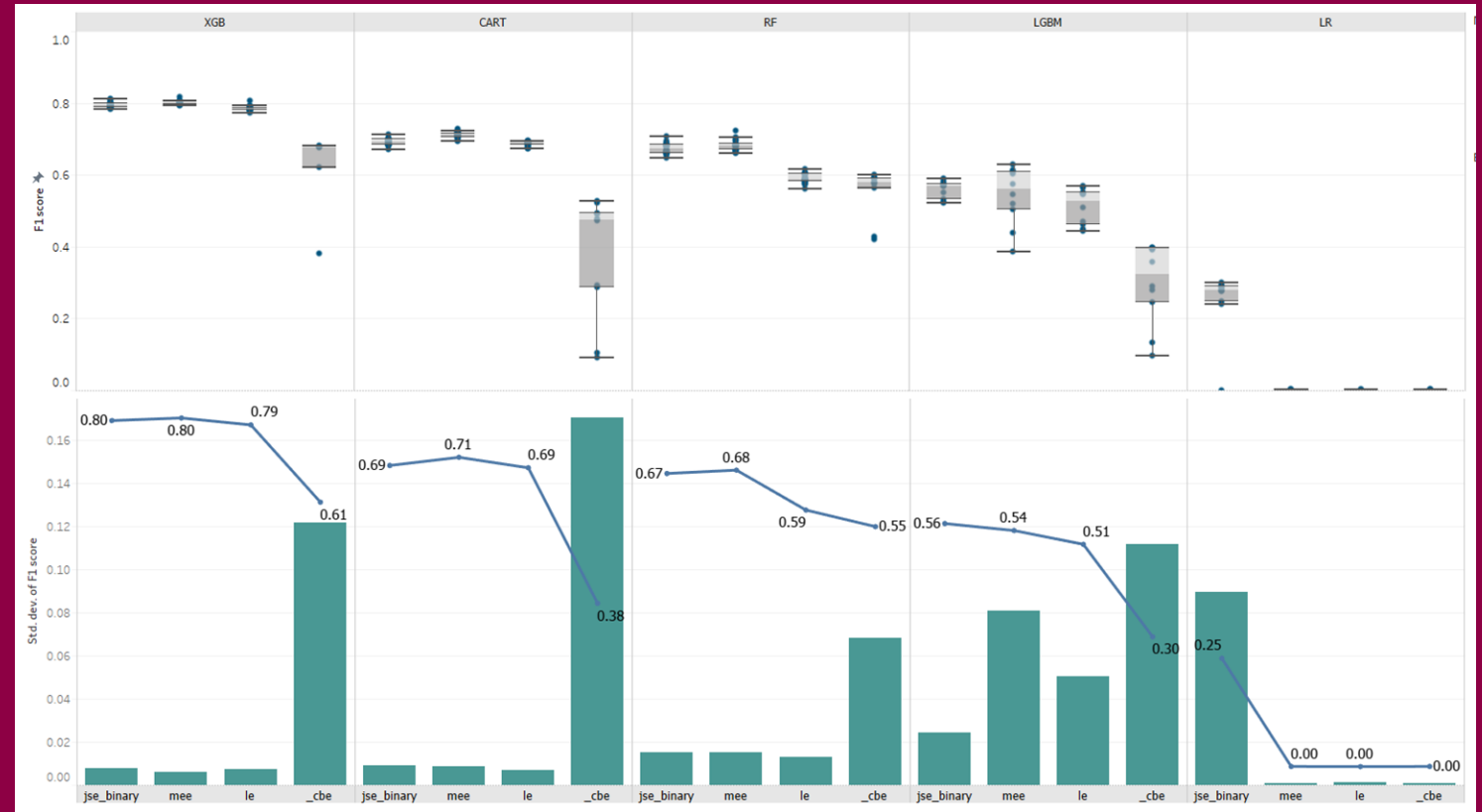
	Darbo pavadinimas	Atlikimo terminai
1.	<p><u>Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):</u></p> <p>1.1. Disertacijos tyrimo objekto detalizavimas. 1.2. Atlikti būdų klasifikatorių veikimo optimizavimui analitinę apžvalgą. 1.3. Nustatyti (identifikuoti) mokslines problemas, kylančias uždaviniuose, susijusiuose su klasifikavimo kokybės optimizavimu, o taip pat ir naudojant giliuosius neuroninius tinklus. 1.4. Tyrimo tikslo suformavimas.</p>	<p>2021 m. gruodžio mėn. – 2022 m. vasario mėn. 2021 m. gruodžio mėn. – 2022 m. spalio mėn. 2022 m. kovo mėn. – 2022 m. spalio mėn. 2022 m. kovo mėn. – 2022 m. spalio mėn.</p>
2.	<p><u>Mokslinio tyrimo vykdymas:</u></p> <p>2.1. Tyrimo metodikos sudarymas: 2.1.1. Tyrimo metodikos išskeltiems uždaviniams spręsti parinkimas; 2.1.2. Teorinio ir empirinio tyrimų suplanavimas pagal pasirinktą metodiką.</p> <p>2.2. Teorinis tyrimas: 2.2.1. Klasifikatorių efektyvumo galimybių tyrimas optimizuojant mokymo aibės taškų parinkimą. 2.2.2. Giliųjų neuroninių tinklų panaudojimo galimybių optimaliai mokymo aibei rasti tyrimas.