

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: **проф. д.н. Даниела Ананиева Орозова,** Тракийски университет - Стара Загора в качеството ми на член на научното жури, съгласно Заповед № РД-38-199/26.04.2024 г. на Ректора на Софийския университет. Представеният комплект материали е в съответствие с правилника за условията и реда за придобиване на научни степени и заемане на академични длъжности в СУ „Св. Климент Охридски“ и включва следните документи: Заявление от Кристин Александрова до ръководител катедра за откриване на процедура за защита на дисертационен труд; Дисертационен труд на английски език; Автореферат на български и на английски език; Декларация за авторство; Справка за изпълнение на минималните национални изисквания; Справка за публикациите в Scopus; Доклад от научния ръководител за готовност на докторанта; StrikePlagiarism доклад, Протокол за проверка на оригиналността на дисертационен труд и Становище от научния ръководител; Списък с публикации; Заповед за зачисляване; Удостоверение за обучение като задочен докторант във ФМИ на СУ; Заповед за отчисляване на докторанта; Вътрешна рецензия от проф. Цунижев; Протокол от катедрен съвет, за предварително обсъждане на дисертационния труд; Дипломи за завършена бакалавърска и магистърска степен на обучение; Автобиография; Справка от информационна система - Авторите, Файлове на двете научни публикации по темата на дисертационния труд.

1. Обща характеристика на дисертационния труд и представените материали

Автор на дисертационния труд е Кристина Илиева Александрова – докторант в задочна форма на обучение към катедра „Компютърна информатика“ на Факултета по математика и информатика при СУ „Св. Климент Охридски“, с научен ръководител проф. д-р Мария Нишева-Павлова.

Дисертационният труд на Кристина Александрова 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) е в общ обем от 181 страници (основен текст и приложения). Текстът е структуриран в седем глави, библиография и три приложения. Включва 66 фигури и 12 таблици. Текстът би спечелил, ако е приложен списък на съкращенията и пояснения на използваните термини.

В работата първо се разглеждат пропуски, идентифицирани от предишни научни трудове, разглеждащи системите за подпомаган живот от обкръжаващата среда с общо предназначение, наричани AAL (Ambient Assisted Living) системи, като устойчиво осигуряване на защита и сигурност на данните в разрастваща се система. Това са фактори, които възпрепятстват широкото разпространение и употреба на този тип системи. Идентифицирани са функционалните изисквания на такава AAL система, съобразена с изискванията на Общият регламент за защитата на личните данни (GDPR). Чрез преизползване на съвременен мидълуер, наличен и поддържан в платформи с отворен код, се създава прототип на системата. Потокът от данни и събития от наличните сензори и интелигентни устройства, относно ежедневните дейности на наблюдаваното лице, се използва за обучение на модели, базирани на невронни мрежи и асоциативни правила. Техните прогнози и изводи позволяват на потребител, който се грижи за човека, наблюдаван от системата, да разбира сигналите и да игнорира фалшиви резултати.

На страница 10 от дисертационния труд са дефинирани три изследователски въпроса и на тяхна основа е формулирана хипотезата: *„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“.

Целта на изследването е „да докаже хипотезата като цяло и в нейните производни части, като отговори на зададените изследователски въпроси“.

В **глава 1 - увод** се представя историята и актуалността на проблема и е обоснована необходимостта от провеждане на изследванията. Дефинират се изследователските въпроси, хипотезата и целта на дисертационния труд.

В **глава 2** е направен литературен обзор на системите за подпомаган живот от обкръжаващата среда (AAL systems). Първо са класифицират типове системи, базирани на тяхната функционалност (умни домове, интелигентни асистенти, преносими и роботизирани системи за помощ). Следва разглеждане на AAL systems, базирани на използваните технологии (RFID и NFC технологии за уникално идентифициране и обмен на данни с радиовълни, безжични сензорни мрежи, преносими сензори и други технологии). Накратко са представени най-често срещаните алгоритми за машинно обучение за откриване на отклонения и прогнозиране на човешкото поведение.

