

Становище

на дисертационен труд на тема:

ИЗПОЛЗВАНЕ НА СЕМАНТИЧНИ ТЕХНОЛОГИИ В ЕДИН КЛАС СОФТУЕРНИ ПРИЛОЖЕНИЯ

С автор: Даниела Иванова Петрова

Член на Научното жури: проф. д. т. н. Андон Димитров Лазаров, ВВМУ, Варна

1. Актуалност на разработения в дисертацията проблем

В съвременния информационен свят на Internet, web технологиите и приложенията за обработка на огромни масиви от данни, все по-актуален е въпросът за разработване на ефективни алгоритми за класифицирането, структурирането и съхраняването на данни. Например, семантичният web, определящ общия формат и обмен на данните се прилага за добиване на знание от текст (text-mining – ТМ), анализа на настроения (sentiment analysis – SA) и др. В този смисъл, разработването на инструменти за семантичен анализ на текст на български език е актуална задача, авторско решение на която се предлага в настоящия дисертационен труд с основна цел: създаването на алгоритми и за анализ на текст и извличане на мнения (настроения, чувства) в текстове на български език. За постигане на поставените цели и задачи са приложени емпирични методи, включващи научен обзор и анализ на публикации, доклади, книги и резултати от дейности; теоретични методи, включващи индукция, анализ, синтез, сравнение, извличане на общи характеристики, обобщение и др. Приложен е обобщен алгоритъм за извличане на мнение от коментари в текст на български език, който асемблира метод на опорните вектори - Support Vector Machines (SVM) и рекурентни невронни мрежи и мета класификатор случайна гора (Random Forest-RF), в комбинация с метод на лексикона.

2. Приноси в дисертацията: научно-приложни приноси

- Създадени са два набора от данни с мнения и коментари на български език (с приблизително по 100 000 коментара всеки), използвани за целите на дисертационния труд с възможност за по-нататъшно изследване и анализ.
- 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.
- Предложен е модифициран вариант на алгоритъма за създаване на набор данни и предварителната обработка на текстове в него, съгласно специфичните особености на българския език.

- Осъществен е сравнителен анализ на точността на различни методи и алгоритми за анализ на мнения чрез класификация на текстове на български език.
- На база на извършените в дисертацията експериментални изчисления е предложен подобрен по точност обобщен алгоритъм, който представлява комбинация от алгоритми за извличане на мнения от коментари на български език (att_SVM+biLSTM+lex_RF).

Приносите в дисертационния труд могат да се интерпретират като обогатяване на съществуващите знания, както и приложение на научните постижения в практиката.

Докторантът представя списък с публикации по дисертационния труд, който се състои от седем заглавия, четири от които в съавторство, четири - индексирани в SCOPUS, както и справка за нуко-метричните показатели на публикациите – 140 точки, чийто брой съществено надхвърля изискваните 30 точки при дисертационен труд.

3. Критични бележки

Авторефератът отразява пълно съдържанието на дисертационния труд. На негова база могат да се направят следните забележки и препоръки. В текста се откриват: правописни грешки (препинателни знаци, пълен член); стилни грешки (...носят значение (има значение), единичен юнит..., вместо терминът „настроение“ на коментара може да се използва терминът „оценка“ или „знак“ на коментара, което означава коментар с положителна оценка (знак), коментар с отрицателна оценка).

4. Мотиви и заключение

На базата на изпълнените цел и задачи и получените резултатите с научно-приложна стойност давам своята положителна оценка на дисертационния труд, като предлагам на дисертанта Даниела Иванова Петрова да се присъди образователната и научна степен „доктор“ по професионално направление 5.3 „Комуникационна и компютърна техника“ и докторска програма „Автоматизирани системи за обработка на информация и управление“.

ЧЛЕН НА ЖУРИТО: проф. д. т. н (А. Д. Лазаров)

Дата: 29. 12. 2023 г.