



**Vilnius  
universitetas**

# **PROPAGANDOS APTIKIMAS IR KLASIFIKAVIMAS SOCIALINĖS MEDIJOS PRANEŠIMUOSE NAUDOJANT HIBRIDINIO GILIOJO MOKYMOSI METODUS**

Doktorantė: Ieva Rizgeliene

Vadovas: prof. dr. Darius Plikynas

Mokslo kryptis: N009 Informatika

Studijų laikotarpis: 2022 – 2026 m.

Studijų metai: 2023/2024



# Tyrimo objektas ir tikslas

Vilniaus  
universitetas

## Objektas:

- socialinėje medijoje publikuojami pranešimai informacinio karo kontekste ir su jais susijusi informacija;
- hibridiniai giliojo mokymosi algoritmai propagandos požymių aptikimui.

## Tikslas:

Sukurti hibridinį giliojo mokymosi metodą, skirtą propagandos požymių aptikimui socialinės žiniasklaidos pranešimuose lietuvių kalba.

# Visų studijų planas ir jo vykdymo suvestinė

Vilniaus  
universitetas

Studijų metai	Egzaminai	
	Planas	Įvykdyta
I (2022/2023)	2	2
II (2023/2024)	2	1
III (2024/2025)		
IV (2025/2026)		
Iš viso:	4	3

**LIKS** LIETUVOS  
KOMPIUTERININKŲ  
SĄJUNGA

## SERTIFIKATAS

PAŽYMIME, KAD VU DMSTI DOKTORANTĖ

*Ieva Rizgelienė*

2023 m. rugsėjo 28-29 d. dalyvavo  
„Kompiuterininkų dienų 2023“ organizuojuose užsiėmimuose:

- išklausė „Išmaniosios visuomenės, skaitmeninės Lietuvos“ plenarinio posėdžio pranešimus;
- pristatė savo darbą „Propagandos, dezinformacijos ir melagingos informacijos autorių bei skleidėjų aptikimas socialiniuose tinkluose naudojant mašininio mokymosi metodus“ Mokslinėje kompiuterininkų konferencijoje;
- dalyvavo 6 akad. val. mokymuose „Informacinės technologijos, saugumas, komunikacija, duomenų analizė, blokų grandinės“ [bit.ly/KoDi-23md](https://bit.ly/KoDi-23md);
- atliko 12 akad. val. trukmės savarankišką darbą tema „Blokų grandinių ir išskirstytų įrašų technologijų įvaldymas“.

Užsiėmimų ir mokymų trukmė vertinama **dviem studijų kreditais**.

2023 m. spalio 24 d.  
Nr. KoDi2023-002

Saulius Maskeliūnas  
LIKS Tarybos pirmininkas

Dokumentą elektroniniu parašu  
pasirašė SAULIUS MASKELIŪNAS  
Data: 2023-10-24 15:30:58

# Visų studijų planas ir jo vykdymo suvestinė

Vilniaus  
universitetas

Egzaminai		
Planas	Įvykdyta	Būklė
Mašininis mokymas	Egzamino pavadinimas: Mašininis mokymas Data: 2023-03-02	Išlaikytas
Informatikos ir informatikos inžinerijos tyrimo metodai ir metodika	Egzamino pavadinimas: Informatikos ir informatikos inžinerijos tyrimo metodai ir metodika Data: 2023-06-27	Išlaikytas
Fundamentalieji informatikos ir informatikos inžinerijos metodai	Egzamino pavadinimas: Fundamentalieji informatikos ir informatikos inžinerijos metodai Data: 2024-01-24	Išlaikytas
Bendruosius gebėjimus stiprinančios veiklos	“Kompiuterininkų dienos 2023” Data: 2023m. rugsėjo 28-29d.	1.5 kredito.

# Visų studijų planas ir jo vykdymo suvestinė

Vilniaus  
universitetas

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse		Nacionalinėse		Su citav.rodikliu			Be citav. Rodiklio		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė	Planas	Įvykdyta	Būklė
I (2022/2023)			1	1						
<b>II (2023/2024)</b>	<b>1</b>			<b>1</b>						
III (2024/2025)	1				1					
IV (2025/2026)	1				1					
Iš viso:	3	0	1	2	2	0				

## Dalyvavimas konferencijose

Skaitmeninės humanitarikos seminaras, VU, 2022-11-23. Lietuva, Vilnius	Pristatymo pavadinimas: Propogandos automatinio atpažinimo mašininio mokymo metodais tyrimas ir socialinio poveikio modeliavimas. Autoriai: Ieva Rizgelienė, Darius Plikynas.	Seminaras
13th Data analysis Methods for Software Systems Conference, 2022 12 02. Lietuva, Druskininkai	Pristatymo pavadinimas: “Automated Propaganda Detection Using Deep Learning Methods: a Pilot Systematic Review Research”, Autoriai: Ieva Rizgelienė, Darius Plikynas.	Nacionalinė
Kompiuterininkų dienos – 2023, 2023-09-28. Lietuva, Kaunas.	Pristatymo pavadinimas: “Propagandos, dezinformacijos ir melagingos informacijos autorių bei skleidėjų aptikimas socialiniuose tinkluose naudojant mašininio mokymosi metodus”, Autoriai: Ieva Rizgelienė.	Nacionalinė

