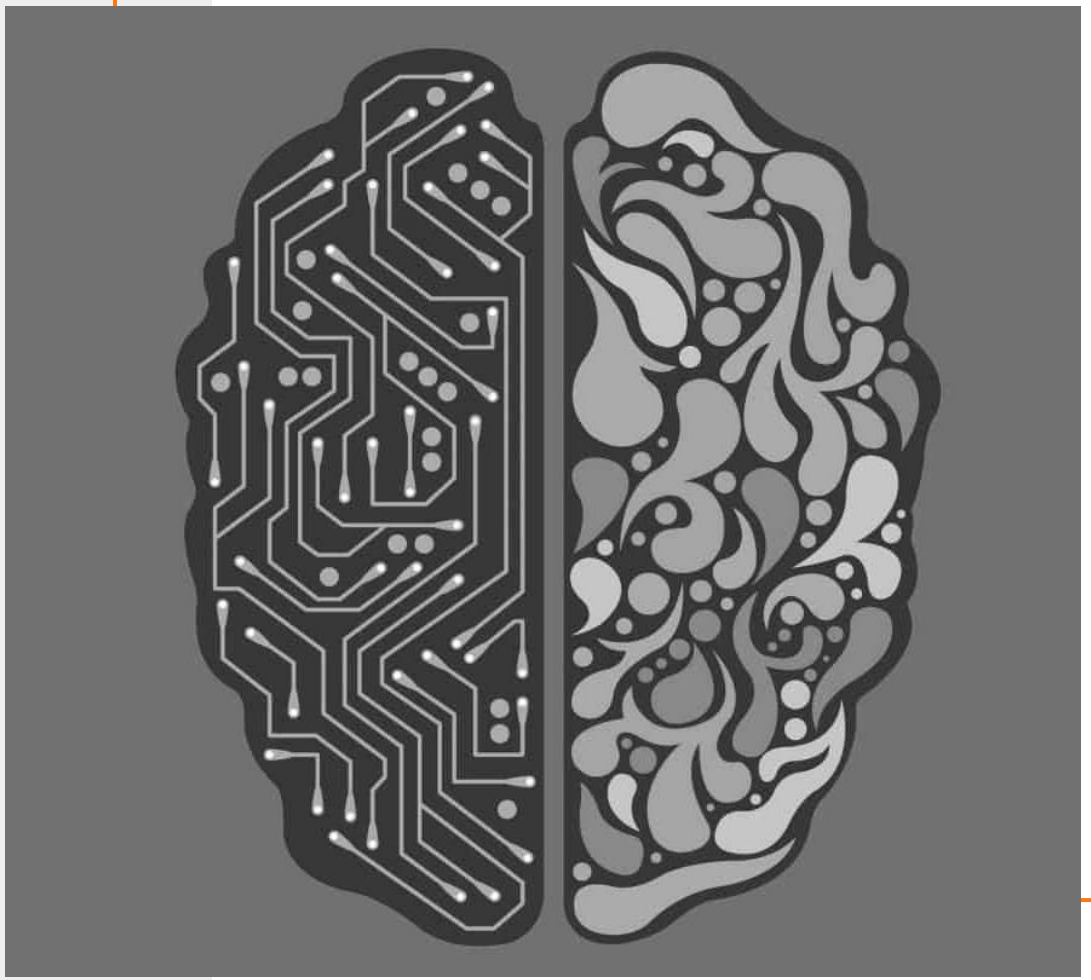
A magic tool or totally overrated?

Machine learning for both offense and defense in cyber security

Cas Bilstra

17 Nov. 2021 - 19:00



Who am I?

- Cyber security specialist at EYE security
- Background in both data science and cyber security
- Specialised in attacking Machine Learning (ML) models

- In my free time I like to work on cars
- For holidays I enjoy going to the mountains

