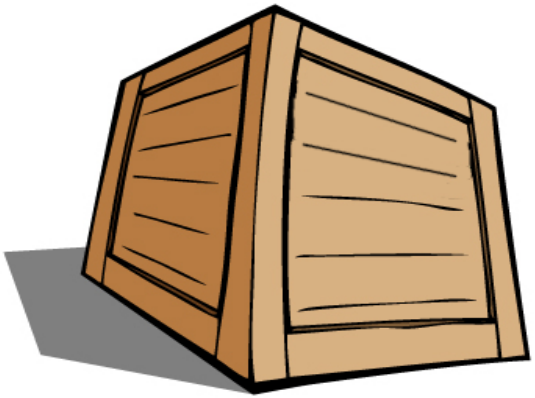
When Malware Is Packing Heat

Davide Balzarotti and Giovanni Vigna

EURECOM
Sophia Antipolis

lastline

USENIX Enigma 2018



Packing

Packing

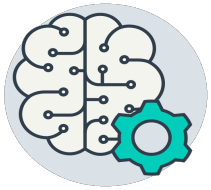
Researchers often have a limited understanding of the complexity of runtime packers

Researchers often have a limited understanding of the complexity of runtime packers

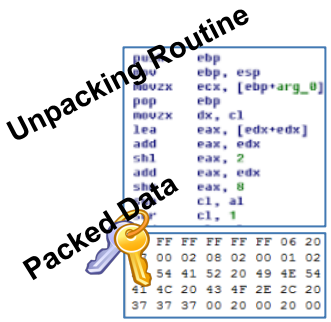
AV software often mis-classify benign packed samples as malicious

Researchers often have a limited understanding of the complexity of runtime packers

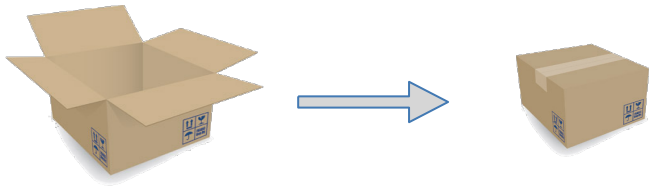
AV software often mis-classify benign packed samples as malicious

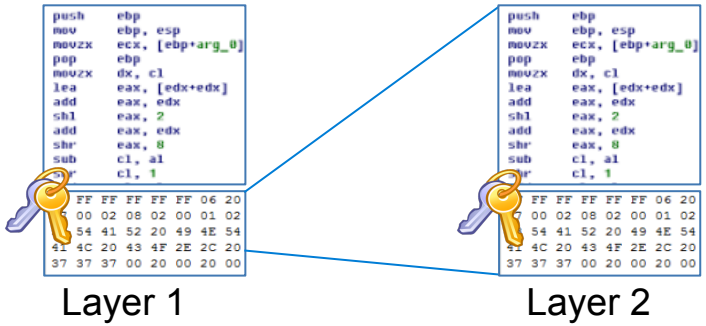


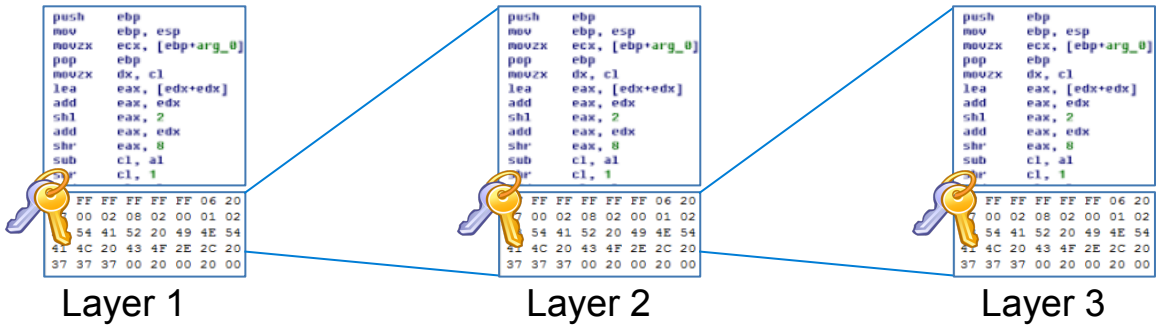
We all love ML, but in the presence of packing it just learns the wrong thing

