ABSTRACT

Application of multimodal neural networks in solving problem of labeling bug reports

Author: Łukasz Chmielowski, MSc, Eng.

Under the supervision of:

Prof. Robert Burduk, DSc, PhD, Eng.

Michał Kucharzak, PhD, Eng.

In large scale software development organizations, there are often deployed complex processes related to the handling of software bug reports. Such documents about system malfunctions usually contain a title and a description of the discrepancy in the behavior of the solution in relation to the expected one in the form of natural language. In addition, system information is attached in raw format or processed by tools that analyze it. The main problem is assigning a problem report to the proper organizational unit and determining whether a given bug report is related to a security risk, memory issue, or performance issue. The data used in the experimental studies comes from a set of reports of bugs from telecommunications operators or were reported inside a company which develops telecommunications equipment including software for it. Dossier concerns base transceiver stations (BTS).

The research confirmed the thesis stated that *there exists a method for automated assignment of a software bug report to appropriate development group, responsible for resolving the bug, which outperforms well-known methods for bug report assignment.* A literature review was conducted. One of the important conclusions was to define the research gap which indicates the drawbacks of the currently used validation techniques and to propose an alternative one devoid of the above-mentioned negative features. Experiments using protocols utilizing different ways of

building training and test sets showed a significant difference in results. An analysis of the possibility of using explainable artificial intelligence (XAI) has been performed.

A number of studies were carried out, both related to preprocessing, vectorization of source data and at various levels of the company's structure. As part of the implementation of the doctorate, a novel method of assigning reports in the context of the organization's composition has been proposed. On the basis of research on the effectiveness of selected machine learning algorithms, a decision was made, and a pilot solution was implemented in the company. Due to the fact that it was possible to collect both the predictions of the machine learning model and the decisions of humans under certain conditions, it was decided to conduct a comparative analysis of the results.

The applicability of multimodal neural networks as well as other ways of solving the problem of assigning reports was investigated. The key aspect was to prepare an appropriate representation of the input data as well as to design the required multimodal neural network architecture. The conducted studies showed the superiority of this method compared to the reference methods used.

The attestation about implementation works at the company is placed in Appendix C. It contains information about the ongoing work related to the introduction of a multimodal neural network. Its first versions are already implemented. Currently, work is underway related to a better selection of features and dealing with missing data. In addition, it provides information on other implementations related to software bug reports.

Łukasz Arendowski