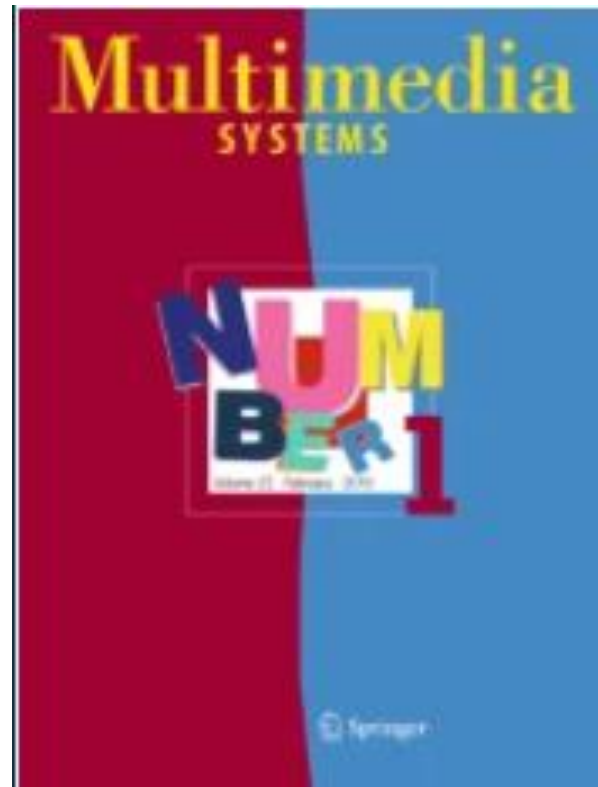
# Sisteminės apžvalgos straipsnis

A combined PRISMA review on machine learning research to analyze the authors, content and social impact of fake news, disinformation and propaganda in social media

D. Plikynas , I. Rizgelienė, and G. Korvel

Vilniaus  
universitetas

- Straipsnis paruoštas ir pridotas žurnalui Multimedia Systems.
- Laukiama atsakymo.



- Impact factor 3.9
- COMPUTER SCIENCE, THEORY & METHODS - (Q1)
- COMPUTER SCIENCE, INFORMATION SYSTEMS - SCIE(Q2)

# Dalyvavimas tarp. konferencijoje

Vilniaus  
universitetas

## Comparative Analysis of Various Data Balancing Techniques for Propaganda Detection in Lithuanian News Articles

Ieva Rizgeliienė and Gražina  
Korvel

The 11th IEEE  
Workshop on Advances  
in Information,  
Electronic and Electrical  
Engineering AIEEE'2024  
Valmiera, Latvia, May  
31 –June 1, 2024



# Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus  
universitetas

Darbo pavadinimas	Atlikimo terminai	Pastabos
<p>1. Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):</p> <p>1.1 Disertacijos tyrimo objekto detalizavimas.</p> <p>1.2 Gilių neuroninių tinklų mokymu grįstų metodų apžvalga kompiuterinės lingvistikos ir natūralios kalbos apdorojimo kontekste, siekiant aptikti propagandos požymius.</p> <p>1.3 Metodų apžvalgos palyginimas, apibendrinimas ir pateikimas disertacijos analitinės dalies aprašyme.</p> <p>1.4 Tyrimo tikslo suformavimas.</p>	<p>2022 m. lapkričio mėn. – 2023 m. vasario mėn.</p>	<p>Įvykdyta.</p>
<p><b>Mokslinio tyrimo vykdymas:</b></p>	<p>2023 m. kovo mėn. – 2023 m. rugsėjo mėn.</p>	<p>Įvykdyta.</p>
<p><b>2.1. Tyrimo metodikos sudarymas:</b></p> <p>2.1.1. Uždavinių, skirtų tyrimo tikslui pasiekti, suformulavimas.</p> <p>2.1.2. Tyrimo metodikos išsikeltiems uždaviniams spręsti parinkimas.</p> <p>2.1.3. Teorinio ir empirinio tyrimų suplanavimas pagal pasirinktą metodiką.</p>		



# Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus  
universitetas

<b>2.2. Teorinis tyrimas:</b>  2.2.1. Socialinės medijos naujienų pranešimų anglų kalba duomenų repozitoriumo panaudojimo galimybių analizė skirta gilių neuroninių tinklų apmokymo ir testavimo tyrimui.  2.2.2. Ekspertinės semantinės anglišų tekstinių duomenų analizės tyrimas skirtas nustatyti propagandos tipus.	2023 m. spalio mėn. – 2024 m. vasario mėn.	Įvykdyta
2.2.3. Gilaus mokymo (pvz. BERT, GPT, LSTM bei CRM ) ir semantinės analizės metodų pritaikymo galimybių analizė, skirta propagandos požymių aptikimui tekstuose ir jų klasifikavimui.  2.2.4. Propagandos reitingavimo galimybių tyrimas.	2024 m. kovo mėn. – 2024 m. rugsėjo mėn.	Vykdoma
<b>2.3. Empirinis tyrimas:</b>  2.3.1. Remiantis teoriniu 2.2.1. tyrimu, socialinės medijos naujienų pranešimų lietuvių kalba duomenų repozitoriumo sudarymas skirtas gilių neuroninių tinklų apmokymui ir testavimui, siekiant aptikti propagandą naujienų pranešimuose.  2.3.2. Remiantis teoriniu 2.2.2. tyrimu, lietuviškų naujienų tekstų repozitoriumo duomenų ekspertinė semantinė bei sentimentų analizė, nustatant propagandos požymius ir juos klasifikuojant.	2024 m. spalio mėn. – 2025 m. vasario mėn.	Vykdoma

# Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus  
universitetas

2.3.2. Remiantis teoriniu 2.2.2. tyrimu, lietuviškų naujienų tekstų repozitoriumo duomenų ekspertinė semantinė bei sentimentų analizė, nustatant propagandos požymius ir juos klasifikuojant.		
2.3.3. Remiantis teoriniu 2.2.3. tyrimu, gilaus mokymo BERT, GPT, LSTM bei CRM ir semantinės analizės metodų pritaikymas lietuviškų tekstų propagandos požymių aptikimui tekstuose ir jų klasifikavimui.  2.3.4. Naujienų pranešimų propagandos reitingavimo metodikos sukūrimas lietuviškiems tekstams, panaudojant 2.2.4. tyrimo rezultatus.	2025 m. kovo mėn. – 2025 m. rugsėjo mėn.	Vykdoma
<b>2.3.5. Programinio įrankio/aplikacijos sukūrimas naujienų pranešimų automatinei propagandos analizei.</b>	2025 m. spalio mėn. – 2026 m. gruodžio mėn.	
2.4. Gautų duomenų analizė, apibendrinimas, išvadų parengimas:  2.4.1. Teorinio tyrimo apibendrinimas.  2.4.2. Empirinio tyrimo apibendrinimas.  2.4.3. Rezultatų apibendrinimas, esminių rezultatų išskyrimas bei išvadų parengimas.	2025 m. gruodžio mėn. – 2026 m. kovo mėn.	

# Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus  
universitetas

<b>2.4. Gautų duomenų analizė, apibendrinimas, išvadų parengimas:</b>	2025 m. gruodžio mėn. – 2026 m. kovo mėn.	
2.4.1. Teorinio tyrimo apibendrinimas.		
2.4.2. Empirinio tyrimo apibendrinimas.		
2.4.3. Rezultatų apibendrinimas, esminių rezultatų išskyrimas bei išvadų parengimas.		
<b>Atskirų daktaro disertacijos dalių (tyrimo metodikos, rezultatų, ginamų teiginių, išvadų ir kt.) parengimas:</b>	2025 m. kovo mėn. – 2026 m. birželio mėn.	
1. Tikslų, uždavinių, tyrimo metodikos, ginamųjų teiginių patikslinimas.		
2. Analitinės disertacijos dalies parengimas.		
3. Teorinės disertacijos dalies parengimas.		
4. Eksperimentinės disertacijos dalies parengimas.		
5. Bendrųjų išvadų formulavimas.		
Daktaro disertacijos parengimas ir svarstymas padalinyje	2026 m. birželio mėn.	
Daktaro disertacijos gynimas	2026 m. rugsėjo mėn.	



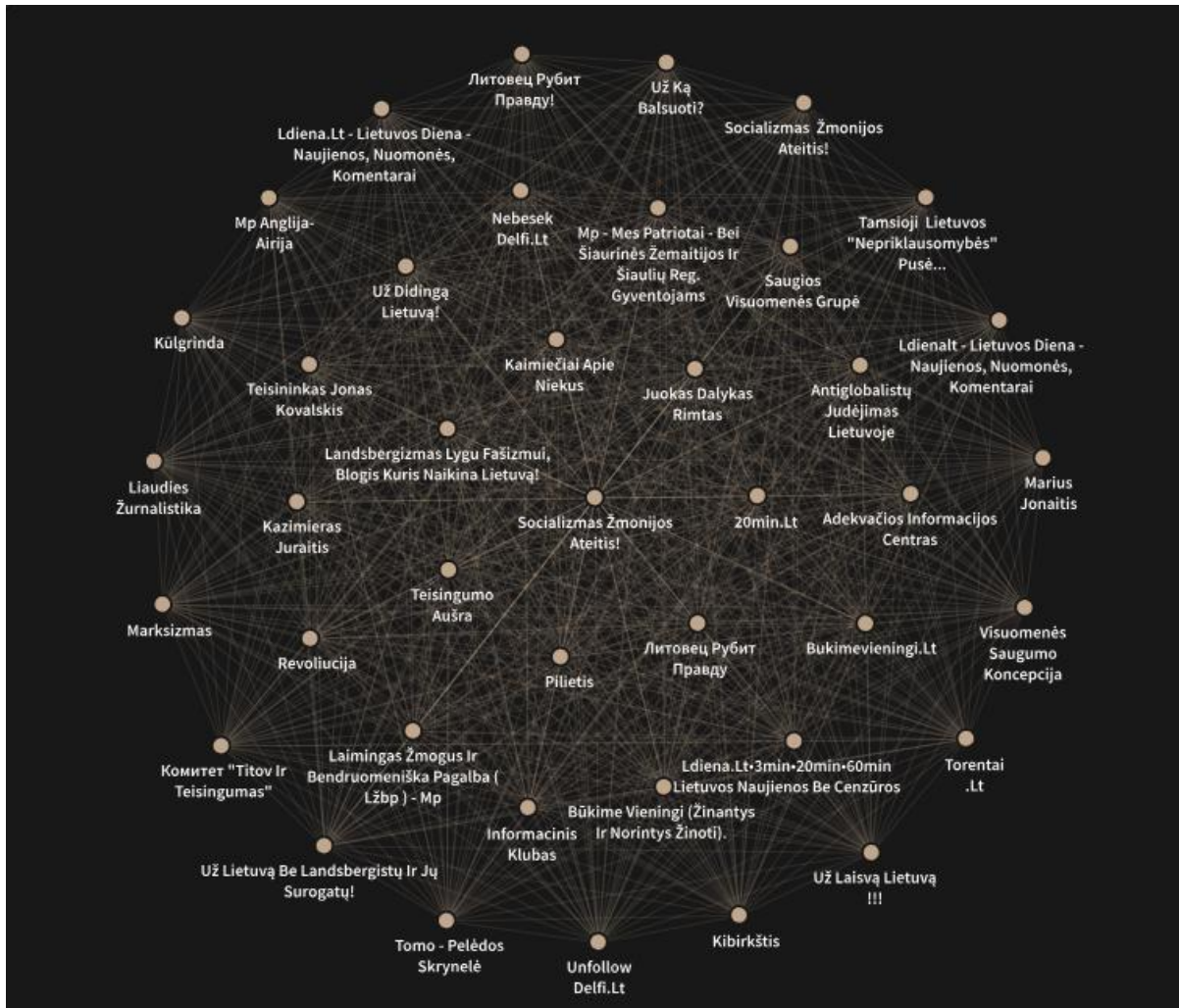
**Vilnius  
universitetas**



# **PROPAGANDOS TEKSTYNO SUDARYMAS LIETUVIŲ KALBA**

# Atrinkti šaltiniai

Vilniaus  
universitetas



- Nepatikimi informacijos šaltiniai atrinkti remiantis LRT atliktu žurnalistiniu tyrimu, bendradarbiaujant su „DebunkEU.org“, kurio metu nustatyti šaltiniai, skleidžiantys propagandą ir dezinformaciją.
- Atrinktų tinklapių visi straipsniai buvo atsiųsti, šiuo metu anotuojami tekstai iš dviejų šaltinių, vėliau planuojama pridėti daugiau.
- Taip siunčiamas LRT.lt portalas, kaip patikimas šaltinis.

