

Review

for the procedure of defending PhD thesis:

“Semantic Technologies in eHealth (Application of Machine Learning Technologies in Supporting Independent Living of the Elderly and Disadvantaged)”

by

Kristin Ilieva Aleksandrova, PhD student in

The field of Higher Education: **4. Natural Sciences, Mathematics and Informatics**

The professional field: **4.6. Informatics and Computer Sciences**

Doctoral Program: **“Information systems” - Knowledge-based systems,**

Department: **“Computer Informatics”,**

Faculty of Mathematics and Informatics (FMI),

Sofia University “St. Kliment Ohridski” (SU),

Review is prepared by **professor, PhD, Vladimir Todorov Dimitrov** – retired, FMI, SU “St. Kliment Ohridski”, as a member of the scientific committee, appointed with Order № ПД-38-199 / 26.04.2024 of the Rector of Sofia University.

1. General characteristics of the dissertation and accompanying materials

The dissertation: consists of 181 pages and is divided in the following chapters : 1. Introduction – 6 pages; 2. Literature Review – 12 pages; 3. Concept of the AAL System – 8 pages; 4. System architecture and implementation – 37 pages; 5. GDPR Requirements and compliance – 16 pages; 6. Behavioural predictions – 63 pages; 7. Conclusion – 9 pages; and Declaration of Authorship – 2 pages; References – 12 pages; Figures – 3 pages; Tables – 1 page; and three appendixes: Appendix A. OpenRemote’s original docker-compose file deploy.yaml – 1 page; Appendix B. Python script for data formatting for ARM algorithms – 1 page; и Appendix C. Example ruleset in need of rule cleanup – 2 pages.

Referenced bibliography consists of 100 titles – all in English, published during 1990-2024. 16 of them are published in the Internet.

2. Personal data about the candidate

In 2017, Kristin Alexandrova graduated with a bachelor's degree in "Computer Science", in 2019 she defended her master's degree in "Artificial Intelligence", and from 2020 to 2024 she is a doctoral student in the doctoral program "Information Systems - Knowledge-based systems" at FMI at SU "St. Kliment Ohridski".

In the period 2014-2017, the candidate was an assistant at FMI at SU "St. Kliment Ohridski" on "Introduction to Programming", "Object-Oriented Programming" and "Data Structures and Programming".

In the period 2015-2016, Kristin Alexandrova has been on an internship at SAP Labs Bulgaria EOOD, and then on an employment contract from 2017-2022, after which she continues at SAP SE - Walldorf, Germany, where is working now.

I know Kristin Alexandrova as PhD student, as an assistant in the FMI at the SU "St. Kliment Ohridski" and her work, including the development of her dissertation work, the review subject. My personal impressions of Kristin Alexandrova are positive in a professional sense.

3. Content analyses of the scientific and scientific-applied achievements of the PhD student, contained in the dissertation and the publications to it, included in the procedure

Research problems set to the PhD student are formulated as follow:

1. Can we create a cost-conscious Ambient Assisted Living (AAL) System? When talking about cost aware there are several parameters on which this can be judged, which we strive to optimize. This includes the computational resources for running the AAL System and for training the ML behavioral models; the cost of development and support of the system and models; the cost per user of the system for initial onboarding to the system; the cost per person for providing a functional AAL system and up to date machine learning models.

2. Can we create a data privacy compliant AAL System, more specifically a GDPR-aware one? Also, what would be the implication of GPPR on the trained machine learning algorithms?

3. Would an AAL System benefit from enhancing it with a personalized machine-learning algorithm, trained on the collected data that aims to identify outliers in the person's behavior and raise the appropriate alert to their caretaker? If so, what would be the best approach to tackle the problem, without compromising on the cost of the system and the established data privacy requirements?

The following hypothesis was formulated for the research activity:

“We can reuse open-source smart home middleware software to create a cost- and data privacy-aware AAL system, extend it with machine learning algorithms in a useful manner and prove that association rule mining (ARM) algorithms can be used for human behavioral recognition and they would be the better choice compared to standard outlier detection approaches for an AAL system as they are overall cheaper, easier to conform to data privacy regulations and they have explainable results.”

Brief content of the dissertation work:

1. Introduction. The introduction presents: research motivations; topic formulation; the research tasks; hypothesis and research objective. The content of the dissertation is presented.
2. Literature Review. AAL systems are classified here based on their functionality and based on the used technologies; some machine learning models applicable in the context of the present study are noted.
3. Concept of the AAL System. The use case of the AAL system and the target person for assistance are defined. The categories of data required for this class of systems have been investigated.
4. System architecture and implementation. Initially, the criteria for selecting suitable middleware for the purposes of the study are described. After comparing the criteria, OpenRemote was selected. A brief description of the product is presented, and the possibilities for expanding the architecture in the direction of expanding on the tasks of the dissertation are discussed.
5. GDPR Requirements and compliance. The chapter is devoted to the requirements of the standard in the processing and storage of personal data.
6. Behavioral predictions. The chapter discusses self-learning models for detecting human behavior. The two approaches chosen by the PhD student are based on neural networks and associative rules, which are implemented on the chosen prototype. The implementations have been investigated with test data from OpenRemote and the results are summarized.
7. Conclusion. This chapter presents the main contributions of the study, which we discuss below. The author's vision for further research is presented, as well as the approvals made for the obtained results.

Appendix A. Contains the deploy.yaml file for OpenRemote configuration for docker version composition.

