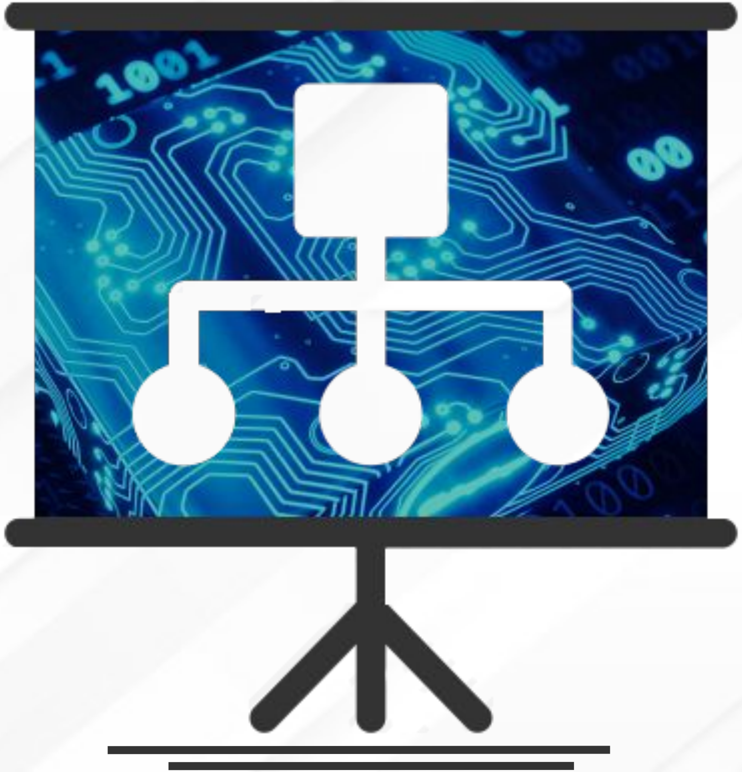
IEEE CSR

Cyber Security and Resilience

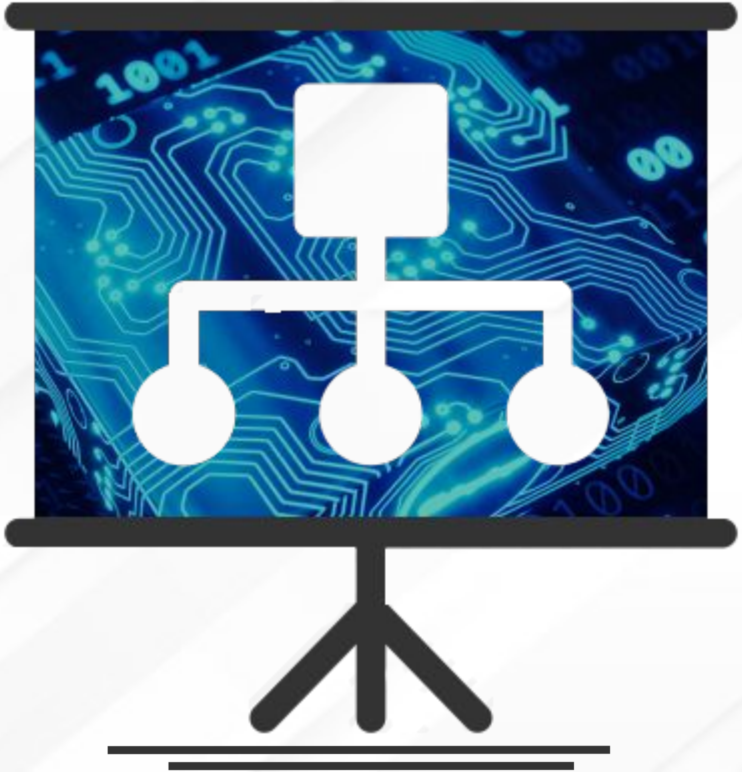
Cream Skimming the Underground

Identifying Relevant Information Points from Online Forums

Felipe Moreno-Vera, Mateus Nogueira, Cainã Figueiredo, Daniel S. Menasché, Miguel Bicudo, Ashton Woiwood, Enrico Lovat, Anton Kocheturov, and Leandro Pfleger de Aguiar



1. Introduction
2. Dataset
3. Empirical Findings
4. Cream Skimming Methodology
5. Experimental Results
6. Conclusion



1. Introduction
2. Dataset
3. Empirical Findings
4. Cream Skimming Methodology
5. Experimental Results
6. Conclusion

Motivation

Exploitation of vulnerabilities in the wild poses significant threat to Internet ecosystem

Need for early detection of weaponization and tentative exploitation

Context

Underground hacking forums: privileged info

- **PoC:** Tutorials and demos
- **weaponization:** development of exploits
- **exploitation:** tentative use of those exploits in the wild

Monitoring these forums allows for tracking

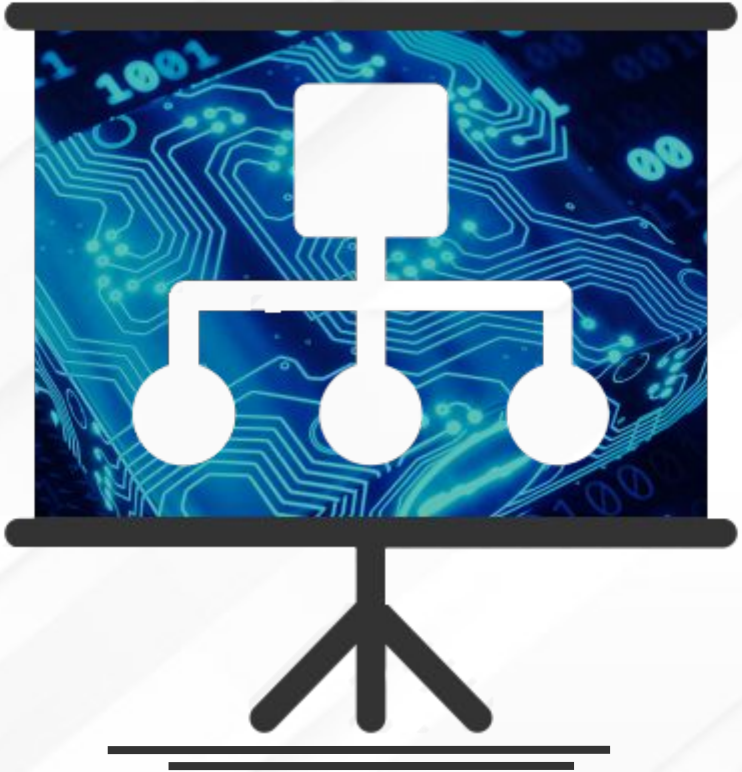
- exploit prices
- their usage
- demand and targets

Our Key Contributions

- Methodology to analyze, explore and identify significant information, and classify discussions.
- **Exploitation analysis:** How are CVEs used in the wild?
    - Monetary profits in hacking communities
    - Delays between publication of CVEs and discussion
- **Threat classifier:** How to know what is discussed?
    - Decision-based classifier for assessing threat maturity
    - Interpretation of results from decision trees

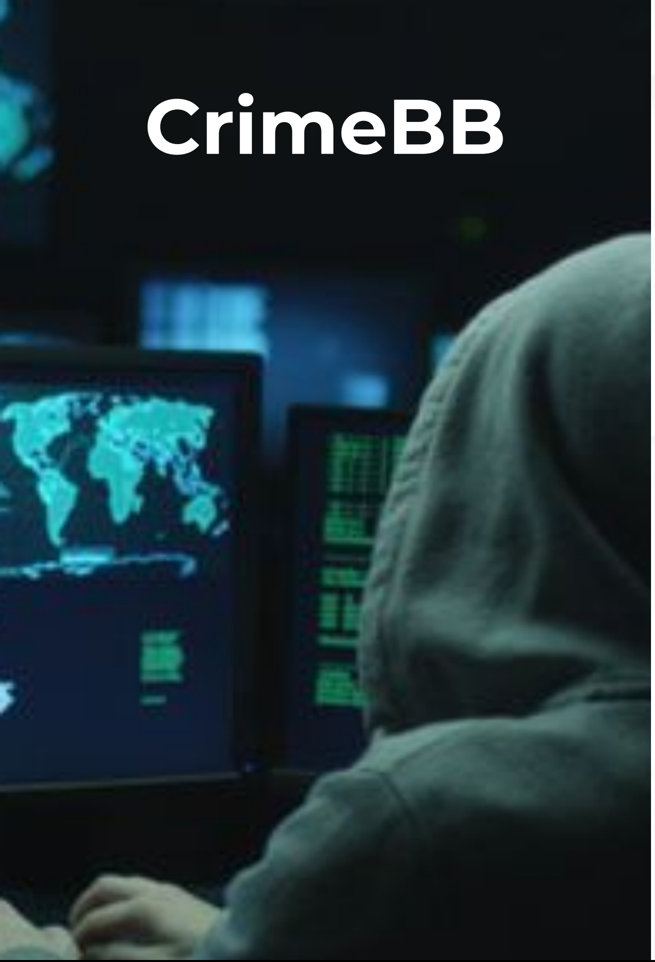
Previous Work

- Previous work focus on the **weaponization**
  - … but **exploitation** in the wild has received much less attention
- **EPSS** and Expected Exploitability aim at **exploitability** in the wild
  - … but they **do not cite CrimeBB**



