

РЕЦЕНЗИЯ

от доц. д-р инж. Цветослав Станиславов Цанков,
Факултет по технически науки, катедра ККТ,
Шуменски университет „Епископ Константин Преславски“

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

в област на висшето образование: 5. Технически науки
професионално направление: 5.3. Комуникационна и компютърна техника
докторска програма: Автоматизирани системи за обработка на информация и
управление

Автор на дисертационния труд: инж. Нели Ананиева Арабаджиева - Калчева

Тема на дисертационния труд: „Изследване на методи и алгоритми за
класификация на текст в машинното обучение“

1. Общо описание на дисертационния труд и на приложените към него материали

Дисертационният труд на маг. инж. Нели Ананиева Арабаджиева - Калчева представлява завършен научно изследователски труд в конкретна област от познанието, а именно в научната специалност „Автоматизирани системи за обработка на информация и управление“. В него задълбочено са изследвани и анализирани методите и алгоритмите за класификация на текст в машинното обучение.

Дисертационният труд се състои от 154 страници. Тезата е развита в 5 глави и включва: актуалност на темата, целта и задачите на изследването, теоретични научни изследвания, експериментални изследвания, заключение, научни и научно-приложни приноси, списък на публикациите и литература. В края на дисертационния труд са двете приложения. Използвани са 38 фигури, 7 таблици и 132 литературни източника. Основните моменти по дисертационния труд са представени в 4 научни публикации.

От изложените факти се вижда, че маг. инж. Нели Ананиева Арабаджиева - Калчева е преминала пълноценно обучение в докторантура, а дисертационният ѝ труд съответства на научната специалност и на изискванията на Закона за развитието на академичния състав в Република България (ЗРАСРБ), Правилника за приложението на ЗРАСРБ и процедурните правила на Технически университет – Варна.

2. Тема и актуалност на дисертационния труд

Актуалността на изследвания в дисертацията проблем произтича от следните фактори:

Първо – машинното обучение, като модерно направление в областта на изкуствения интелект става изключително популярно, благодарение на успешни проекти на учени от цял свят;

Второ – при навлизането на Интернет във всеки дом, в момента за бизнеса особена важност има информацията за мнението на потребителите при ревюта, онлайн реклами, отзиви за услуги и много други;

Трето – дори и при множеството научно-изследователски разработки на методи и технологии в машинното обучение, в тази област винаги са налице множество възможности за иновации.

От посочените фактори следва убеждението, че темата на дисертационния труд е актуална, разглежданите в нея въпроси са значими, а получените резултати са иновативни, полезни и обвързани с тематиката на изследването.

3. Познаване състоянието на проблема

За формирането на целта и задачите на дисертационния труд е направен задълбочен литературен обзор, в който са включени 132 литературни източника, от които 98 са на латиница. Голяма част от литературните източници са с най-съвременните постижения, като са помогнали на докторанта да работи по това модерно направление в областта на изкуствения интелект – машинното обучение.

Анализът на литературните източници, направен в Глава 1 показва, че докторантът задълбочено познава състоянието на изследвания в дисертационния труд проблем. Формулирани са цел, задачи, обект и предмет на дисертационния труд. Постигнато е съответствие между темата на дисертационния труд, поставената научна цел и задачите на изследването.

4. Методика на изследване

За реализиране на поставените задачи докторантът е направил теоретични изследвания на методи и алгоритми за класификация на текст в машинното обучение. Методите за машинно обучение са използвани обосновано и в пълно съответствие с целта и задачите на дисертационния труд.

Докторанта използва методика за провеждане на експериментално изследване, включващо описание на реда на провежданите дейности за изпълнение на поставените цели и задачи, избор на софтуерни средства и подбор на данни.

В дисертационния труд са реализирани научни изследвания за изпълнение на целите и задачите, които завършват със сравнителен анализ и обобщена оценка на създадените класификационни модели.

Научните изследвания са проведени коректно, поставените задачи в дисертацията са адекватни на целта и са решени в необходимия обем и на високо научно ниво, като са използвани съвременни методи за анализ на данни.

5. Приноси на дисертационния труд

В резултат от проведените в съответствие с целта и задачите на дисертационния труд теоретични и експериментални изследвания се свеждат до научни и научно-приложни приноси.

1) Научни приноси:

- основен научен принос е разработването на модифициран алгоритъм на Наивния Бейсов класификатор, използващ разпределението на Лаплас;
- статистически е доказано, че наблюдаваната разлика в точността на класификатора Гаусов Наивния Бейсов класификатор и предложението Лапласов Наивния Бейсов класификатор върху избраната база данни не е резултат от случайния начин на нареждането на примери, използвани за обучение и тестване, а от различното класификационно поведение на алгоритмите;
- анализирани са осем алгоритъма за класификация на текст в машинното обучение с техните предимства и недостатъци: Гаусов Наивен Бейсов класификатор, Бернулиев Наивен Бейсов класификатор, Мултиномиален Наивен Бейсов класификатор, метод на K най-близък съсед, метод на опорните вектори, Дърво на решението, ансамблов метод Случайна гора и ансамблов метод Адаптивен усилвател.