Appendix B. Contains a Python script to format data for the ARM algorithms.

Appendix C. Contains an exemplary set of cleaning rules.

The contributions of the dissertation work are presented in three categories: scientific, scientific-applied and applied. They are presented below as presented by the PhD student.

Scientific Contributions

1. Analytical review of the state of art of the field of AAL Systems, the currently existing types of systems, commonly solved problems, and methodology.
2. Analytical review of GDPR and a summarization of the functional requirements an AAL system needs to implement in order to ensure data privacy. The same can easily be translated and applied to different space of systems.
3. Modifications and improvements of machine learning algorithms for the purpose of recognizing human behaviour patterns. Comparison of performance, accuracy, and applicability in the proposed AAL system of neural network and association rule mining algorithms.
4. Development of a methodology to general AAL system creation.
5. Discovery and validated recognition of an optimized Apriori implementation, that provides faster, and as accurate results compared to FPGrowth for behavioural pattern recognition.
6. Proposed method for the transformation of recoded as timeseries daily activities into labelled daily transactions for the purpose of rule generation.

Scientific-Applied Contributions

1. Proposal for a machine learning enhanced AAL system, that uses already existing sensors and devices in a person's home to create a model that recognizes human behavioural routines and can alert to outliers to them.
2. Proposed and implemented architecture of an AAL system, enhanced with machine learning algorithms that ensures data privacy compliance by running as a containerized solution in an isolated network.
3. Experiments on approximately 15 datasets from the CASAS collection of datasets, resulting in a conclusion that ARM algorithms are more cost-conscious and easier to maintain when applied to and AAL system for behavioural pattern recondition.
4. Concept and requirements for the creation of AAL systems focused on the overall costeffectiveness of implementation, long-term support, and cost per person.

Applied Contributions

1. Prototype implementation of an AAL system, enhanced with machine learning models.
2. Prototype implementing the functional requirements of a GDPR-compliant system.

Notes on the contributions:

- Regarding scientific contribution 4, the PhD student has presented an approach to creating an AAL system, but not a methodology.
- Regarding scientific and applied contributions, the research and development base, namely OpenRemote and its architecture, should be considered, so that "predecessors" and "architecture" are understood in this context. Based on these notes, a scientific-applied contribution does not go beyond the framework indicated here.

Taking in account the above findings, I generally accept the presented research contributions.

3. Approbation of the results

The dissertation is based on the next two publications:

1. Aleksandrova, K., Using General-Purpose Instead of Domain-Specific Middleware Platforms for the Creation of an Ambient Assisted Living System. CEUR Workshop Proceedings, ISSN (online): 1613-0073, Vol. 3191, 2022, pp. 237-252. SJR 0.202 (2022).
2. Aleksandrova, K., The Right to Erasure and its Implication on AAL Systems. CEUR Workshop Proceedings, ISSN (online): 1613-0073, Vol. 2933, 2021, pp. 227-235. SJR 0.228 (2021).

After checking and from the presented reference, it can be concluded that:

- a) The scientific works meet the minimum national requirements (under Article 2b, Paragraphs 2 and 3 of the ZRASRB) and, accordingly, the additional requirements of the SU "St. Kliment Ohridski" for the acquisition of an educational and scientific degree "doctor" in the scientific field and professional direction of the procedure;
- b) The results presented by the candidate in the dissertation work and related scientific works do not repeat those from previous procedures for acquiring a scientific title and academic position;

c) There is no proven plagiarism in the submitted dissertation and scientific works under this procedure.

4. Quality of the abstract

The abstract consists of 45 pages divided into five chapters and an appendix as follows: 1. Introduction – 5 pages; 2. System architecture and implementation – 6 pages; 3. GDPR requirements and implementation – 3 pages; 4. Prediction of behavior - 13 pages; 5. Conclusion – 11 pages; Literature - 4 pages. It reflects the main achievements of the research.

There is abstract presented also in English of 41 pages with almost identical content to the Bulgarian version.

The desire for the two abstracts to be identical, almost at the level of sentences, has led to the effect of the automatic translation from English to Bulgarian.

The abstract meets all the requirements for its preparation and correctly presents the results and content of the dissertation work.

5. Critical remarks and recommendations

The presentation of the dissertation allowed the doctoral student to avoid a number of terminological problems that occur in the text of the abstract in Bulgarian, as I noted above.

The assessment by the doctoral student of the contributions of the dissertation work exceeds the actual achievements, which I have noted in the relevant section.

6. Conclusion

After getting acquainted with the dissertation work presented in the procedure and the scientific works accompanying it and based on the analysis of their significance and the scientific and applied scientific contributions contained in them, I confirm that the presented dissertation work and the scientific publications to it, as well as the quality and the originality of the results and achievements presented in them, meet the requirements of ZRASRB, the Regulations for its application and the relevant Regulations of SU "St. Kliment Ohridski" for the candidate's acquisition of the educational and scientific degree "doctor" in the scientific field 4. Natural sciences, Mathematics and Informatics and professional direction 4.6. Informatics and Computer Science. In particular, the candidate satisfies the minimum national requirements in the professional field and no plagiarism has been found in the scientific works submitted for the competition.

Based on the above, I **recommend** the scientific jury to award Kristin Ilieva Aleksandrova an educational and scientific degree "doctor" in scientific field 4. Natural sciences, mathematics and informatics, professional direction 4.6. Informatics and Computer Science.

07.06.2024 г.

Sign:

(professor, doctor Vladimir Dimitrov)