</p> <p>2.3. Empirinis tyrimas: 2.3.1. Sudarytų metodų pritaikymas praktinių uždavinių sprendimui. 2.3.2. Gautų duomenų analizė, rezultatų apibendrinimas, išvadų parengimas.</p>	<p>2022 m. kovo mėn. – 2022 m. spalio mėn. 2022 m. kovo mėn. – 2022 m. spalio mėn.</p> <p>2022 m. lapkričio mėn. – 2023 m. spalio mėn. 2022 m. lapkričio mėn. – 2023 m. spalio mėn.</p> <p>2024 m. kovo mėn. – 2024 m. spalio mėn. 2024 m. spalio mėn. – 2025 m. vasario mėn.</p>
3.	<p><u>Atskirų daktaro disertacijos dalių (tyrimo metodikos, rezultatų, ginamų teiginių, išvadų, ir kt.) parengimas:</u></p> <p>3.1. Tikslų, uždavinių, tyrimo metodikos, ginamųjų teiginių patikslinimas; 3.2. Analitinės disertacijos dalies parengimas; 3.3. Teorinės disertacijos dalies parengimas; 3.4. Eksperimentinės disertacijos dalies parengimas; 3.5. Bendrųjų išvadų formulavimas.</p>	<p>2024 m. spalio mėn. – 2025 m. vasario mėn. 2024 m. kovo mėn. – 2025 m. rugpjūčio mėn. 2024 m. kovo mėn. – 2025 m. rugpjūčio mėn. 2024 m. kovo mėn. – 2025 m. rugpjūčio mėn. 2024 m. kovo mėn. – 2025 m. rugpjūčio mėn.</p>
4.	Daktaro disertacijos parengimas ir svarstymas padalinyje	2025 m. rugsėjo mėn.
5	Daktaro disertacijos gynimas	2025 m. lapkričio mėn.