[LRT tyrimas. Lietuvos „penktoji kolona“: Rusijos propaganda platina šeimos gynėjai, sektos ir knygu apie Staliną leidėjai - LRT](#)

# Propagandos ir dezinformacijos temos

Vilniaus  
universitetas

1. Pabėgėlių/migrantų krizės eskalavimas.
2. Karo Ukrainoje dezinformacija.
3. Lietuvos respublikos šmeižimas.
4. Lietuvos ginkluotųjų pajėgų finansavimo šmeižimas.
5. Nepasitikėjimas Lietuvos institucijomis.
6. Pasitikėjimo Vakarų institucijomis mažinimas.
7. Vakarų civilizacijos pabaigos idėjos skatinimas (vertybių išsigimimas, LGT..).
8. Efektyvaus valdymo pavyzdžio (pvz. Minskas, Maskva, Pekinas) skatinimas.
9. Vašingtono hegemonijos' pabaigos teorija.
10. Naujosios pasaulio tvarkos idėjos propagavimas.
11. Litvinizmas.

# Propagandos technikos

Vilniaus  
universitetas

1. **Emocinė raiška** - Bandytas sukelti stiprius jausmus.
2. **Whataboutism, Red herring, Straw Men** – Išsisukinėjimas, oponento pozicijos, jo teiginių menkinimas, dėmesio nukreipimas kitur.
3. **Supaprastinimas** – Daroma prielaida, kad yra viena problemos priežastis, neanalizuojant problemos sudėtingumo.
4. **Neapibrėžtumas** - Vartojama neaiški kalba, kad auditorija žinutę galėtų interpretuoti savaip.
5. **Apeliavimas į autoritetą** - Cituojami garsūs, žinomi autoritetai, kurie remia propagandisto idėją.
6. **Mojavimas vėliava** - Pastangos pateisinti veiksma remiantis patriotiškumu.
7. **Sekimas iš paskos** - Apeliavimas į bandos jausmą bandant įtikinti auditoriją prisijungti ir imtis veiksmų.
8. **Abejojimas** - Šmeižtas, abejonės dėl kieno nors patikimumo.
9. **Reductio ad hitlerium** - Siekiama įtikinti nepritarti veiksmui, idėjai, nurodant, kad tai populiaru tarp grupių, kurių tikslinė auditorija nekenčia.
10. **Pakartojimas** - Tekste kartojama ta pati žinutė.



# Anotavimo procesas

**Ar šis tekstas  
yra  
propaganda?**

- Taip
- Ne
- Propagandos citavimas
- Neįmanoma nustatyti



**Jei TAIP**

- Pasirenkama propagandos tema/temos

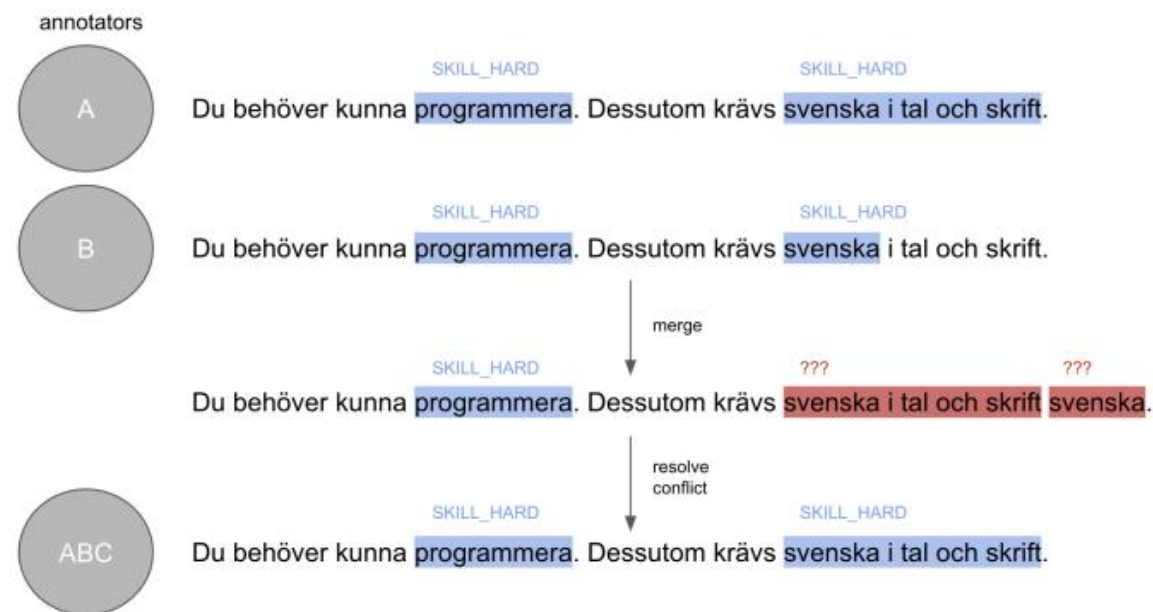
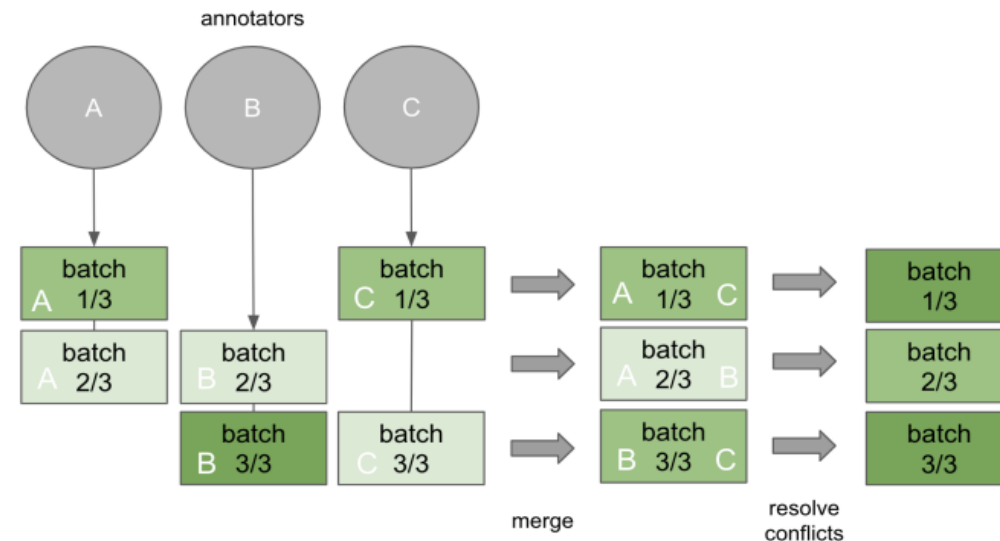