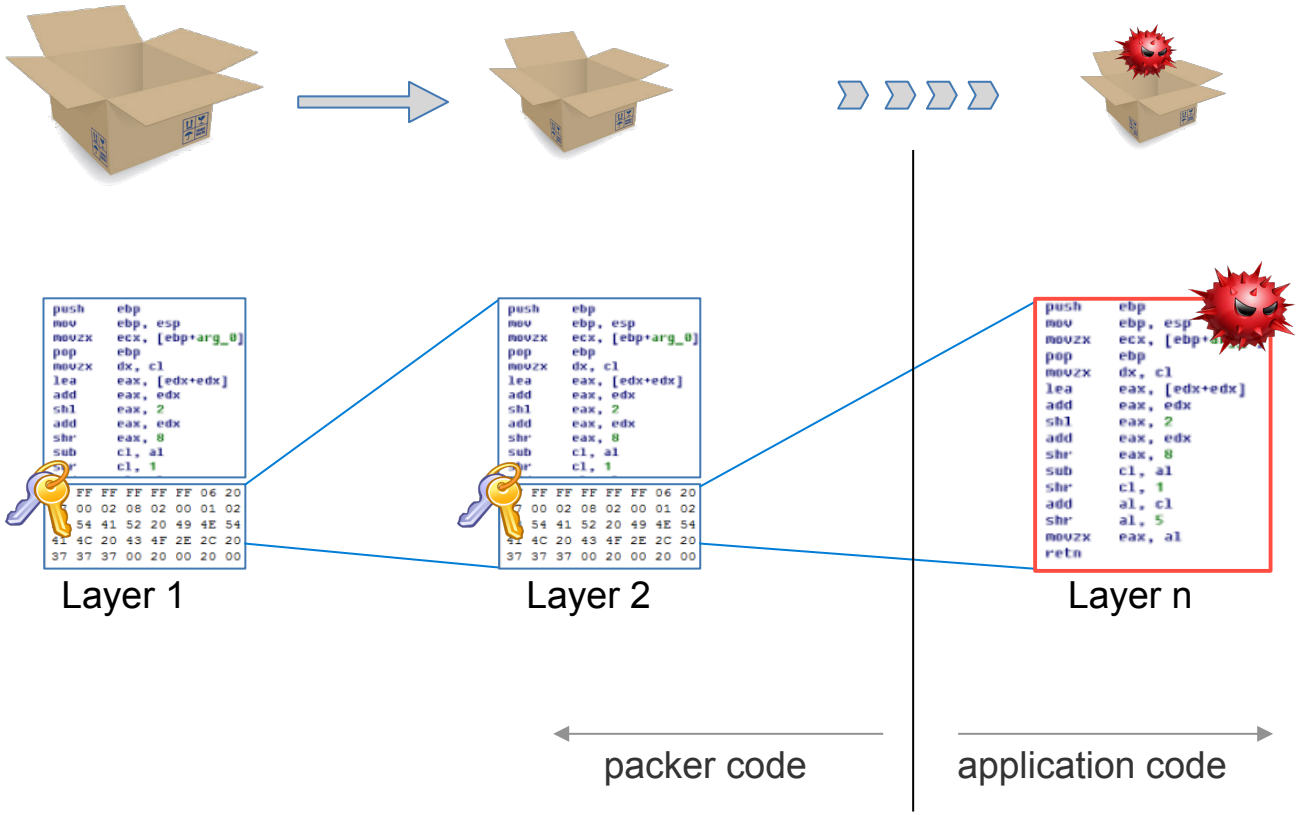

Unpacking Routine
Packed Data
ebp
ebp, esp
ecx, [ebp+arg_0]
ebp
dx, cl
eax, [edx+edx]
eax, edx
eax, 2
eax, edx
eax, 8
cl, al
cl, 1
FF FF FF FF FF 06 20
00 02 08 02 00 01 02
54 41 52 20 49 4E 54
4C 20 43 4F 2E 2C 20
37 37 37 00 20 00 20 00

