



РЕЦЕНЗИЯ

За нуждите на публична защита на дисертационен труд

Изготвена от: проф. д.ик.н. Антон Герунов, катедра „Икономика и управление по отрасли“, Стопански факултет на Софийския университет „Св. Климент Охридски“

На дисертация на тема: Приложение на методите на машинното самообучение за решаване на бизнес задачи

Дисертант: Сергей Сергеев Вичев

1. Обща информация за кандидата

Сергей Сергеев Вичев завършва висшето си образование в Университета за национално и световно стопанство (УНСС) с отлични резултати – първоначално като бакалавър (2012), а след това и като магистър (2018) по бизнес администрация. Зачислен е в Стопанския факултет през 2021 като редовен докторант в ПН 3.8 Икономика и е отчислен с право на защита през 2024 г. със заповед РД-20-1364/29.07.2024. Извън академичните си търсения Сергей Вичев има и богат практически опит – като анализатор на данни, основател и управляващ директор на компании, а към настоящия момент и като главен научен директор в консултантска компания с фокус върху изкуствения интелект. Отчитам богатия практически опит като отлична основа, но и като ценно допълнение към научния профил на докторанта.

2. Характеристика на дисертационния труд

Представената дисертация се състои от 155 стр. основен текст и списък с информационни източници (библиография). Структурата на дисертационният труд е логична, като резултатите са разделени в три глави, които последователно изпълняват поставените цели и задачи. От структурна гледна точка е удачно решението отделните глави да имат поставени свои задачи, както и да завършват с преглед и обобщение на основните резултати. **Увода** съдържа основните необходими реквизити – актуалност на темата, цел и задачи, предмет и обект на изследването. Отчитаме липсата на явно изведена научна хипотеза в уводната част, което би било полезно, макар и хипотезите да се подразбират достатъчно ясно в самите изследователски глави.

Първа глава „Теоретични основи“ е по същество обзорна. Тя очертават теоретичния фундамент, на който стъпва последващото изследване. Разглеждат се основните понятия и се дават подробни сведения за големите езикови модели (ГЕМ). Прави се обзор и на приложенията на този клас модели в услуга на бизнеса, като дисертантът се спира с особено внимание върху мултимодалност, генерация с обогатено извличане на информация и специализиращото обучение (fine-tuning). Главата създава в читателя увереност, че докторантът познава задълбочено основни теоретични и практически постановки в областта на дисертацията. Това се потвърждава и от големия обем актуални научни изследвания в библиографията. В предишен вариант на дисертацията Първа глава беше значителна по-обем, но във финалната версия е осезателно съкратена (от 82 стр. на 54 стр.), което я прави по-четима, а и по-полезна за изследователи и практики.

Втора глава „Методика за изграждане на автоматизирана система за решаване на задълбочени бизнес проблеми върху информация от база данни“ представя общото описание на архитектурата на разработената система. Разглеждат се методите за прогнозиране с ГЕМ, подходите за подобрене на вграждащи (embedding) модели, както и подробна блок схема на предложената автоматизирана система. Въпросната архитектура е значителна по обем, като докторантът детайлно и добросъвестно разработва всеки от нейните компоненти. На някои дисертантът включва сравнително големи блокове код в основния текст, но много от тях по същество не са с приносен характер. Бих препоръчал блоковете код (или поне по-голямата част от тях) да бъдат изнесени в нарочно приложение. Това важи особено за по-несъществени от тях като използвани версии на Python библиотеки (стр. 101), общ код наличен в документацията на съответните библиотеки и пакети (пр. стр. 103, но и други), код за импортиране на използване библиотеки (стр. 106) и т.н. В много случаи описанието на процедурите и параметрите на налични в документацията на самата библиотека (често това е Langchain) и препоръчвам дисертантът да реферира към първоизточника вместо същия текст да залегне в дисертацията. В особена сила това важи за описанието на аргументите на използваните функции.