В **глава 3** е дадена кратка дефиниция на концепцията на AAL системата и личността на потребителя, която лежи в основата за създаването на прототипа и очакванията към системата. Оценяват се вече публично достъпните набори от данни, които симулират достатъчно близо личността на потребителя и могат да се използват за обучение и оценка на методи за машинно обучение.

В **глава 4 „Архитектура и имплементация на системата“** е представена първа версия на прототипа, базирана на избория софтуер OpenRemote. Предлага се разширение на архитектурата на системата с алгоритъм за машинно обучение. Разглежда се процеса на създаването на прототип на AAL система, която подпомага изпълнението на ежедневни дейности на възрастните хора с помощта на персонализирани сензори и комбинация от правила, такива че системата да е съобразена с лицето, за което е предназначена.

Глава 5 „ОРЗД Изисквания и изпълнение“ разглежда подробно възможностите за прилагане на общия регламент за защитата на данните при изграждане на AAL система, която има за цел да комбинира ненаатрапчиво събиране на данни за потребителите с алгоритми за машинно обучение, за да извлече информация за физическото и психическото състояние за лицето, за което се грижи системата. Изведени са 15 изисквания към прототипа на AAL системата и нейното поведение. Това позволява

количествено измерване на съответствието с GDPR и защитата на поверителността на данните.

В глава 6 „Прогноза на поведение“ е представен подход за прилагане на различни алгоритми за машинно обучение, които създават поведенчески модели. Тези модели се обучат от стандартен набор от данни, съдържащ различни дейности от ежедневието, и могат да различават нормалното стереотипно поведение на определен човек или отклонение от очакваното поведение. Към проблема се прилагат и се оценяват два основни подхода: невронни мрежи (двата алгоритъма Автоенкодери и LSTM) и алгоритми за извличане на асоциативни правила (Argioi и FPGrowth) при анализ на човешкото поведение. Използвани са 15 набора от данни от колекцията на CASAS, за да сравнят подходите и работата на четирите алгоритъма.

В глава 7 „Заклучение“ е направена дискусия върху поставените изследователски въпроси, отбелязани са получените приноси и са формулирани перспективи за бъдещо развитие на тематиката.

В три приложения са дадени допълнителни материали и справки, свързани с проведените изследвания.

Литературният обзор и анализ на актуалното състояние на научните изследвания в областта, направен във втора глава от дисертационния труд, както и проучените литературни източници и направени изводи ми дават основание да заключа, че докторантката познава задълбочено проблемите, свързани с приложение на технологии за машинно обучение за подпомагане на живота на възрастни хора и хора в неравностойно положение. Предложена е система, която може да намери широко приложение, поради възможностите за реална употреба, персонализация, ниска цена на поддръжка и съответствие с принципите за поверителност и сигурност на данните за потребителите.

Реализираното изследване в дисертационния труд изисква задълбочени знания и научна и практическа квалификация, които авторът ѝ безспорно притежава. Това личи от лекотата, с която докторантката аргументира и излага основните тези в отделните глави на дисертационния труд и от представените публикации.

2. Данни и лични впечатления за кандидата

Кристин Александрова завършва средното си образование в природоматематическа гимназия “Гео Милев”, гр. Стара Загора през 2013 г. След това от 2013 до 2017 г. завършва бакалавърска степен на обучение в СУ “Св. Климент

Охридски”, Факултет по Математика и Информатика, специалност Компютърни науки. През 2019 г. получава магистърска степен, специалност Изкуствен Интелект, при Факултет по математика и информатика на Софийския университет „Св. Климент Охридски“. Кристина Александрова от 2020 г. работи като продуктов мениджър към SAP SE от 2022 г. до сега, а преди това в САП Лабс България ЕООД.

Работи като хоноруван асистент във Факултет по математика и информатика на СУ “Св. Климент Охридски” от 2014 г. до 2017 г., провежда упражнения по дисциплините: „Увод в програмирането“, „Обектно ориентирано програмиране“ и „Структури от данни и програмиране“.