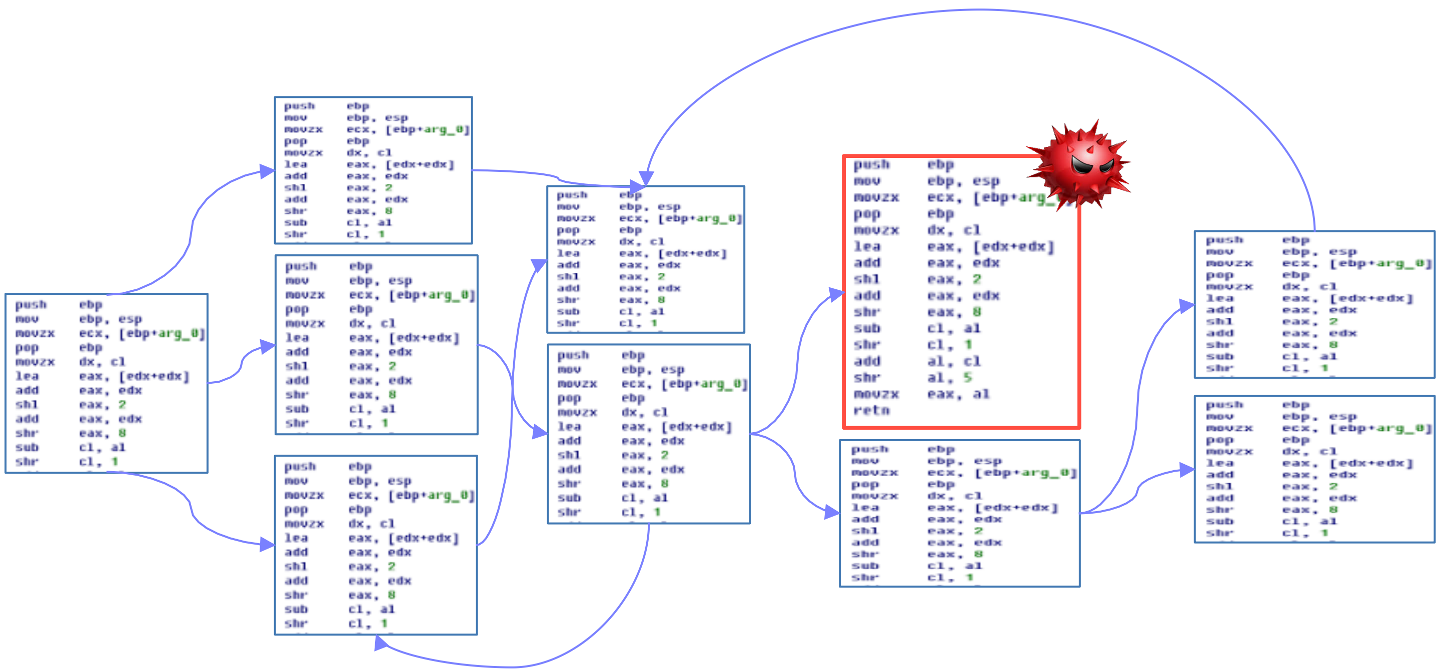
Layer 1

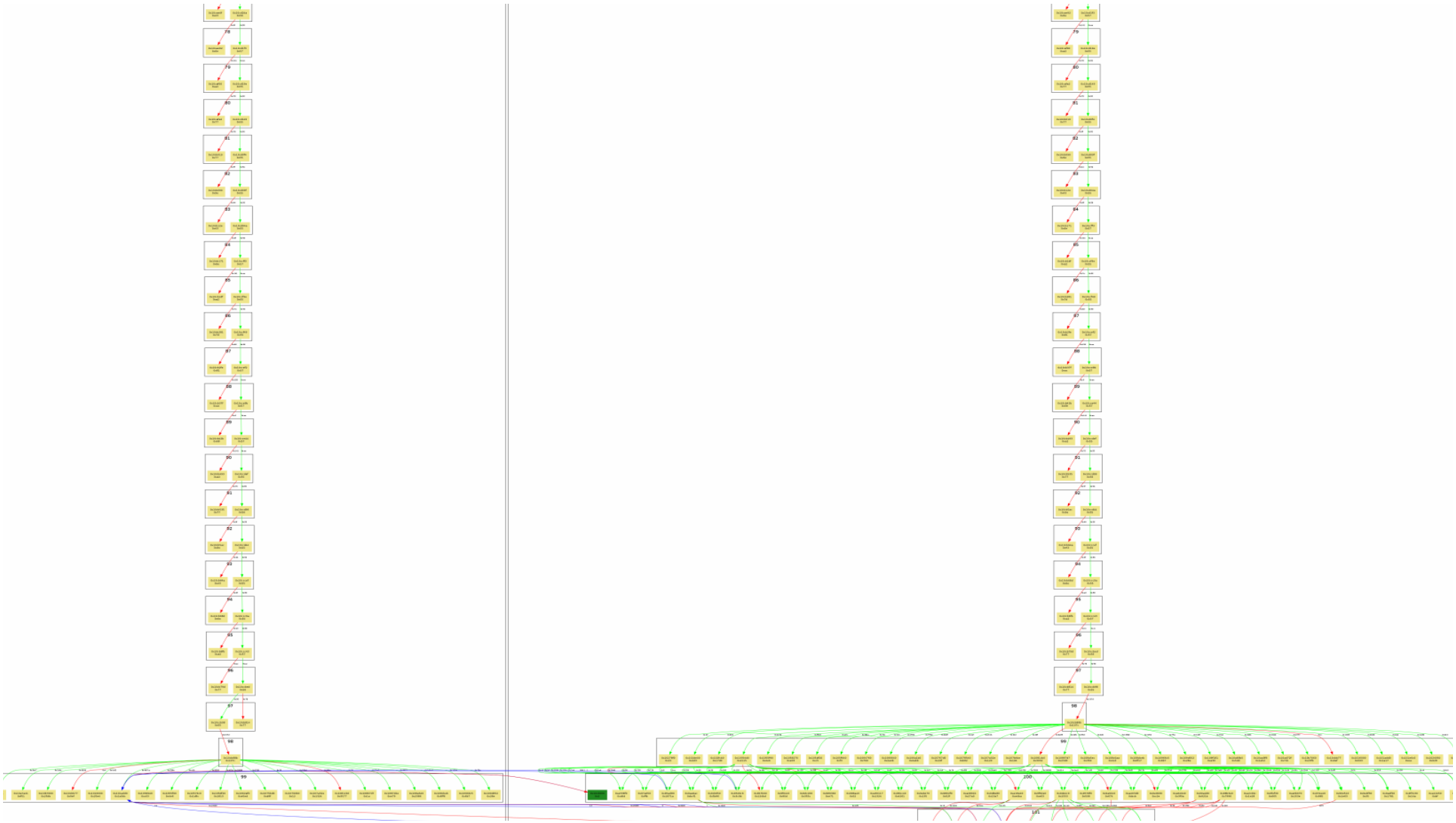



push ebp
mov ebp, esp
movzx ecx, [ebp+arg_0]
pop ebp
movzx dx, cl
lea eax, [edx+edx]
add eax, edx
shl eax, 2
add eax, edx
shr eax, 8
sub cl, al
shr cl, 1
FF FF FF FF FF 06 20
00 02 08 02 00 01 02
54 41 52 20 49 4E 54
4C 20 43 4F 2E 2C 20
37 37 37 00 20 00 20 00
Layer 1
Layer 2


push ebp
mov ebp, esp
movzx ecx, [ebp+arg_0]
pop ebp
movzx dx, cl
lea eax, [edx+edx]
add eax, edx
shl eax, 2
add eax, edx
shr eax, 8
sub cl, al
shr cl, 1
FF FF FF FF FF 06 20
00 02 08 02 00 01 02
54 41 52 20 49 4E 54
4C 20 43 4F 2E 2C 20
37 37 37 00 20 00 20 00
Layer 1
Layer 2
Layer 3


push ebp
mov ebp, esp
movzx ecx, [ebp+arg_0]
pop ebp
movzx dx, cl
lea eax, [edx+edx]
add eax, edx
shl eax, 2
add eax, edx
shr eax, 8
sub cl, al
shr cl, 1
add al, cl
shr al, 5
movzx eax, al
retn
Layer 1
Layer 2
Layer n
packer code
application code


push ebp
mov ebp, esp
movzx ecx, [ebp+arg_
pop ebp
movzx dx, cl
lea eax, [edx+edx]
add eax, edx
shl eax, 2
add eax, edx
shr eax, 8
sub cl, al
shr cl, 1
add al, cl
shr al, 5
movzx eax, al
retn



Complexity Classes

[Class I] *a single unpacking routine is executed before transferring the control to the unpacked program*

[Class II] *multiple unpacking layers are executed sequentially and lead to the original code at the end*

