To:
Saint Petersburg State University

Dr Ivan Luković, Full Professor
University of Belgrade
Faculty of Organizational Sciences
Chair of information Systems
E-mail: ivan.lukovic@fon.bg.ac.rs
Phone: +38163568749
ORCID: 0000-0003-1319-488X

Belgrade, 18-MAY-2023

Subject: External reviewer report of a Ph.D. thesis

As a nominated external reviewer and a member of the dissertation council established for the defense of the dissertation prepared by **Ershov Vasilii Alekseevich** for the degree of a Candidate of Sciences, Scientific Specialty 2.3.5, Mathematical and software support for computing systems, complexes and computer networks; dissertation title: **EVOLVING INTELLIGENT SYSTEMS**, I submit here the following:

REPORT
about the Ph.D. thesis: Evolving intelligent systems, proposed by
Ershov Vasilii Alekseevich

The following main contributions of this Ph.D. thesis can be identified. First, the candidate proved a need for a high-level approach to the process of building evolving intelligent systems. As such, he proposed a generic framework for building evolving intelligent systems that is based on machine learning techniques. Second, he developed a new family of interpretable metrics for the quality assessment of Automatic Speech Recognition (ASR) systems that consider a wide range of possible changes of user requirements. He compared some well-known metrics with the proposed one and found evident advantages. Third, the author considered various aspects of the complexity of interpretation of the models supported by modern intelligent systems. He selected a class of intelligent systems based on decision tree ensembles and then proposed a new model interpretation technique for such systems, named MonoForest. Fourth, the author considered a problem of balancing the accuracy and speed of an intelligent system not during model training, but during its operation based on the requirements of a particular user. Using the MonoForest framework, the author developed special algorithms to represent decision tree ensembles in a form optimal for the operation of such class of systems.

The author clearly declares mostly the theoretical and, in some extent, practical values of his Ph.D. thesis. As a theoretical contribution in the first part of his thesis, the author proposes the framework of building metrics that makes it possible to adapt ASR systems for specific application scenarios, as well as to determine the general direction of improvement of the whole system. In the second part, the author advocates that the methodology of interpretation of decision trees allows analyzing the results of model application and working on systematic improvement of the whole system. Also, representation of decision tree assemblies in the form of linear equations opens a new way of theoretical analysis of model performance and allows developing new algorithms to analyze this

Јове Илића 154, 11000 Београд, Србија, Тел.: (011) 3950-800, Факс: (011) 2461-221
ПИБ: 100383934, Матични број: 07004044, Текући рачун: 840-1344666-69
Е пошта: dekanat@fon.bg.ac.rs; Посетите: www.fon.bg.ac.rs

33-06-631 от 23.05.2023

type of decision functions. To demonstrate practical values, the author considers in the thesis several examples and gives a discussion of possibilities to apply his approach on a much wider class of intelligent systems, aimed not only at ASR, but also covering other problem domains of an interest. The authors validated his approach by its application on the selected examples and public data sets in the domain of ASR.

After the introductory section, which gives motivation for the research work, the Ph.D. report is organized into 3 chapters, followed by the concluding section, and then 3 appendices. Chapter 1 presents the main object of research: evolving intelligent systems and gives a discussion of how such systems function in practice and what tasks have already been solved and what need to be solved to build them. In this chapter, the author advocates that the mathematical formalizations of learning algorithms are not sufficient to develop applied systems based on artificial intelligence. In this regard, he proposes a new high-level architecture for the functioning of such systems and discusses how this architecture allows applied professionals to determine how an entire artificial intelligence-based service functions together, and for scientists, he discusses how practical tasks relate to current areas of research in the field of artificial intelligence. In Chapter 2, the author considers the problem of assessing the quality of an evolving intelligent system. The author proposes the approach of a quality assessment of an evolving intelligent system for an ASR system. He advocates that the proposed approach and ideas used here can be naturally adapted to other intelligent systems, while the proposed mathematical models and results are specific to ASR systems. The main contribution of this chapter is the theoretical development and implementation of a proof of concept on practical examples of the MERa framework, proposed by the author. In Chapter 3, the author discusses a requirement that for each relatively broad class of models, the methodologies, and specific tools for evolving intelligent systems need to be developed. In this chapter the author selected, as an example, intelligent systems based on ensembles of decision trees as decision functions. For such systems, he developed the MonoForest framework to optimize the application rate of trained decision tree ensembles and give the user more control over the model application process, as well as opening new ways of interpreting decision tree ensembles.