1. Introduction
2. Dataset
3. Cream Skimming Methodology
4. Empirical Findings
5. Experimental Results
6. Conclusion

CrimeBB



Dataset Description

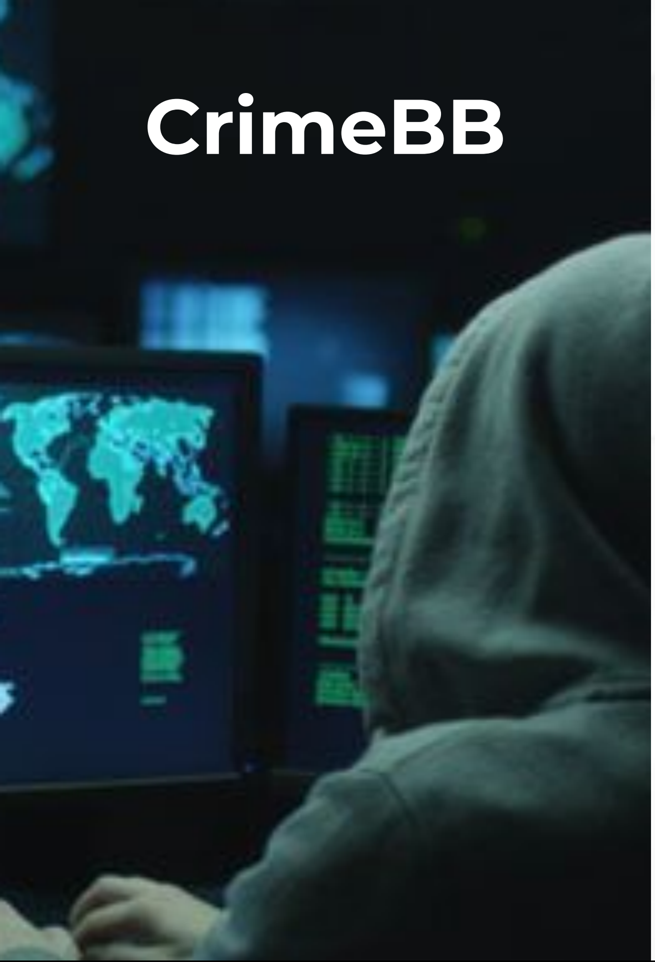
Made available by Cambridge Cybercrime Centre

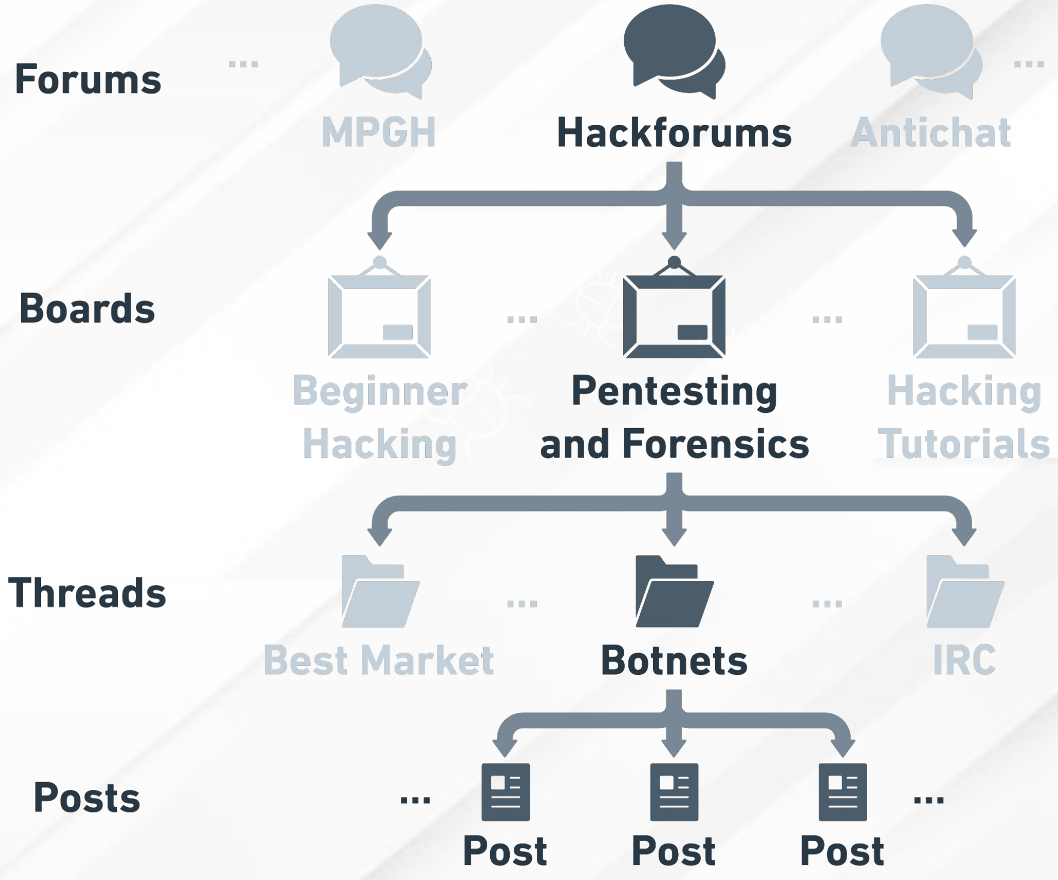
Contains data scraped from multiple underground forums (16 studied)

Organized in forums, boards, threads and posts

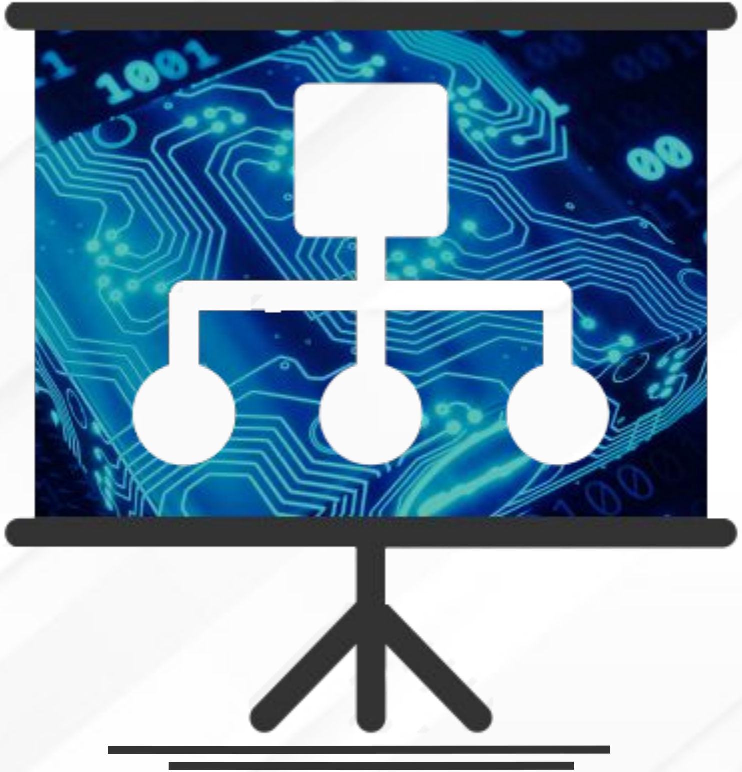
Provide about 54,512,094 lines of textual information.