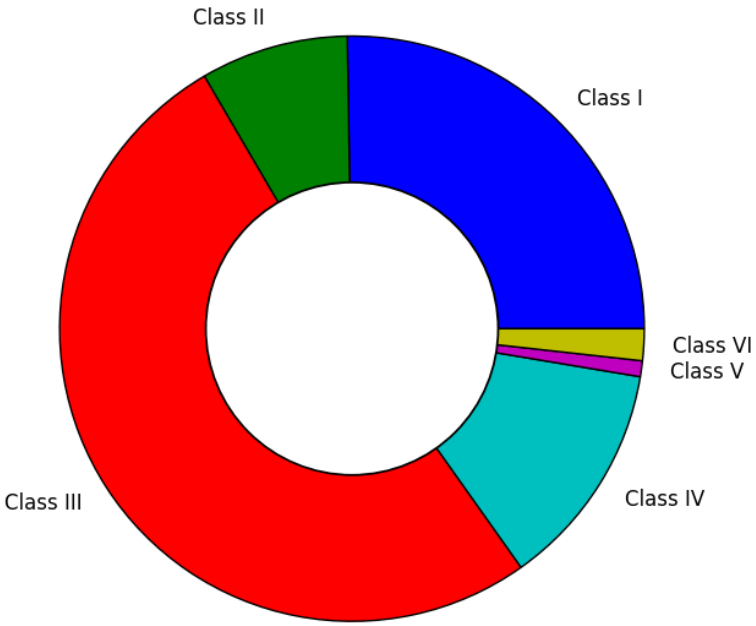
[Class III] *intermediate layers are executed in loops*

[Class IV] *the packer code is interleaved with the execution of the unpacked program*

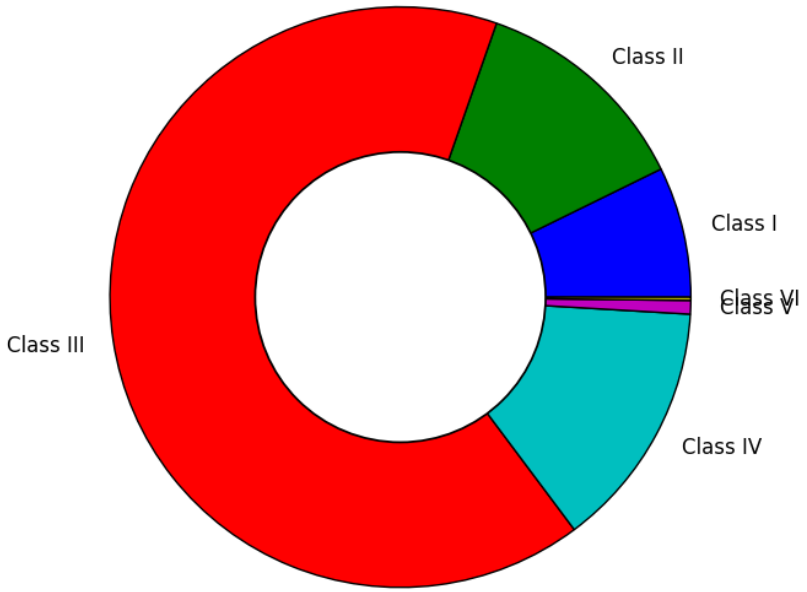
[Class V] *pieces of the original program are unpacked on-demand*

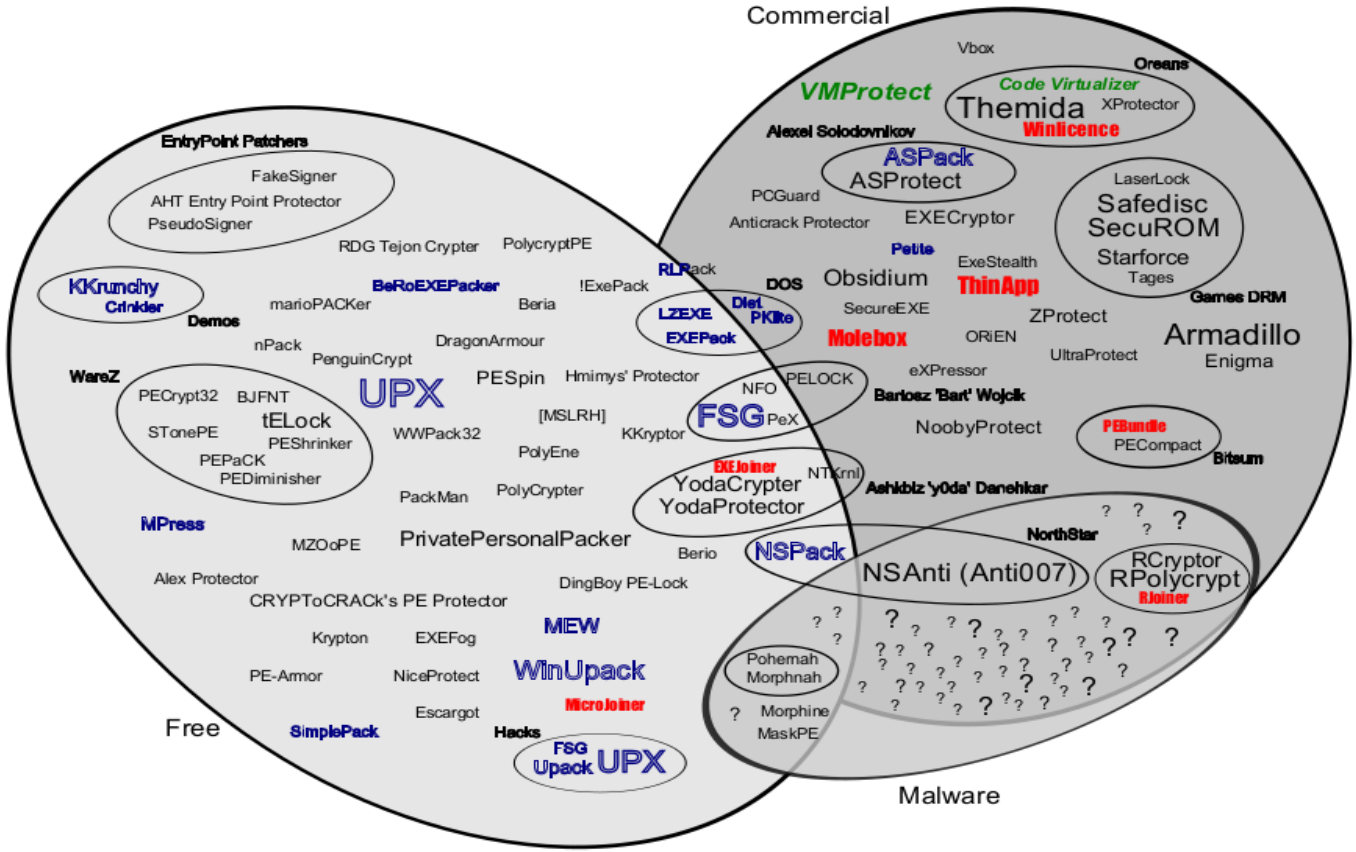
[Class VI] *only a single fragment of the original program (as little as a single instruction) is unpacked in memory at any moment in time*

Off-The-Shelf Packers



Custom Malware Packers





Why Does Packing Matter?

