

## РЕЦЕНЗИЯ

от проф. дмн Галя Ангелова, ИИКТ-БАН  
на дисертационен труд за придобиване на образователната и научна степен „доктор“  
по професионално направление 4.6 „Информатика и компютърни науки“,  
докторска програма: “Софтуерни Технологии” – Откриване на знания  
на тема „Интелигентен агент за диалог на естествен език с отчитане на контекста“  
с автор Момчил Емилов Хардалов

Съгласно заповед № РД-38-600/28.10.2022 г. на Ректора на Софийски университет „Св. Климент Охридски“ (СУ) участвам в Научното жури за присъждане на образователната и научна степен (ОНС) „доктор“ на Момчил Емилов Хардалов, задочен докторант във Факултета по математика и информатика (ФМИ) на СУ през 2017-2021 г., отчислен с право на защита. Темата на дисертацията е насочена към предизвикателствата за създаване на висококачествени „разговорни агенти“ за нуждите на практически приложения: обучения, обяснения към клиенти, отговори на въпроси, съветници в диалогов режим, виртуални асистенти. Тази тема е измежду най-актуалните направления в човеко-машинните интерфейси, които непрекъснато разширяват представите ни за взаимодействие между човека и компютъра. Автоматизацията на интерфейса и надеждността на агента са от първостепенно значение за качествено обслужване на клиенти и потребители. Така че първото положително впечатление от дисертацията идва от амбициозния избор на актуална съвременна област, в която ще се развият приложения през сравнително дълъг период в бъдещето.

Чл. 26, ал. 2 и 3 от Закона за развитие на академичния състав в Република България изисква кандидатите за получаване на ОНС „доктор“ по специалност 4.6 „Информатика и компютърни науки“ да покриват минимални изисквания от 50 точки по показател А (представяне на докторски труд) и 30 точки по показателите група Г. Момчил Хардалов представя справка за изпълнението на минималните национални изисквания по специалност 4.6, която съдържа списък от четири научни статии на английски език, в които той е съавтор. Статиите са публикувани в периода 2018-2020 г., представят резултатите от дисертационния труд, всички са индексирани от Скопус като три от тях имат SJR-ранк. Точките събрани от публикационната дейност (над 108) надхвърлят трикратно изискванията на НАЦИД за минимален праг от 30 точки по показателите група Г при защита на докторска дисертация. С това формалните условия за удовлетворяване на националните критерии са изпълнени и може да се пристъпи към защита на представения дисертационен труд. В първа глава са посочени още две статии, които представят резултати свързани с труда. Общо са намерени 77 цитата на статии, свързани с дисертацията. При търсене в Google Scholar на 14/01/2023 г. за Момчил Хардалов излизат 20 статии по компютърна лингвистика, всички на английски и в престижни международни издания, общо с 413 цитирания (включват се самоцитирания, но все пак бройката е внушителна). Този профил представя г-н Хардалов като зрял специалист по автоматична обработка на естествения език.

Целите на дисертацията са (*i*) създаване на ефективни подходи за изграждане на разговорни агенти, обработващи свободен текст с отчитане на контекста на

специфични задачи като чатботове за обслужване на клиенти, и (ii) създаване на нови езикови ресурси, подпомагащи разработването на интелигентен агент за многоезиков диалог, който може да генерира дълги отговори (напр. статии), за разлика от обичайните къси отговори. Поставени са задачите да се проектират, разработят и оценят (i) компонент разбиращ естествен език, който идентифицира потребителското намерение и разпознава какво е релевантно за неговите слотове; (ii) алгоритъм за синтез на изказвания на базата на външни източници на знания; (iii) цялостни генеративни модели за чатботове с приложение в клиентската поддръжката, водещи диалог с несрещани досега изречения и (iv) система за многоезиков и междуезиков диалог за приложение в разговорен агент.

Дисертационният труд съдържа 154 страници и две приложения А и В на 8 страници. Структуриран е в 6 глави и включва 18 фигури, 39 таблици, библиография от 350 заглавия на английски език, благодарности и списък съкращения.

### **Съдържание и приноси на дисертацията**

Първа глава (Въведение) убедително мотивира актуалността и важността на изследванията и въвежда основното понятие: *“Разговорен агент, наричан още чатбот, е компютърна програма, която се опитва да генерира отговори, подобни на тези, които човек би използвал по време на диалог.”* Агентът има поне три функционалности: *отговаряне на въпроси* (синтезира кратки, директни отговори на потребителски запитвания като използва натрупана информация, извлечени от различни източници); *завършване на задачи* например резервация на ресторант, насрочване на срещи; и *социален чат*: да разговаря свободно и да предоставя полезни препоръки.

Втора глава (Обзор на свързани разработки) е компактно и съдържателно резюме на съвременни подходи за създаване на разговорни агенти. Разгледани са отделните компоненти: разбиране на темата (намерението) на заявката и идентификация на релевантни слотове в диалога, извличане на отговора с фокус върху откриване на вече проверени твърдения, както и техники за реализация на диалог и комбиниране на отговори. Разгледани са и езикови ресурси, нужни за изграждане на разговорни агенти.

