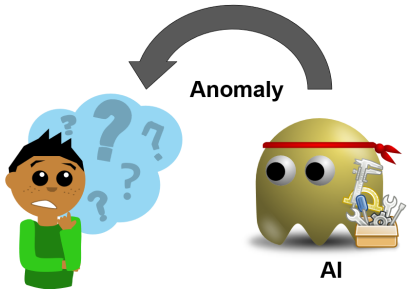
Chair of Network Architectures and Services
Department of Informatics
Technical University of Munich

Thesis M.Sc.

IDP

Explainable AI for Anomaly Detection

Motivation



Anomaly detection is a common tool to secure networks and to spot security incidents. The use of machine learning instead of signature-based approaches offers benefits. Examples are automated training of benign models and the detection of unseen attacks (zero days). However, a drawback especially of neural network-based models is the lack of interpretability. In such cases, the cause for the detection of an anomaly is not given and cannot easily be found. This complicates the analysis of an anomaly through humans or algorithms that work on top of the anomaly detection. As such, the detection of the underlying attack or the classification as false positive might not be possible.

Topic

The goal of this thesis is to create anomaly detection models that allow an interpretation of the result. In addition to the detection of an anomaly, further information about the input that caused the anomaly should be given. The information may contain the causing feature and how it derivates from the expectation. This goal should be achieved through combination of preprocessing data and application of existing machine learning techniques.

Your Task

- Analysis of machine learning models that allow interpretation of anomaly causes
- Implementation of a demonstrator
- Evaluation based on provided datasets

Requirements

- Basic network knowledge
- Ability to write maintainable code
- Knowledge in Machine Learning & Python

Sources

- [1] K. Amarasinghe, K. Kenney and M. Manic, "Toward Explainable Deep Neural Network Based Anomaly Detection," 2018 11th International Conference on Human System Interaction (HSI), Gdansk, Poland, 2018, pp. 311-317, doi: 10.1109/HSI.2018.8430788.

Contact

Christian Lübben luebben@net.in.tum.de
Lars Wüstrich wuestrich@net.in.tum.de
Holger Kinkelin kinkelin@net.in.tum.de
http://go.tum.de/910206