

РЕЦЕНЗИЯ

върху дисертационен труд за получаване на образователната и научна степен

“Доктор”

по докторантска програма

“Автоматизирани системи за обработка на информация и управление”

професионално направление

5.3. *“Комуникационна и компютърна техника”*

на тема

" Използване на семантични технологии в един клас софтуерни приложения "

с автор инж. Даниела Иванова Петрова

от доц. д-р Ирена Вълва, катедра КСТ, РУ “А. Кънчев” Русе

1. Актуалност на разработвания в дисертационния труд проблем в научно и научно-приложно отношение. Степен и мащаб на актуалността на проблема и конкретните задачи, разработени в дисертацията.

Дисертационният труд е посветен на изследвания в интересно и актуално направление на компютърните науки – обработката и използването на непрекъснато нарастващите обеми информация от увеличаващия се брой потребители в интернет пространството, търсенето на методи и средства за категоризиране, описание, подреждане и класифициране на тази информация, така че неструктурираните текстови данни да бъдат лесно използвани. Целта на дисертанта е създаването на алгоритми и подходи за извличане на мнения (настроения, чувства) и анализ на текстове на български език - нещо, което липсва или е в оскъдни количества в публикуваните до момента и открити от докторанта научни проучвания в България.

Свободният текст представлява неструктурирани данни, които лесно се възприемат и обработват от хората, но значително по трудно и сложно от машините. За сметка на това се генерира непрестанно огромен обем текстово

съдържание, което е незаменим източник на знания и информация. Това налага все по-голямата актуалност на алгоритмите и методите за ефективната обработка и оценка на това съдържание. Естествено тези алгоритми трябва да работят автоматично и без субективната намеса на човешки фактор. Поради спецификата на различните езици очевидно не може да се говори за универсалност на подходите и алгоритмите и е необходимо да се търсят такива, подходящи в случая за анализ на текстови данни на български.

2. Познава ли дисертантът състоянието на проблема и оценява ли творчески литературния материал.

Дисертантът е приложил списък на 115 цитирани литературни източника, включващ статии в престижни специализирани списания и научни форуми и книги в предметната област на дисертацията. Включени са както класически и основополагащи, така и най-съвременни публикации по темата. Изследвани са и са приложени съвременни алгоритми от много актуалното в последните години направление, касаещо изкуствения интелект и машинното обучение. Вижда се, че докторантът е придобил задълбочени познания в разглежданата област и притежава необходимата теоретична основа за успешното разработване на дисертационния труд.

3. Избраната методика на изследване може ли да даде отговор на поставената цел и задачи на дисертационния труд.

Формулираните цел и задачи съответстват на изискванията за дисертационен труд за придобиване на образователната и научна степен "доктор". Избраната методика е класическата - с обзор и анализ на съществуващи решения, описание и анализ на подходящите за изследването алгоритми, събиране и подготовка на данните за експеримента и експериментално доказване и показване на приложимостта на предлаганото решение. Описан е подробно и поэтапно процесът на обработка и подготовка на експерименталните данни, като описанието е придружено със схеми и таблици за по ясна представа. Предложен е алгоритъм за почистване и предварителна обработка на текст на български език. В последната глава са показани резултатите от експериментални

изследвания с различни алгоритми от машинно обучение за оценка на коментари на български език.

Избраната методика е подходяща и съответства на поставените цел и задачи.

4. Кратка аналитична характеристика на естеството и оценка на достоверността на материала, върху който се градят приносите на дисертационния труд.

Дисертационният труд съдържа 143 страници, включително 23 фигури и 53 таблици. Оформен е като увод, четири глави, заключение, списък с използвана литература от 115 заглавия, 10 от литературните източници са на български, от български автори, останалите - на латиница, списък на публикациите на дисертанта, както и известните цитирания, списък на фигурите, списък на таблиците и приложения. Дисертантът достатъчно убедително последователно и систематично е показал целия процес на обработка, подготовка, анализ, приложени алгоритми, резултати и оценката им в своя дисертационен труд, така че нямам съмнения за достоверността на материала.

5. В какво се заключават научните или научно-приложните приноси на дисертационния труд? Формулиране и обосноваване на нов научен проблем (област). Формулиране и обосноваване на нова теория (хипотеза). Доказване с нови средства на съществуващи нови страни в съществуващи научни проблеми и теории. Създаване на нови класификации, методи на изследване, нови конструкции, технологии, препарати. Получаване и доказване на нови факти. Получаване на потвърдителни факти. Характер на приносите за внедряване: методи, конструкции, технологии, препарати, схеми и т.н. Каква е, конкретно, значимостта на тези приноси за науката и практиката?

В дисертационният труд са посочени 5 приноса:

1. Създадени са два набора от данни с мнения и коментари на български език (с приблизително по 100 000 коментара всеки), които могат да се използват и за бъдещи научни изследвания.

