

Opponent Review

for the master thesis

"Text classification of customer inquiries via e-mail"

Reviewer: Prof. Dr. Phillipp Torkler
Student: Armin Weigold
Matriculation No: 00629413

Summary

Mr. Weigold's thesis aims to evaluate the performance of classifying customer emails with support requests to shorten response times to requests. The thesis was performed at the company Check24, an online service for the comparison and order of insurance, finance and public providers for gas and electricity. A labeled dataset of around 400k emails and 10 classes was provided by Check24. Text classification performances are evaluated for random forests, support vector machines, graph neural networks and feed forward neural networks.

Review

The thesis is structured in a coherent and logical fashion, adheres to common scientific standards, and research objectives are stated. The evaluation of the chosen classifiers is driven by strong class imbalances in the dataset.

The thesis could be improved by the following points:

- The subsection related work is too brief and could have been more detailed.
- Using subsets or random splits of the dataset could have been a potential way to address the memory issues of the Tf-idf vectorization rather than completely avoiding Tf-idf.

- The evaluation would benefit by applying cross-validation. Only a single 20/80% split was used for evaluation.
- Class imbalance problem is not addressed at all and the reported results are clearly affected by class imbalance effects.
- More explorative work could have been performed on the dataset.
- 4.3.2.1 mentions too long computation times for GNNs and thus a smaller dataset was used only for GNNs. A compute node on campus could have been used to address computation issues and to make an unbiased comparison the same datasets should be used for all classifiers.
- To foster good scientific practice, the code could have been provided via common repositories like GitHub.

In conclusion, I recommend rating the thesis as *fair*.

Deggendorf, 01.09.2023

Prof. Dr. Phillipp Torkler