- There are many misconceptions around ML

After this webinar, you

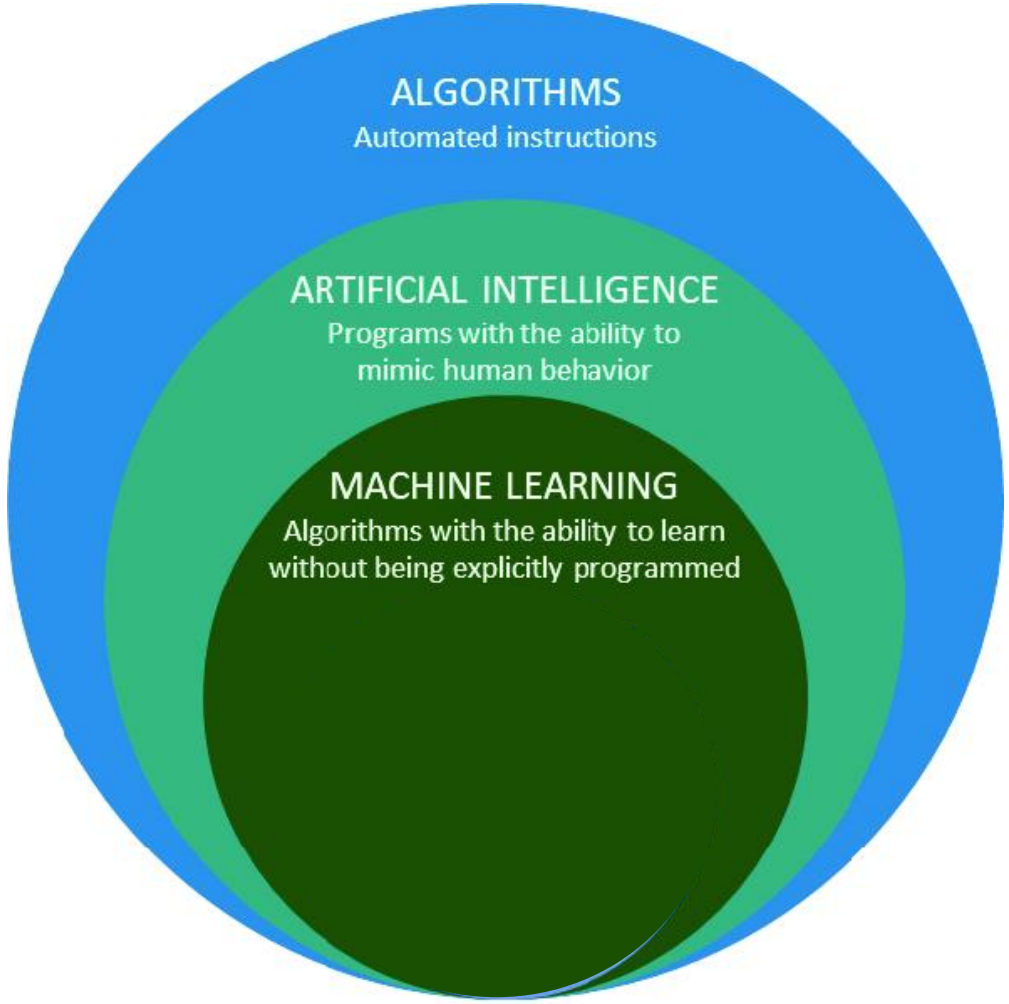
- Have an understanding of the **fundamentals** behind ML
- Know the strengths and weaknesses of ML

- Know how ML can be used as both a defensive and offensive mechanism in cyber security

01. What is Machine Learning?

Artificial intelligence (AI) and ML

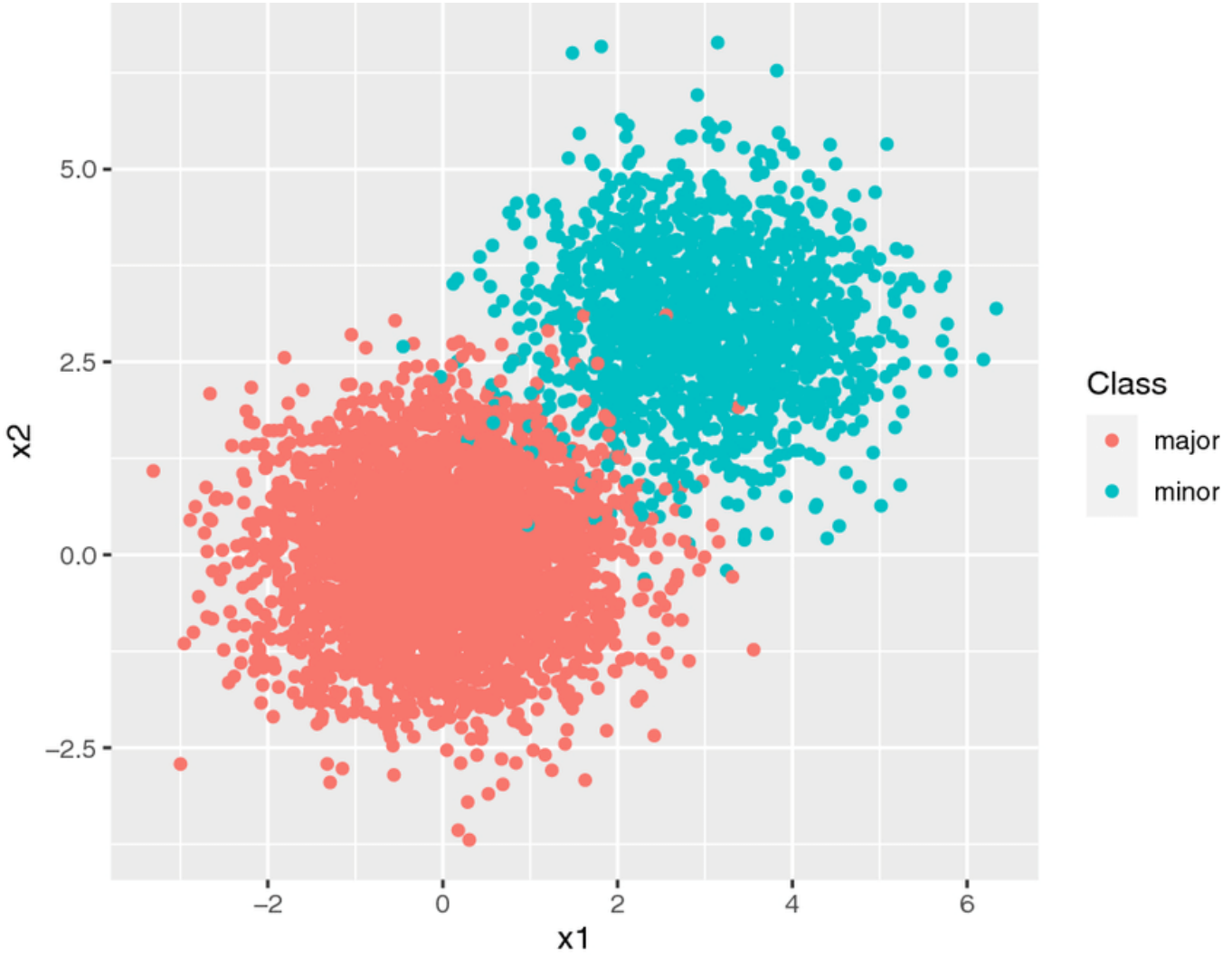
- Algorithms - “recipe” on how to solve problems
- Artificial Intelligence – type of algorithms that can perceive its environment and autonomously make decisions that maximize the chance of it achieving its goals
- Machine Learning algorithms – train a model that can make decisions without being told how to make these decisions. Based on large quantities of data
- Example: chess computer.
    - Set of rules vs large set of historical games