The research results and contributions presented in the Ph.D. thesis are sound, well justified, and valuable. The thesis contains all necessary components required for such kind of research work. Obviously, the author invested a lot of his efforts in proposing novel approaches, methods, algorithms, and finally giving theoretical bases to create high capacities for its further research and possible practical applications. The author demonstrated a solid public visibility of the reported research results through the papers published in the two journals, as well as at the two conferences. The author also invested a significant effort in analysis of related works. However, I have expected in this Ph.D. thesis report a special section (or chapter) devoted to a detailed systematic related work analysis only. In this form of a report, it is somehow not as explicit as it should be. The author was expected to give a critical analysis of the related works, with a clear message, what it is already done in the research area, and what he proposes as a contribution to the state of the art, in comparison to the already reported contributions.

The introduction is well structured and brings valuable information about the main results and contributions of the author's Ph.D. thesis.

Also, I would like to point out some minor concerns about the thesis. The first, about a special section devoted to the related works I have already mentioned. Also, it was not clear from the introduction if MERa is a concept and tool proposed by the author and his advisor, or it is referenced from some other sources. It became a little bit clearer to me in the middle of Chapter 2, but it was not explicitly said in the text, particularly in the beginning of the thesis, as it was expected by my viewpoint. I believe that the author was supposed to introduce MERa with a specific section, giving detailed information about previous development of MERa, and what are the research and development values added just in this thesis.

In some segments, the style of presentation was truly mathematically oriented. For example, Lemma 1 and Lemma 2 were introduced without any prior discussion. Just at the end of their proofs, the author just briefly discussed the role of those lemmas. From my point of view, it should have been done earlier in the text, with more details.

In Figure 2.4, it is not clear if x-axis is annotated with WER or with LER, as in the following text the author referred to LER, while in the figure it is WER.

Chapter 3: "In this paper, we focus on intelligent systems that use ensembles of decision trees as decision functions." –> The author is supposed to give there a strong rationale for such a selection. Why just ensembles of decision trees are selected? There are plenty of methods, and I have expected here to see clear criteria, why such a decision is made in this research. What are the consequences of such a decision for the validity of the whole research? It is not discussed in this chapter, but it is crucial to be.

Page 86: "The first problem can largely be considered as solved at this point." –> Nowhere before the author mentions what is an exact problem. He just listed the three important functionalities, but he did not discuss at all what was a problem and in what aspects the problem existed.

Page 87: "The remaining two problems, interpretation and effective application of the model, are addressed in this chapter." --> Also, it was not qualified before, in what aspects the interpretation and effective application of some models were real problems.

One important question arose to me while I read the thesis, was how Section 2 and Section 3 were mutually related? I had a strong impression that this thesis was just split into the two, almost independent parts, the first (presented in Chapter 2) devoted to the problem of assessing the quality of evolving intelligent systems, and the second (presented in Chapter 3) devoted to the optimization problem of the application rate of the trained decision tree ensembles. What is the exact common value that gives a context to both results, given in Sections 2 and 3, respectively? From my point of view, the author was supposed to give a general context, which makes all those results recognizable as cumulative contributions to the same general goal.

The author put a lot of his attention to the discussion of future work perspectives in his thesis. Just to notify here that Conclusion was also supposed to give an overview of future work in more details. The author should have presented here a general plan of his or someone else's future research, as by defending a Ph.D. thesis the author declares himself capable of being an advisor in some future research activities.

I can say that those my findings do not significantly influence the evident main values and contributions reported in this Ph.D. thesis. By all facts given in this my report, I believe that the author did a great job with high potential for developing future research results and publications, based on the presented Ph.D. thesis.

Conclusion

In conclusion, I declare that the dissertation with the title **Evolving intelligent systems** proposed by **Ershov Vasilii Alekseevich** meets all the requirements established by Saint-Petersburg State University for Doctor of Science degree, and Ershov Vasilii Alekseevich deserves the award of the degree of Doctor of Sciences in Scientific Specialty 2.3.5, Mathematical and software support for computing systems, complexes and computer networks.

Ivan Luković