CrimeBB

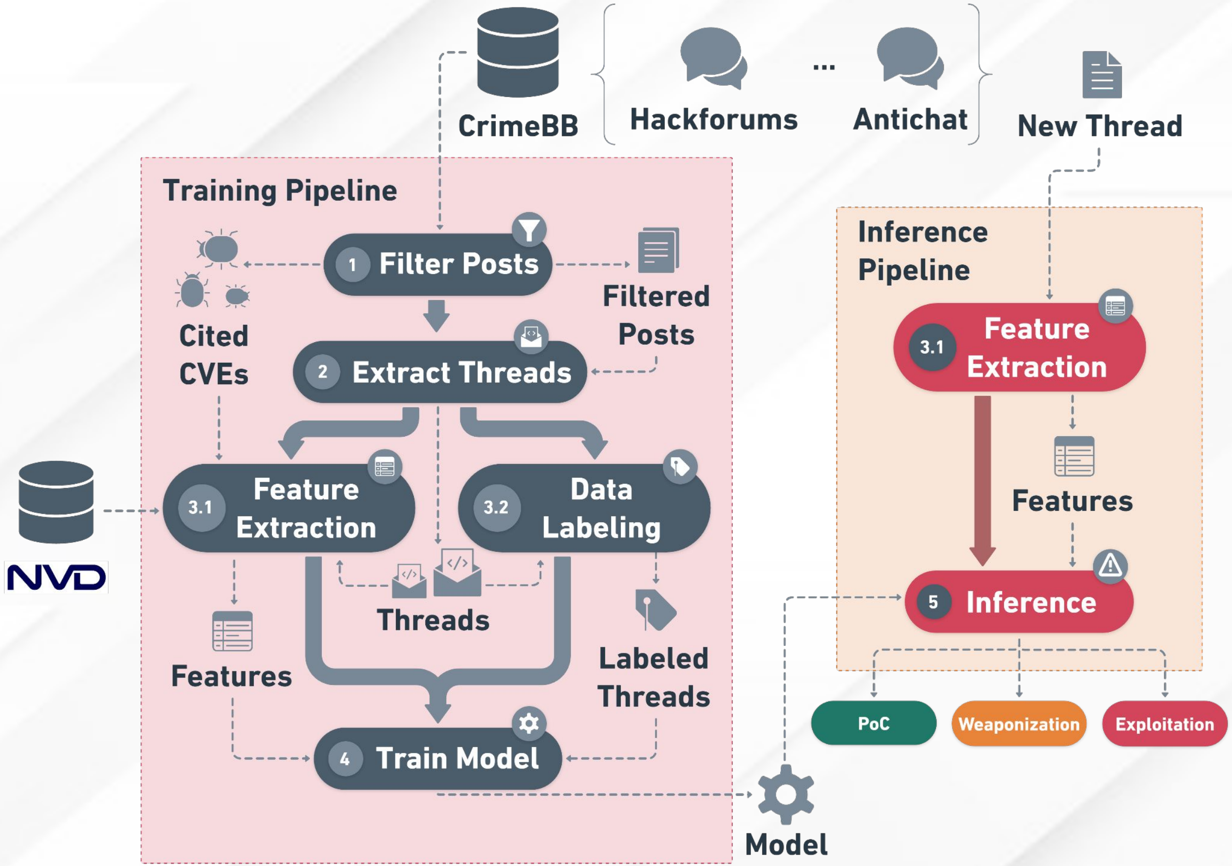




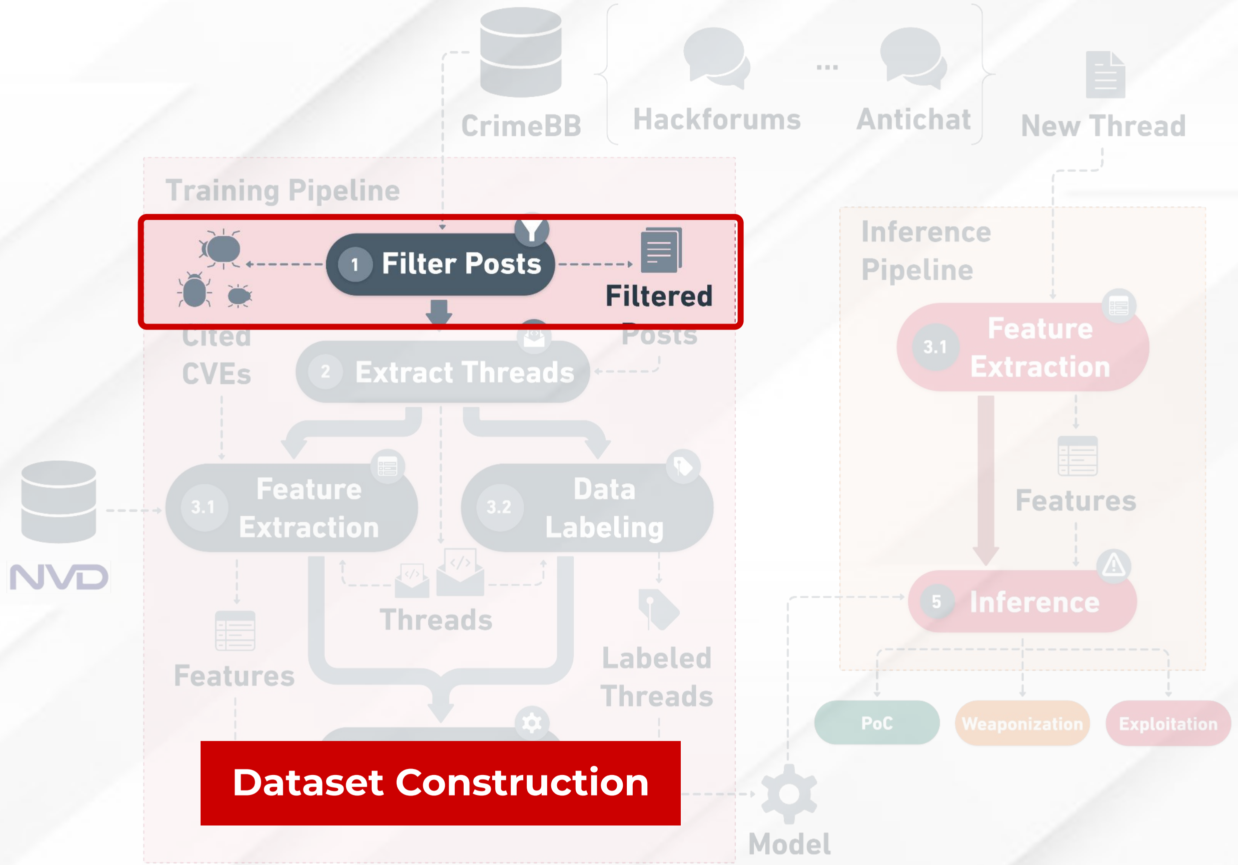
Forums
...
MPGH
Hackforums
Antichat
...
Boards
Beginner Hacking
...
Pentesting and Forensics
...
Hacking Tutorials
Threads
Best Market
...
Botnets
...
IRC
Posts
...
Post
Post
Post
...



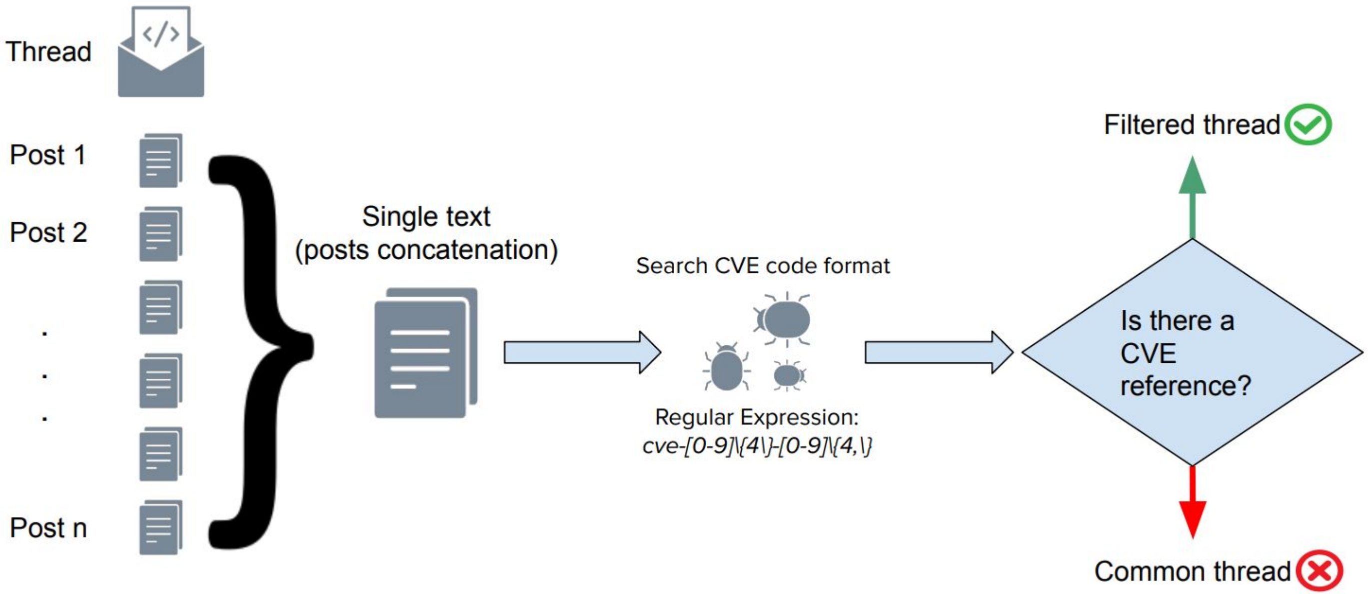
Pipeline


CrimeBB
Hackforums
...
Antichat
New Thread
Training Pipeline
1 Filter Posts
Cited CVEs
Filtered Posts
2 Extract Threads
3.1 Feature Extraction
3.2 Data Labeling
NVD
Threads
Features
Labeled Threads
4 Train Model
Model
Inference Pipeline
3.1 Feature Extraction
Features
5 Inference
PoC
Weaponization
Exploitation