Dalinai atliktas teorinis tyrimas

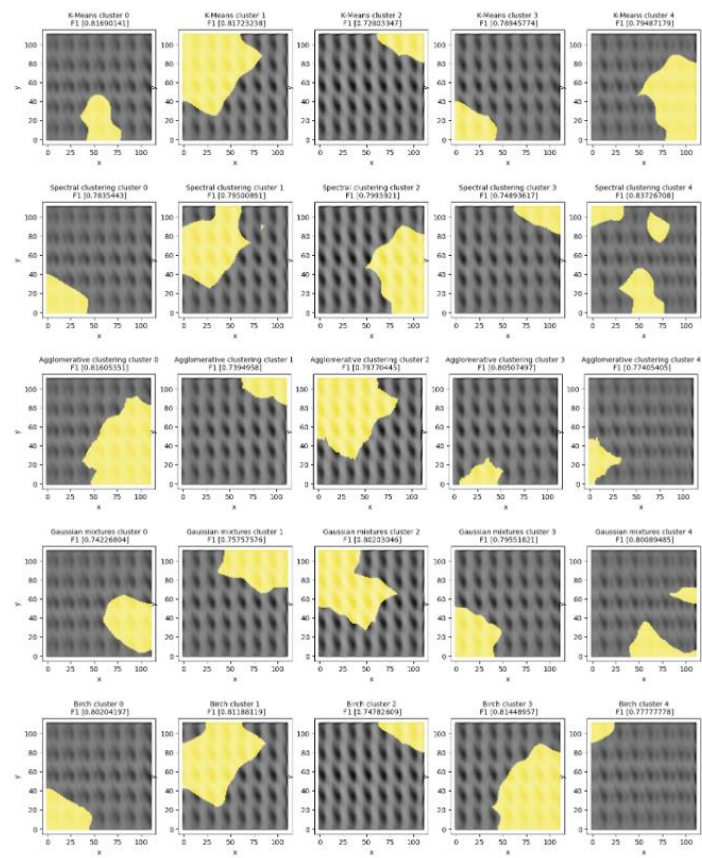
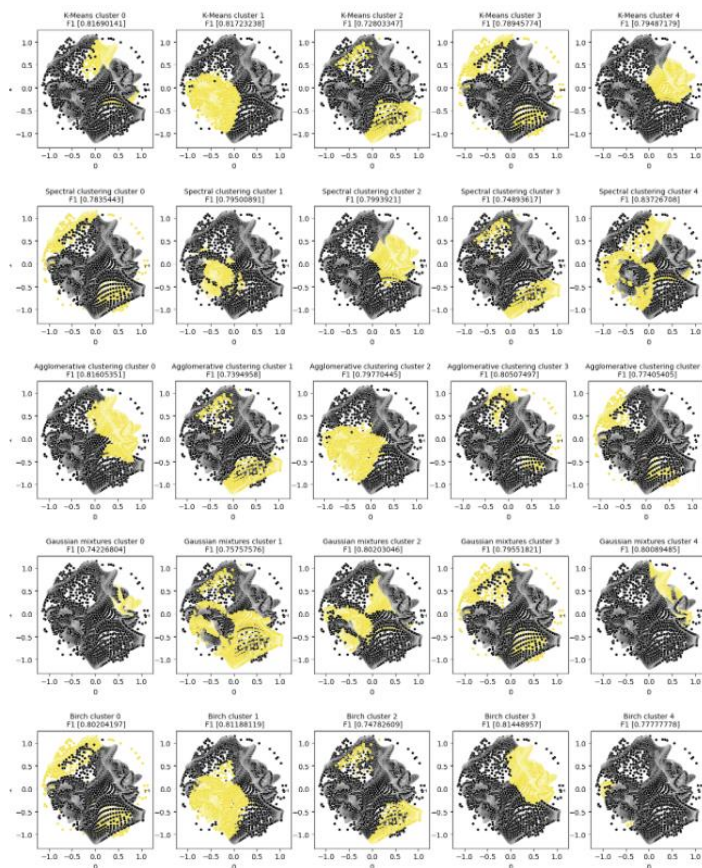
Vilniaus
universitetas

- Atliktas teorinis tyrimas apie nesubalansuotos aibės optimizavimą naudojant *undersample* ir *oversample* metodus.
- Atliktas empirinis tyrimas panaudojant *Kmeans* metodą, tam kad būtų galima sukurti kuo homogeniškesni duomenų klasteriai, kuriems vėliau pritaikytas *undersample* metodas.
- Atliktas teorinis ir empirinis tyrimas siekiant nustatyti ar kategorinių duomenų kodavimas turi įtakos klasifikavimo kokybei.
- Dalinai atliktas teorinis ir empirinis SOM (self-organizing-map) panaudojimas mokymo aibės optimizavimui

Atliktas teorinis ir empirinis tyrimas siekiant nustatyti ar kategorinių duomenų kodavimas turi įtakos klasifikavimo kokybei.



Dalinai atliktas teorinis ir empirinis SOM (self-organizing-map) panaudojimas mokymo aibės optimizavimui



KITO PUSMEČIO DARBO PLANAS

- Teorinis tyrimas:
 - Klasifikatorių efektyvumo galimybių tyrimas optimizuojant mokymo aibės taškų parinkimą.
 - Giliųjų neuroninių tinklų panaudojimo galimybių optimaliai mokymo aibei rasti tyrimas.
 - Straipsnio rengimas periodiniam mokslo žurnalui International Journal of Computers, Communications & Control (IJCCC) su cituojamumo faktoriumi.
-



**Vilnius
universitetas**

Ačiū už dėmesį!

Klausimai?