Глави трета, четвърта и пета съответно представят резултатите, постигнати при решаване на основните задачи поставени пред дисертационния труд: разбиране на потребителската заявка на естествен език, намиране на отговора, и реализация на диалога с потребителя. Представените решения не са обединени в единен прототип на разговорен агент (което по принцип може да се очаква поради заглавието на дисертацията – “агент” в единствено число). Шеста глава съдържа заключението, описание на приносите и планове за бъдеща работа.

В трета глава (Семантичен анализ на потребителски изказвания) е представен обогатен невронен езиков модел Transformer-NLU за едновременно извършване на класификация на намерения и запълване на слотове, който използва предварително обучените BERT и RoBERTa. Моделът е тестван за английски език над корпусите ATIS (заявки за информация относно самолетни полети) и SNIPS (записи от система за откриване на намерения при персонални гласови асистенти). Въвежда се слой с обединяване на вниманието за класифициране на намеренията, който използва цялостно представяне на

входното изречение. Подобренията са в три посоки: за откриване на намерение се използва обединено представяне от последните скрити състояния за всички токъни; моделът различава наименованите единици благодарение на главните букви (допълнителни характеристики) и на слоя за запълване на слот се подават предсказаният вектор с вероятностното разпределение на разпознатите намерения, последните скрити представяния на BERT и характеристиките на думата. Тестването е извършено с премахване и добавяне на новите компоненти, за да се оцени приносът на предложените подобрения към резултатите. В сравнение с предишни разработки, Transformer-NLU показва по-добри резултати от други модели, базирани на многослойни невронни мрежи.

Четвърта глава (Извличане на отговори от външни източници на знания) представя резултати, свързани с намиране на отговора на постъпилата потребителска заявка. Извличането на вече публикувани статии или други готови документи, които съдържат информация за правилния отговор, осигурява на агента възможност да генерира вече проверени твърдения. Изследвани са два сценария: първи, отговаряне на въпроси с множествен избор, с ползване на извлечени контексти с доказателствени параграфи от външна база от знания и втори, извличане на отговор от колекция с дълги обяснения (документи или статии). При първия сценарий са разгледани невронни модели, които използват външно знание в многоезична среда от езици с богати ресурси (английски) и малко ресурси (български). Най-удобни за експериментиране са корпусите от двойки въпрос-отговор, за които се търсят подходящи обяснителни пасажки. Освен корпуса RACE с над 87 000 въпроса на училищно ниво с множествен избор на английски език е създаден нов корпус на български език с над 2600 въпроса. Използван е многоезичен BERT модел, предварително обучаван над английски или SlavicBERT над славянски езици (български, чешки и полски статии в Уикипедия, както и новини на руски език). За български е получена най-добра точност над 42%. Интересно е наблюдението, че предварителното обучение над славянски езици причинява „катастрофално забравяне“ на научения английски синтаксис и семантика (а обратното не е вярно). При втория сценарий извличането на отговор се преплита с проверка на факти: за даден потребителски коментар да се провери дали изказаното твърдение е било вече проверявано. В дисертацията са разгледани туитове от уебсайта Snopes за периода октомври 2017 г. - октомври 2021 г. Създаден е нов голям корпус CrowdChecked с близо 150 хил. уникални двойки <туит> - <ретуит на „човек от тълпата, проверяващ фактите,“ с връзка към доказателствена статия> и други над 200 хил. туита, включени в разговорите. Вложен е значителен труд в ръчно анализиране и структуриране на разговорите (при което се приема, че „човекът от тълпата“ не допуска грешки). Корпусът CrowdChecked се ползва за експерименти по проверка на факти, в които входни данни са двойките <туит> - <статия за проверка на фактите> етикетирани по два начина: според сходството на Жакард и предсказанията на многослойна невронна мрежа. Базовият модел е Sentence-BERT (SBERT) с разширения за независима оценка на лексикални и семантични сходства, оптимизиран температурен параметър във функцията на грешката, обучение с шумни данни с използване на самоадаптивно обучение и допълнителни тегла във функцията на грешката. Моделът постига усреднена средна точност MAP над 90%, по-добра от други от модели в състезанието CLEF 2021.