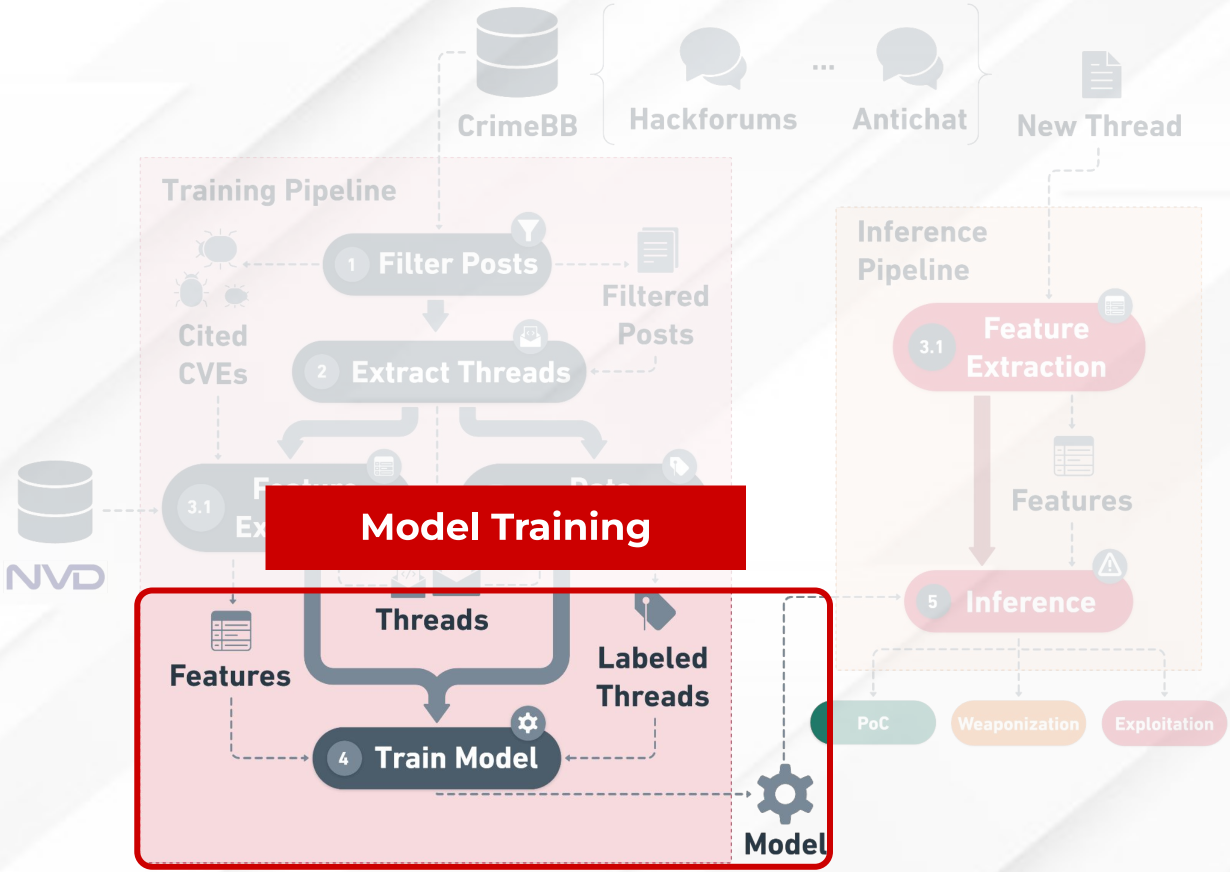
Pipeline


CrimeBB
Hackforums
...
Antichat
New Thread
Training Pipeline
1 Filter Posts
Filtered Posts
Cited CVEs
2 Extract Threads
3.1 Feature Extraction
3.2 Data Labeling
NVD
Threads
Features
Labeled Threads
Inference Pipeline
3.1 Feature Extraction
Features
5 Inference
PoC
Weaponization
Exploitation
Model
Dataset Construction

Data Preparation - Filtering threads



Pipeline


CrimeBB
Hackforums
...
Antichat
New Thread
Training Pipeline
1 Filter Posts
Cited CVEs
Filtered Posts
2 Extract Threads
NVD
3.1
Model Training
Threads
Features
Labeled Threads
4 Train Model
Model
Inference Pipeline
3.1 Feature Extraction
Features
5 Inference
PoC
Weaponization
Exploitation

Data Preparation - Manual Labeling

- HackForums: **3,037 posts** (in **1,162 threads**) cite a CVE
- **Manually labeled threads** by the posts content: **1,067**
- Hackforums: **2,666 posts** (in **1,042 threads**) were labeled
- A total of **8,915** (**969 unique**) CVE codes were found
- In this study, we **focus** only in **Weaponization**, **PoC**, and **Exploitation**

Label	Threads labeled	Threads citing CVE	Posts citing CVE
Weaponization	410	397	891
PoC	247	244	861
Others	195	192	520
Exploitation	107	102	232
Warning	55	55	67
Help	43	42	60
Scam	10	10	35
Total	1,067	1,042	2,666

Data Preparation - Manual Labeling

- HackForums: **3,037 posts** (in **1,162 threads**) cite a CVE
- **Manually labeled threads** by the posts content: **1,067**
- Hackforums: **2,666 posts** (in **1,042 threads**) were labeled
- A total of **8,915** (**969 unique**) CVE codes were found
- In this study, we **focus** only in **Weaponization**, **PoC**, and **Exploitation**

Label	Threads labeled	Threads citing CVE	Posts citing CVE
Weaponization	410	397	891
PoC	247	244	861
Others	195	192	520
Exploitation	107	102	232
Warning	55	55	67
Help	43	42	60
Scam	10	10	35
Total	1,067	1,042	2,666

Data Preparation - Feature Extraction

Corpus	
Document 1	I like cats
Document 2	cats are the best, they are awesome
Document 3	also dogs are nice

Document-Term Matrix

Words	I	like	cats	are	the	best	they	awesome	also	dogs	nice
Document 1	1	1	1	0	0	0	0	0	0	0	0
Document 2	0	0	1	2	1	1	1	1	0	0	0
Document 3	0	0	0	1	0	0	0	0	1	1	1

Doc2Vec

Vectors	x1	x2	x3	x4	x5	x6	x7	...
Document 1	0.35	0.86	1.82	3.48	1.05	10.15	8.63	...
Document 2	0.84	0.45	3.45	4.49	2.64	2.87	13.97	...
Document 3	0.39	1.0	0.98	7.92	5.14	6.19	20.98	...

Bag-Of-Words (1-2-gram)

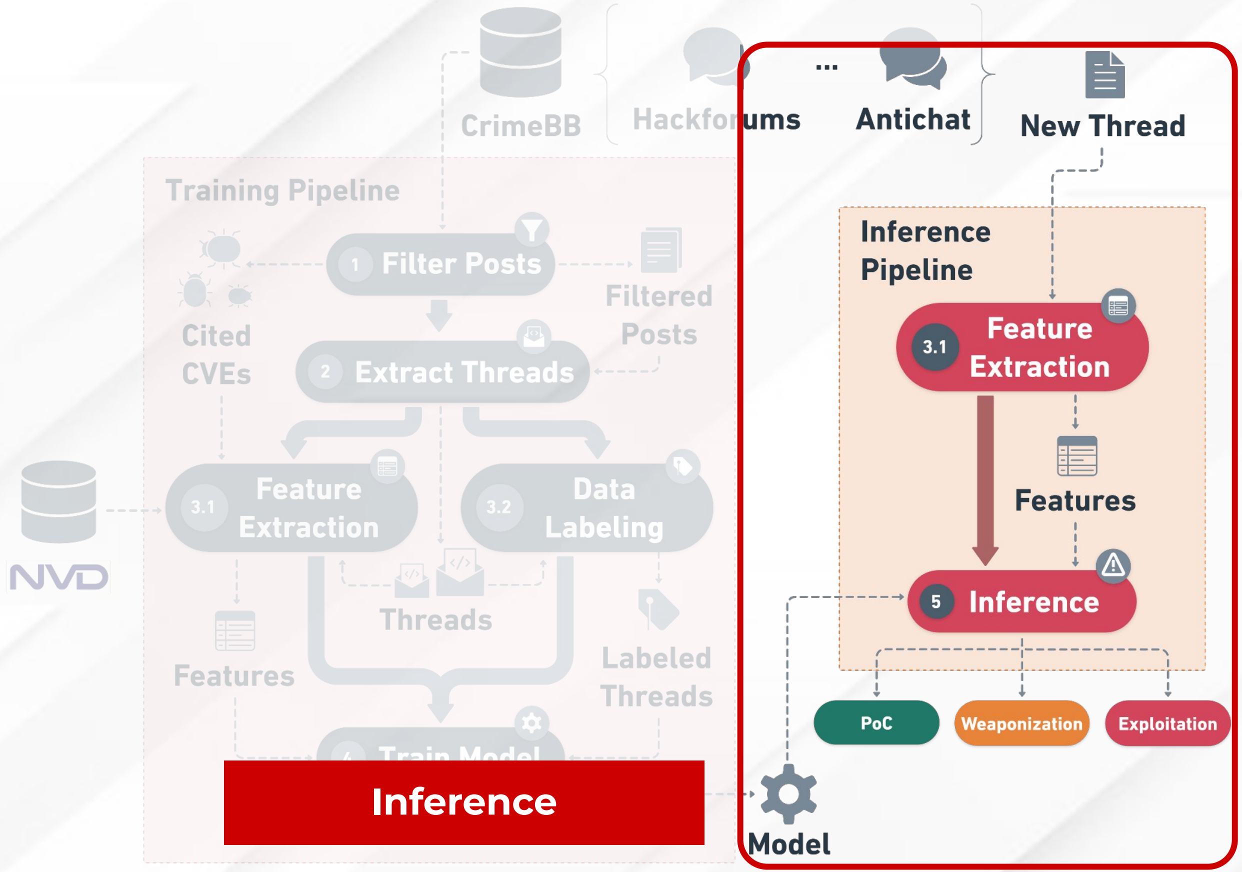
Words	I	like	cats	I like	like cats	are	the	best	they	awesome	cats are	the best	they are	...
Document 1	1	1	1	1	1	0	0	0	0	0	0	0	0	...
Document 2	0	0	1	0	0	2	1	1	1	1	1	1	1	...
Document 3	0	0	0	0	0	1	0	0	0	0	0	0	0	...

