Vilniaus universitetas

Mašininiu mokymusi grindžiami metodai apgaulingoms ir obfuskuotoms kenkėjiškoms programoms generuoti stiprinant kibernetinį saugumą

Doktorantūros pradžios/pabaigos metai: 2023–2027
Studijų metai: 2024–2025 antras pusmetis

Doktorantas: Juozas Dautartas
Vadovas: Dr. Viktor Medvedev

Studijų planas ir jo vykdymo suvestinė

Vilniaus universitetas

Studijų metai	Egzaminai [1]	
	Planas	Įvykdyta
I (2023/2024)	3	3
II (2024/2025)	**1**	**1**
III (2025/2026)		
IV (2026/2027)		
Iš viso:	4	4

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse [2]		Nacionalinėse [3]		Su citav. rodikliu [4]			Be citav. rodiklio [5]		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta [6]	Būklė [7]	Planas	Įvykdyta [6]	Būklė [7]
I (2023/2024)				1						
II (2024/2025)			**1**	**1**						
III (2025/2026)	1				1			1		
IV (2026/2027)	1				1					
Iš viso:	2		1	2	2			1		

Ataskaitinio pusmečio darbo planas ir jo vykdymo suvestinė

Vilniaus universitetas

Egzaminai 2024/2025 (II pusmetis)		
Planas	**Įvykdyta**	**Būklė**
-	-	-

Dalyvavimas konferencijose 2024/2025 (II pusmetis)		
Planas	**Įvykdyta**	**Konferencijos tipas**
-	-	-

Publikacijos 2024/2025 (II pusmetis)			
Planas	**Įvykdyta**	**Būklė**	**Publikacijos tipas**

Tyrimų objektas

Mašininio mokymosi algoritmai C2 sistemų agentams (kenkėjiškoms programoms) generuoti ir obfuskuoti.

