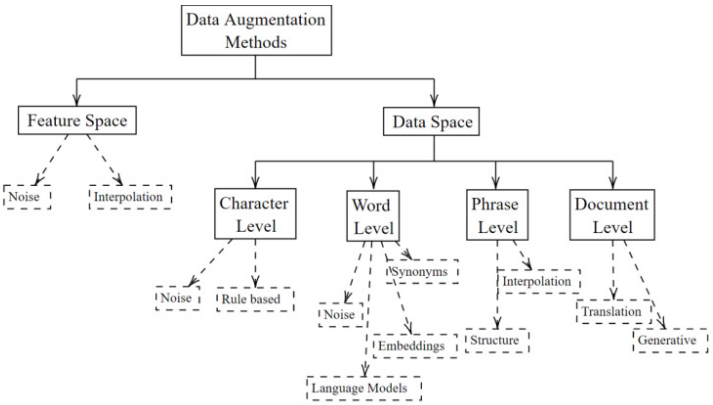
Hessisches Kompetenzzentrum
für Hochleistungsrechnen

CYWARN – Entwicklung von Strategien und Technologien zur Analyse und Kommunikation der Sicherheitslage im Cyberraum II



Project Manager
Markus Bayer

Researchers
Philipp Kühn, Tobias Frey and David Relke

Principal Investigator
Prof. Dr. Christian Reuter

Project Term
2022 - 2023

Clusters
Lichtenberg Cluster Darmstadt

Software
TensorFlow, PyTorch

Institute
Department of Computer Science

University
Technische Universität Darmstadt

Introduction

The increase in complex cyber-attacks illustrates the vulnerability of society and information infrastructure. In addition to technologies for information and IT security, early warning systems and reaction strategies are needed to strengthen civil security. So-called Computer Emergency Response Teams (CERTs) are the central point of contact for preventive and reactive measures in the event of IT security incidents. Due to the confusing information situation in cyber-attacks, the evaluation and target group-specific preparation is a major challenge for these teams. The aim of the CYWARN project is to support CERTs with new strategies and technologies for recording, analysing, and communicating the cyber situation. A demonstrator is created that enables the automated collection of public data sources as well as data evaluation with credibility analysis and information prioritization. The use of resource intensive Machine Learning and especially Deep Learning methods is necessary for the evaluation.

Methods

For the analysis of the data sources textual Deep Learning methods are used. These methods are, for example, enriched by sophisticated Data Augmentation methods to improve their prediction quality. In addition, Active Learning can be used to guide the labelling process by incorporating a sensible strategy. This results in a better performing model with less data. Furthermore, novel Few-Shot-Learning methods are developed for the quick training of new classifiers. These Few-Shot-Learning approaches perform best if they are based on large pre-trained models, that require many resources. As a novel approach,

metadata-based indicators, and explanatory deep learning (XAI) will be combined in order to establish transparency for algorithmic decisions and to investigate its influence, acceptance and user confidence in the algorithms and the entire CYWARN demonstrator, in addition to optimizing the classification quality compared to existing black-box approaches. Additionally, we test the Deep Learning models regarding Adversarial Attacks. The developed Adversarial Examples are then used to make the models more robust about their decision processes.

Results

In the meantime, many results have already been achieved with the help of the Lichtenberg Cluster. In the first phase, we developed and tested several Data Augmentation methods, from which we derived that a textual generation method performs best for artificially increasing a dataset. Based on this we conducted a large-scale data augmentation survey, in which the Lichtenberg Cluster helped to test the various methods. Furthermore, we trained a large language model on a Cybersecurity dataset which builds the basis for further classifiers. This language model is, for example, used in the Few-Shot-Learning setting that was the next goal of this project. Our novel Few-Shot-Learning method reaches state-of-the-art results which are highly relevant in the CYWARN project. In the following phase of the project, we will also develop a novel Adversarial Attack method that is able to find the weak spots of the trained models. The last part of this project is about Active Learning, where a strategy is designed which decides which examples the annotators must label and which not.

Discussion

The current results of the project are very promising. The Data Augmentation survey gives an in-depth overview for all researchers working with textual data and our developed Data Augmentation method produces very good performance improvements. Furthermore, the cybersecurity-specific language model, CySecBERT, is very important for the research field as it can be used with many different tasks and in various contexts. Especially, we show that it is more suitable for cybersecurity tasks than general language models. This language model is also the basis for the Few-Shot-Learning approach that reaches the state-of-the-art results. However, one must acknowledge that the research of Natural Language Processing and Machine Learning in general strives more and more towards much larger models, which may achieve even better performance. For both research directions, the next steps of the project with the Adversarial Attack method and the Active Learning strategy are highly important.

Figures

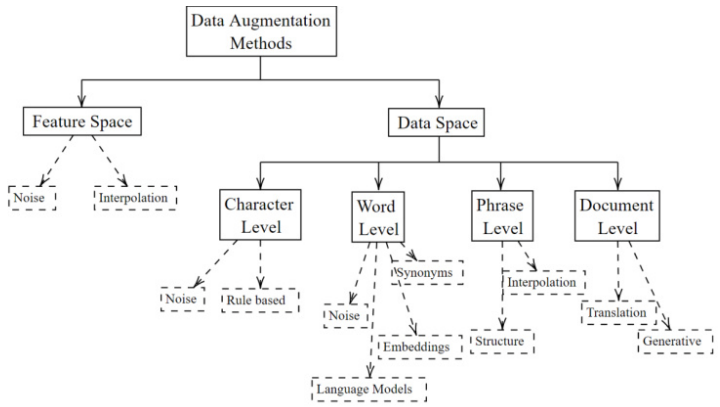


Figure 1: Taxonomy of different data augmentation methods for textual data.

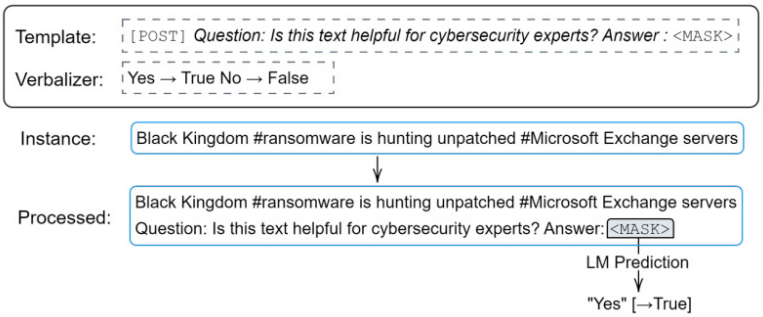


Figure 2: Depiction of on step in the Few-Shot Learning process.

Publications

Bayer, M.; Kaufhold, M. A.; Reuter, C.: "A survey on data augmentation for text classification." ACM Computing Surveys, 55(7), 1-39, (2022). https://doi.org/10.1145/3544558

Bayer, M.; Frey, T.; Reuter, C.: "Multi-Level Fine-Tuning, Data Augmentation, and Few-Shot Learning for Specialized Cyber Threat Intelligence." (2022) https://doi.org/10.48550/arXiv.2207.11076

Bayer, M.; Kuehn, P.; Shanehsaz, R.; Reuter, C.: "CySecBERT: A Domain-Adapted Language Model for the Cybersecurity Domain." (2022) https://doi.org/10.48550/arXiv.2212.02974

Last Update: 2023-04-04 15:36