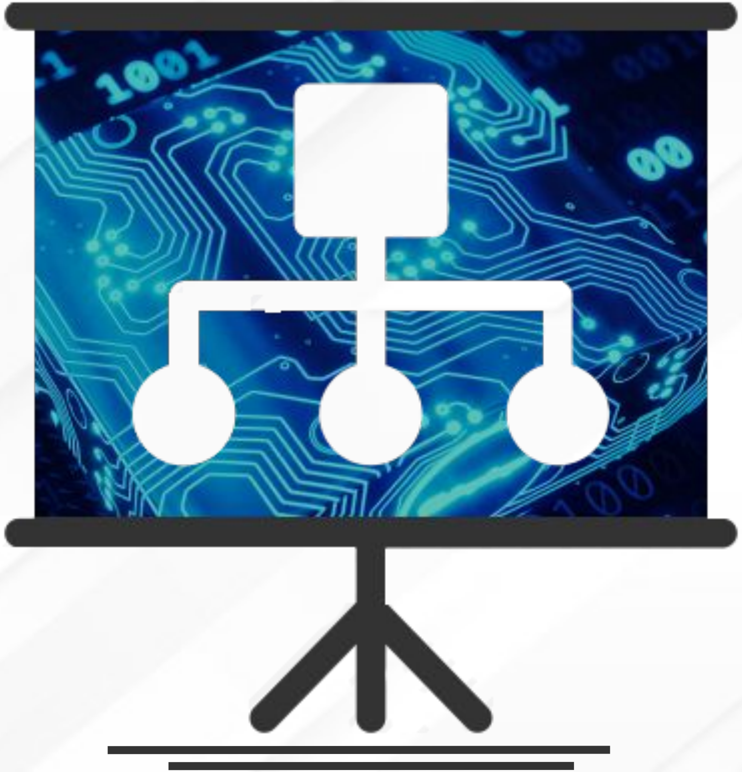
TF-IDF (1-2-gram)

Words	I	like	cats	I like	like cats	are	the	best	they	awesome	cats are	the best	they are	...
Document 1	0	0	0.47	0.62	0.62	0	0	0	0	0	0	0	0	...
Document 2	0	0	0.33	0	0	0.56	0.43	0	0	0	0.43	0.43	0.13	...
Document 3	0	0	0	0	0	0.35	0	0	0	0	0	0	0	...

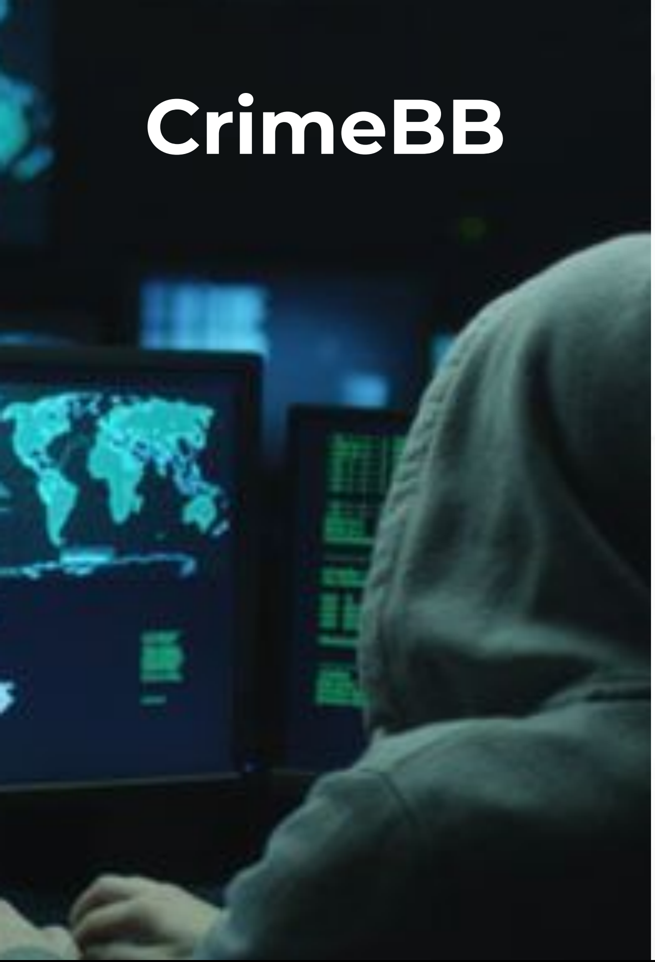
We also apply standard NLP pre-processing techniques, e.g., filtering stopwords and punctuation

Pipeline


CrimeBB
Hackforums
...
Antichat
New Thread
Training Pipeline
1 Filter Posts
Cited CVEs
Filtered Posts
2 Extract Threads
NVD
3.1 Feature Extraction
3.2 Data Labeling
Threads
Features
Labeled Threads
Train Model
Inference
Inference Pipeline
3.1 Feature Extraction
Features
5 Inference
PoC
Weaponization
Exploitation
Model



1. Introduction
2. Dataset
3. Cream Skimming Methodology
4. Empirical Findings
5. Experimental Results
6. Conclusion

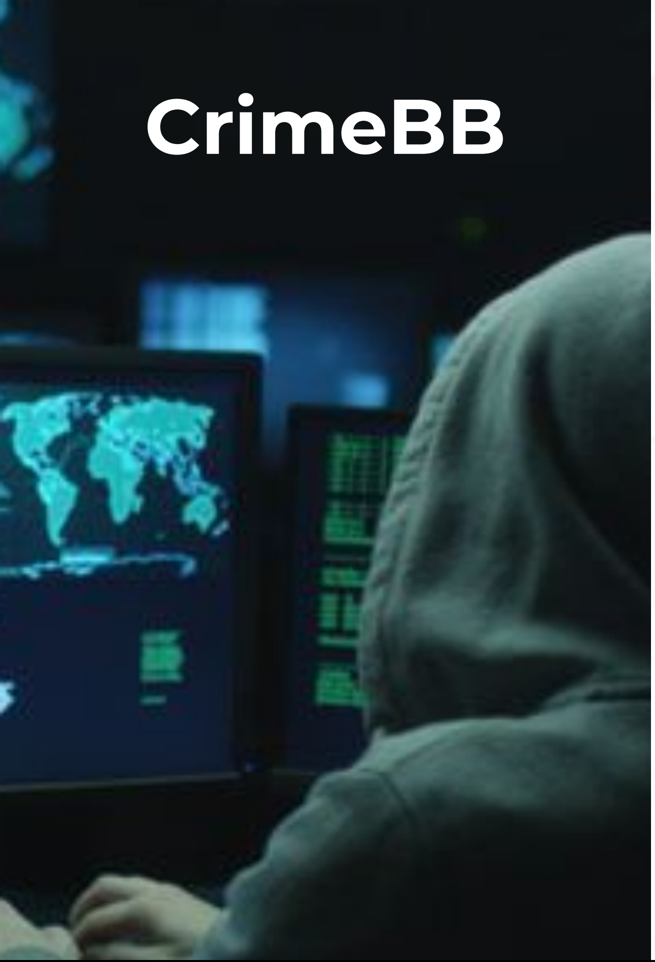


CrimeBB

Forum	#Users	#Boards	#Threads	#Posts
Hackforums	630,331	177	3,966,270	41,571,269
MPGH	478,120	715	763,231	9,363,422
Antichat	79,769	60	242,064	2,449,404
Offensive Community	11,800	58	119,228	161,492
DREADditevelidot	44,631	382	74,098	294,596
RaidForums	29,038	73	33,240	214,856
Runion	16,719	19	16,792	240,632
Safe Sky Hacks	7,433	44	12,956	27,018
The-Hub	8,243	62	11,274	88,753
Torum	3,813	11	4,328	28,485
Kernelmode Forum	1,644	11	3,438	25,825
Germany Ruvvy	2,206	42	2,845	20,185
Garage4hackers	880	31	2,096	7,697
Greysec	728	25	1,630	9,228
Stresser Forum	777	16	702	7,069
Envoy Forum	362	76	454	2,163
Total	1,316,494	1,802	5,254,646	54,512,094

CrimeBB general statistics.

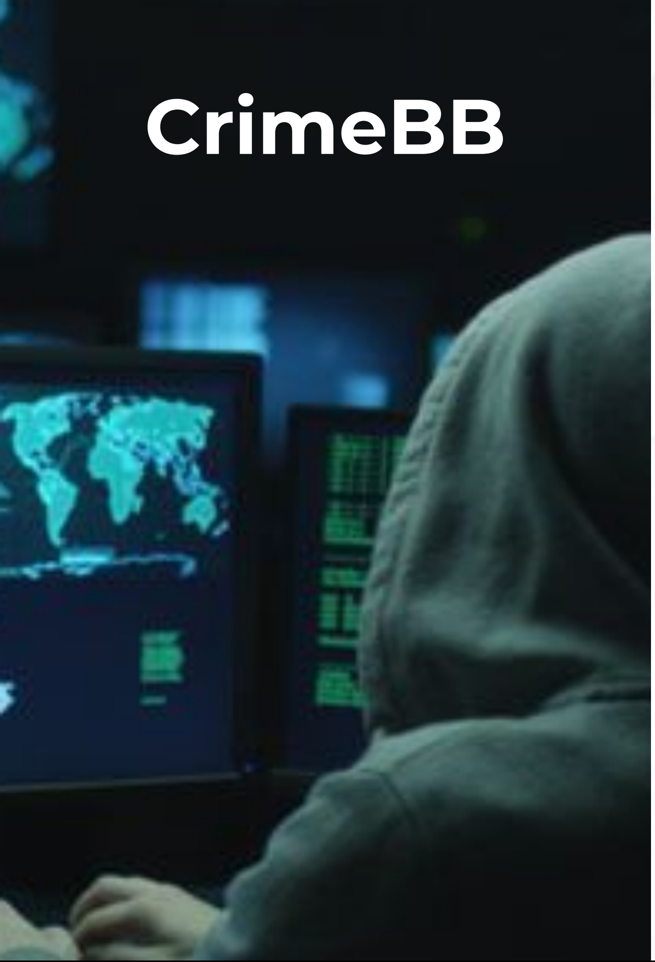
CrimeBB



Forum	#Users	#Boards	#Threads	#Posts
Hackforums	630,331	177	3,966,270	41,571,269
MPGH	478,120	715	763,231	9,363,422
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Safe Sky Hacks	7,433	44	12,956	27,018
The-Hub	8,243	62	11,274	88,753
Torum	3,813	11	4,328	28,485
Kernelmode Forum	1,644	11	3,438	25,825
Germany Ruvvy	2,206	42	2,845	20,185
Garage4hackers	880	31	2,096	7,697
Greysec	728	25	1,630	9,228
Stresser Forum	777	16	702	7,069
Envoy Forum	362	76	454	2,163
Total	1,316,494	1,802	5,254,646	54,512,094

HackForums has about **76% of posts** and **75% of threads** in all CrimeBB dataset.
Also, have about **65.5%** of CVE code citations.

