

## **Рецензия**

**по процедура за защита на дисертационен труд на тема:**

***„Семантични технологии в електронното здравеопазване (Приложение на технологии за машинно самообучение за подпомагане на самостоятелния живот на възрастни хора и хора в неравностойно положение)“***

**за придобиване на**

**образователна и научна степен „доктор“**

**от**

кандидат: **Кристин Илиева Александрова,**

Област на висше образование: **4. Природни науки, математика и информатика**

Професионално направление: **4.6. Информатика и компютърни науки**

Докторска програма: **„Информационни системи – Системи, основани на знания“,**

катедра: **„Компютърна информатика“,**

**Факултет по математика и информатика (ФМИ),**

**Софийски университет „Св. Климент Охридски“ (СУ),**

Рецензията е изготвена от: **професор, доктор, Владимир Тодоров Димитров** – пенсионер, ФМИ, СУ „Св. Климент Охридски“, в качеството ми на член на научното жури, съгласно Заповед № РД-38-199 / 26.04.2024 г. на Ректора на Софийския университет.

### **1. Обща характеристика на дисертационния труд и представените материали**

Представеният дисертационен труд: съдържа 181 страници и се състои се от седем глави: 1. Introduction – 6 страници; 2. Literature Review – 12 страници; 3. Concept of the AAL System – 8 страници; 4. System architecture and implementation – 37 страници; 5. GDPR Requirements and compliance – 16 страници; 6. Behavioural predictions – 63 страници; 7. Conclusion – 9 страници; и Declaration of Authorship – 2 страници; References – 12 страници; Figures – 3 страници; Tables – 1 страница; и три приложения: Appendix A. OpenRemote’s original docker-compose file deploy.yaml – 1 страница; Appendix B. Python script for data formatting for ARM algorithms – 1 страница; и Appendix C. Example ruleset in need of rule cleanup – 2 страници.

Представената библиография се състои от 100 заглавия – всичките на английски, публикувани в интервала 1990-2024 г. От тях 16 са Интернет публикации.

## **2. Данни и лични впечатления за докторантката**

През 2017 г., Кристин Александрова завършва висшето си образование степен „бакалавър“ специалност „Компютърни науки“, през 2019 защитава степен „магистър“ специалност „Изкуствен интелект“, а от 2020 г. до 2024 е докторант в докторска програма „Информационни системи – Системи, основани на знания“ към ФМИ при СУ „Св. Климент Охридски“.

В периода 2014-2017 г. кандидатката е асистент към ФМИ при СУ „Св. Климент Охридски“ по „Увод в програмирането“, „Обектно-ориентирано програмиране“ и „Структури от данни и програмиране“.

В периода 2015-2016, Кристин Александрова е на стаж в САП Лабс България ЕООД, а после и на трудов договор от 2017-2022, след което продължава в SAP SE – Walldorf, Германия, където работи и до сега.

Познавам докторантката като асистент във ФМИ при СУ „Св. Климент Охридски“ и работата ѝ, включително на развитието на дисертационния ѝ труд обект на настоящата рецензия. Личните ми впечатления за Кристин Александрова са положителни в професионално отношение.

## **3. Съдържателен анализ на научните и научноприложните постижения на докторантката, съдържащи се в представения дисертационен труд и публикациите към него, включени по процедурата**

Изследователските задачи поставени пред докторантката са формулирани така:

1. Можем ли да създадем рентабилна AAL система? Когато говорим за рентабилна системи, има няколко параметъра, по които може да се прецени това, и които съответно се стремим да оптимизираме. Това включва изчислителните ресурси за управление на AAL системата и за машинното самообучение на поведенческите модели; разходите за разработка и поддръжка на системата и моделите; цената на потребител на системата за първоначално включване в системата; цената на човек за осигуряване на функционална AAL система и актуални модели за машинно самообучение.
2. Можем ли да създадем AAL система, осигуряваща поверителност на данните, по-точно съобразена с изискванията на Общият регламент относно защитата на

данните (ОРЗД)? Също така какво би било влиянието на ОРЗД върху обучените алгоритми?