Tikslas

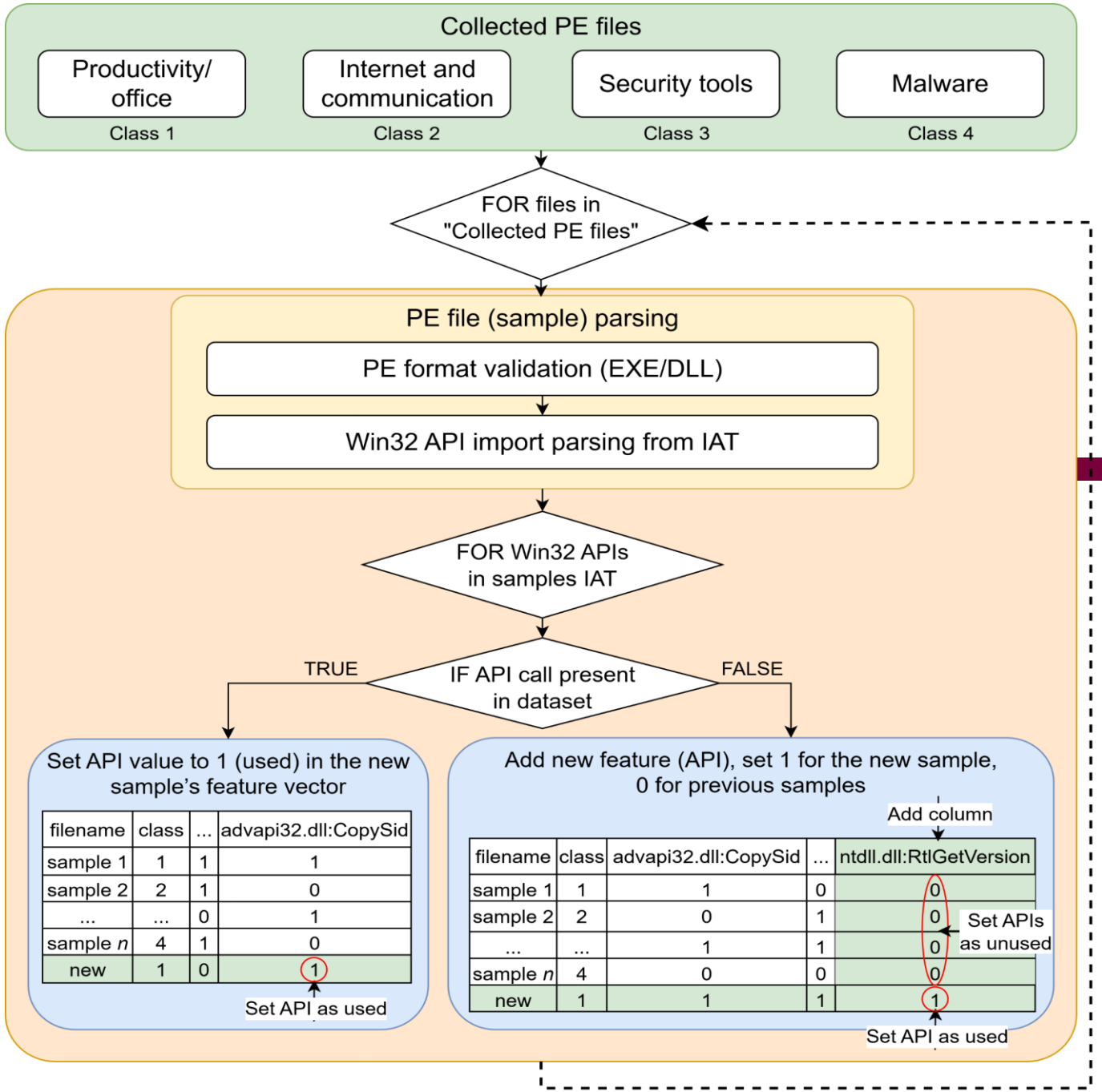
Pasiūlyti ir ištirti naują kenkėjiškos programinės įrangos generavimo metodą, skirtą veiksmingai išnaudoti mašininiu mokymusi pagrįstų kenkėjiškų programų aptikimo sistemų pažeidžiamumą, siekiant padidinti kibernetinių grėsmių atsparumą ir pagerinti bendrą jų aptikimo tikslumą ir patikimumą.

Uždaviniai

- Atlikti išsamią analitinės literatūros apžvalgą, siekiant nustatyti esamus mašininio mokymosi metodus ir būdus, skirtus kenkėjiškoms programoms aptikti, klasifikuoti ir generuoti kenkėjiškas programas.
- Analizuoti antivirusinių ir galinių įrenginių aptikimo ir reagavimo (angl. Endpoint Detection and Responce, EDR) sistemų kenkėjiškos programinės įrangos aptikimo mechanizmus, siekiant suprasti kenkėjiškos programinės įrangos klasifikavimo ypatybes.
- Įvertinti ir išskirti pagrindinius kenkėjiškos programinės įrangos požymius, turinčius įtakos jų aptikimui, siekiant pagerinti kenksmingų programų klasifikavimo tikslumą.
- Sukurti naują arba modifikuoti esamą kenkėjiškų programų (angl. Command and Control framework agents) generavimo metodą, kuriuo siekiama sukurti klaidinančius ir sunkiai aptinkamus C2 sistemų agentų pavyzdžius, galinčius klaidinti mašininiu mokymusi pagrįstus kenkėjiškų programų aptikimo modelius.
- Įvertinti sugeneruotus kenkėjiškų programų pavyzdžius, naudojant įvairius kenkėjiškų programų aptikimo sprendimus ir mašininiu mokymusi grindžiamas kibernetinio saugumo sistemas.
- Ištirti esamas ir pasiūlyti naują strategiją, kaip padidinti mašininio mokymosi modelių atsparumą priešiškoms atakoms.

