

Supervisor report

for the master thesis

“Text classification of customer inquiries via E-Mail”

Student: Armin Weigold

Supervisor: Prof. Dr. Andreas Fischer

Submitted on: May 05th, 2023

Summary

The thesis of Mr. Armin Weigold applies text classification to incoming e-mails in order to preprocess customer requests. Incoming requests are labeled with one of 10 labels, thereby allowing later support workers to react accordingly, and, in summary, reducing turnaround time by customer support. The thesis was carried out in cooperation with the „CHECK24 Vergleichsportaal Energie GmbH“. The corresponding data set consists of real production data. It was provided by the company.

Evaluation of the content

Mr. Weigold compares multiple machine learning approaches in a supervised approach, learning from a data set of more than 400k pre-labeled e-mails. The data set suffers from heavy label imbalance. Unfortunately, this is mostly glossed over by Mr. Weigold. Better handling of the label imbalance would likely have improved the results considerably.

Text classification is a widely investigated area. In that respect, the discussion of related work provided by Mr. Weigold remains at the minimum. The described novelty of his contribution is vague, at best.

The machine learning approaches presented by Mr. Weigold are discussed in appropriate detail. The evaluation of results supports the conclusions drawn by Mr. Weigold, although in some cases the results seem incomparable.

Structure and quality of presentation

The structure of the thesis is clear, following common scientific standards, and the text is readable without major language problems. Mr. Weigold makes good use of figures to summarize key results. Tables are also used appropriately in most cases, although a few tables could have been explained better.

Collaboration with the student

Mr. Weigold worked in a quite self-organized way. While he solved problems on his own, he was not very communicative. A presentation of his results was delayed to the last possible date, leading to some timing issues in the end. Only some of the remarks provided could be included in the final version of the thesis. Nevertheless, Mr. Weigold still presents valid results.

Result

Overall, I rate the thesis as: **3.3 (befriedigend, fair)**

Deggendorf, 9. August 2023

Prof. Dr. Andreas Fischer