Трета глава „Симулации и експерименти с автоматизирани системи за решения от бази данни“ е по същество основната приносна глава на дисертацията. Тук се апробират две различни приложения на предложената методология (наречени „експерименти“). При първия експеримент се апробира предложената архитектура върху сериозната игра SQL Murder Mystery (Canale & Farinetti, 2022¹), като се разглеждат три нейни варианта – агент без памет, агент с

¹ Canale, L., & Farinetti, L. (2022, June). SQL Murder Mystery: a serious game to learn querying databases. In *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 129-138). IEEE.

контекстуална памет и многоагентна система. По същество подчертаваме, че тази игра е по-скоро разработена с обучителна цел и в този смисъл не отговаря на тясната дефиниция на бизнес приложение. Препоръчвам на дисертанта да обоснове защо демонстрираните функционалности и резултатите в рамките на тази обучителна сериозна игра са близки до тези, които бихме видели в реални бизнес казуси. Настоящата аргументация („Използваме този конкретен казус, защото вече е добре измислен и многократно изпробван“, стр. 123) е недостатъчна. Вторият „експеримент“ по същество цели да провери дали специализиращото обучение на модела за вграждане (embedding model) в рамките на предложената от дисертанта архитектура води до подобрени резултати (измерени чрез NDCG метриката). Тук вероятно е постигнат и най-интересният резултат от дисертацията: оказва се, че предварителното обучение на (много малък) модел за вграждане води до значително по-добре резултати спрямо значително по-големи модели, които не са преминали предварително обучение (fine-tuning). Изводът е от несъмнен интерес както за теорията, така и за практиката.

Дисертацията завършва със заключение и набор от научни и научно-приложни приноси.

3. Оценка на получените резултати и научните и приложни приноси

Несъмнено представеният дисертационен труд е оригинална и интересна научна разработка с всички необходими реквизити за докторска дисертация. **Приносните моменти в дисертационния труд са достатъчни като количество и качество и напълно отговарят на изискванията за придобиване на ОНС „доктор“.** Приемам така формулираните приноси, но те следва да бъдат редактирани и прецизирани допълнително.

По същество принос 1 следва допълнително да се коригира. Тъй като вече съществуват набори от сходни методически рамки за извличане на знания (вж. пр. Xie et al., 2024², но и др.) и в този смисъл приносът следва да се прецизира, за да се подчертае какво е новото познание, което дисертацията предоставя. Принос 2 следва да се фокусира върху придобития резултат относно представянето на езикови модели с малка размерност. Принос 3 се състои предимно в сравнителните резултати спрямо рамката ReAct.

Приемам напълно приноси 4 и 5. Тук следва да се подчертае несъмнената полза от разработването на допълнителни артефакти в рамките на дисертационното

² Xie, W., Wu, G., & Zhou, B. (2024). Mag-sql: Multi-agent generative approach with soft schema linking and iterative sub-sql refinement for text-to-sql. arXiv preprint arXiv:2408.07930.

усилие, които да позволят не просто възпроизвеждане на получените резултати, но и надграждане на представените изследвания.

4. Критични бележки и препоръки

Конкретни критичните бележки и препоръки са отправени в т. 1 спрямо съответните глави. Струва си да се отбележат и някои съображения от по-общ характер:

- Дисертацията по същество представлява един чудесен пример за проектираща наука (design science), при която дисертанта създава конкретен приложим артефакт (Dresch et al., 2015)³. Макар че видимо подходът и методиката на работа плътно следват принципите и практиките на проектиращата наука, то това не е изрично споменато в методологичната секция. Бих препоръчал това да се добави като допълнителен нюанс на работата. Приложението на този подход по същество може да представлява и методологически принос в областта.
- Темата на дисертацията се фокусира върху решаване на бизнес задачи, докато съществена част от приносите не са с изрична бизнес насоченост. Подзаглавието вече насочва към „сложни задачи в бази данни“, но това също не е особено ясно дефиниран термин. Бих препоръчал в своите бъдещи изследвания докторантът да разгледа как представените резултати по същество могат да бъдат приложени към конкретни бизнес задачи.
- Ключов въпрос при оценката на алтернативни архитектури е подходът за оценка на постигнатите резултати. Дисертантът се спира върху NDCG, но в литературата и практиката съществуват различни подходи, особени при декомпозираната оценка на различните компоненти от генерацията с обогатено извличане на информация (RAG), вж. пр. Es et al., 2024⁴. Удачно е да има по-разширена обосновка на методологическите съображение относно избора на специфична рамка за оценка.

