FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

Review report of a final thesis

Reviewer:	Rodrigo Augusto da Silva Alves, Ph.D.
Student:	Patricie Petriľáková
Thesis title:	Automatic categorization of job ads
Branch / specialization:	Artificial Intelligence 2021
Created on:	12 June 2023

Evaluation criteria

1. Fulfillment of the assignment

- ▸ [1] **assignment fulfilled**
  [2] assignment fulfilled with minor objections
  [3] assignment fulfilled with major objections
  [4] assignment not fulfilled

The thesis proposes the development of a classification framework specifically designed to automatically classify job advertisements within a web system. The thesis demonstrates a practical approach by utilizing advanced techniques in natural language processing and machine learning. The student successfully fulfills the assigned objectives.

2. Main written part 80 /100 (B)

The thesis is presented in a clear and accessible manner, although there are certain sections that could benefit from a more formal description. I appreciate the overall structure of the thesis and the effort made by the student to ensure its readability. While I would have liked to see a more comprehensive positioning of the thesis within the existing literature, I acknowledge that since it is focused on applied work, this particular limitation is not significant.

3. Non-written part, attachments 95 /100 (A)

The code appears to be comprehensible and sound, as the files are well-organized and easily understandable. Additionally, the inclusion of a "read-me" file greatly assists in the understanding of the overall structure of the codebase.

4. Evaluation of results, publication outputs and awards 90 /100 (A)

The obtained results are satisfactory; however, there is potential for improvement by comparing them with other existing approaches. Another possibility would be to conduct a comprehensive comparative analysis, considering different modern modules, which could yield scientifically valid results that could potentially lead to a publication. In terms of practical application, I believe that the developed solution is applicable in real-world scenarios, which would represent a significant achievement for a bachelor's thesis.

The overall evaluation 90 /100 (A)

The thesis successfully fulfills the requirements, as it is well-organized and exhibits a solid experimental results.

Questions for the defense

How applicable is your model in a real-world scenario, and what are the primary infrastructure challenges associated with its implementation? Additionally, how quickly can it be adapted to an online setting?

The student claims that the model is unbiased. However, it is widely recognized that machine learning classifiers can be influenced by bias introduced during the data collection process. Can you discuss the biases that the model focuses on reducing?

Instructions

Fulfillment of the assignment

Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Main written part

Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies?

Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 52/2021, Art. 3.

Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Non-written part, attachments

Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Evaluation of results, publication outputs and awards

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

The overall evaluation

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.