How does Machine Learning work

- ML algorithms define a mapping from the input (data) to the expected output
- For this they generally need large quantities of data
- They are able to autonomously infer relations/patterns in this data
  - These patterns or relations are sometimes invisible to humans
  - Or: difficult for humans to completely describe these patterns

How does Machine Learning work



Lifecycle of Machine learning

- In a training phase, large quantities of data are processed by ML algorithms to create a model


cat

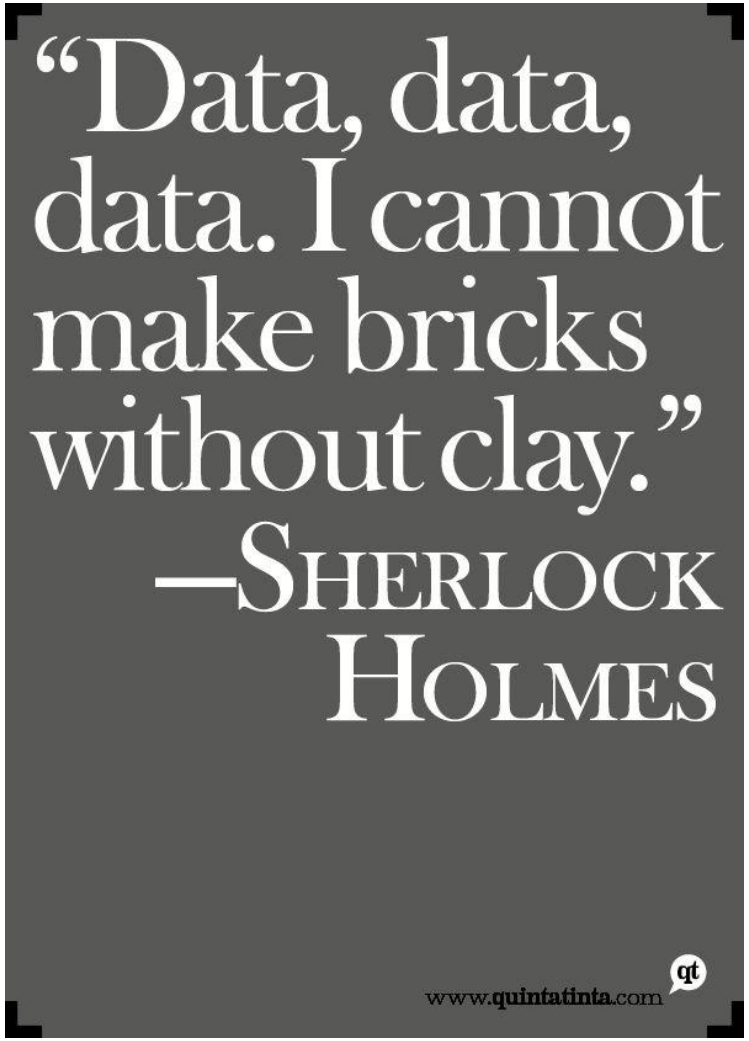

cat


fish


fish

- This model consists of the relations/patterns the algorithm has found in the training data
  - E.g. pictures where ears are present are cats
  - Or: fish are orange
- This model can then be used to classify unseen data as well

The importance of the data



The importance of the data

- The model will capture relations that are present in the data
  - It will not capture relations that are not present in the data!
    - Make life easy for the ML algorithm: provide it with the right data. It is not a magic tool!
    - Sometimes it is necessary to “extend” the data, transforming it with simple operations. Instead of “start” and “end”, provide it with the duration of the event.
  - If faulty relations are present in the data, the model will capture those relations as well (fishes are orange)
- Generally, gathering and preparing the right data is much more difficult than the implementation of the machine learning algorithm itself

02. Pros and cons of Machine Learning

Pros of machine learning

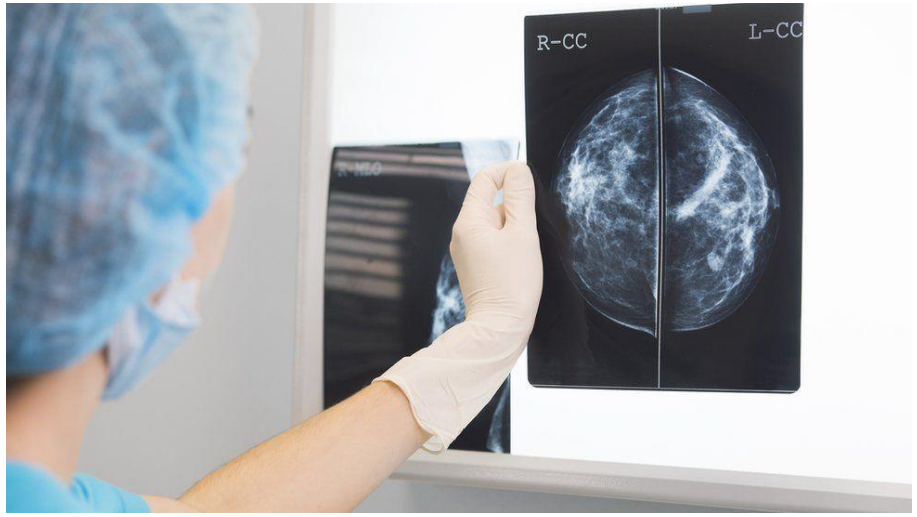
- It can be much better at certain tasks then humans
  - It can take into account much more variables
  - It can find relations or patterns that are invisible to humans
- It can be much faster in achieving certain tasks than humans

