

С Т А Н О В И Щ Е

на проф. д-р Мария Нишева – ФМИ, СУ „Св. Климент Охридски“

за дисертационен труд за придобиване на образователната и научна степен „доктор“ по професионално направление 4.6 „Информатика и компютърни науки“ (Информатика – Изкуствен интелект)

с автор Тодор Васков Цонков

на тема „Изследване на мнения и чувства от текстове от социалните мрежи“

Със заповед № РД 38-99/17.02.2016 г. на Ректора на СУ „Св. Климент Охридски“ съм назначена за член на научно жури във връзка с процедурата за придобиване на ОНС „доктор“ по професионално направление 4.6 „Информатика и компютърни науки“ (Информатика – Изкуствен интелект) от Тодор Васков Цонков с дисертационен труд на тема „Изследване на мнения и чувства от текстове от социалните мрежи“.

1. Общо описание на представените материали

Във връзка с процедурата по защитата на дисертационния труд на Тодор Цонков са представени следните материали:

- дисертационен труд в обем от 101 страници;
- автореферат на дисертационния труд в обем от 32 страници;
- текстове на четири публикации, в които са отразени резултати на автора, свързани с темата на дисертационния труд;
- резюме на доклад, изнесен на Юбилейната конференция „125 години математика и природни науки в Софийски университет „Св. Климент Охридски““ (София, 5-7 декември 2014 г.).

2. Анализ на съдържанието, резултатите и приносите на дисертационния труд

Дисертационният труд е посветен на изследвания в областта на извличането на информация. Основната му цел е създаване на нови методи за откриване на текстове с двойно значение и в частност разработване на алгоритми за установяване на невярна информация, ирония, сарказъм, литота и преувеличения в текстове, публикувани в социалните мрежи. За проверка на резултатите от теоретичните изследвания на автора е създаден прототип на софтуерна система за

извличане на мнения от социалните мрежи, която оценява мненията от намерените текстове, и са проведени експерименти с тази система.

Дисертацията се състои от 101 страници текст, включващ седем глави и списък на използваните литературни източници.

Глава 1 има уводен характер. В нея е представена областта на изследванията и е обоснована значимостта на проблематиката, върху която е фокусирана работата на автора.

Глава 2 съдържа аналитичен обзор на областта „извличане на мнения от текст“ – основни задачи, подходи за решаването им и открити въпроси. Дискутирани са най-съществените проблеми при автоматично откриване на ирония, сарказъм, хипербола и неверни твърдения в текст.

Глава 3 е посветена на разработените от автора на дисертационния труд обобщени алгоритми за автоматично откриване на текстове от социалните мрежи с двойно значение. Представени са предложените алгоритми за откриване на ирония, сарказъм и сатира в текстове на български език.

В глава 4 са разгледани предложените нови алгоритми за установяване на невярна информация и преувеличение в текст. Анализирани са проведени експерименти, които демонстрират по-добри резултати от тези с използване на известните алгоритми.

Глава 5 представя подход за класификация на мнения от текстове на български език, основан на правила за автоматично засичане на слухове от текстове от социалните мрежи.

Глава 6 е посветена на описание на разработения от автора на дисертацията прототип на софтуерна система за автоматично извличане и оценяване на мнения на потребители, публикувани в социалните мрежи.

Заключителната седма глава съдържа обобщение на постигнатите резултати и препоръки за бъдещи изследвания по темата на дисертацията.

Списъкът на използваните литературни източници включва 86 заглавия, голяма част от които са цитирани по подходящ начин в текста на дисертационния труд. Този списък съчетава много добре класически и най-актуални източници в различни направления, свързани с отделните аспекти на темата на дисертацията.

Основните *научни приноси* в дисертационния труд могат да бъдат формулирани както следва:

- Предложени са правила за откриване на ирония, сарказъм, сатира, хипербола и литота в текст, както и за разпознаване на слухове, публикувани в социалните мрежи;
- Дефинирано е множество от характеристики на текстове, съдържащи мнения с двойно значение.

