Research Institute
Cyber Defence
Bundeswehr University Munich

# Report: AI for Cybersecurity Workshop

## Moderation and Organization:

Norbert Pohlmann, Professor in the Computer Science Department for information security and director of the Institute for Internet Security at the University of Applied Sciences Gelsenkirchen, Germany.

## Speakers/Panelists:

- Ammar Alkassar, Commissioner for Innovation, Saarland
- Ulla Coester, Partner Wegesrand, Lecturer at the Hochschule Fresenius

The workshop started with a presentation held by Prof. Dr Norbert Pohlmann and covered the fundamental basics of machine learning, classification and AI. With various examples from the areas of online-banking, passive authentication and automotive, the great value of machine learning (MI) and artificial intelligence (AI) for cyber security applications was clearly explained. In the second part of the workshop, political and ethical aspects of using AI for cyber security as well as rising security threats for machine learning were discussed in detail in a panel.

## Summary of the presentation:

### Main Targets

- Increasing detection rate for attacks
- Support of cyber security experts in sensing of events and provision of recommendations for action
- Improvement of existing solutions for defense of attacks

### Important definitions

- Data science: extraction of knowledge from data
- AI: is a branch of computer science which deals with the simulation of intelligent behavior in computers and transfers intelligent behavior into algorithms
- Machine learning: artificial generation of knowledge from experience (data). The resulting generalizations can be applied to new data.

### Success factors

- Quality of data is crucial to achieve trustworthy results
- High quality and secure sensors are necessary
- Establishment of data pools

- More data from smart devices, IOT, networks (log data), cars….
- Improvement of hardware and communication speed
- Usage of parallelization
- Support by government

## Classification

- Some classification algorithms were explained. First an example of what an attack looks like and how can we protect them. Example: increasing the detection rates of attacks in a network and how we need to continuously change our model to adapt.
- What can we do in that area? Finding security relevant events to analyze that information and generate a recommendation of actions. Also improve the existing security solutions. Example, risk based and adaptive authentication. Use the normal behavior of user as passive authentication, etc. We can measure it and use it as an additional factor of authentication.
- Defining AI. Translating intelligence behavior in algorithms. Replicate human like behavior in algorithms. Showed the diagram where CS - DS - AI - ML - DL evolves passing through data science.
- Explained the paradigm of ML, you will get whatever you feed the algorithm.
- Success factor explained in terms of performance of systems and algorithms.
- The output of models was mentioned such as classification, numerical values or binary values. How results can be.

## Machine Learning

- Classification under supervised and unsupervised.
- Goals of both. It could be regression and classification. Example, spam detection. Models to use could be SVM, kNN.
- SVM is explained more in depth. First during training time. Goes through the email spam detection example.
- Now unsupervised learning. Searching for patterns in data. Models used are mostly clustering. K means, hierarchical clustering procedures. How it works? We have any data, and have to find distance between points, define number of clusters (k) and start by randomly assigning elements to clusters. Brief explanation and examples showed with k means and hierarchical clustering procedures.

## Deep learning

- Explanation of what it is. An important aspect is that allows incomplete and noisy data to be processed.
- Showed architectures such as: CNN and LSTM.
- It’s important to have knowledge which algorithms has to be used for which topic

## Application examples of AI for security

- For online banking. We don’t want to just detect the attack but we want to see if there is an increased risk situation. To create prevention data, input data is used from news, media, emails, webpages, national vulnerability databases…
- Application implemented: alert system using AI for online banking. The analysis involving ANN, and other models were displayed along with results and comparisons of all the models. Some sort of early detection system.
- 2. Application analyzed: Passive authentication – XignQR : data is acquired from accelerometer and position sensors, which are included in smart phones

- A few more examples were given: malware classification, threat intel, siem, etc.

### Attacks to ML

- A few aspects were introduced such as the quality of implementation. Also it was mentioned that state of the art security measures need to be used and security goals need to be guarded such as: confidence, integrity, availability.
- One attack explained: manipulation of training data. Incorrect labeled data is injected into the training process. That can lead to wrong classification.
- Manipulation of traffic signs. Some adversarial mentioned without saying the word adversarial but showed a paper where the 80 sign could be misclassified as stop sign.

### Further challenges

- Attackers can also use AI. Vulnerability search, social engineer (chat bots), deep fake videos. Example with Obama talking and Putin smiling.
- General challenges: Personal data, GDPR.
- Discrimination. Sex, origin of individuals as bias factor.
- Traceability of results is important: Results from AI must be understand as recommendation for the user. Traceability supports trustworthiness which gives the user the confidence to use AI systems
- Some research questions addressing availability of data. Currently 95% is in USA and 4% in EU
- Can we use technology from other countries without sovereignty?
- We need a powerful AI infrastructure to maintain digital sovereignty

### Questions / Comments

- Importance of working on European level instead of national level. Some countries did not deliver the money they promised. Example, Germany promised 3 billions and invested only 500k. Additionally including non-EU countries such as Switzerland and Israel.
- Data needs to be good but sometimes we cannot change that. So we need to evaluate that the quality of the data is good.
- Reducing the information we process.

## Summary of the panel discussion:

### Political issues

- Political motivation in Europe vs. China, Russia, USA
- AI development on European scale rather than national
- More coordination and budgets for AI in Europe (including non EU countries – e.g. Switzerland and Israel)
- Target and purpose for AI in Europe compared to China and US should be defined
- China invests large sums of money in AI but sometimes the goal is related to surveillance unlike what we do in the EU
- Generating large European programs with clear targets for AI
- A lot of focus and priority for AI in Saarland

### Capability of ML and AI

- Depend on data availability, evaluation and reduction
- Would a machine be able to win a chess game?
- AI example: deep blue against a grand chess master

- Strong improvement in the past 10 years in e.g. autonomous driving - same expected in CD (cyber defense)
- Autonomous cars & liability? – Build up trust is very important and a matter of time
- AI for instance is better to analyze large amounts of data compared to humans (e.g. Law enforcement in Baden-Württemberg)

## Security of ML

- How to evaluate quality of data?
- Preventing adversarial attacks on neural networks
- Manipulation of data as possible surface for attacker
- How to improve trustworthiness
- Data generation by machines
- Traceability of AI systems (how is decision made?)
- Possibility of overruling AI decision
- Trust in AI
- Do we have a chance to control systems in the future?
- AI will be a game changer of cyber security

## Ethics aspects for AI

- How do we want to live in the future?
- The paradox of face recognition in San Francisco, CA; from starting with the technology and now moving to stop it
- Can we use person related data to reach higher goals with AI (e.g. protection of a power plants and critical infrastructure)?
- Who is supposed to make the decision?
- Problem of bias in information (which attack is more important? e.g. from US or China)
- Can ethics be included in AI models? (possible solution: human in the loop)
- Ethics depend on the culture
- Ethical issues are not addressed so far – needs to be addressed by politicians, industry and society
- Learn from human decision making
- Who will be responsible for the decision AI has been made? E.g. autonomous driving or autonomous strike back
- Problem of bad marketing for AI
- Trust in technology is below 100%
- It is also a question of experience
- Better explanation of AI might improve acceptance
- Create a label / AI Seal (e.g. TÜV in Germany)
- Show benefits of AI