- Dynamic analysis techniques (e.g., sandboxes) have been introduced to deal with packing…
- …but static analysis techniques are more efficient!

An Experiment

- Benign programs from Windows OSs (XP, Vista, 7, NT)
  - 7983 samples
- Packed with 4 different packers
  - 16663 samples
- Submitted to VirusTotal
  - Looking for 10+ detections
- See: http://sarvamblog.blogspot.com/2013/05/nearly-70-of-packed-windows-system.html

Results

- UPX: 0% False Positives 
- BEP: 72.78% False Positives 
- NsPack: 98.72% False Positives 
- Upack: 99.88% False Positives 

Packing = Malware?

- False Positives

- Dataset Pollution

How Did We Get Here?

- Machine Learning has been increasingly used to perform malware detection
- The misclassification of packed binaries is the result of learning the wrong thing…

- Let's take a step back!

What Is Machine Learning?

- “Machine learning explores the study and construction of algorithms that can learn from and perform predictive analysis on data”

https://en.wikipedia.org/wiki/Machine_learning

Why Machine Learning?

- Supports data analysis
- Supports characterization
- Supports classification

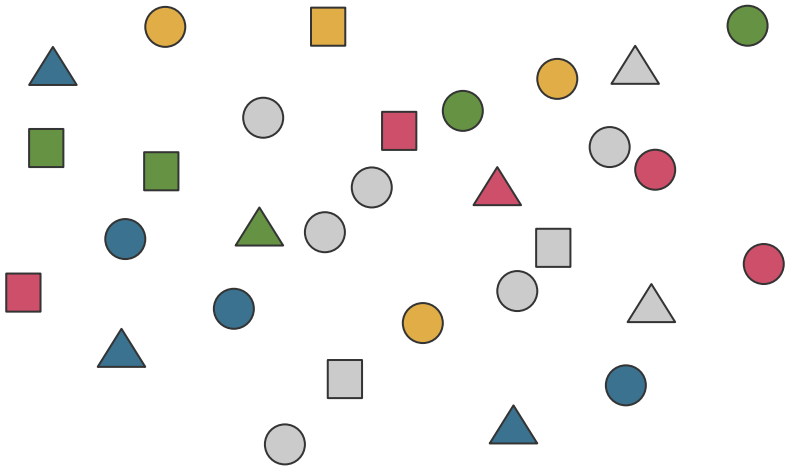
Machine Learning

?

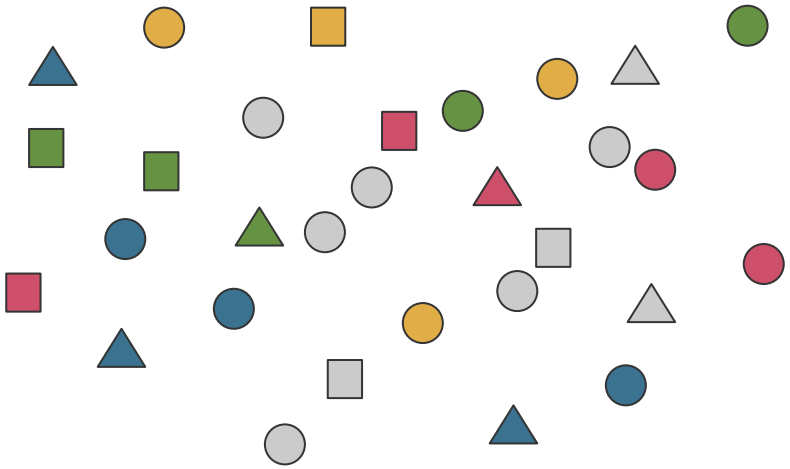
Round?

Has >3 sides?

…



Machine Learning



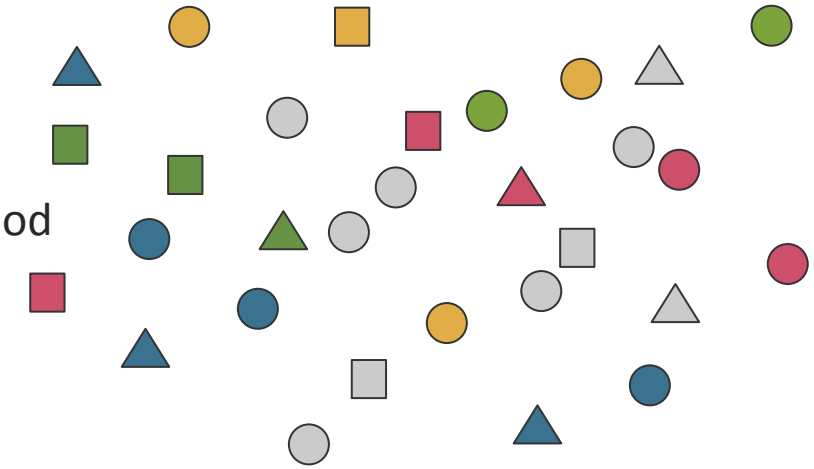
Machine Learning

?