Table 1
CrimeBB general statistics.



CrimeBB

Forum	#Users	#Boards	#Threads	#Posts
Hackforums	630,331	177	3,966,270	41,571,269
MPGH	478,120	715	763,231	9,363,422
Antichat	79,769	60	242,064	2,449,404
Offensive Community	11,800	58	119,228	161,492
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]596
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]856
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]632
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]018
The-Hub	8,243	62	11,274	88,753
Torum	3,813	11	4,328	28,485
Kernelmode Forum	1,644	11	3,438	25,825
Germany Ruvvy	2,206	42	2,845	20,185
Garage4hackers	880	31	2,096	7,697
Greysec	728	25	1,630	9,228
Stresser Forum	777	16	702	7,069
Envoy Forum	362	76	454	2,163
Total	1,316,494	1,802	5,254,646	54,512,094

Antichat has about **4.49% of posts**, **4.6% of threads**, and the remaining **34.4%** of CVE citations.

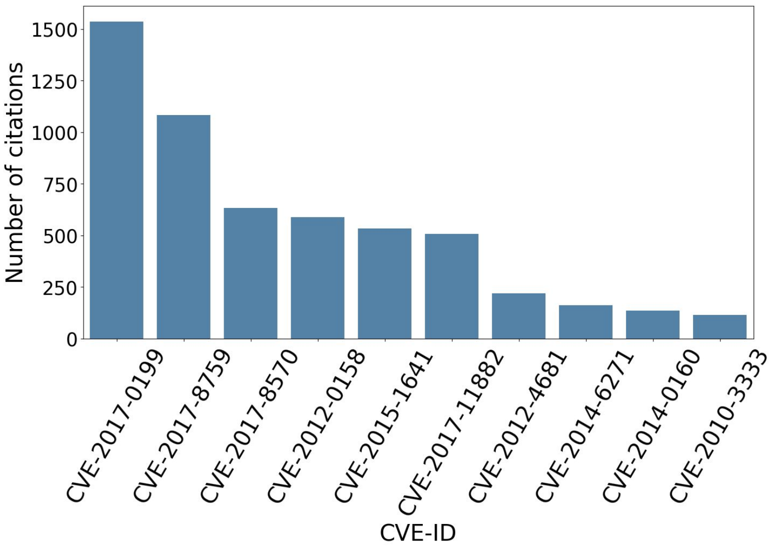
Table 1
CrimeBB general statistics.

Top 10 CVEs Cited in Posts

CVE-2017-0199 affects Microsoft Office: remote execution of arbitrary code

top CVEs cover a wide time horizon: from 2010 to 2017

top 3 most cited CVEs are most recent

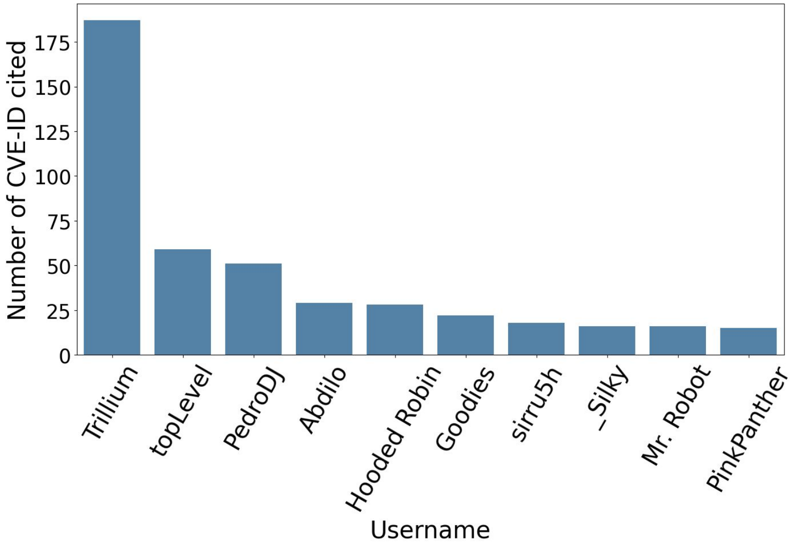


Top 10 Users Citing CVEs in Posts

*User **Trillium** mentioned the largest number of CVEs across various posts*

Seems to be advertising exploits

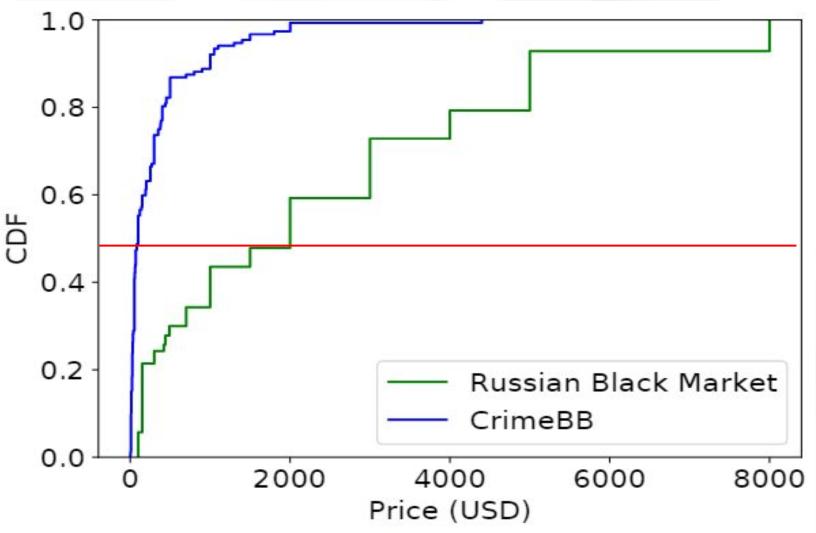
We found some names repeated across forums, but hard to match



CrimeBB vs Russian Market

How do prices of artifacts sold at CrimeBB compare against Russian Market?

- *Russian Market (ACM CCS, Luca Allodi)*
- ***Prices at Russian Market are larger***
    - *Median value at CrimeBB < 100 USD and > 2000 USD at Russian Market*
- ***Why?***
    - *Russian Market is closed market*
    - *Admission control to enter the market*
    - ***Artifacts sold at Russian market are more valuable***



CDF of hacking tools prices

CrimeBB vs Russian Market

How do delays at CrimeBB compare against Russian Market?

Delay definition:

CrimeBB

date post at CrimeBB - date NVD published CVE

Russian Market

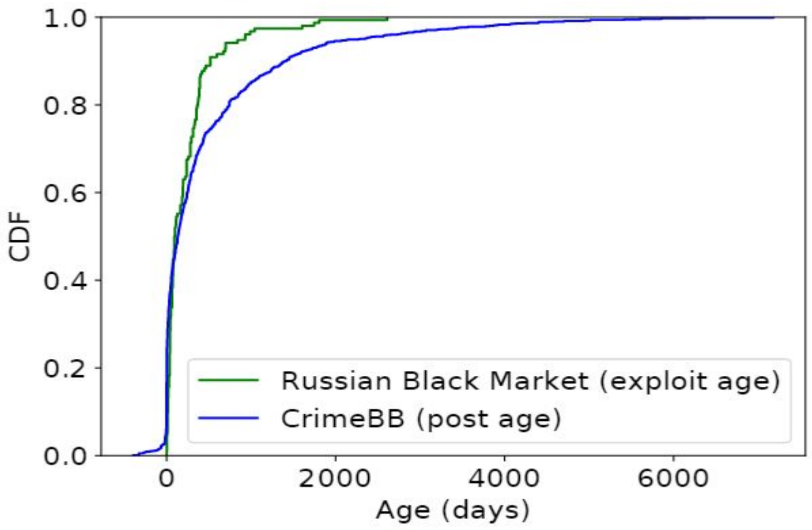
date exploit publication at market - date NVD published CVE

Delays at CrimeBB are larger: why?