Tekstas  
anotuojamas,  
pažymint  
propagandos  
fragmentus.

# Kryžminio anotavimo procesas

- Tekstus anotuoja trys ekspertai.
- Ekspertai pirmiausiai suanotuoja tik žymint ar tekstai yra propaganda ar ne. Kas savaitę apie 150 tekstų.
- Tuomet atrenkama 30 propagandinių tekstų.
- Kiekvienas ekspertas anotuoja po 20 tekstų kiekvieną savaitę, taip, kad kiekvienas tekstas būtų suanotuotas dviejų ekspertų.
- Kiekvieną savaitę ekspertai susitinka poromis ir aptaria anotacijas, bei susitaria dėl galutinės teksto anotacijos.

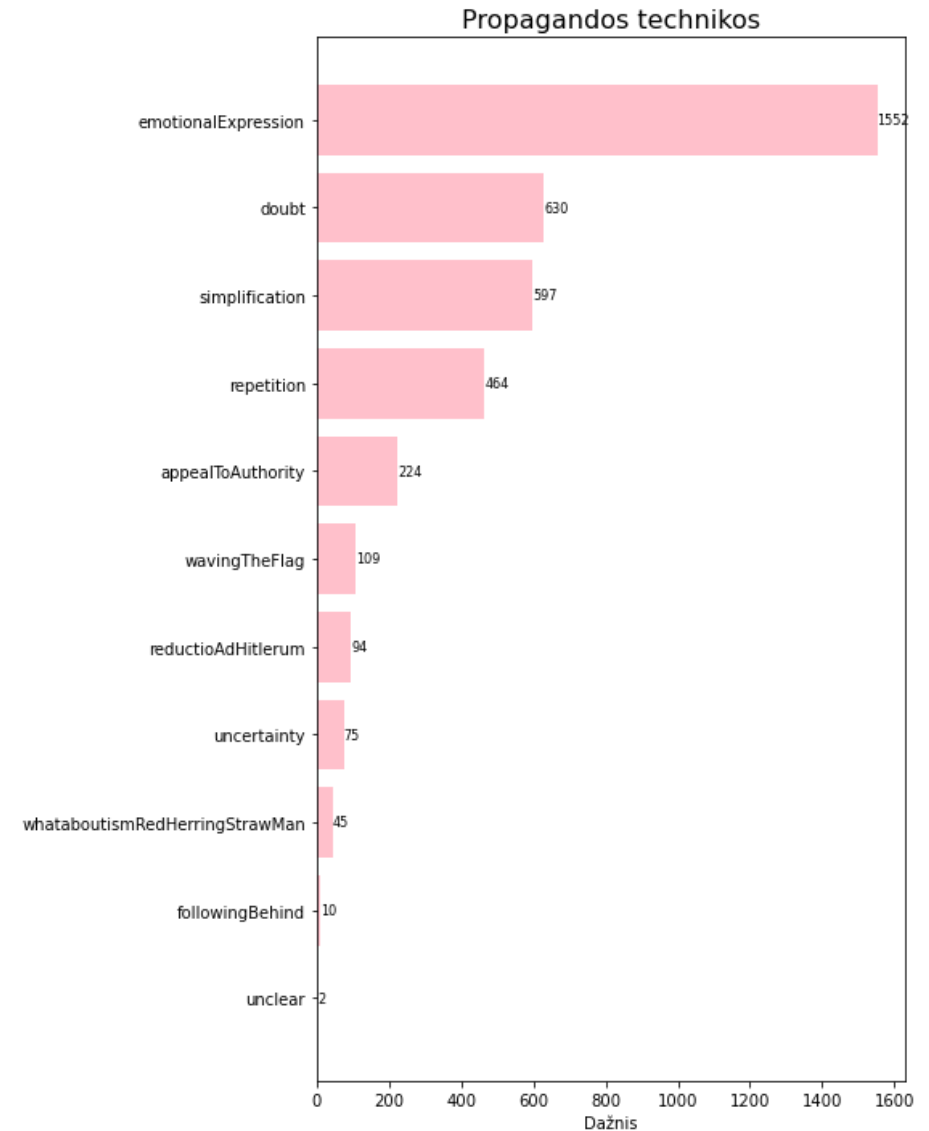
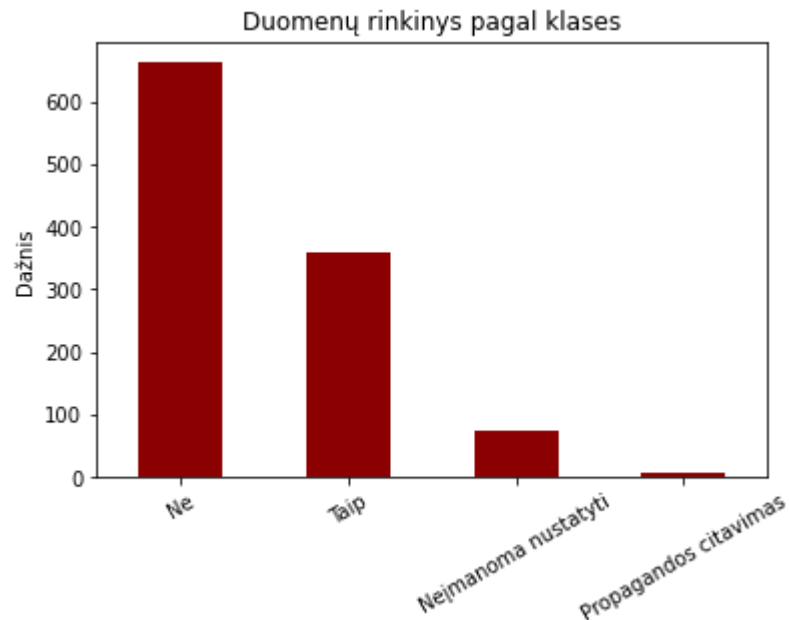
Stollenwerk, F., Öhman, J., Petrelli, D., Wallerö, E., Olsson, F., Bengtsson, C., Horndahl, A., & Gandler, G. Z. (2023). Text Annotation Handbook: A Practical Guide for Machine Learning Projects. *ArXiv*. /abs/2310.11780