Подчертавам, че отправените коментари не намаляват стойността на получените резултати и следва да се считат предимно като потенциални насоки за подобрене на представения дисертационен труд.

³ Dresch, A., Lacerda, D. P., Antunes Jr, J. A. V., Dresch, A., Lacerda, D. P., & Antunes, J. A. V. (2015). *Design science research* (pp. 67-102). Springer International Publishing.

⁴ Es, S., James, J., Anke, L. E., & Schockaert, S. (2024, March). Ragas: Automated evaluation of retrieval augmented generation. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations* (pp. 150-158).

5. Оценка на публикационната дейност и съответствие с националните минимални наукометрични изисквания

Докторант Сергей Вичев представя **доказателства за значителна публикационна дейност** по темата на дисертационния труд:

- Sergey Vichev, Angel Marchev, Reasoning capabilities of large language models on information extracted from databases using Text-to-SQL, PROCEEDINGS OF THE 49TH INTERNATIONAL CONFERENCE "APPLICATIONS OF MATHEMATICS IN ENGINEERING AND ECONOMICS", 2025, doi:<https://doi.org/10.1063/5.0247481>, Ref, IF (- 2025), SCOPUS, SJR (- 2025)
- Sergey Vichev, A Multi-Agent Framework for Large Language Model-Driven SQL Analysis, KSI Transactions on KNOWLEDGE SOCIETY, vol:XV, issue:Nr. 1, 2024, ISSN (print):1313-4787
- Vichev, Sergey, Marchev, Angel, RAGSQL: Context Retrieval Evaluation on Augmenting Text-to-SQL Prompts, International IEEE Conference proceedings, IS, 2024, doi:10.1109/IS61756.2024.10705186, Ref, SCOPUS, SJR (- 2024)
- Vichev, Sergey, Application of K-means Clustering Algorithm with RFM for Customer Segmentation in Logistics Sector., Vanguard Scientific Instruments in Management, issue:19, 2023, ISSN (print):ISSN: 1314-0582, Ref, др.(CEEOL)

Тук следва да се отбележи факта, че от приведените четири публикации, две са индексирани в световноизвестната база данни с научна информация Scopus, което показва високото ниво на проведените изследвания.

Допълнително е представена и Справка за изпълнение на критериите и изискванията към докторските тези за получаване на образователна и научна степен „доктор“ на Стопанския факултет, СУ „Св. Климент Охридски“. От нея е видимо, че освен публикациите, докторантът има 4 доклада на национални научни школи и семинари, както и още 3 доклада на международни научни форуми в страната. Извън това докторантът преподава и курса Бази данни във финансите, което свидетелства и за неговия преподавателски опит.

На база на представените документи може да се достигне до заключението, че докторант Сергей Вичев не просто отговаря, но **значително надхвърля националните минимални наукометрични изисквания** за придобиване на ОНС „доктор“ (чл. 12 от ЗРАСРБ, чл. 35 от Правилника за прилагане на ЗРАСРБ и чл. 1а, ал. 1 от ППЗРАСРБ).

6. Оценка на автореферата

Авторефератът подробно и точно представя основните резултати от дисертацията.

7. Заключение

Дисертационният труд представлява задълбочено и добросъвестно научно изследване, което ясно демонстрира дълбоките теоретични познания на докторанта, както и уменията за провеждане на самостоятелно научно изследване на високо ниво. Текстът съдържа научни резултати, които представляват оригинален принос и отговарят на минималните изисквания на Закона за развитие на академичния състав в Република България (ЗРАСРБ) и Правилника за прилагане на ЗРАСРБ. Смятам, че представената дисертация напълно отговаря на изискванията на чл. 6, ал. 3 от ЗРАСРБ и давам своята **положителна оценка на представения труд**, като убедено препоръчвам на уважаемото научно жури да присъди ОНС „доктор“ на докторант Сергей Сергеев Вичев в ПН 3.8 Икономика.

Рецензент:

Проф. д.ик.н. Антон Герунов

София, 16.06.2025 г.