Per pusmetį gauti rezultatai

- Dalyvauta doktorantų vasaros mokykloje Italijoje Toskanos regijone „Advanced Course On Data Science & Machine Learning“ 2025 m. birželio 9–13 d. (8 ETCS).
- Dalyvauta SANS instituto mokymuose Amsterdame „Red Teaming Tools – Developing Windows Impplants, Shellcode, Command and Control“ 2025 m. liepos 21–26 d.
- Baigti Maldev academy nuotoliniai mokymai „Malware development course“ apie kenkėjiškos programinės įrangos kūrimą.
- Sudarytas unikalus Windows PE failų (.exe, .dll) rinkinys, kuris skiriasi nuo tradicinių, plačiai naudojamų rinkinių. Vietoj vienos apibendrintos „gerybinės“ klasės, šiame rinkinyje gerybinės programos yra suskirstytos į kelias funkcines kategorijas (pvz., biuro įrankiai, interneto ir tinklo programos bei saugumo įrankiai), o kenkėjiškai programinei įrangai yra išskiriama viena klasė. (Duomenys publikuoti MIDAS nacionaliniame duomenų archyve: (MIDAS). 10.18279/MIDAS.265677)
- Pasiūlytas metodas ir atlikti eksperimentai, kurie yra skirti ištirti ir pademonstruoti naują kenkėjiškos programinės įrangos (angl. malware) aptikimo sistemų (ML klasifikatorių) klaidinimo būdą. Šie tyrimai yra aktualūs „Red Team“ operacijoms, siekiant testuoti ir stiprinti organizacijų apsaugos sistemas.
- Parengta publikacija recenzuojamam mokslinių konferencijų leidiniui IEEE Conference Proceedings. Iki š. m. spalio 3 d. publikacija bus įteikta konferencijai „2026 International Conference on Advances in Artificial Intelligence and Machine Learning (AAIML 2026)“.



Vilniaus universitetas

Dataset collection pipeline

Surinktas duomenų rinkinys

- Duomenys yra viešai publikuoti MIDAS duomenų bazėje: Nacionalinis atviros prieigos mokslinių tyrimų duomenų archyvas (MIDAS). 10.18279/MIDAS.265677

Table 3.1: Contents of the collected dataset

Class	Class name	Number of samples
1	Productivity / Office	216
2	Internet, Communication, Network	203
3	Security	214
4	Malware	219
Total:		**852**

