**FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE**

Review report of a final thesis

Reviewer:	Ing. Luděk Kopáček, Ph.D.
Student:	Tomáš Bánhegyi
Thesis title:	Product Compatibility Detection from Product Description
Branch / specialization:	Knowledge Engineering
Created on:	5 June 2022

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 reviewed bachelor thesis aims to leverage Natural Language Processing techniques of Artificial Intelligence to extract the compatibility information from the text description of each product.

The content of the presented bachelor thesis is aligned with the primary assignment to identify the compatibility information from the free text description. The only small difference I noticed is a missing review of the question Answering, which was specified as a subgoal in the assignment.

However, as the author explained in conclusion, after discussion with the supervisor, they decided to omit the Question Answering with the justification that "using named entity recognition and relationship extraction is sufficient to fulfill the main objective, extraction of product compatibility." I agree with this justification, and I state that the reviewed bachelor thesis fulfilled the assignment.

2. Main written part   80 / 100 (B)

The theoretical part is covered in chapter 2, where Mr. Banhegyi reviewed Natural Language Processing techniques, specifically Named Entity Recognition and Relationship Extraction. The scope of chapter 2 corresponds to the practical part. However, the descriptions of certain topics are brief and deserve more focus and detail. E.g., there is a missing description of inputs in figure 2.8 on page 9.

Chapter 3 examines alternative datasets and the reasons why and how the author created his own dataset.

Chapter 4 covers model architecture. The introduction of this chapter contains an example of the Spacy pipeline adapted from the cited source. However, the chapter

doesn't contain any diagram of the model used for the practical part. Therefore, a reader has to imagine the Model architecture from the description in chapters 4.1 and 4.2. A visual representation of the described model would significantly help the reader understand this part of the thesis.

Chapters 5 and 6 cover implementation details, training, and evaluation. Subchapter 6.4 deserves a better explanation of the changes that led to the REL model's significant improvements in table 6.5.

The final conclusion in chapter 7 summarizes the achieved results, goal attainment, and discussions about possible future improvements.

The literature citations are well used in the work and are relevant.

Overall, the written part of the presented bachelor thesis is well-structured and comprehensible, but on the other hand, too brief and could have better presented the accomplishments of the practical part.

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

The attachment included in the thesis contains the electronic version of the thesis, used modeling data, Python source code, trained models for named entity recognition (NER) and relationship extraction (REL). The attachment is well structured and contains "readme.txt" files that simplify the viewer's orientation. The "requirements.txt" file with the related Python dependencies supports the reproducibility of the achieved results. In addition, the employed technology and tools are appropriate to the thesis topic.

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

The motivation for detecting product compatibility relationships from text data was obviously to reduce manual work to extract the compatibility information in a structured way.

From this perspective, the presented model can be used in practice when it's retrained to a particular domain. The structured compatibility data is a highly valuable additional input for subsequent models, e.g., product recommender to recommend mainly compatible products.

The additional benefit of the work is in creating the benchmark dataset, which can be used to compare further models designed for the same type of task.

The achieved F-score of 62.3% for the classification of the compatibility relationship is a very good starting point and could be further improved, as the author described in the conclusion chapter.

The overall evaluation   90 /100 (A)

In summary, the presented bachelor thesis fulfilled the assignment. The selected topic is relevant, and the presented results can be used in a real application. Moreover, the author put a lot of effort into the practical part, including data preparation and model training which is well documented in the attachment of the thesis.

Therefore, I recommend this thesis for defense and I propose the grade of A.

Questions for the defense

Could you illustrate the used model architecture (e.g., as a Spacy pipeline) and briefly describe it?

The results of the REL model significantly improved when a "new dataset" was used (table 6.5 on page 27). Could you describe the changes you made in the dataset and why they improved the results?

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.