2. Разработен е списък от стоп думи, който може да се използва за предварителната обработка на текстове в бъдещи анализи и изследвания на текстове на български език.

3. Предложен е подобрен вариант на алгоритъма за създаване на набор данни и предварителната обработка на текстовете в него, съгласно спецификите на българския език.

4. Осъществен е сравнителен анализ на точността на различни методи и алгоритми за анализ на мнения чрез класификация на текстове на български език.

5. На база на извършените в дисертацията експериментални изчисления е предложен подобрен по точност подход, който представлява комбинация от алгоритми за извличане на мнения от коментари на български език (att_SVM+biLSTM+lex_RF).

Бих класифицирала приносите като:

1 научен – предложена е модификация на алгоритъм за създаване на набори от данни и предварителна обработка на текстовете в него, в съответствие със спецификите на българския език.

4 научно-приложни и приложни, които касаят създадените набори от данни и оценката на резултатите от експерименталните изследвания.

6. Може ли да се оцени и в каква степен дисертационния труд и приносите представляват лично дело на дисертанта?

Публикациите на автора дават основание да се заключи, че приносите в труда са лично дело. Според описанието на изследването и направените анализи и оценки нямам повод за съмнение, че дисертационният труд е лично дело на дисертанта.

7. Преценка на публикациите по дисертационния труд: брой, характер на изданията, в които са отпечатани. Какво е отражението им в науката – използване и цитиране от други автори, в други лаборатории, страни и пр.?

Общият брой на публикациите по дисертацията е седем. Четири от тях са в съавторство с ръководителя и три са самостоятелни, четири са индексирани в SCOPUS. Има една публикация в българско издание, две в международни

форуми в България и четири в чужбина в международни конференции, като всички са на английски език. Показано е едно цитиране на публикациите на автора от други автори.

8. Резултатите от дисертационния труд използвани ли са вече в научната и социалната практика? Има ли постигнат пряк икономически ефект и пр.? Документи, на които се основава твърдението.

Няма данни и не са приложени документи към дисертацията, доказващи това, освен апробацията с изнесените доклади на международни конференции, участие в два научно изследователски проекта на ТУ Варна и едно цитиране.

9. Мотивирани препоръки за бъдещо използване на научните и научно-приложните приноси: какво и къде да се внедри.

Тематиката е изключително интересна и актуална и затова смятам, че работата по практическото приложение на научно-приложните приноси на труда би трябвало да продължи. Интересно е другите подобни решения и проекти в същата област какви алгоритми използват и дали биха използвали предложеното от дисертанта.

10. Авторефератът направен ли е съгласно изискванията, правилно ли отразява основните положения и научните приноси на дисертационния труд?

Авторефератът, състоящ се от 39 страници, е направен съгласно изискванията и отразява основните положения и приносите на дисертационния труд. (В точка 11 имам забележка към оформлението му!)

11.Критични бележки по дисертацията, включително и по литературната осведоменост на кандидата.

Тъй като писах предварителната рецензия и забележките ми по същество вече са отстранени, мога да направя няколко технически забележки, които реално не се отразяват на качествата на дисертационния труд:

- При оформлението на дисертационния труд има малка техническа неточност – на стр. 116 е фигура 22, а след това на стр.118 е фигура 19. Предполагам е техническа грешка.

- Липсва обръщение към фигурите в текста, на някои места в текста има двоеточие и следва фигурата, но като цяло не си личи къде е текстът, касаещ отделните фигури.
- Някои от таблиците са реално фигури.
- Много често има смесване на текст на английски и български – фигура 7 например.
- На фигура 16, в графиката, не са оразмерени координатните оси.
- В таблица 39 има две диаграми, в които текстовете са не четими.
- Не е описано в дисертационния труд как се решават проблемите с граматическите и синтактични грешки, примерно некоректно изписани думи или написани с латински букви?
- В автореферата е сменена номерацията на таблиците и фигурите, но в текста са останали старите номера при обръщенията към таблиците, а обръщения към фигурите липсват.

12. Други въпроси, по които рецензентът счита, че следва да вземе отношение.

От дисертацията и приложените към нея документи се вижда, че дисертантът е положил много труд и е работил самостоятелно над проблемите. Нямам допълнителни въпроси или коментари.

13. Заключение с ясно становище да се даде ли или не научна степен.

Въз основа на предоставените ми материали, постигнатите резултати и приноси, считам, че дисертационния труд отговаря на изискванията на Закона за развитието на академичния състав в Република България и Правилника за научни степени на ТУ-Варна.

Предлагам на членовете на научното жури да присъдят на инж. Даниела Петрова образователната и научна степен “доктор” в Професионално направление 5.3. „Комуникационна и компютърна техника“.

12.01.2024 г.

.....

/доц. д-р Ирена Вълва /