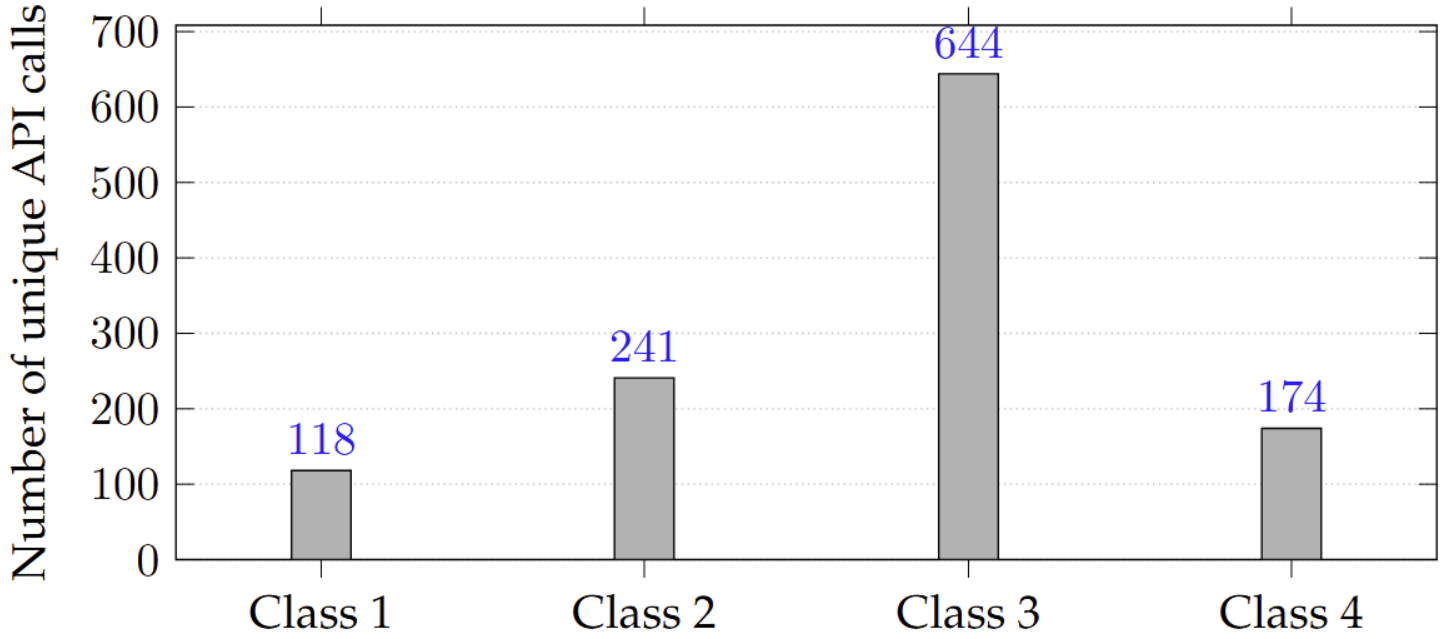
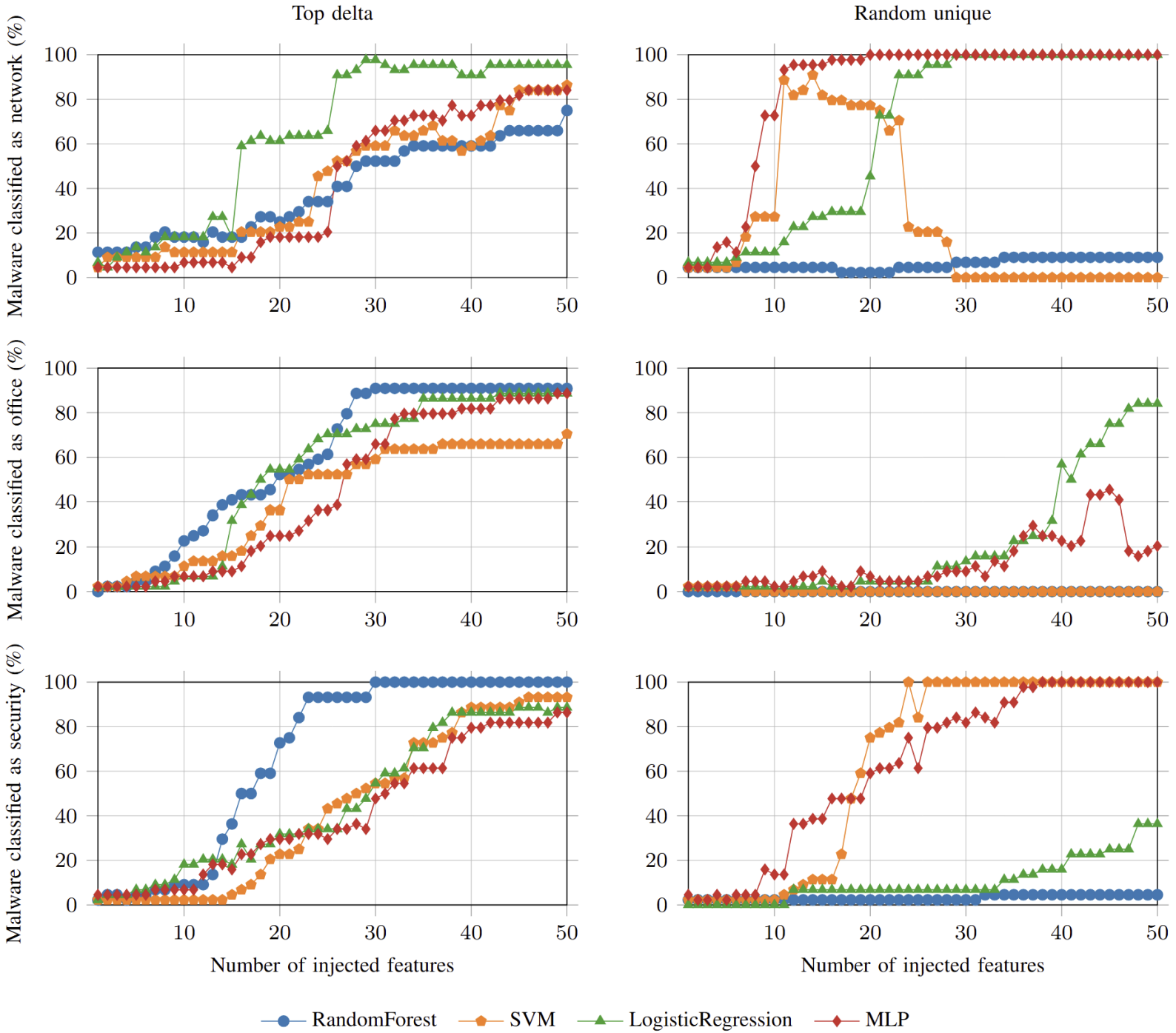
