

Bachelor Thesis Review

Faculty of Mathematics and Physics, Charles University

Thesis author František Kmječ
Thesis title Methods of User-Assisted Summarization of Meetings
Year submitted 2023
Study program Computer Science
Study branch General Computer Science

Review author Ing. Zdeněk Kasner Reviewer
Department Ústav formální a aplikované lingvistiky

Overall

good OK poor insufficient

| | good | OK | poor | insufficient |
|--|------|----|------|--------------|
| Assignment difficulty | X | X | | |
| Assignment fulfilled | X | | | |
| Total size <i>... text and code, overall workload</i> | X | | | |

The thesis of František Kmječ tackles the task of automatic minuting, i.e. capturing the core points from an online meeting. The thesis presents two core contributions: (1) a web application for interacting with inputs and outputs of automatic minuting models and (2) a set of experiments with training and evaluating automatic minuting models.

The first contribution of the thesis is *Minuteman* – a web application for interactive automatic minuting. The interface of the application is built on a collaborative text editor in which the users can interact in real time with voice transcriptions and generated minutes. As a second contribution, the student describes experiments with finetuning and prompting Transformer language models which are used as a backend for *Minuteman*. Both the web application and the models are evaluated in live meetings.

The assignment combines software engineering with research work. For the research part, the student had to get familiar with the current techniques of automatic summarization and minuting using neural Transformer-based language models, which is a difficult and currently unsolved topic undergoing rapid changes in terms of available models and model performance. Although more advanced models proposed by the author were not able to beat the baseline, the fault was mostly in unsuitable training data quality and the baseline model was successfully deployed for the web application backend. The web application then consists of multiple open software frameworks appropriately adapted and combined for the task in question.

I consider all the points in the assignment as fulfilled, generally very consistently and thoroughly. Given the extensiveness of implementation and originality of the experiments, the scope of the thesis exceeds the requirements for a bachelor thesis. On top of that, I appreciate that the work has led to follow-ups, namely a system demonstration paper for *Minuteman* submitted to the IJCNLP-AACL 2023 conference and the baseline minuting model submitted for the AutoMin 2023 Shared Task. Therefore, I recommend this thesis for defense and suggest nominating the thesis for a special award.

Thesis Text

good OK poor insufficient

| | good | OK | poor | insufficient |
|---|------|----|------|--------------|
| Form <i>... language, typography, references</i> | X | | | |
| Structure <i>... context, goals, analysis, design, evaluation, level of detail</i> | X | X | | |
| Problem analysis | X | | | |

| | | | | |
|--|---|--|--|--|
| Developer documentation | X | | | |
| User Documentation | X | | | |
| <p>The thesis is written in fluent and almost flawless scientific English; the text is readable and well-structured. The problem is made clear from the beginning and the theoretical part gives the necessary background for the following experiments.</p> <p>My only major issue is with the evaluation. Overall, the evaluation is not very thorough (even though it is spread between two chapters). The automatic evaluation in Section 5.5 looks a bit hurried (including Table 5.3 with unexplained abbreviations) and the accompanying text suggests that the best model was selected by subjective observations instead of (quite minor) differences in the results. On top of that, the observations and feedback from the users using the <i>Minuteman</i> system, presented in the next chapter, are not accompanied by any quantifiable results. Given the focus of manual evaluation – which I agree with – I would expect also a proper error analysis on a subset of examples or live meetings, preferably with multiple annotators and multiple backend models.</p> <p>As for the minor issues, the large language models introduced in Section 5.4.2 are not properly described in the theoretical part, but this can be understood by the speed of recent developments in this area. More space could be also devoted to introducing a large number of utilized software frameworks (Jitsi, Flask, RabbitMQ, TorchServe, Etherpad), that the reader may not be familiar with.</p> <p>Overall, I consider the form and structure of the thesis appropriate and well done. The student has done a very good job of summarizing the research area and introducing the experiments, generally with the right level of detail. The insights from user feedback suggest that the main limitation of automatic interactive minuting is a different style of summaries produced by the model compared to what users request from meeting minutes, namely contextualized outcomes instead of a literal summary of the meeting. These insights, together with the problem analysis and design of the experiments, are useful for follow-up work in the area.</p> | | | | |

Thesis Code

good OK poor insufficient

| | | | | | |
|--|---|---|--|--|--|
| Design | <i>... architecture, algorithms, data structures, used technologies</i> | X | | | |
| Implementation | <i>... naming conventions, formatting, comments, testing</i> | X | | | |
| Stability | | X | | | |
| <p>In the implementation part, the student has created a comprehensive framework for interactive automatic minuting. The software is documented in the appendix of the thesis and the code is publicly available on GitHub. I appreciate that it was built using other open-source frameworks and libraries (Jitsi for online meetings, Etherpad for collaborative editing, etc.) and that it can be reproducibly built as a Docker image.</p> <p>At first, I was not able to access the live demo of the application at https://minuteman.kmjec.cz/ but the problem has been promptly solved after communication with the student. I was also able to install and run the Docker image according to the Github README. The quality of code is high and it is apparent that it has been continuously worked on over a longer period of time.</p> | | | | | |

Overall grade Excellent
Award level thesis Yes

Date 15.8.2023

Signature