Пета глава (Усъвършенствани методи за разговор) представя резултати в областта на управление на диалога и генерация на неизползвани досега нови изказвания на агента, без да се разчита на външни източници. Прототипите за автоматизирано обслужване на клиенти се обучават над корпуси от твитове и отговори от реални фирмени разговори, проведени при диалог с клиентите. Направени са експерименти с разговори в Twitter с два типа агенти: които извличат информация или използват невронни мрежи за синтез на отговори на въпроси. При оценката, извършена над около 50хил. двойки въпрос-отговор, са използвани метрики, оценяващи припокриване на думи и семантика. Проличава по-доброто качество на отговорите синтезирани от генеративните модели, базирани на невронни мрежи, понеже те успяват да генерират граматически правилни и подходящи отговори на зададените им въпроси. От друга страна, при недостатъчно примерни въпроси и отговори за обучение по определена тема, агентите не се справят добре. Предложен е подход за ранкиране на кандидати за следващ отговор, който включва обучение на класификатор за оценка на качеството на двойка въпрос-отговор. В тази глава е представен и нов многоезиков корпус Exams от въпроси и отговори на 16 езика на ниво училищни знания по 24 предмета, който може да се използва в многоезикови сценарии за обучение на системи въпрос-отговор. Корпусът е изграден основно над енциклопедични (проверени) факти от Уикипедия и може да се използва за разработка на многоезикови модели, които могат да разсъждават с наученото знание. Доколкото става ясно от посочените теми за бъдеща работа, Exams може би (според автора) е база фактологична информация, която може да стане основа за обучение на интелигентни модели способни да обясняват съхранени в модела знания.

Шеста глава (Заклучение и планове за бъдеща работа) съдържа резюме на резултатите и изброява оригиналните научни и научно-приложни приноси на дисертацията. Насоките за бъдеща работа съдържат планове за подобряване на моделите (за да се увеличи капацитетът им, да се осигури ефективна актуализация, предпазване от злонамереност, постигане на обяснимост и др.). В тази глава липсва раздел за „апробация“, в който да се изброят публични представяния на резултатите от дисертационния труд на конференции и семинари.

## **Обсъждане**

Дисертацията представя технологични резултати получени с разширени модели и обогатени техники, базирани върху най-съвременните невронни среди за обработка на текста. Създадени са и разнообразни нови езикови ресурси за експериментиране с предложените разширени модели, които изненадват читателя с усилията вложени за събиране на корпуси от твитове или училищни тестове. Поради сбитото представяне на голям обем техническа информация редица въпроси остават без отговор, например сравнително простият пример за слот в статията „Enriched Pre-trained Transformers for Joint Slot Filling and Intent Detection“ отваря въпросът за скалируемостта на подхода към разбиране на по-сложни заявки/намерения и слотове. Друг въпрос в тази глава касае ролята на главните букви при разпознаване на наименованите единици, като имаме предвид, че при запис от гласови асистенти те могат да бъдат поставени неправилно. Има и въпроси по показаните генерирани отговори, напр. защо само IR-моделът в Табл. 5.5 на стр. 82 употребява „please“? Като цяло може да се препоръча включване в труда

на повече текстови примери за решенията, предложени за разгледаните в глави трета, четвърта и пета задачи, или отделно приложение на дисертацията с примери на разбрани намерения и слотове, извлечени отговори и генерирани реплики в диалози.

Приемам научните и научно-приложни приноси на автора така, както са декларирани в шеста глава на дисертационния труд. Според мен най-важното постижение е навлизането на кандидата в „горещата“ проблематика на използване на невронни модели за обработка на свободен текст и приобщаването му към световната практика за участие в международни конференции, интернационални авторски колективи и установени тематични състезания.

Текстът на дисертацията е стегнат и конкретен, подходящо организиран с ясно и логично разделяне по глави и с препратки във всяка глава към списък от авторски публикации, в които се представени резултатите. Авторефератът отразява коректно съдържанието на труда.

Приемам, че г-н Хардалов е равноправен съавтор на публикациите, представящи резултатите на дисертационния труд. Без съмнение той е вложил много голям дял от усилията по конструкция на новосъздадените езикови ресурси.

Не откривам плагиатство в дисертационния труд, който се базира върху текстовете на научни публикации на изброените съавтори.

Познавам г-м Хардалов от студентските му години в магистърската програма ИИОЗ на катедра „Софтуерни технологии“. Приятно е да се наблюдава израстването на млад специалист по компютърна лингвистика, още повече с вкус към обработка на български текст и създаване на езикови ресурси за славянски езици.

## Заклучение

Законът за развитието на академичния състав в Република България изисква докторският труд да съдържа научни или научно-приложни резултати, които представляват оригинален принос в науката. Дисертацията трябва да показва, че кандидатът притежава задълбочени знания по съответната специалност и способности за самостоятелни научни изследвания. Тези изисквания са изпълнени за дисертационния труд на Момчил Хардалов и научните статии, в които са апробирани получените резултати. Големият брой цитирания на публикациите, свързани с дисертацията, както и на останалите научни трудове на кандидата, е доказателство за широк международен отзвук и интерес.

На тези основания ще гласувам положително за присъждане на степента и убедено предлагам на уважаемото Научно жури да **присъди на Момчил Емилов Хардалов образователната и научна степен "доктор"** по професионално направление 4.6 „Информатика и компютърни науки“.

17 януари 2023 г.

Член на Научното жури:

проф. дмн Галя Ангелова