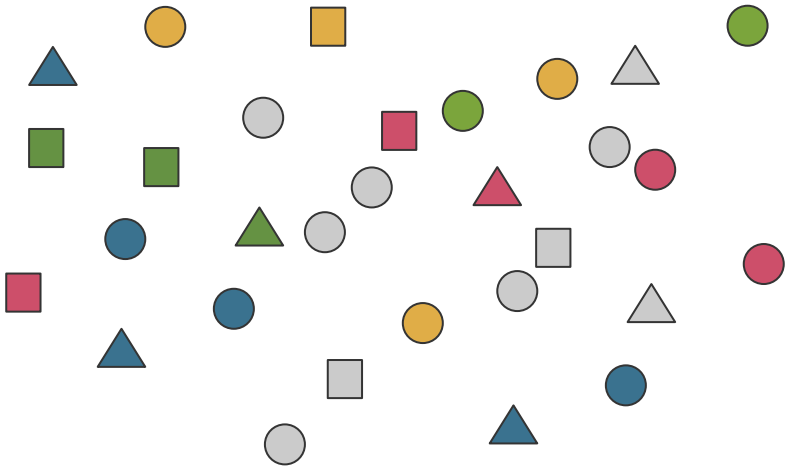
Reds are bad

Blues, greens, oranges are good

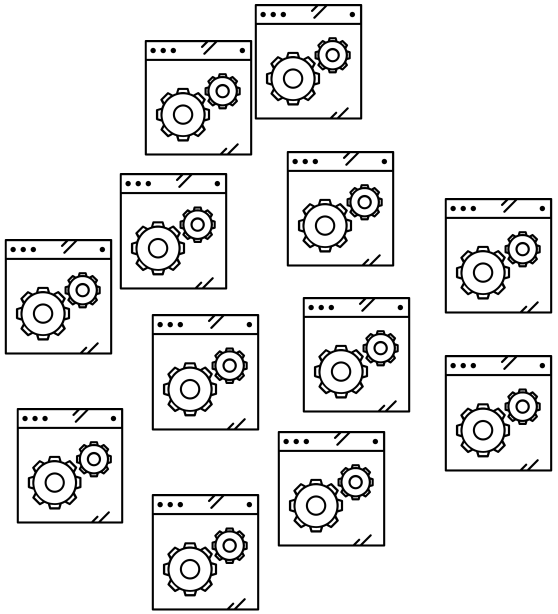
What about greys?