Schaker	ELO	Tijd
Magnus Carlsen	2882	30 jaar
AlphaZero	3500	4 uur

Success story

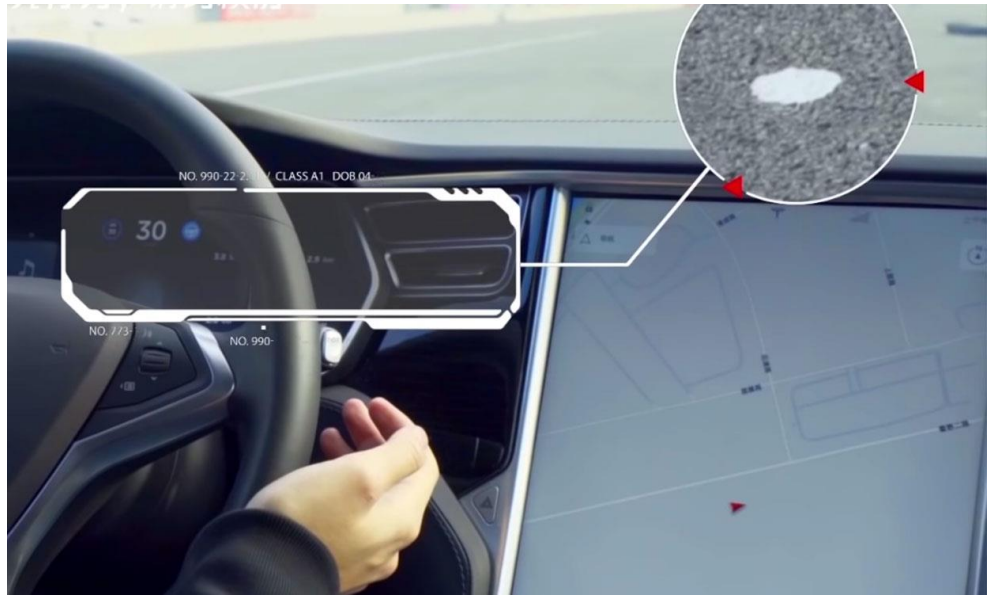
Breastcancer research Imperial College Londen

- X-ray images of 29.000 women
- 1.2% decrease in “False Positives”
- 2.7% decrease in “False Negatives”
- Replaces 2 radiologists



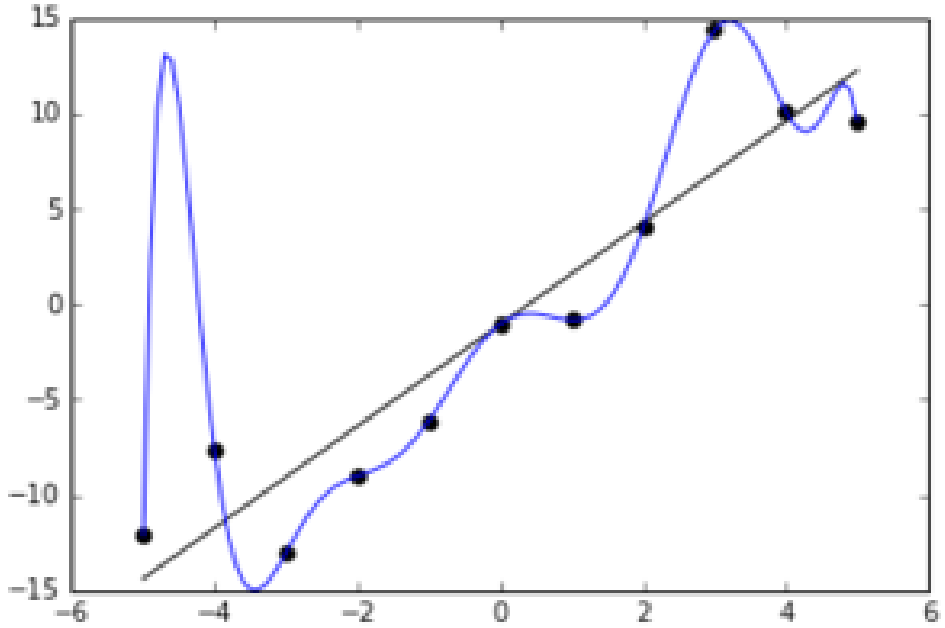
Pitfalls of Machine Learning

- Your ML model is trained to accurately describe the training data
  - Is the training data representative for ALL data that can be fed to the model?
  - Adversarial Examples



