RAFAŁ CHĄCPNIK

Abstract of the doctoral thesis

„Applications of Simplified Brain-Computer Interfaces in Cybersecurity and Emotion Recognition"

The main purpose of the thesis is to present the original solution to the scientific problem of exploring the applications of simplified brain-computer interfaces in cybersecurity and emotion recognition, especially in terms of evaluating the potential of such interfaces within the field of stress detection and subject recognition. It has been done by building stress detection and subject recognition classifiers based on the data provided by the Neurosky MindWave Mobile 2 interface and evaluating the results, with cybersecurity applications in mind, and what the metrics of such classifiers are. Considering the participation of human participants, the research has been approved by the Committee for Research Ethics.

Two classifiers have been built to evaluate the possibilities of using a simplified EEG interface for stress detection in a reproducible manner, by fine-tuning the parameters of the model, exploring the preprocessing approaches, and evaluating the classifier by its metrics and the improvement ratios for all available EEG bands, along with the Attention and Meditation features, to define the impact of a specific feature on the performance of the model.

The classifier built upon the "Stress" dataset achieved the F1 score of 0.921606 with precision of 0.857198 and recall of 0.996541, while the classifier built upon the "Login" dataset achieved the F1 score of 0.945356 with precision of 0.901603 and recall of 0.993748. The analysis of the confusion matrices produced by both models indicates that it is possible to use simplified EEG interfaces in stress detection, as both models tend to skew the results into false-positive area, which in terms of cybersecurity makes it more desirable than skewing into false-negative predictions.

Two other classifiers have been built to evaluate the possibilities of using the simplified EEG interface for subject recognition. Contrary to the stress detection recognition study, these two classifiers were, respectively, binary model calculated for every subject based on the Fast Tree classifier, compared to the one-for-all multi-class model based on the LightGBM classifier. The methodology used to develop these classifiers was similar to that in the stress detection study.

Both classifiers have been evaluated by calculating their metrics and improvement ratios for all included features. The binary classifier achieved the F1 score of 0.927333 with accuracy of 0.984822, while the multi-class model achieved the macro accuracy of 0.929150 and micro accuracy of 0.929454. Such measure values indicate that it may be more viable to create a model per subject in commercial setup, as the system would not have to relearn every subject when there is a new subject joining the group of users.