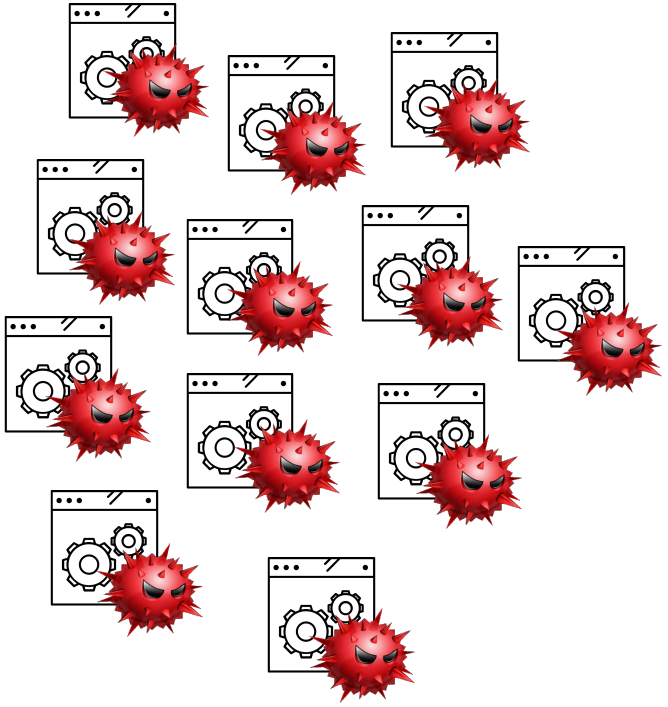


Machine Learning

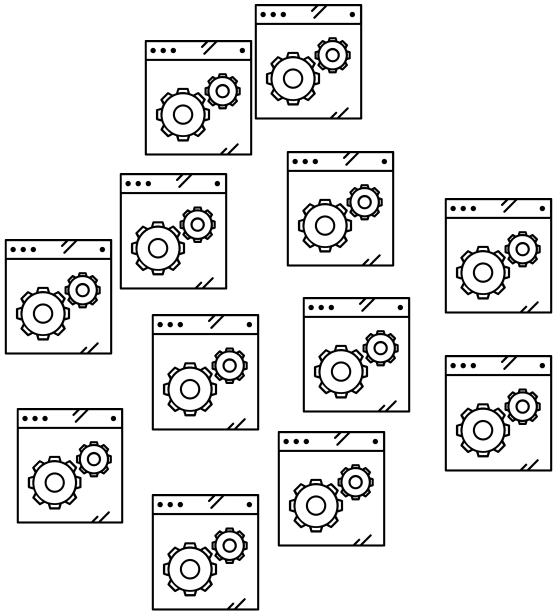


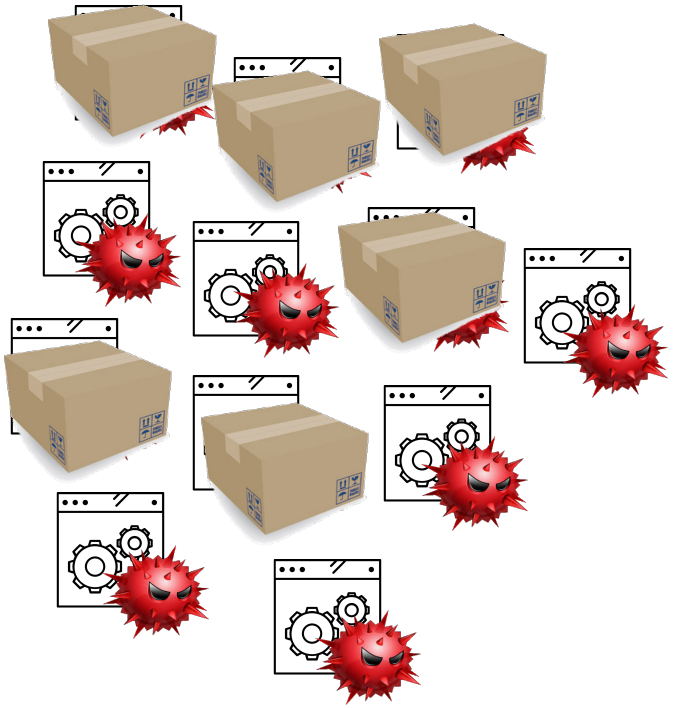
Pitfalls in Machine Learning



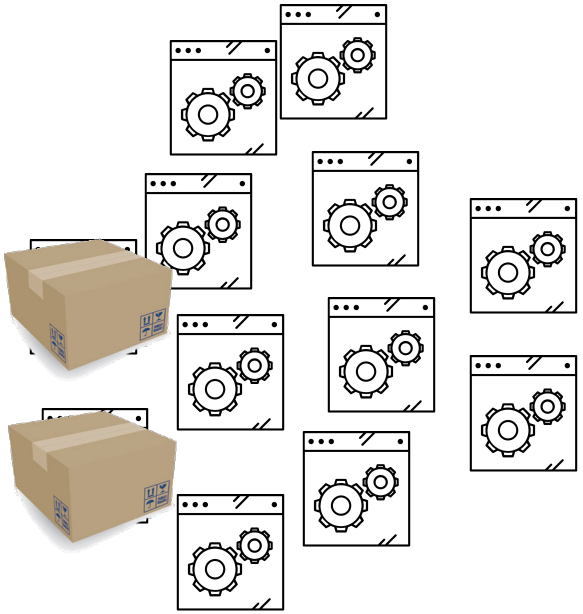


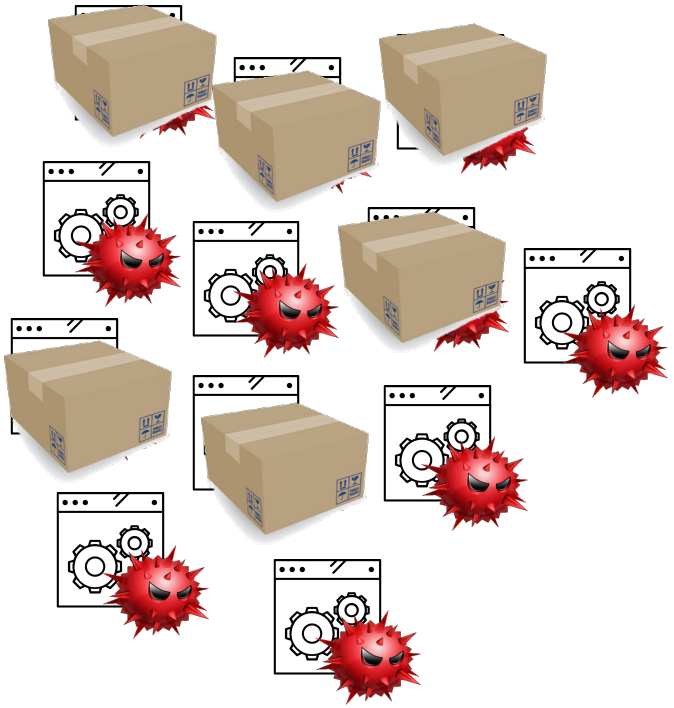
Pitfalls in Machine Learning



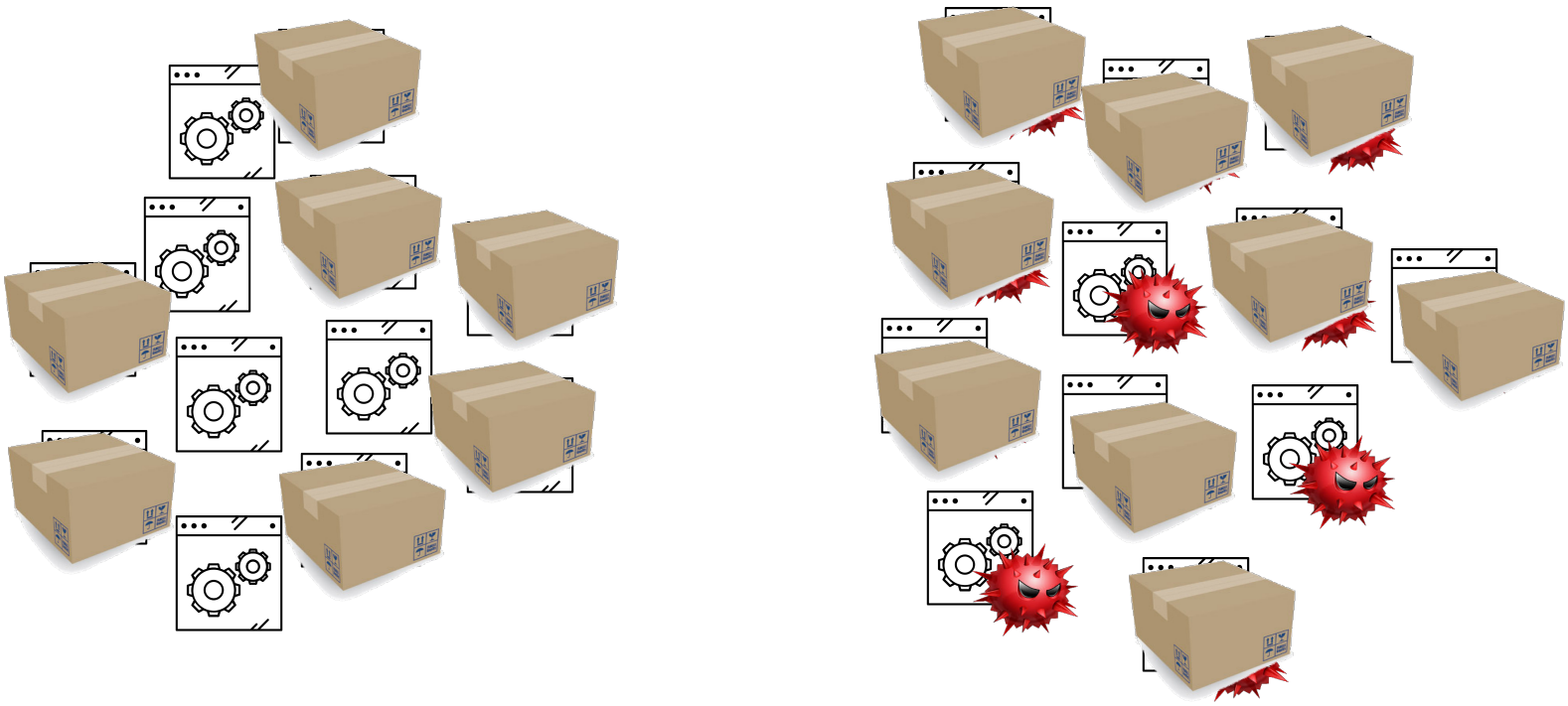


Pitfalls in Machine Learning

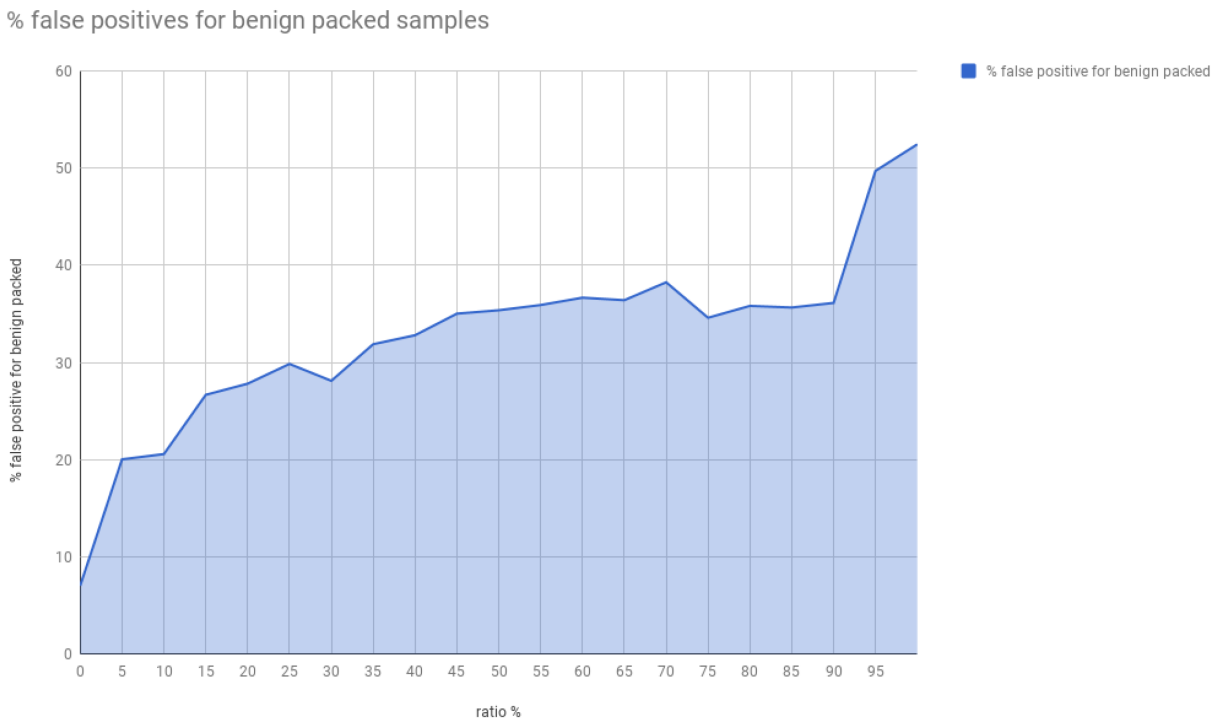




Pitfalls in Machine Learning



Another Experiment

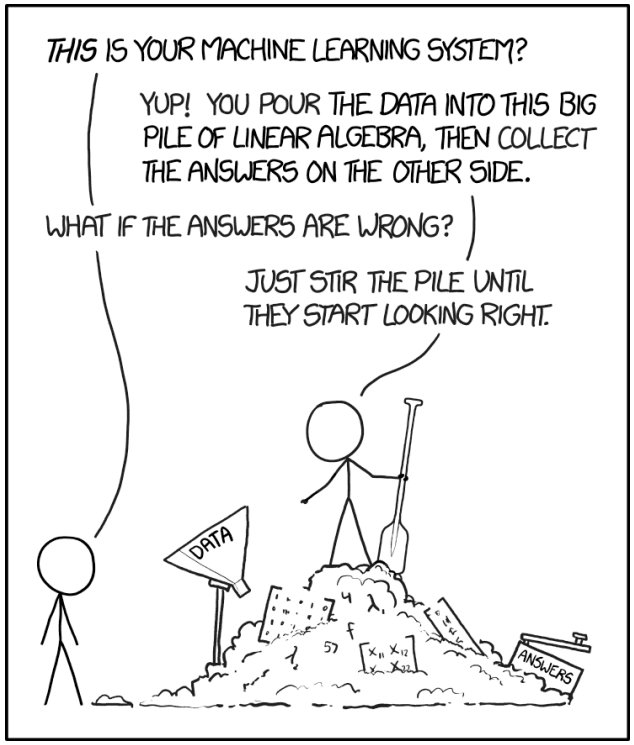


Insight: When most of malware is packed, *packing* is what is actually learned

Conclusions

- Applying machine learning to packed malware might lead to the detection of packing (and not the detection of malicious behavior) resulting in false positives
  - De-sensitization caused by false positives
  - Pollution of datasets
- Sophisticated dynamic unpacking and analysis is necessary

Questions?



process by Roman from the Noun Project
Machine learning picture: https://xkcd.com/1838/