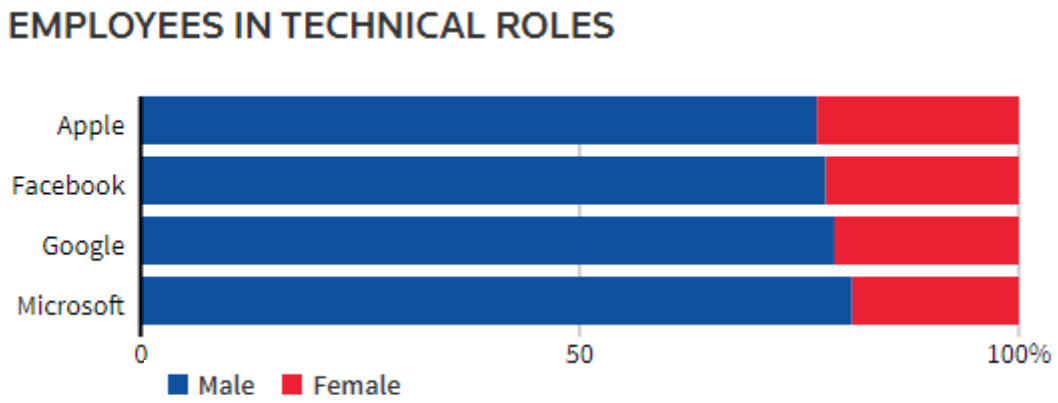
Pitfalls of Machine Learning

- Your ML model is trained to accurately describe the training data
  - It thereby tries to find a “perfect” mapping from input to output
    - This may result in a “over-eager” model where it is so focused on making perfect predictions that it loses generality. This is called **overfitting**



Pitfalls of Machine Learning

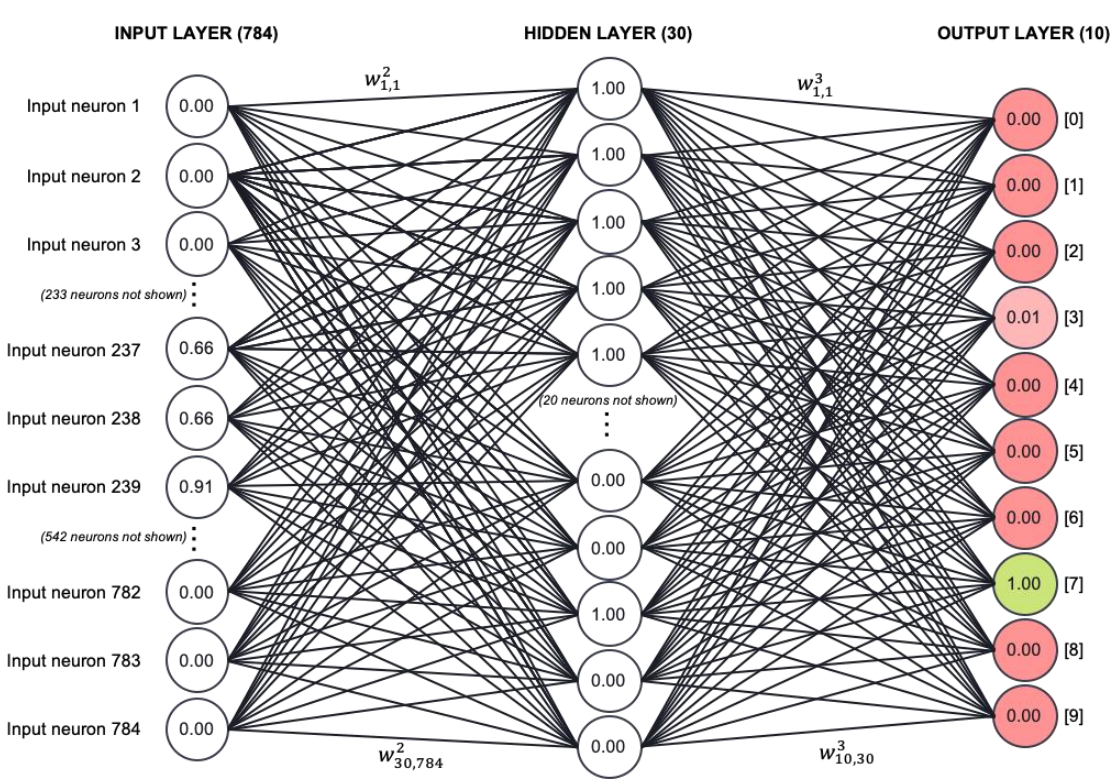
- If your dataset discriminates, your model will discriminate too!
- Amazon trained a ML system to speedup the solicitation process
- Input:
  - Resumes of candidates 2004-2014
  - Selected candidates

