

СТАНОВИЩЕ

върху дисертационен труд за придобиване на ОНС „Доктор” към професионално направление 5.3. „Комуникационна и компютърна техника” по 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”

Изготвил становището: доц. д-р инж. Марияна Цветанова Стоева

1. Актуалност на разработвания в дисертационния труд проблем

Тематиката на дисертационния труд е в актуалната област на машинното обучение, която се развива усилено поради възможностите и да се прилага в много и различни приложения. Най-популярните, от които са класирани на страници на Google, обработка на изображения, разпознаване на спам, плагиатство, персонализиране на реклами и много други.

Целта на настоящата дисертация е 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. Постигането на тази цел е реализирано чрез пет задачи. Направен е анализ на съвременните методи и алгоритми за класификация на текст, предложен е модифициран алгоритъм на Наивния Бейсов класификатор, използващ разпределението на Лаплас, реализирани са множество експерименти за изследваните на резултатите при класификация на български и английски текст и влиянието на обема на данните върху техните характеристики. Актуалността на направените в дисертационния труд изследвания се потвърждава и ще може да се ползва от постоянно увеличаващите се области на приложения на машинното обучение.

2. Научно-приложни и приложни приноси на дисертационния труд

Приемам приносите, които са представени в дисертационния труд, в следната редакция.

НАУЧНИ ПРИНОСИ

1. Разработване на модифициран алгоритъм на Наивния Бейсов класификатор, използващ разпределението на Лаплас, наречен Лапласов Наивен Бейсов класификатор.

НАУЧНО – ПРИЛОЖНИ ПРИНОСИ

2. Статистическо доказване, че разликата в точността на класификатора Гаусовия Наивен Бейсов класификатор и предложения Лапласов Наивен Бейсов класификатор не зависи от примерите, използвани за обучение и тестване, а от различното класификационно поведение на двата алгоритъма.
3. Анализ на осем алгоритъма за класификация на текст в машинното обучение с техните предимства и недостатъци

ПРИЛОЖНИ ПРИНОСИ

4. Класификация на текстове на български автори с използване на пет различни класификатори, анализирани са резултатите и е изследвано влиянието на обема от данни върху точността, прецизността, чувствителността и F-measure на класификаторите.
5. Експериментални изследвания, сравняващи резултатите на предложения алгоритъм на Лапласов Наивен Бейсов класификатор, с осем други алгоритми като е изследвано влиянието на обема от данни върху характеристиките им.

3. Мнения, препоръки и критични бележки

В дисертационния труд е извършена голяма изследователска работа на методи и алгоритми за класификация на текст. Направени са огромно количество експерименти, сравняващи резултатите на различните методи за класификация на английски и български текстове. Предложен е алгоритъм наречен Лапласов Наивен Бейсов класификатор, чиито резултати са сравнени с тези от други осем метода за класификация. Забележките ми са от технически характер.

4. Заключение - оценка на дисертационния труд

Оценката ми за цялостната работа на докторанта е положителна. Считам, че представеният материал отговаря на изискванията на Закона за развитие на академичния състав в Република България и на Правилника за приложение на ЗРАСРБ.

Предлагам на научното жури да приеме резултатите и приносите в дисертационния труд и да присъди на докторанта Нели Ананиева Арабаджиева - Калчева ОНС „Доктор” в област на висше образование – 5.Технически науки, професионално направление 5.3 „Комуникационна и компютърна техника” по докторска програма „ Автоматизирани системи за обработка на информация и управление ”

25.07.2020г.

Изготвил становището:

доц.д-р инж. Марияна Цветанова Стоева