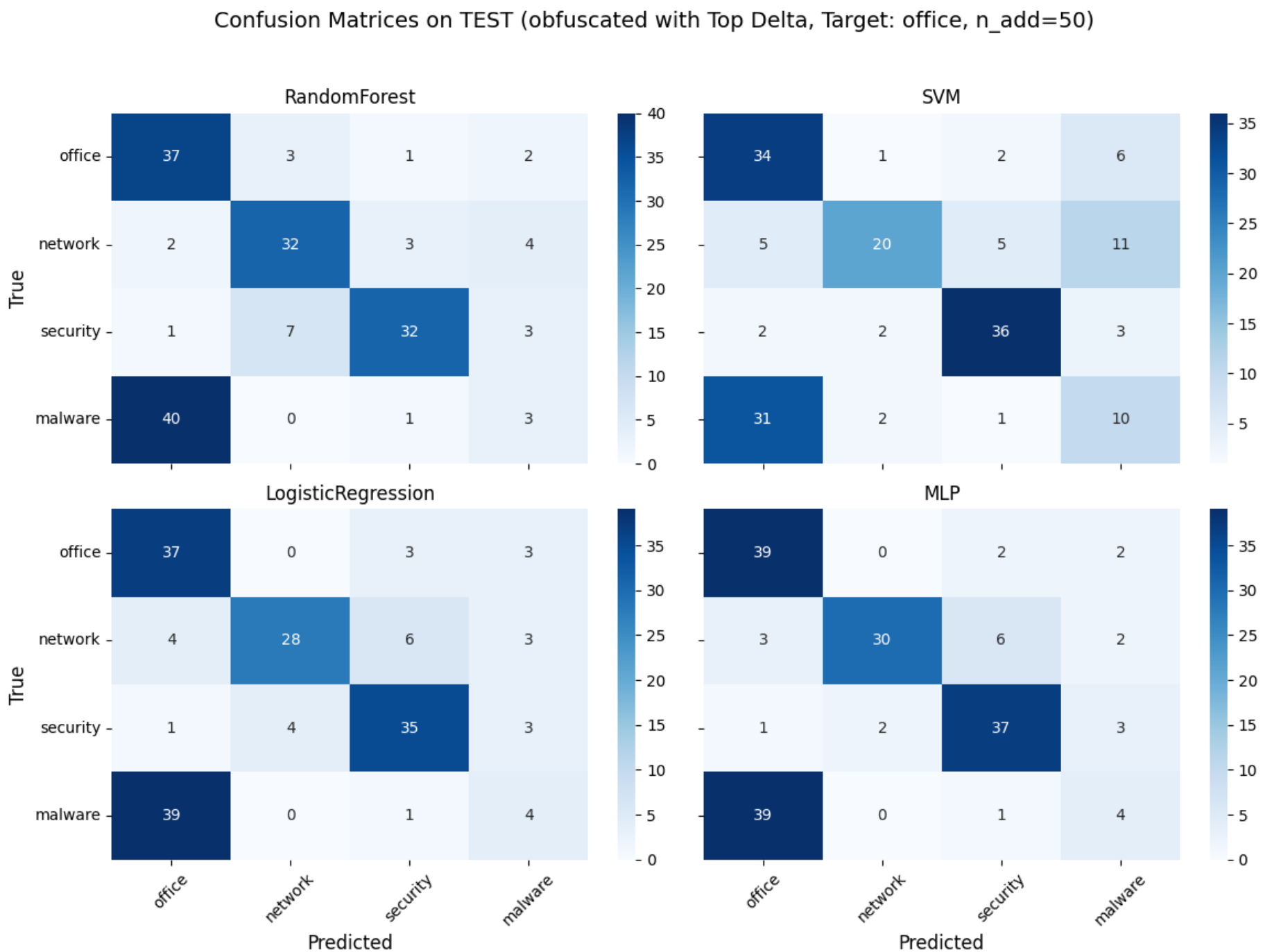