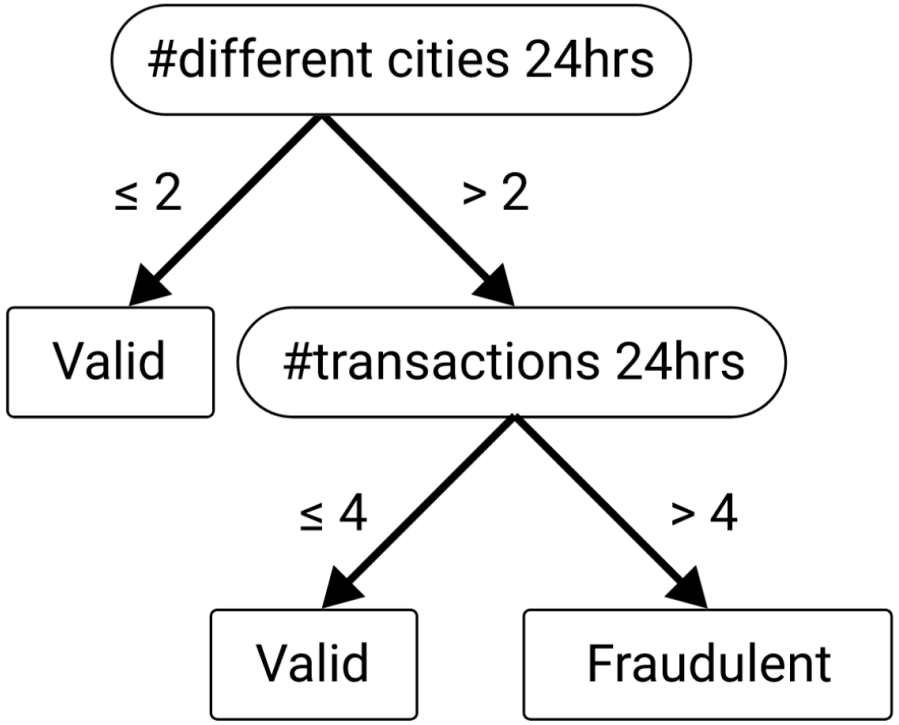


Pitfalls of Machine Learning

- Your ML model finds relations that are generally hard to find for humans
  - For this it may use very complex structures, which are not interpretable for humans.
  - Some laws require the decisions of autonomous systems to be explainable to the end-user.

Pitfalls of Machine Learning





03. Defensive Machine Learning in cyber security

Defensive machine learning

- Based on the assumption that attacker behavior will be different from legitimate behavior
- So, we can recognize it by observing abnormal behavior

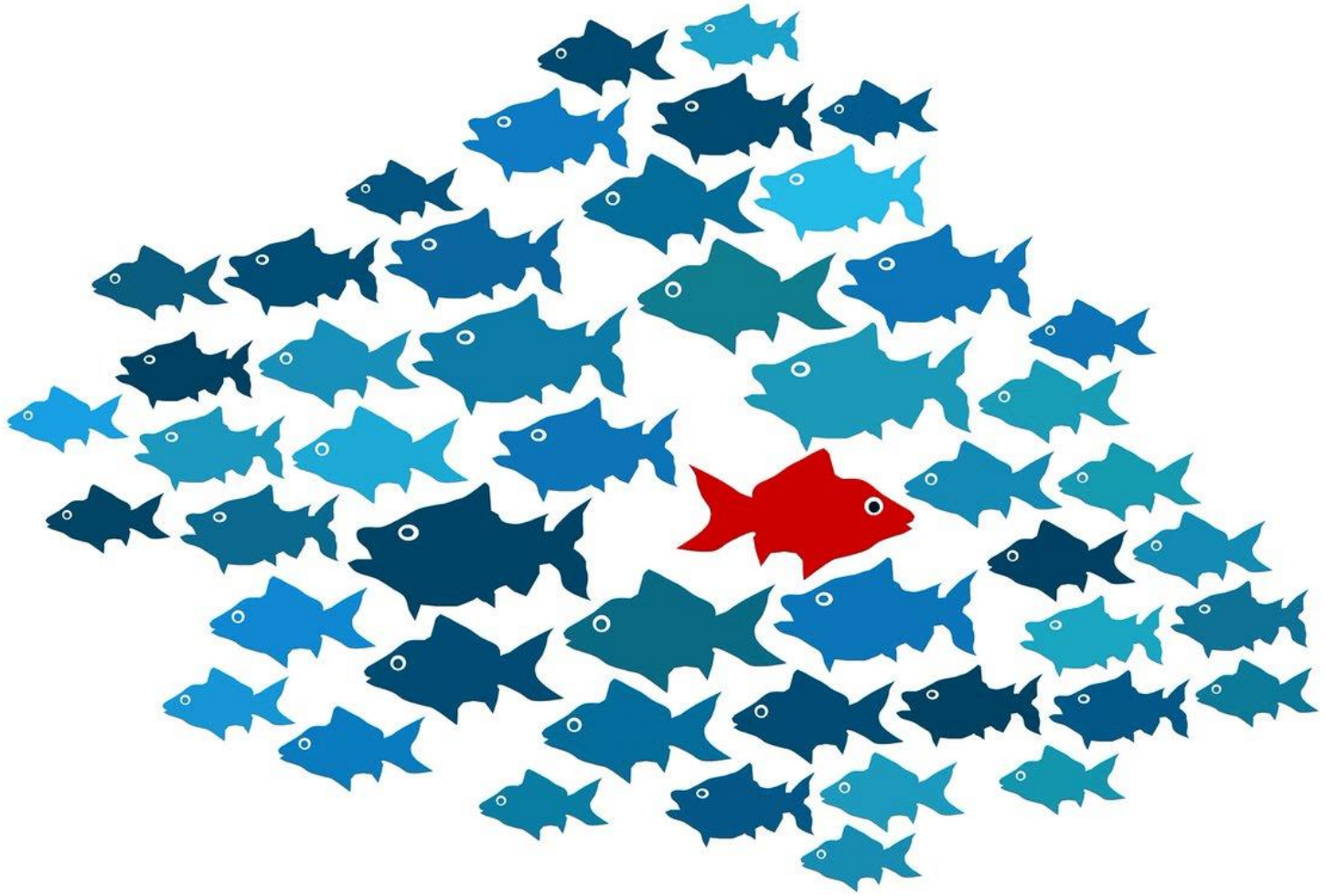
Types of defensive mechanisms based on data

- Signature-based
  - Through rules you define unacceptable behavior
  - Traditional Antivirus solutions use this
  - High probability that the events you identify are indeed malicious
  - You need a very big number of rules to capture all malicious behavior
  - Requires a lot of time investment for defining all the rules