Не познавам лично докторантката. От публикуваните материали и предоставени документи, получените от мен впечатления за Кристина Александрова са изцяло положителни.

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

Приемам основните резултати и приноси, посочени от докторантката на страници 155 и 156 от дисертационния труд. Приносите имат научен, научно-приложен и приложен характер.

Направен е аналитичен обзор на състоянието в областта на AAL системите. Анализирани и обобщени са функционалните изисквания, които AAL системата трябва да изпълнява, за да гарантира поверителността на данните. Разработена е методология за създаване на AAL система с общо приложение.

Модификации са алгоритми за машинно самообучение с цел разпознаване на модели на човешкото поведение. Създадена е оптимизирана реализация на алгоритъм Argioi, който предоставя по-бързи и еднакво точни резултати в сравнение с алгоритъм FPGrowth за разпознаване на поведенчески модели. Предложен е метод за трансформиране на прекодирани като времеви редове дейности в етикетирани ежедневни транзакции с цел генериране на правила.

Направено е сравнение на производителността, точността и приложимостта в предложената AAL система на невронни мрежи и алгоритми за извличане на асоциативни правила. Проведени са експерименти, които водят до заключението, че алгоритмите за извличане на асоциативни правила са по-рентабилни и по-лесни за

поддръжка, когато се прилагат към AAL система за предсказване на поведенчески модели.

Предложена е AAL система, която използва съществуващи сензори и устройства, интегрирана с алгоритми за машинно самообучение, която гарантира съответствие с поверителността на данните и работи като контейнерно решение в изолирана мрежа.

Създадена е концепция и изисквания към AAL системи, фокусирани върху общата цена на внедряването, дългосрочната поддръжка и разходите.

Реализиран е прототип на AAL система, разширена с модели за машинно обучение, разпознаващи стереотипно поведение, базирано на събраните от системата данни и изпълняващ функционалните изисквания на система, съвместима с GDPR.

4. Аprobация на резултатите

Резултатите от дисертационното изследване са представени в 2 публикации в рецензирани издания: доклади от международни научни конференции през 2021 и 2022 г., индексирани в научната база данни Scopus, със SJR. Публикациите са на английски език. Всяка от двете публикации носи по 30 точки и се надхвърлят 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“. Освен това е представена справка за още една публикация, което е изпратена за рецензиране. Трите публикации са самостоятелни.

Добро впечатление правят обема и задълбочеността в публикациите, отразяващи основните аспекти на разглежданата в дисертацията проблематика.

След подробен преглед на представените научни трудове установих, 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 кандидата 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 на български и английски език коректно представят резултатите и съдържанието на дисертационния труд. В автореферата не са представени глава 2 и глава 3 от дисертационния труд, поради техния обзореи характер.

6. Критични бележки и препоръки

Не откривам пропуски в структурирането на съдържанието, точността и пълнотата в описанието, като и в оформянето на представените документи към дисертационния труд. Документите показват успешна и резултатна изследователска активност на Кристина Александрова.

Към докторантката има следните два въпроса:

- Къде и как се съхраняват обработваните данни от сензорите и интелигентните устройства при работа на системата в реални условия и могат ли да имат по-широка употреба в бъдещи изследвания?
- Достъпен ли е създаденият прототип и как може да бъде получен достъп?

Системи от този тип имат потенциал да повишават качеството на живота на възрастните хора и да осигуряват жизненоважна подкрепа в ежедневните им дейности. Препоръчвам на докторантката да продължи работата си по тематиката и да развие текущите изследвания. С цел запознаване на по-широка научна общност с постигнатите резултати, добре е те да бъдат представени в специализирани издания с импакт фактор.

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 процедурата научни трудове.

Въз основа на гореизложеното, **препоръчвам** на научното жури да присъди на **Кристин Илиева Александрова** образователна и научна степен „доктор“ в 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“.

6.06.2024 г.

Изготвил рецензията:

(проф. д.н. Даниела Орозова)