2) Научно-приложни приноси:

- основен научно-приложен принос на дисертацията е класификацията на текстове на български автори. Проведени са експериментални изследвания с Бернулиев Наивен Бейсов класификатор, Мултиномиален Наивен Бейсов класификатор, метод на K най-близък съсед, метод на опорните вектори и Дърво на решението върху данни представляващи извадка от стихотворения на български автори, живели и творили през 19 и 20 век. Изследвано е влиянието на обема от данни върху точността, прецизността, чувствителността и F -мярка на класификаторите. Направен е сравнителен анализ на резултатите от експериментите;
- проведени са експериментални изследвания с предложението модифициран алгоритъм на Наивния Бейсов класификатор, използващ разпределението на

Лаплас и алгоритмите: Гаусов Наивен Бейсов класификатор, Бернулиев Наивен Бейсов класификатор, Мултиномиален Наивен Бейсов класификатор, метод на K най-близък съсед, метод на опорните вектори, Дърво на решението, ансамблов метод Случайна гора и ансамблов метод Адаптивен усилвател с потребителски мнения на английски език. Изследвано е влиянието на обема от данни върху точността, прецизността, чувствителността и F-мярка на класификаторите. Направен е сравнителен анализ на резултатите от експериментите.

Приносните моменти имат научно-приложен и приложен характер със значимост на новост в разглежданата проблематика и представляват доказване с нови средства на съществени нови страни в съществуващи научни проблеми и теории и получаване и доказване на нови факти.

Обобщавайки същността и съдържанието на дисертационното изследване, приносът на докторанта се изразява в изследване влиянието на обема от данни при класификацията на текст на български и английски език върху различни методи и алгоритми за класификация в машинното обучение.

6. Публикации по дисертационния труд

Основните резултати, получени в дисертационния труд, са представени в четири публикации. Една от публикациите е подготвена самостоятелно на английски език, а три са на български език, от които две са в съавторство с научния ръководител и една с колега от ТУ-Варна. Едната публикация е в чужбина, от представени резултати на международна конференция. Две публикации са отпечатани в списания и една публикация в Годишник на ТУ-Варна.

Публикациите са направени в периода 2016 г. – 2018 г. и съдържат най-съществените моменти от изследването, проведено от маг. инж. Нели Ананиева Арабаджиева - Калчева. Всички те са дали възможност на научната общественост да се запознае с дисертационния труд на докторанта.

7. Авторство на получените резултати

Проведеното в дисертационния труд изследване е тясно свързано с цялата научно-изследователска дейност на докторанта. Резултатите от дисертационни труд са апробирани и потвърждават достоверността на възприетите от автора подходи. Структурата на дисертационния труд позволява да се проследи логиката и причинно-следствените връзки при изследването, при което не остават съмнения относно достоверността на представения научен материал. Всичко това показва, че горепосочените приноси са лично дело на маг. инж. Нели Ананиева Арабаджиева - Калчева.

8. Автореферат и авторска справка

Представеният автореферат на дисертационния труд е в обем от 44 страници. Той е подготвен съгласно изискванията и представя в синтезиран вид същността на дисертационния труд като отразява основните приноси на кандидата и публикации по дисертационния труд.

9. Препоръки и забележки

Към дисертационния труд следва да се направят някои препоръки и забележки. При част от надписите във фигурите и таблиците може да бъде избегнат английския език. Десетичният разделител при стойностите в таблиците и графиките е точка, поради което следва да се обърне внимание на софтуера за статистическа обработка и представяне на резултатите.

Силно препоръчвам на маг. инж. Нели Ананиева Арабаджиева - Калчева в бъдеще да публикува резултати от научно-изследователската си работа в списания с импакт-фактор или импакт-ранг.

10. Заключение

Докторантът е осъществил пълноценно своето обучение в докторантура и има възможности за самостоятелна научна работа. Посочените забележки към дисертацията са редакционни и стилни и лесно ще бъдат избягвани в бъдещата му научно-изследователска работа. Това не влияе върху значимостта на научните и научно-приложните приноси на дисертационния труд.

Дисертационният труд има качеството на завършено научно-приложно изследване по актуален и важен проблем с публикувани резултати, като отговаря на изискванията за присъждане на образователната и научна степен „Доктор“, според ЗРАСРБ, Правилника за приложението на ЗРАСРБ и на процедурните правила на Техническия университет - Варна.

Предвид гореизложеното давам положителна оценка на представения ми за рецензиране дисертационен труд. Препоръчвам на научното жури да присъди образователната и научна степен „Доктор“ на маг. инж. Нели Ананиева Арабаджиева - Калчева в професионално направление 5.3. Комуникационна и компютърна техника, научна специалност „Автоматизирани системи за обработка на информация и управление“.

21.07.2020 г.
гр. Шумен

Подпис:
/доц. д-р Ц. Цанков/