Types of defensive mechanisms based on data

- Anomaly-based
  - Use machine learning to create a model of normal behavior – both for endpoints and accounts/cloud
  - Anything that does not fit in the model of what is “normal” may be malicious activity
  - Able to detect attackers that use previously unknown tools/vulnerabilities/methods to infiltrate and move
  - Raises a lot of False Positives: situations where the ML signals legitimate user behavior as abnormal (malicious)

Anomaly-based



Types of anomalies

1. **Completely different behavior**

   Previously unknown programs are installed and started

2. **Contextually different behavior**

   10 different mails are sent before 07.00 AM

3. **Collectively abnormal behavior**

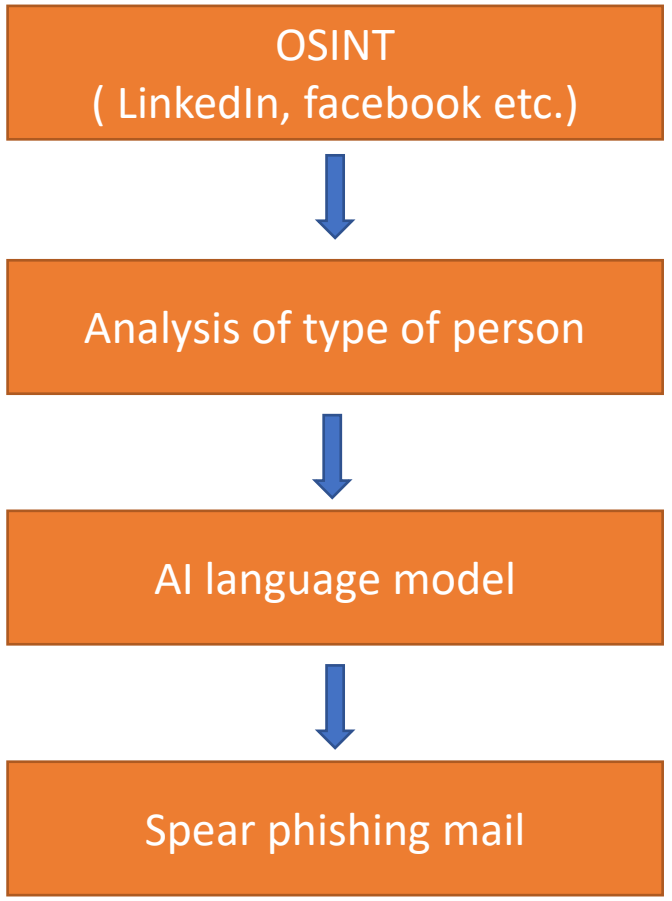
   3 laptops start the same process at exactly the same moment in time

Defensive – Phishing detection

1. Recognizes the logo of Amazon on this phishing page
2. Checks if the domain is registered by Amazon

04. Offensive Machine Learning in cyber security

Offensive - Spear phishing



Review Output From OpenAI's GPT3

Dear John,

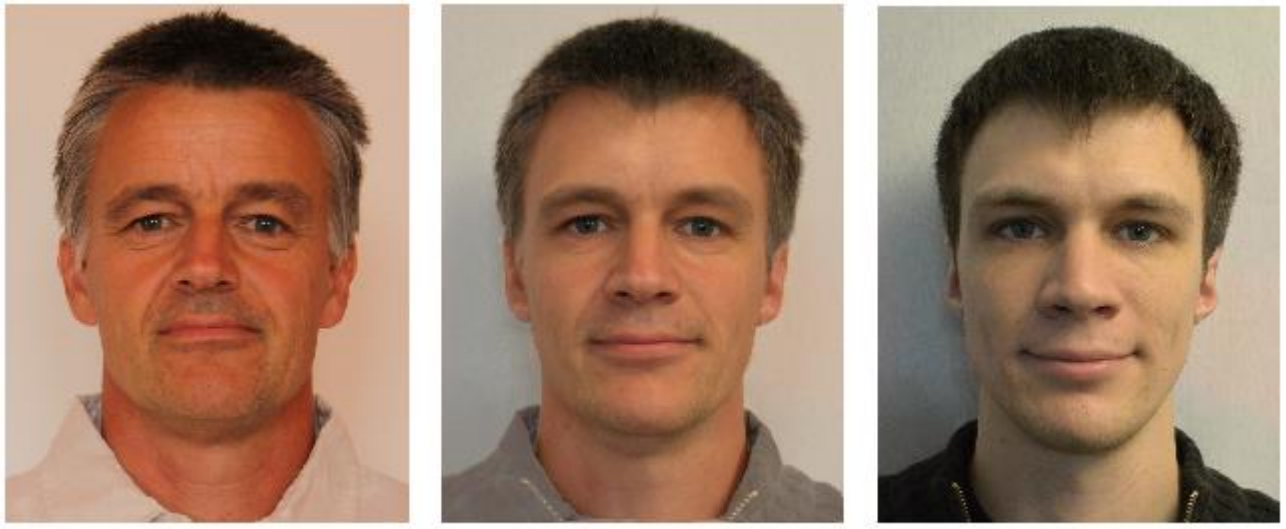
I hope this email finds you well. I am writing to you on behalf of our company, Company B, to inquire if you are interested in exploring a career opportunity with us.

Company B is the leading company that develops and markets software, services, and devices. We are known for our innovations in computer software, hardware, and for providing cloud-based services that empower every person and every organization on the planet to achieve more.

