

Short report

on the thesis of Momchil Emilov Hardalov

“Intelligent Context-Aware Natural Language Dialogue Agent”,
presented for awarding the educational and scientific degree “doctor”
in Professional area 4.6 Informatics and Computer Science

Doctoral program: “Software Technologies” – Knowledge Discovery,
Faculty of Mathematics and Informatics (FMI),
Sofia University St. Kliment Ohridski (SU)

by Prof. Dr. Maria Nisheva-Pavlova, SU, FMI

Pursuant to Order ПД 38-600 /28.10.2022 of the Rector of SU I was appointed a member of the scientific jury for the defense of the submitted thesis in professional area 4.6 Informatics and Computer Science, entitled “Context-Aware Conversational Agent for Multi-Source and Multilingual Task-Oriented Dialogue”.

1. General characteristics of the dissertation and the presented materials

The dissertation is written in English. It contains 163 pages of text, including six chapters, list of references with 350 titles and two appendices.

In addition to the dissertation, the following documents are also presented:

- abstract in Bulgarian and abstract in English;
- folder with personal documents of the doctoral student, containing
 - curriculum vitae,
 - copy of diploma for awarded Bachelor’s degree,
 - copy of diploma for awarded Master’s degree,
 - reference for the exams passed and the ECTS credits accumulated within the doctoral studies,
 - list and copies of publications on dissertation results,
 - list of noticed citations of the dissertation publications,
 - documents according to the procedure for verifying the originality of the dissertation,

- report on the fulfillment of the minimum national requirements under Art. 26, para. 2 and 3 of the Law for the Development of Academic Staff in the Republic of Bulgaria.

2. Applicant data

Momchil Hardalov has a Bachelor's degree, completed at TU - Sofia, and a Master's degree, completed at SU – FMI, Master's program in Information Retrieval and Knowledge Discovery. He has significant professional experience in developing projects in the fields of information retrieval, natural language processing and machine learning. For more than seven years, he has served as a teaching assistant in Information Retrieval, Knowledge Discovery in Data and Knowledge Discovery in Text for several Master's programs of FMI.

I have no immediate personal impressions of the applicant's work, apart from his fruitful participation as a co-supervisor and thesis consultant of students from the Master's program in Artificial Intelligence at FMI.

3. Analysis of the content, results and contributions of the doctoral thesis

The doctoral thesis consists of eight main parts: six chapters and two appendices.

The introductory first chapter presents in brief the research area, motivates and formulates the goals and specific research tasks of the dissertation. The structure of the dissertation is outlined and the content of its remaining chapters is summarized. A list of incomplete bibliographic data of the candidate's publications on dissertation results is included as well.

In the second chapter, which has the characteristics of an analytic overview, a wide variety of state-of-the-art models of conversation agents are discussed.

In Chapter 3, a new natural language understanding method that jointly solves the tasks of intent detection and slot filling is presented and analyzed.

In Chapter 4, a number of approaches are proposed for curating answers from external knowledge sources.

Chapter 5 is devoted to exploring a series of advanced conversation methods with an emphasis on methods that rely on additional information to determine an appropriate response to a given user question.

The concluding sixth chapter contains a formulation of the thesis' contributions and directions for future research on the topic.

The main scientific and applied scientific contributions of the doctoral thesis of Momchil Hardalov may be formulated as follows:

- A new pre-trained language model is proposed for joint modeling of intent detection and slot filling tasks.
- An original approach for training from noisy data that uses using self-adaptive learning and additional weights in the error function is developed and investigated. A performance improvement is achieved compared to alternative methods based on training on fully manually annotated data.
- An end-to-end approach to solving the task of multiple-choice reading comprehension for low-resource languages is designed.
- A new approach for ranking candidate responses from conversational agents is developed and evaluated.
- A new Bulgarian corpus for multiple-choice reading comprehension is created.
- A large-scale dataset for detecting previously fact-checked claims is created.

The doctoral thesis makes an excellent impression with the scope, depth and argumentation of the presentation. The field of research is modern and complex, and achieving significant results in it requires serious fundamental and specialized training, constant and intensive work. The achieved results are original and significant and fully correspond to the defined goals.

4. Publications on the doctoral thesis

The results obtained in the doctoral thesis are presented in a total of six papers, published as follows:

- one – in the series *Lecture Notes in Computer Science* (SJR 0.283 for the year of publication),
- one – in the journal *Information* (SJR 0.353 for the year of publication),
- one – in a volume of proceedings of the International RANLP Conference (SJR 0.244 for the year of publication),
- one – in an ACM edition, referenced and indexed by Scopus,

- two – in non-refereed online editions.

All papers on this list are co-authored, with declarations of the author teams being presented, in which the personal contribution of the individual co-authors is indicated – the main contribution of the applicant in five of the papers and equal contribution of all co-authors in one paper.

In this way, the minimum national requirements under Art. 26, para. 2 and 3 of the Law for the Development of Academic Staff in the Republic of Bulgaria and the additional requirements of SU for the acquisition the educational and scientific degree “doctor” in professional area 4.6 Informatics and Computer Science have been fulfilled and significantly exceeded.

5. Abstract

The abstract meets all the requirements for its preparation and fully and accurately presents the topic, purpose, content, achieved results and contributions of the dissertation.

6. Critical Remarks and Recommendations

I have no significant critical comments on the form and content of the doctoral thesis, as well as on the formulation of the main results and contributions stated in the thesis and the accompanying scientific papers. I would recommend that for the introductory sections of the individual chapters of the thesis, more appropriate titles related specifically to their content be chosen.

7. Summary

Summing up, I consider that the doctoral thesis of Momchil Hardalov satisfies the requirements of the national regulations and the specific conditions and requirements of Sofia University St. Kliment Ohridski. Its author has achieved significant research results that make an original contribution to the chosen field of study. My assessment of the dissertation, the abstract, the publications and the scientific contributions of their author, Momchil Hardalov, is **positive**.

Therefore, **I advise the honorable scientific jury to award to Momchil Emilov Hardalov the educational and scientific degree “doctor” in professional area 4.6 Informatics and Computer Science.**

Sofia, January 14, 2023

Prof. Dr. Maria Nisheva-Pavlova