# Duomenų rinkinys šiuo metu

Vilniaus  
universitetas

- Suanotuota tekstų, tik propagandos aptikimui – **1105**.
- Suanotuota tekstų, propagandos technikų aptikimui – **236**.
- Šiuo metu anotuojami tekstai iš dviejų skirtingų šaltinių.





**Vilniaus  
universitetas**

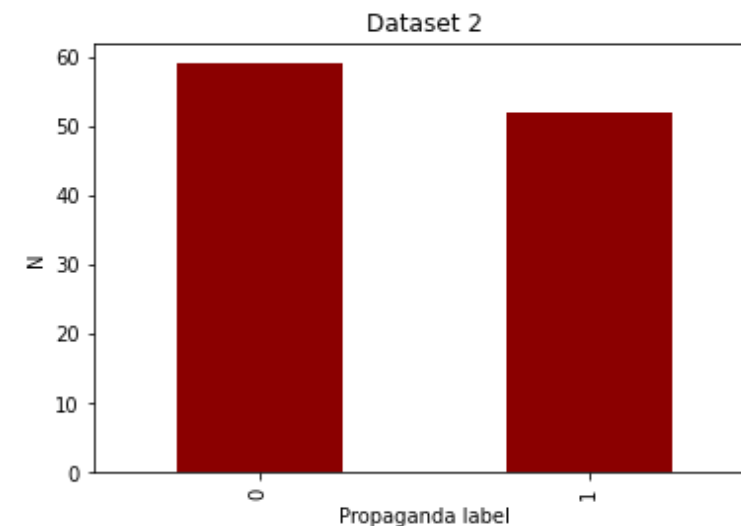
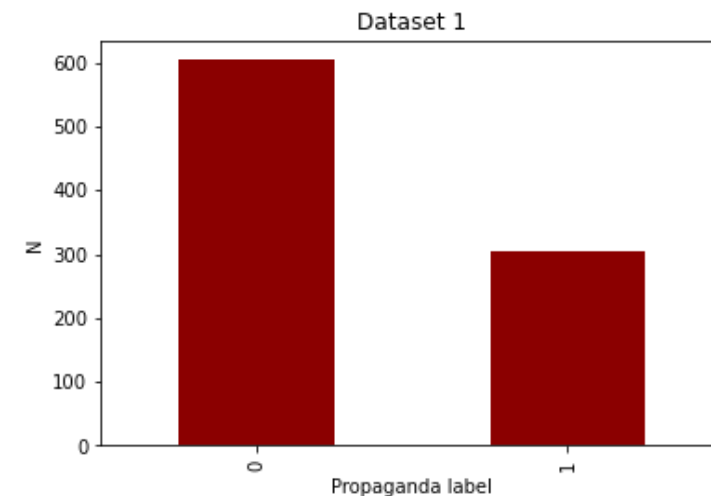
---

# **Duomenų balansavimo metodų lyginamoji analizė**

# Duomenų rinkiniai tyrimo metu

Vilniaus  
universitetas

- Iš viso 1021 straipsnių.
- Naudojami du duomenų rinkiniai, kiekvienas duomenų rinkinys yra sudarytas iš skirtingų šaltinių (naujienų portalų).
- Duomenų rinkiniai anonimizuoti ir pavadinti: dataset 1 ir dataset 2.
- Pirmasis duomenų rinkinys: 604 ne propagandinių straipsnių ir 306 propagandinių straipsnių.
- Antrasis duomenų rinkinys: 59 ne propagandiniai straipsniai ir 52 propagandiniai straipsniai.