Като основни *научно-приложни приноси* в дисертационния труд може да бъдат посочени:

- Осъществена е експериментална проверка на точността на предсказване на различни алгоритми за установяване на двойно значение на мнения от социалните мрежи;

- Осъществена е експериментална проверка на точността на предсказване на предложения алгоритъм за разпознаване на слухове от социалните мрежи.

Приложените приноси на дисертацията може да бъдат обобщени както следва:

- Разработен е прототип на софтуерна система за извличане на мнения от социалните мрежи, която оценява мненията, съдържащи се в намерените текстове.

Мнението ми дисертационния труд като цяло е положително. Изложението е добре структурирано, ясно и аргументирано, илюстрирано с подходящи фигури и таблици. Избраната методика е адекватна на предмета, целите и задачите на изследването. Постигнатите теоретични и приложни резултати са оригинални и значими и напълно съответстват на поставените цели.

3. Публикации, които отразяват дисертацията. Отражение на резултатите на дисертацията в трудове на други автори

Резултатите, получени в дисертационния труд, са отразени в четири статии, публикувани както следва:

- една – в сборник с доклади на Балканската конференция по информатика ВСИ 2015;
- една – в научно списание, издавано от ТУ-София;
- две – в сборници с доклади на национални научни конференции.

Три от публикациите са в съавторство с научния ръководител на докторанта и една е в съавторство с друг докторант на научния ръководител. Нямам съмнение в съществуващия личен принос на Т. Цонков в тези публикации.

Не са посочени цитирания на публикации на Тодор Цонков по резултати от дисертационния му труд.

4. Критични бележки и препоръки

Като основна слабост на дисертацията бих посочила необрежното оформяне на текста ѝ и наличието на значителен брой фрази, които се нуждаят от внимателно прецизиране. Това в частност се отнася и до формулировката на някои от приносите на автора. Текстът изобилства от граматически и машинописни грешки.

Изложението на т. 2.8 обединява по неподходящ начин абстракции от различни нива, които са представени твърде лаконично и неинформативно. Допуснати са концептуални и терминологични неточности.

Не са посочени пълните библиографски данни на много от цитираните литературни източници. Същото следва да се отбележи и по отношение на списъка на публикациите на Тодор Цонков, представящи резултати от работата му по дисертацията. Заглавието с номер 3 от списъка на публикациите по темата на дисертационния труд по-скоро би трябвало да попадне в списъка на изнесените научни доклади. Съществена част от литературните източници, включени в частта „Използвана литература“, фактически не са цитирани в текста на дисертацията. Не

са цитирани по установения начин и повечето включени в текста таблици и фигури. Представени са ръкописите на статиите, отразяващи резултати от дисертационния труд, вместо техните публикувани варианти.

Като цяло посочените недостатъци са преодолени и не оказват съществено влияние върху преценката за значимостта на изследванията на автора и получените резултати.

5. Автореферат

Авторефератът е изготвен в съответствие с изискванията на Правилника за условията и реда за придобиване на научни степени и за заемане на научни длъжности на ФМИ – СУ и отразява компактно и точно съдържанието на дисертационния труд.

6. Заключение

Въз основа на изложеното по-горе смятам, че дисертационният труд напълно удовлетворява изискванията на ЗРАСРБ, ПЗРАСРБ и Правилниците за условията и реда за придобиване на научни степени и за заемане на научни длъжности на СУ „Св. Климент Охридски“ и ФМИ. **Препоръчвам на уважаемото научно жури да присъди на автора му Тодор Васков Цонков образователната и научна степен „доктор“ по професионално направление 4.6 „Информатика и компютърни науки“ (Информатика – Изкуствен интелект).**

София, 03.05.2016 г.

Член на научното жури:

/проф. д-р М. Нишева/