

ФКСТ 44 - НСА - 049

05.11.2021г.



СТАНОВИЩЕ

върху дисертационен труд за придобиване на образователна и
научна степен „доктор“

Автор на дисертационния труд: Камен Николаев Стойков

Тема на дисертационния труд: „ОБРАБОТКА, АНАЛИЗ И ИЗВЛИЧАНЕ НА
ЗНАНИЯ ОТ ГОЛЕМИ МАСИВИ ДАННИ НА ЕСТЕСТВЕН ЕЗИК ЧРЕЗ МАШИННО
ОБУЧЕНИЕ“

Член на научното жури: доц. д-р Милен Йорданов Петров, ФМИ на СУ

Професионално направление: 5.3. Комуникационна и компютърна техника

Основание за представяне на становището: участие в състава на научното жури по
защита на дисертационния труд съгласно Заповед № ОЖ-5.3.30/22.07.2021г. на
Ректора на ТУ-София.

**1. Актуалност на разработвания в дисертационния труд проблем в научно и
научноприложно отношение. Степен и нива на актуалността на проблема и
конкретните задачи, разработени в дисертацията.**

Обработка, анализ и извличане на знания от големи масиви данни на естествен
език чрез използване на съвременни методи и техники от изкуственият интелект и
машинно обучение са една от най-бурно развиващите се области в съвременните
компютърни науки и информационни технологии. Дисертационният труд е посветен на
методите за обработка на големи масиви от неструктурирани данни и извличане на
знания от тях с използване на подходящи структури от данни, методи, алгоритми и
софтуерни архитектури за постигане на оптимални резултати.

С оглед на това може да заключим, че темата на представената дисертация е в
изключително **актуалната област на комуникационната и компютърна техника**
посредством изследване и прилагане на алгоритми и машинно обучение.

**2. Степен на познаване състоянието на проблема и творческа
интерпретация на литературния материал.**

Кандидата е много добре запознат с областта на изследване – това показват
цитираните и реферирани 139 литературни източника, предимно на чужд/английски
език, от реномирани списания и статии в конференции, включващо и достатъчен брой
публикации – научни и технологични от последните 10 години.

3. Съответствие на избраната методика на изследване и поставената цел и задачи на дисертационния труд с постигнатите приноси – Обект на изследването в дисертационния труд са големи масиви от неструктурирани данни като специално внимание се обръща на обработването на текстови документи на естествен език с правен характер. Предмет на дисертационния труд е изследване на методи и алгоритми за обработка на текстови документи на естествен език чрез машинно обучение с цел създаване на модел за обработка на съдържанието на документите.

Избраната методологията на изследване е подходяща за постигане на целите на дисертационния труд, а именно посредством използване на следните научни методи:

- Проучване на литературни източници в областта на обработка на текстови документи на естествен език чрез машинно обучение;
- Проектиране на методология за обработка на текстови документи на естествен език с правен характер чрез машинно обучение;
- Проектиране и имплементация на методи и алгоритми за машинно обучение за извличане на знания и изграждане на модел за структурата и съдържанието на текстови документи на естествен език с правен характер;
- Провеждане на експеримент за оценка, валидация и верификация на предложената методология и създадения модел
- Статистически методи за обработка и представяне на получените експериментални данни.

4. Приноси на дисертационния труд: приемам приносите, на докторанта, като ги характеризирам на научно-приложни и приложни. Бих препоръчал да се обобщят приносите по следният начин: **Научни-приложни:** 1. Предложена е методология за обработване на неструктурирани данни под формата на правни документи от вида декларации за поверителност с използване на машинно обучение; 2. Предложен е метод подход за формиране на обучаващо множество данни чрез автоматично аотиране на декларации за поверителност посредством използване на невронни мрежи и клъстеризация; 3. преглед и анализ на методи и алгоритми за всеки от етапите на подготовка, аотиране и класификация на декларации за поверителност; и **Приложни приноси:** 4. Имплементирани и експериментално са изследвани алгоритми за всеки от етапите на подготовка, аотиране и класификация на декларации за поверителност; 5. Дефинирани са подходяща конфигурация и параметри на модел за класификация на декларации за поверителност, с които се постига висока точност на класификацията и ниска изчислителна сложност.

5. Публикациите по дисертационния труд - резултатите от работата са публикувани в общо 6 реномирани научни издания, от които всичките са на английски език. Брой статии, представени по конкурса са шест на брой, от които три в научни списания и три в сборници на научни конференции. Като две от статиите са самостоятелни, а в три от статиите е първи съавтор. В случая смятам, че същественият

принос на докторанта е видим.

6. Мнения, препоръки и бележки. Авторефератът е добре направен и точно отразява дисертационния труд. Имам забележка, че няма приложен списък с цитираната литература в автореферата. Като научен труд, подлежащ на самостоятелно разпространение, отделно от дисертацията, следва да се допълнят източниците, които са използвани. Към представената дисертация нямам съществени забележки, тя е старателно оформена и добре подредена.

Като забележка в оформлението на документите - на три от статиите, съответно [3], [5] и [6] от "Списък на публикации по дисертационния труд", не са посочени страници на отпечатване. На публикация [6] не е посочено и заглавие. По справка от приложената бележка се вижда, че заглавието на статията е "Analysing privacy policies". Следва подобен вид справки да се изготвят по-прецизно.

От представеното до тук е видно, че разглежданата дисертация напълно отговаря на изискванията на ЗРАСРБ, Правилника за прилагане на ЗРАСРБ, и съответните вътрешни правилници за присъждане на образователната и научна степен „доктор”. Отправените коментари, забележи и препоръки по никакъв начин не намаляват приносите на докторанта по същество. Не познавам докторантът лично от преди конкурса и нямам лични впечатления, извън мнението ми, съставено по представените ми документи – а именно, че колегата е високо квалифициран специалист със задълбочени познания в областта на дисертацията.

В заключение давам **положителна** оценка на представената дисертация и предлагам на уважаемото научно жури, по обявената процедура за защита на докторска дисертация в Технически университет – гр. София, да **присъди** образователна и научна степен доктор в **професионално направление: "5.3. Комуникационна и компютърна техника"** на автора на дисертационния труд **Камен Николаев Стойков**.

04.11.2021 г

гр. София

Подпис:

/доц. Милен Петров/