3. Би ли имала полза AAL система от подобряването ѝ с персонализиран алгоритъм за машинно обучение, обучен върху събраните данни, който има за цел да идентифицира извънредните стойности в поведението на лицето и да отправи подходящ сигнал към техния гледач? Ако е така, какъв би бил най-добрият подход за справяне с проблема, без да се прави компромис с цената на системата и установените изисквания за поверителност на данните.

За изследователската дейност е формулирана следната хипотеза:

„Можем да преизползваме междинен софтуер за интелигентен дом с отворен код, за да създадем AAL система, съобразена с разходите и поверителността на данните, да я разширим с алгоритми за машинно самообучение по полезен начин и да докажем, че алгоритмите за извличане на асоциативни правила (ARM) могат да се използват за разпознаване на човешкото поведение и те биха били по-добрият избор в сравнение със стандартните подходи за откриване на отклонения, тъй като те като цяло са по-евтини, по-лесно изпълняват изискванията за поверителност на данните и имат обясними резултати.“

Кратко съдържание на дисертационния труд:

1. Introduction. Във въведението са представени: мотивация за изследванията; формулировка на темата; изследователските задачи; хипотеза и цел на изследванията. Има представено съдържанието на дисертационния труд.
2. Literature Review. Тук са класифицирани AAL системите на база тяхната функционалност и на база използваните технологии; отбелязани са някои модели за машинно самообучение приложими в контекста на настоящето изследване.
3. Concept of the AAL System. Дефиниран е потребителския случай на AAL системата както и целевото лице за асистиране. Изследвани са категориите данни необходими за този клас системи.
4. System architecture and implementation. Първоначално са описани критериите за подбор на подходящ междинен софтуер за целите на изследването. След сравнение по критериите е избран OpenRemote. Представено е кратко описание на продукта, а също така са дискутирани възможностите за разширение на архитектурата в посока разширение по задачите на дисертацията.

5. GDPR Requirements and compliance. Главата е посветена на изискванията на стандарта при обработката и съхраняването на лични данни.
6. Behavioural predictions. В главата са разгледани модели със самообучение за откриване на човешкото поведение. Двата избрани подхода от докторантката са базирани на невронни мрежи и асоциативни правила, които са реализирани към избрания прототип. Реализациите са изследвани с тестови данни от OpenRemote и резултатите са приведени.
7. Conclusion. В тази глава са представени основните приноси на изследването, които разглеждаме по-долу. Представена е визията на авторката за по-нататъшните изследвания както и апробациите направени за получените резултати.

Appendix A. Съдържа файла `deploy.yaml` за конфигурация на OpenRemote за композиция на докер версия.

Appendix B. Съдържа скрипт на Python за форматиране на данни за ARM алгоритмите.

Appendix C. Съдържа примерно множество от правила за почистване.

Приносите на дисертационния труд са представени в три категории: научни, научно-приложни и приложни. По-долу са изложени така както са представени от докторантката.

### **Научни приноси**

1. Аналитичен обзор на състоянието в областта на AAL системите, съществуващите в момента видове системи, често решавани проблеми и методология.
2. Аналитичен обзор на ОРЗД и обобщение на функционалните изисквания, които AAL системата трябва да изпълни, за да гарантира поверителността на данните. Същото може лесно да се преизползва и приложи към други видове системи.
3. Модификации и подобрения на алгоритми за машинно самообучение с цел разпознаване на модели на човешкото поведение. Сравнение на производителността, точността и приложимостта в предложената AAL система на невронни мрежи и алгоритми за извличане на асоциативни правила.
4. Разработване на методология за създаване на AAL система с общо приложение.
5. Откриване и валидирано разпознаване на оптимизирана реализация на Apriori, която предоставя по-бързи и еднакво точни резултати в сравнение с FPGrowth за разпознаване на поведенчески модели.

6. Предложен е метод за трансформиране на прекодирани като времеви редове ежедневни дейности в етикетиранни ежедневни транзакции с цел генериране на правила.