Russian Market is closed market

Exploits are published at Russian market and activity ceases

***At CrimeBB, continuous discussion** of exploitation strategies*



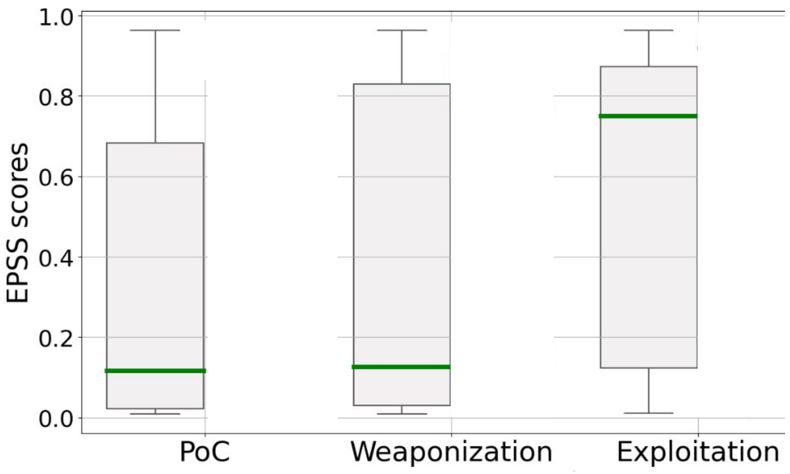
CDF of the difference in days between CrimeBB citation and NVD publish date

EPSS (Exploit Prediction Scoring System)

How risk depends on maturity?

EPSS: *probability that vuln will be exploited in the next 30 days*

Finding 1) *Risk grows with maturity*



EPSS (Exploit Prediction Scoring System)

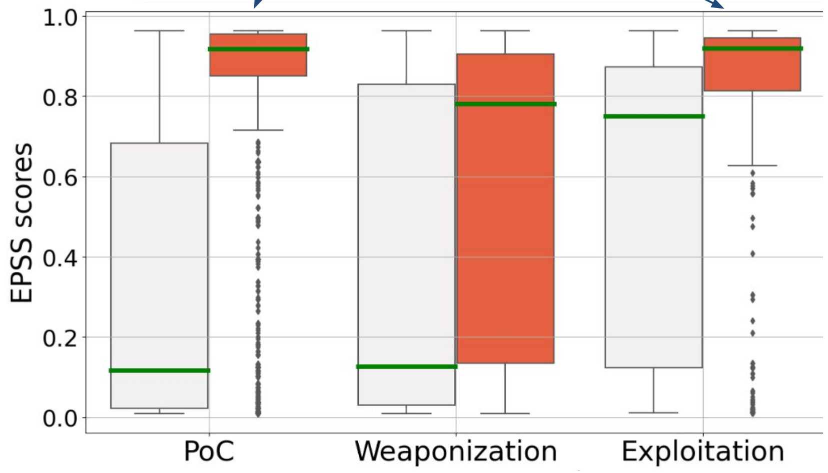
How risk depends on maturity?

EPSS: *probability that vuln will be exploited in the next 30 days*

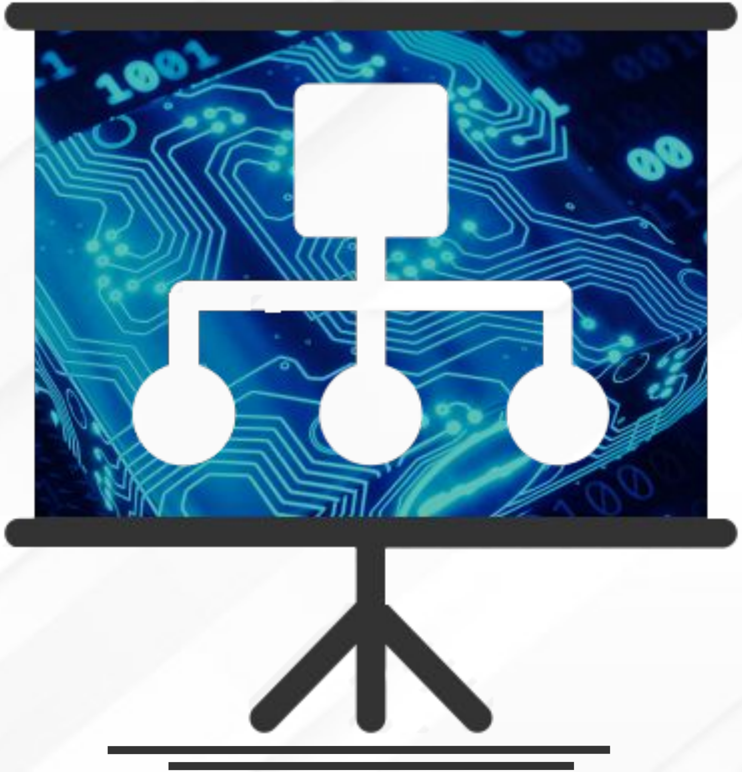
Finding 1) *Risk grows with maturity*

Finding 2) *Most cited vulns are riskier*

more citations = higher risk



in gray, single sample per CVE identifier (does NOT account for # citations)
in red, one sample per CVE citation (accounts for # citations)



Experimental Setting

- Train and test split
  - 75% and 25%, respectively
  - Stratified split in order to preserve the original distribution
- Evaluation metrics
  - Accuracy, Precision, Recall, and F1-score
- Hyperparameters tuning
  - Grid Search
  - 5-fold Stratified Cross-Validation on the training set

	Text encoding	Target classes	Accuracy	Precision	Recall	F1
DT	BoW	PoC, Weaponization, Exploitation	0.71	0.71	0.72	0.70
DT	TF-IDF	PoC, Weaponization, Exploitation	0.73	0.73	0.74	0.72
DT	doc2vec	PoC, Weaponization, Exploitation	0.74	0.74	0.74	0.73
DT	BoW	Exploitation vs Non-exploitation	0.85	0.86	0.85	0.85
DT	TF-IDF	Exploitation vs Non-exploitation	0.91	0.91	0.91	0.91
DT	doc2vec	Exploitation vs Non-exploitation	**0.92**	**0.93**	**0.92**	**0.92**
DT	BoW	PoC vs Non-PoC	0.75	0.75	0.75	0.75
DT	TF-IDF	PoC vs Non-PoC	0.77	0.78	0.77	0.77
DT	doc2vec	PoC vs Non-PoC	0.70	0.71	0.70	0.70
DT	BoW	Weaponization vs Non-weapon.	0.68	0.68	0.68	0.68
DT	TF-IDF	Weaponization vs Non-weapon.	0.63	0.64	0.63	0.62
DT	doc2vec	Weaponization vs Non-weapon.	0.59	0.59	0.59	0.59

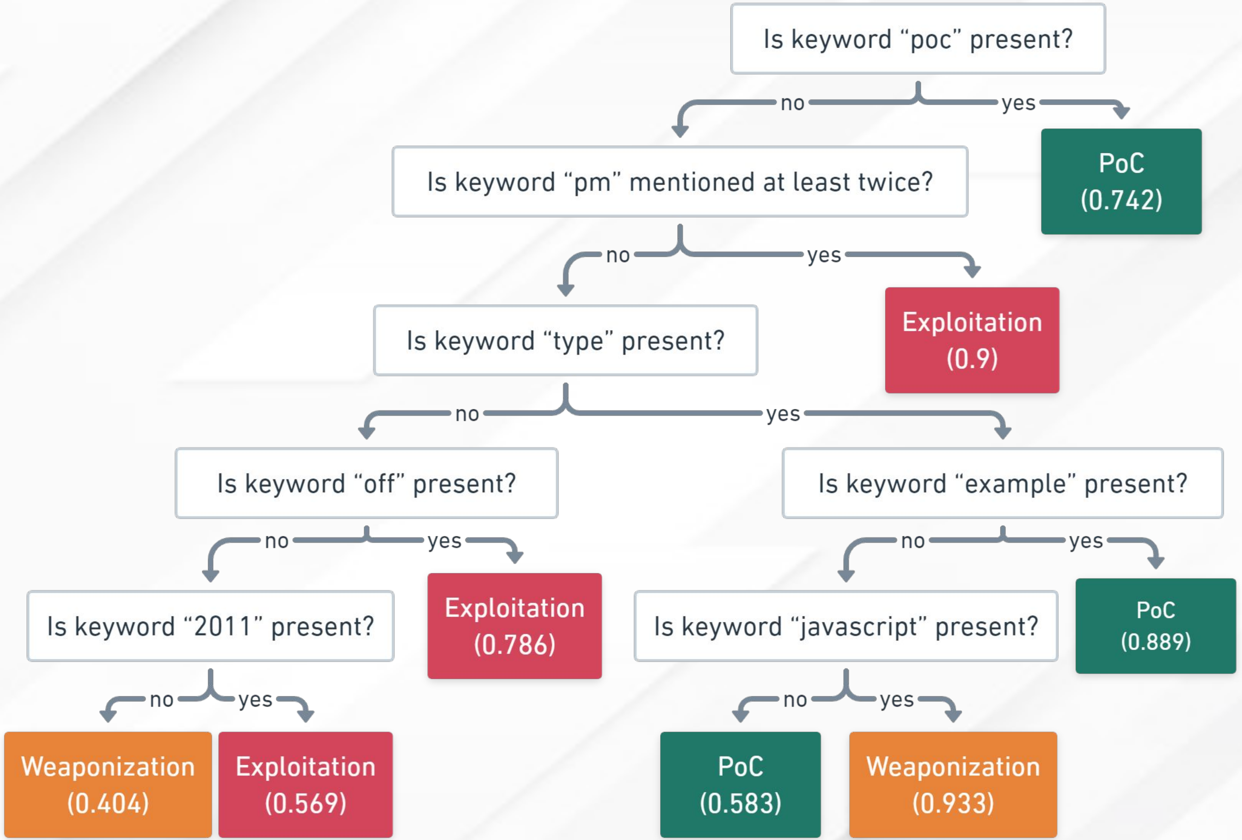
Decision tree performance: easier to distinguish exploitation from rest

