

С Т А Н О В И Щ Е

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“

Член на научното жури: доц. д-р инж. Виолета Тодорова Божикова

1. Актуалност на разработвания в дисертационния труд проблем

С навлизането на интернет във всеки дом, класификацията на текст успешно се прилага при идентификация на тема, жанр, автори, откриване на спам, при отчитане мнението на потребителите за услуги, стоки, при прогнозиране на цени и др. Класификацията на текст има широко приложение в сферата на финансите, в лингвистиката, за разкриване на плагиатство и др. Всеки съвременен потребител на смартфон, таблет и на компютър волно и/или неволно ползва приложения за класификация на текст: при търсене на информация, при машинен превод, при проверка на правопис и др. Алгоритми за класификация на текст в машинното обучение успешно разпознават спам, предотвратяват измами, персонализират реклами, дават препоръки на телевизионните зрители в реално време, оптимизират разходите на бизнеса.

Дисертационният труд анализира осем алгоритъма за класификация на текст в машинното обучение и предлага оригинален алгоритъм, явяващ се модификация на Наивния Бейсов класификатор, чрез използване на разпределението на Лаплас. Разработката и използването на алгоритми за класификация на текст в машинното обучение са едни от най-модерните направления в Изкуствения Интелект, със засилен научен и практически интерес, особено през последните години. В този смисъл актуалността на разработвания в дисертационния труд проблем в научно и научно-приложно отношение е безспорна.

2. Научно-приложни и приложни приноси на дисертационния труд

Приемам приносите на дисертационния труд, но систематизирани като научни, научно-приложни и приложни по следния начин:

Научни приноси

1. Предложен е модифициран алгоритъм на Наивния Бейсов класификатор, използващ разпределението на Лаплас. Статистически е доказана неговата по-висока точност, в сравнение с класическия Наивен Бейсов класификатор, използващ разпределението на Гаус.

Научно – приложни приноси

1. Проведени са експериментални изследвания с Бернулиев Наивен Бейсов класификатор, Мултиномиален Наивен Бейсов класификатор, метод на К най-близък съсед, метод на опорните вектори и Дърво на решението върху данни представляващи извадка от стихотворения на български автори. Изследвано е влиянието на обема от данни върху точността, прецизността, чувствителността и F-measure на класификаторите и е направен е сравнителен анализ на резултатите от експериментите.

Приложни приноси

1. Анализирани са осем алгоритъма за класификация на текст в машинното обучение: Гаусов Наивен Бейсов класификатор, Бернулиев Наивен Бейсов класификатор, Мултиномиален Наивен Бейсов класификатор, метод на К най-близък съсед, метод на опорните вектори, Дърво на решението, ансамблов метод Случайна гора и ансамблов метод Адаптивен усилвател. Разгледани са техните предимства и недостатъци.

2. Проведени са експериментални изследвания с предложения 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 К 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-measure на класификаторите. Направен е сравнителен анализ на резултатите от експериментите.

Предложените приноси имат потенциала да обогатят натрупаните теория и практика в областта на класификация на текст в машинното обучение, а софтуерните реализации и получени резултати могат успешно да се използват в обучението.

В дисертационния труд са включени 4 публикации. Една от тях е самостоятелна, а в останалите три докторантката е водещ автор. С този списък от публикации, докторантката покрива Научнометричните критерии за придобиване на образователна и научна степен „Доктор“ в област „Технически науки“.

Авторефератът е в необходимия обем и надлежно отразява целите, задачите и основните разработки в дисертационния труд, съответства на изискванията на Правилника за условията и реда за придобиване на научни степени в Технически университет – Варна.

3. Мнения, препоръки и критични бележки

Дисертационният труд показва голям обем на извършени научни изследвания. Препоръчвам докторантката да продължи работата си по тази актуална тематика, като насочи усилията си към доразвиване на разработките, към публикуване на научни резултати във форуми, индексирани в Scopus и към последващо научно израстване в областта на комуникационната и компютърна техника. Забележката ми е за по-прецизно и по-кратко формулиране на приносите.

4. Заключение - оценка на дисертационния труд

Оценката ми за цялостната работа на докторантката е изцяло положителна. Дисертационният труд е с ясно изразени научни, научно-приложни и приложни приноси в областта на машинното обучение и удовлетворява изискванията на Закона за развитие на академичния състав в Р. България, Правилника към него, и Правилника за условията и реда за придобиване на научни степени в Технически университет – Варна.

Във връзка с гореизложеното, предлагам на уважаемото научно жури да присъди на инж. Нели Калчева образователно-научната степен „доктор“ по научно направление 5.3 „Комуникационна и компютърна техника“.

20.07.2020г.

Член на научното жури:

/доц. д-р инж. Виолета Божикова/