# Tyrimo planas

## Eksperimentas nr.1

### Duomenų rinkinio nr. 1 balansavimas

1. Be balansavimo
2. SMOTE
3. SMOTENN
4. SMOTOMEK
5. ADASYN

Mokymas ir testavimas naudojant  
10 fold cross validation

## Eksperimentas nr.2

### Duomenų rinkinio nr. 1 balansavimas

1. Be balansavimo
2. SMOTE
3. SMOTENN
4. SMOTOMEK
5. ADASYN

Apmokymas naudojant duomenų  
rinkinį nr. 1, testavimas naudojant  
duomenų rinkinį nr. 2

# Rezultatai (TF-IDF)

## Eksperimentas nr.1

Balancing technique	Metric	Model		
		LR	SVM	RF
No balancing	Mean ACC	0.72	0.72	0.75
	Std. dev. (ACC)	0.05	0.05	0.05
	Mean F1	0.57	0.58	0.62
	Std. dev. (F1)	0.08	0.08	0.06
SMOTE	Mean ACC	0.88	0.89	0.86
	Std. dev. (ACC)	0.04	0.03	0.03
	Mean F1	0.88	0.89	0.84
	Std. dev. (F1)	0.04	0.03	0.04
SMOTE-ENN	MEAN ACC	0.88	<b>0.94</b>	0.93
	Std. dev. (ACC)	0.07	0.06	0.06
	Mean F1	0.53	<b>0.85</b>	0.8
	Std. dev. (F1)	0.15	0.14	0.19
SMOTE-TOMEK	Mean ACC	0.88	0.89	0.85
	Std. dev. (ACC)	0.04	0.03	0.03
	Mean F1	0.88	0.89	0.87
	Std. dev. (F1)	0.04	0.03	0.03
ADASYN	Mean ACC	0.89	0.89	0.86
	Std. dev. (ACC)	0.03	0.03	0.03
	Mean F1	0.89	0.89	0.85
	Std. dev. (F1)	0.03	0.03	0.03

Balancing technique	Propaganda	Not propaganda
No balancing	306	604
SMOTE	604	604
SMOTE-ENN	238	35
SMOTE - TOMEK	604	604
ADASYN	634	604

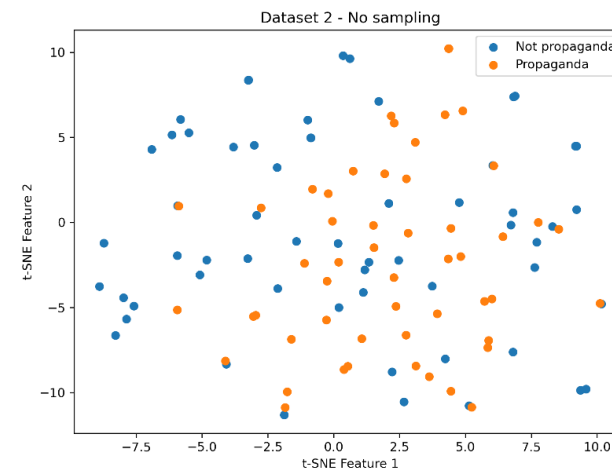
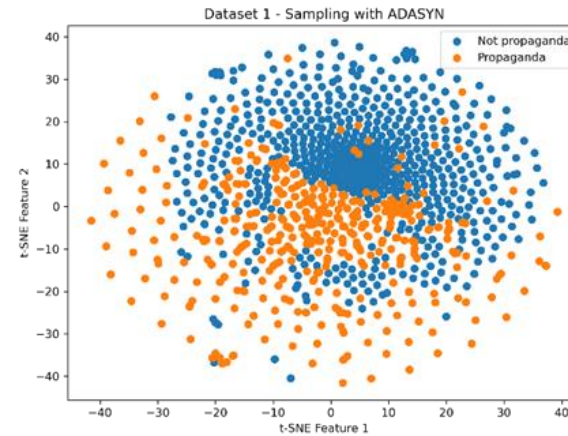
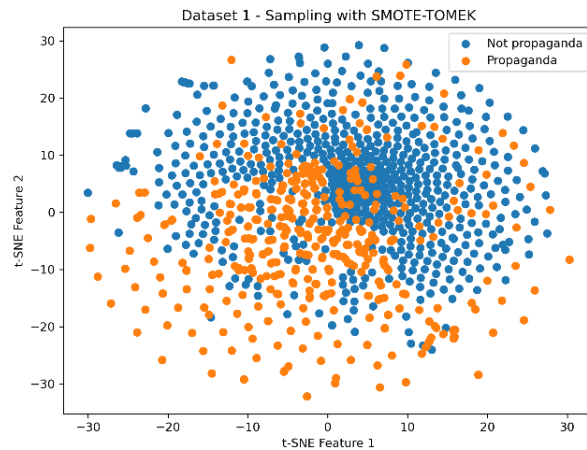
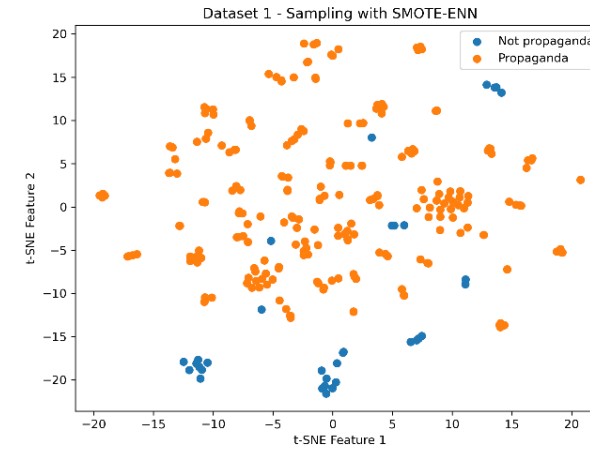
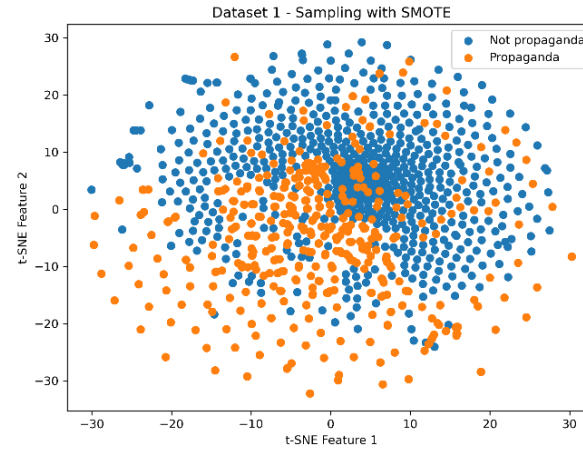
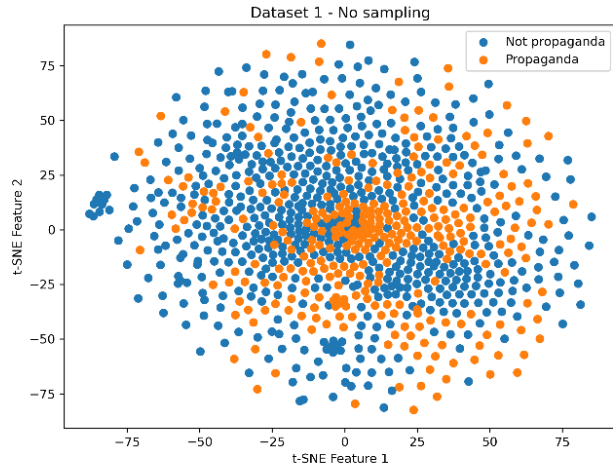
## Eksperimentas nr.2