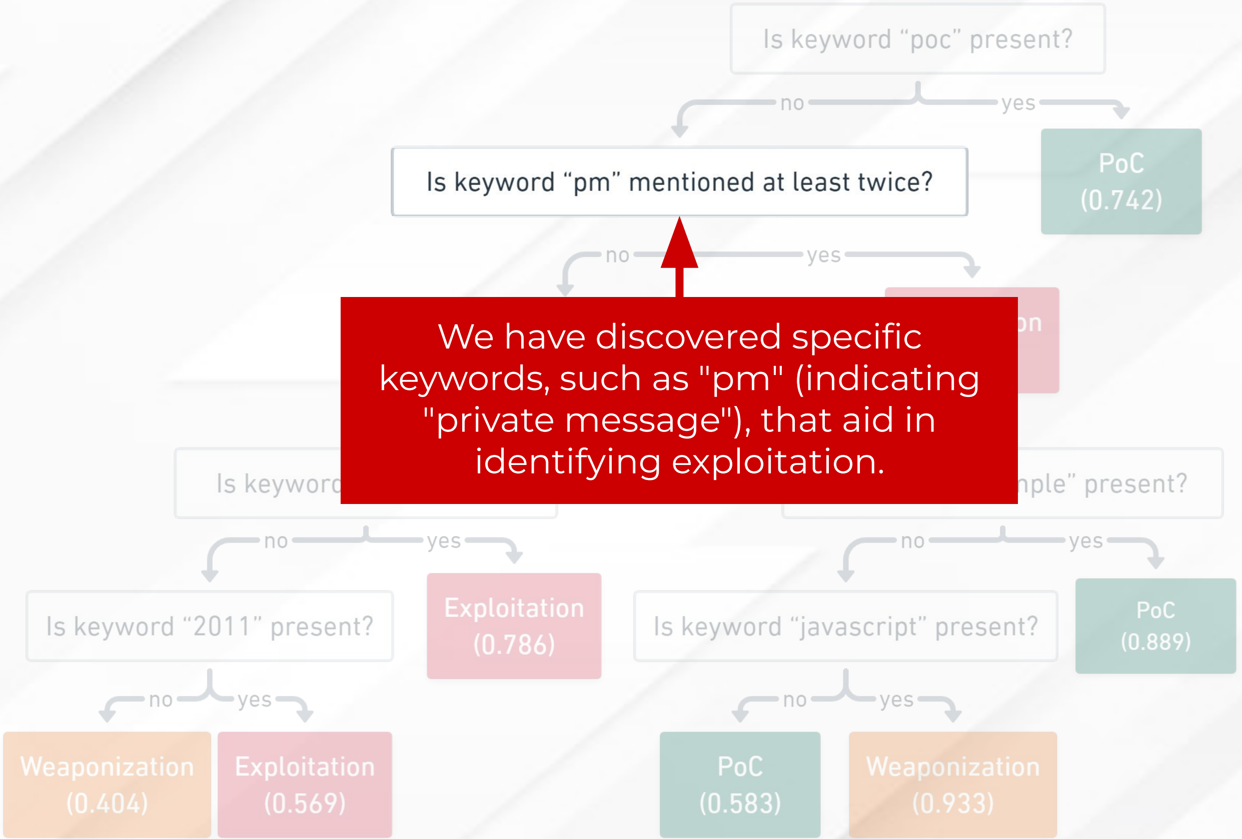
	Text encoding	Target classes	Accuracy	Precision	Recall	F1
DT	BoW	PoC, Weaponization, Exploitation	0.71	0.71	0.72	0.70
DT	TF-IDF	PoC, Weaponization, Exploitation	0.73	0.73	0.74	0.72
DT	doc2vec	PoC, Weaponization, Exploitation	0.74	0.74	0.74	0.73
DT	BoW	Exploitation vs Non-exploitation	0.85	0.86	0.85	0.85
DT	TF-IDF	Exploitation vs Non-exploitation	0.91	0.91	0.91	0.91
DT	doc2vec	Exploitation vs Non-exploitation	**0.92**	**0.93**	**0.92**	**0.92**
DT	BoW	PoC vs Non-PoC	0.75	0.75	0.75	0.75
DT	TF-IDF	PoC vs Non-PoC	0.77	0.78	0.77	0.77
DT	doc2vec	PoC vs Non-PoC	0.70	0.71	0.70	0.70
DT	BoW	Weaponization vs Non-weapon.	0.68	0.68	0.68	0.68
DT	TF-IDF	Weaponization vs Non-weapon.	0.63	0.64	0.63	0.62
DT	doc2vec	Weaponization vs Non-weapon.	0.59	0.59	0.59	0.59

Decision tree performance: easier to distinguish exploitation from rest

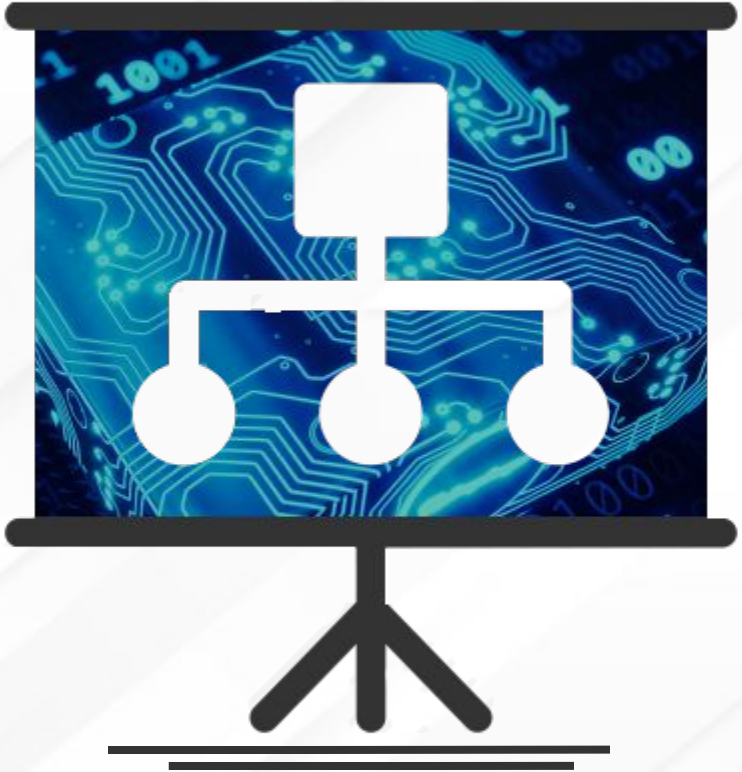
	Text encoding	Target classes	Accuracy	Precision	Recall	F1
DT	BoW	PoC, Weaponization, Exploitation	0.71	0.71	0.72	0.70
DT	TF-IDF	PoC, Weaponization, Exploitation	0.73	0.73	0.74	0.72
DT	doc2vec	PoC, Weaponization, Exploitation	0.74	0.74	0.74	0.73
DT	BoW	Exploitation vs Non-exploitation	0.85	0.86	0.85	0.85
DT	TF-IDF	Exploitation vs Non-exploitation	0.91	0.91	0.91	0.91
DT	doc2vec	Exploitation vs Non-exploitation	**0.92**	**0.93**	**0.92**	**0.92**
DT	BoW	PoC vs Non-PoC	0.75	0.75	0.75	0.75
DT	TF-IDF	PoC vs Non-PoC	0.77	0.78	0.77	0.77
DT	doc2vec	PoC vs Non-PoC	0.70	0.71	0.70	0.70
DT	BoW	Weaponization vs Non-weapon.	0.68	0.68	0.68	0.68
DT	TF-IDF	Weaponization vs Non-weapon.	0.63	0.64	0.63	0.62
DT	doc2vec	Weaponization vs Non-weapon.	0.59	0.59	0.59	0.59

Decision tree performance: easier to distinguish exploitation from rest


Is keyword “poc” present?
no
yes
PoC (0.742)
Is keyword “pm” mentioned at least twice?
no
yes
Exploitation (0.9)
Is keyword “type” present?
no
yes
Is keyword “off” present?
Is keyword “example” present?
no
yes
Exploitation (0.786)
no
yes
PoC (0.889)
Is keyword “2011” present?
Is keyword “javascript” present?
no
yes
Weaponization (0.404)
Exploitation (0.569)
no
yes
PoC (0.583)
Weaponization (0.933)



We have discovered specific keywords, such as "pm" (indicating "private message"), that aid in identifying exploitation.



Conclusion

- We were able to **identify**, **filter**, and **extract** pertinent information related to **CVEs**
    - Early detection of potential threats
- It is feasible to **train** a classifier to **infer** the **maturity level of threads**
- **White-box decision trees** allow understanding the inferences and explain outputs.
- **Best performance** in distinguishing **exploitation** from **other categories**

Thanks! Any questions?

felipe.moreno@ppgi.ufrj.br

THANKS!

Any Questions?