Our commitment is to help every person and every organization on the planet to achieve more. We do this by creating the best devices and services that enable people and businesses to do and achieve more.

...

Offensive - Face morphing attack



(a) Subject 1 (b) Morph (c) Subject 2

Subject 1 and 2 are both recognised in the middle picture

Offensive - Adversarial attack

classified as

Stop Sign

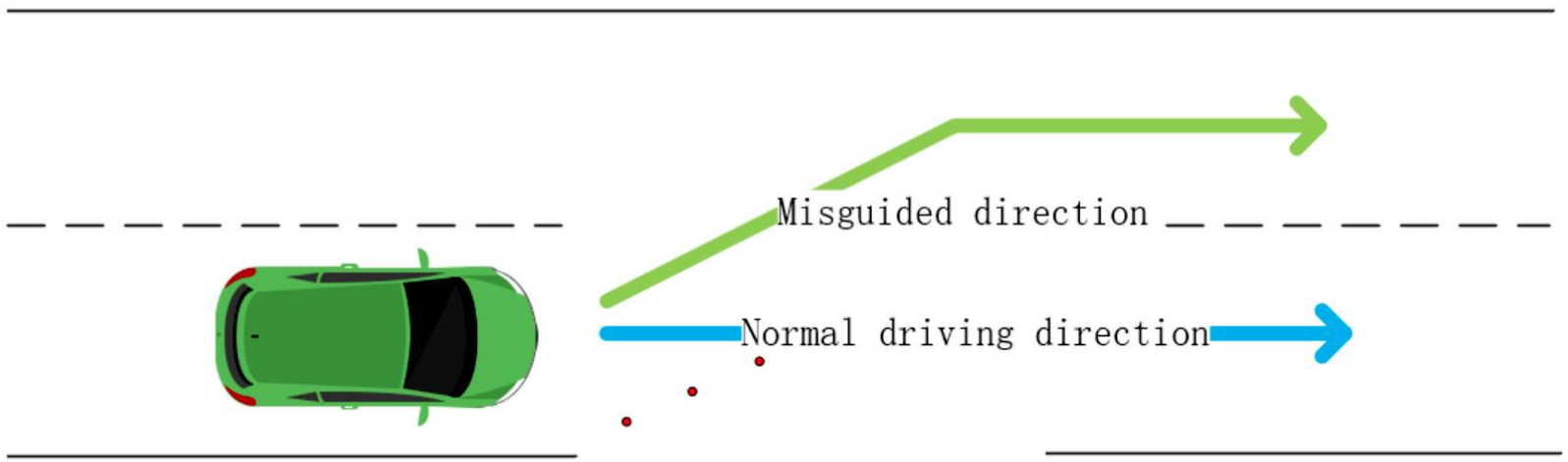
+

=

classified as

Max Speed 100

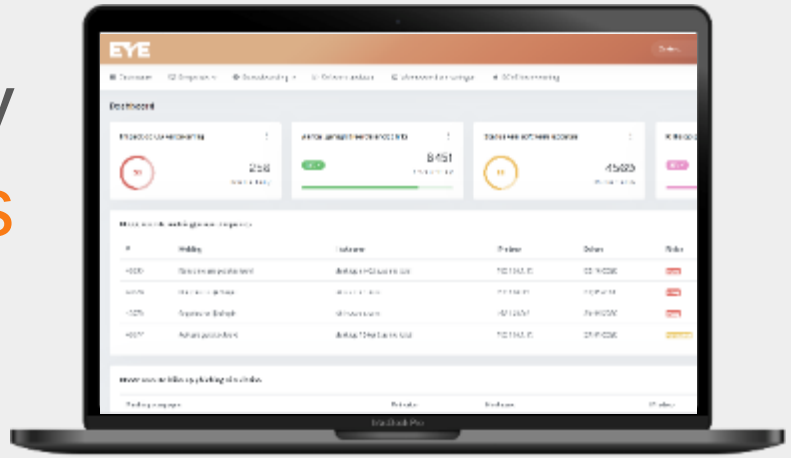
Offensive - Adversarial attack



Machine learning at EYE

Services

Security
Measurements



Cyber
Insurance

Monitoring & detection

Awareness training

24/7 Response

Advise

Cyber Insurance

EYE service overview

- Monitoring of cloud environment (SIEM)
  - Microsoft 365
  - Google Workspace
- Endpoint Detection and Response (EDR) software
- All alerts will be monitored 24/7 by our SOC
- CERT: Very quick reaction times should an incident occur

How we use Machine Learning at EYE

- We use a hybrid approach
  - Our technology partners provide us with 1000’s of rules signalling malicious activities
  - On top of that, we use ML-based systems that are able to identify abnormal behavior
  - A cyber security specialist investigates all reports, ensuring no unnecessary business interruptions while maintaining quick response times
- We have research projects with the TU Delft to build ML systems on top of the data we collect

Questions

For further questions or feedback, feel free to contact met at cas.bilstra@eye.security

How to develop a Machine Learning system

1. Gather the correct data
2. Prepare the data (remove columns which may lead to inaccurate mappings)
3. Choose a suitable type of Machine Learning algorithm
4. Train your model using this algorithm on a part of the data
5. Test (on an unseen part of the data) how the model performs
6. If happy with the result, continue. Else, adjust the settings of the algorithm and continue from step 3.
7. Feed unseen data to the model and have to model make decisions autonomously
8. Mileage may vary

Limitations

Chihuahua or cupcake?