Figure 3.1: Number of unique API calls for each class.

CLASSIFICATION PERFORMANCE METRICS FOR ML DETECTORS (CLEAN TEST)

Model	Class	Precision	Recall	F1-Score	Support
RF	Office	0.92	0.86	0.89	43
	Network	0.73	0.78	0.75	41
	Security	0.86	0.74	0.80	43
	Malware	0.82	0.93	0.87	44
SVM	Office	0.81	0.79	0.80	43
	Network	0.80	0.49	0.61	41
	Security	0.82	0.84	0.83	43
	Malware	0.67	0.91	0.77	44
LR	Office	0.86	0.86	0.86	43
	Network	0.85	0.68	0.76	41
	Security	0.78	0.81	0.80	43
	Malware	0.82	0.93	0.87	44
MLP	Office	0.89	0.91	0.90	43
	Network	0.91	0.73	0.81	41
	Security	0.79	0.86	0.82	43
	Malware	0.85	0.91	0.88	44


Top delta
Random unique
Malware classified as network (%)
Malware classified as office (%)
Malware classified as security (%)
0
20
40
60
80
100
10
20
30
40
50
Number of injected features
RandomForest
SVM
LogisticRegression
MLP


Confusion Matrices on TEST (obfuscated with Top Delta, Target: office, n_add=50)
RandomForest
SVM
LogisticRegression
MLP
True
Predicted
office
network
security
malware
37 3 1 2
2 32 3 4
1 7 32 3
40 0 1 3
34 1 2 6
5 20 5 11
2 2 36 3
31 2 1 10
37 0 3 3
4 28 6 3
1 4 35 3
39 0 1 4
39 0 2 2
3 30 6 2
1 2 37 3
39 0 1 4

Kito pusmečio darbo planas

- Sudalyvauti tarptautinėje konferencijoje *„2026 International Conference on Advances in Artificial Intelligence and Machine Learning (AAIML 2026)“* Chuo University, Japonija, Tokijas; 2026 m. kovo 20–22 d.
- Tęsti pasiūlyto mašininio mokymosi metodo vystymą, skirtą tiksliniam kenkėjiškos programinės įrangos obfuskavimui.
- Praplėsti esamą duomenų rinkinį skirtą tiksliniam kenkėjiškos programinės įrangos obfuskavimui, įtraukiant daugiau kategorijų ir duomenų.
- Atlikti papildomus empirinius tyrimus, siekiant panaudoti ne tik mašininio mokymosi metodus, bet ir giliojo mokymosi metodus kenkėjiškos programinės įrangos obfuskavimui.
- Publikacijos rašymas mokslo leidiniui, turinčiam cituojamumo rodiklį Clarivate Analytics Web of Science duomenų bazėje.

KLAUSIMAI

Juozas Dautartas
juozas.dautartas@mif.stud.vu.lt