### **Научно-приложни приноси**

1. Предложение за подобрена, в сравнение с предшествениците си, по отношение на цена и сигурност на данните, AAL система, която използва вече съществуващи сензори и устройства в дома на човек, за да създаде модел, който разпознава човешките поведенчески стереотипи и може да предупреждава за отклонения от тях.
2. Предложена и внедрена архитектура на AAL система, подобрена с алгоритми за машинно самообучение, която гарантира съответствие с поверителността на данните, като работи като контейнерно решение в изолирана мрежа.
3. Експерименти върху приблизително 15 набора от данни от колекцията от набори от данни CASAS, които водят до заключението, че алгоритмите на ARM са по-рентабилни и по-лесни за поддръжка, когато се прилагат към AAL система за предсказване на поведенчески модели.
4. Концепция и изисквания за създаване на AAL системи, фокусирани върху общата цена на внедряването, дългосрочната поддръжка и разходите на човек.

### **Приложни приноси**

1. Реализация на прототип на AAL система, разширена с модели за машинно самообучение, разпознаващи стереотипно поведение, базирано на събраните от системата данни.
2. Прототип, изпълняващ функционалните изисквания на система, съвместима с GDPR.

Забележки по приносите:

- Относно научен принос 4, докторантката е представила поход за създаване на AAL система, но не и методология.
- Относно научно-приложните приноси трябва да се има в предвид базата на изследванията и разработката, а именно OpenRemote и архитектурата ѝ, така че в този контекст да се разбират „предшественици“ и „архитектура“. Въз основа на тези бележки, научно-приложен принос не излиза извън посочените тук рамки.

Като се отчитат посочените по-горе констатации, приемам като цяло представените приноси от изследването.

#### **4. Аprobация на резултатите**

Представеният дисертационен труд е оформен на базата на следните две публикации:

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).

След проверка и от представената справка може да се заключи, че:

- а) научните трудове отговарят на минималните национални изисквания (по чл. 2б, ал. 2 и 3 на ЗРАСРБ) и съответно на допълнителните изисквания на СУ „Св. Климент Охридски“ за придобиване на образователна и научна степен „доктор“ в научната област и професионално направление на процедурата;
- б) представените от кандидатката резултати в дисертационния труд и научни трудове към него не повтарят такива от предишни процедури за придобиване на научно звание и академична длъжност;
- в) няма доказано по законоустановения ред плагиатство в представения дисертационен труд и научни трудове по тази процедура.

#### **5. Качества на автореферата**

Авторефератът се състои от 45 страници разпределени в пет глави и приложение както следва: 1. Въведение – 5 страници; 2. Архитектура и имплементация на системата – 6 страници; 3. ОРЗД изисквания и изпълнение – 3 страници; 4. Прогноза на поведение – 13 страници; 5. Заключение – 11 страници; Литература – 4 страници. В него са отразени основните постижения на изследването.

Представен е автореферат и на английски език от 41 страници с почти идентично съдържание на българския вариант.

Стремежът двата автореферата да са идентични едва ли не на ниво изречения е довело до ефекта на автоматичния превод от английски на български.

Авторефератът отговаря на всички изисквания за изготвянето му и представя коректно резултатите и съдържанието на дисертационния труд.

## **6. Критични бележки и препоръки**

Представянето на дисертационния труд е позволило на докторантката да избегне редица терминологични проблеми, които се срещат в текста на автореферата на български език както съм отбелязал по-горе.

Оценката от докторантката на приносите на дисертационния труд надхвърлят реалните постижения, което съм отбелязал в съответния раздел.

## **7. Заключение**

След като се запознах с представените в процедурата дисертационен труд и придружаващите го научни трудове и въз основа на направения анализ на тяхната значимост и съдържащи се в тях научни и научноприложни приноси, потвърждавам, че представеният дисертационен труд и научните публикации към него, както и качеството и оригиналността на представените в тях резултати и постижения, отговарят на изискванията на ЗРАСРБ, Правилника за приложението му и съответния Правилник на СУ „Св. Климент Охридски“ за придобиване от кандидата на образователната и научна степен „доктор“ в научната област 4. Природни науки, математика и информатика и професионално направление 4.6. Информатика и компютърни науки. В частност кандидатът удовлетворява минималните национални изисквания в професионалното направление и не е установено плагиатство в представените по конкурса научни трудове.

Въз основа на гореизложеното, **препоръчвам** на научното жури да присъди на Кристина Илиева Александрова образователна и научна степен „доктор“ в научна област 4. Природни науки, математика и информатика, професионално направление 4.6. Информатика и компютърни науки.

07.06.2024 г.

Изготвил рецензията: .....  
(професор, доктор Владимир Димитров)