Balancing technique	Metric	Model		
		LR	SVM	RF
No balancing	Accuracy	0.56	0.56	0.6
	F1 Score	0.45	0.45	0.51
SMOTE	Accuracy	0.7	0.59	0.63
	F1 Score	0.69	0.51	0.52
SMOTE-ENN	Accuracy	0.47	0.47	0.47
	F1 Score	0.32	0.32	0.32
SMOTE-TOMEK	Accuracy	0.7	0.59	0.61
	F1 Score	0.69	0.51	0.55
ADASYN	Accuracy	<b>0.72</b>	0.59	0.61
	F1 Score	<b>0.71</b>	0.51	0.55



# T-SNE vizualizacija (TF-IDF)

Vilnius  
universitetas



# Didelių (daugiakalbių) kalbos modelių įterpinių naudojimas. Mokymas ir testavimas su skirtingais duomenų rinkiniais (Tik Eksperimentas nr.2)

Vilniaus  
universitetas

## MiniLM-L12-v2

Balancing technique	Metric	Model		
		LR	SVM	RF
No balancing	Accuracy	0.77	<b>0.79</b>	0.72
	F1 Score	0.77	<b>0.79</b>	0.7
SMOTE	Accuracy	0.75	0.76	0.76
	F1 Score	0.75	0.76	0.75
SMOTE-ENN	Accuracy	0.63	0.66	0.66
	F1 Score	0.61	0.65	0.63
SMOTE-TOMEK	Accuracy	0.75	0.76	0.73
	F1 Score	0.75	0.76	0.72
ADASYN	Accuracy	0.76	0.76	0.7
	F1 Score	0.76	0.76	0.69

## Distiluse-base-multilingual-cased-v2

Balancing technique	Metric	Model		
		LR	SVM	RF
No balancing	Accuracy	0.7	<b>0.74</b>	0.71
	F1 Score	0.7	<b>0.74</b>	0.7
SMOTE	Accuracy	0.72	0.68	0.74
	F1 Score	0.72	0.68	0.75
SMOTE-ENN	Accuracy	0.46	0.46	0.46
	F1 Score	0.32	0.32	0.32
SMOTE-TOMEK	Accuracy	0.72	0.73	0.72
	F1 Score	0.72	0.73	0.72
ADASYN	Accuracy	0.72	0.72	0.71
	F1 Score	0.71	0.72	0.71

# Kito pusmečio darbo planas

- Sudalyvauti bent vienoje tarptautinėje konferencijoje.
- Tęsti duomenų rinkimą.
- Išlaikyti dalyko “Natūralios kalbos apdorojimas” egzaminą.
- Tęsti didelių kalbos modelių taikymo galimybių analizę propagandos aptikimui lietuvių kalbai.
- Išbandyti duomenų augmentacijų technikas skirtas tekstinių duomenų augmentavimui (žodžių, fragmentų: trynimas, įterpimas, apkeitimas vietomis, sinonimų naudojimas) anglų kalbos propagandos tekstynui SemEval, siekiant išsiaiškinti ar duomenų augmentavimas gali pagerinti tam tikrų propagandos technikų atpažinimo rezultatus.
- Pastebėjus, kad duomenų augmentacija pagerina propagandos technikų aptikimą anglų kalbos atveju, pradėti lietuviškų žodžių sinonimų žodyno sudarymą.
- Atlikti Meta duomenų rinkinio analizę (propagandos sklaidos socialiniuose tinkluose analizė).

# Kiti planai

- Suderintas laikas stažuotei Estijos Tartu universitete (Computational Social Science Group). Preliminarus numatytas laikas 2024m. Birželio mėn. ir spalio mėn.
- Stažuotės metu yra numatyta kartu su prof. Rajesh Sharma parengti bendrą publikaciją ir atlikti tyrimą, skirtą propagandos sklaidos socialiniuose tinkluose analizei (naudojant Meta duomenis).
- Pateikta projekto paraiška jaunųjų tyrėjų konkurse - tarpdisciplininio tyrimo (Informatika, Teisė, Komunikacija ir informacija) vykdymui. Projekto tema: *Mašininio mokymosi taikymo galimybės kovos su dezinformacija ir propaganda priemonėms: teisinis reglamentavimas ir krizių valdymas viešajame sektoriuje*



**Vilnius  
universitetas**

---

# Klausimai