С Т А Н О В И Щ Е

Относно: дисертационен труд за присъждане на образователна и научна степен „доктор“, в професионално направление 3.8 Икономика, ДП-Аналитични изследвания върху данни (Data Science), СУ „СВ. КЛИМЕНТ ОХРИДСКИ“

Автор на дисертационния труд: *Сергей Сергеев Вичев*

Тема на дисертационния труд: *Приложение на методите на машинното самообучение за решаване на бизнес задачи - Многоагентни автоматизирани системи за решаване на сложни задачи в бази данни, с помощта на големи езикови модели.*

Основание за представяне на становището: участие в състава на научното жури по защита на дисертационния труд, съгласно Заповед РД38-220/24.04.2025 на Ректора на Софийски университет „Св. Климент Охридски“

Дисертантът Сергей Сергеев Вичев се е обучавал като редовен докторант в ДП „Аналитични изследвания върху данни (Data Science)“, към катедра „Статистика и иконометрия“, професионално направление 3.8 „Икономика“, от 2021 г. до 2024 г. Успоредно с това той развива своята професионална реализация в областта на софтуерните системи и анализа на данни (data science), което допринася за практико-приложните качества на труда.

Трудът е с общ обем от 171 стр., като се следва структура в общо от три глави. Използвани са общо 118 литературни източника, предимно на английски език. Задължителните атрибути за подобен вид изследване са логично дефинирани. Определено темата има висока степен на актуалност и значимост в контекста на развитието на системите, базирани на изкуствен интелект. Трудът е структуриран в следните глави: 1) теоретични основи; 2) методика на изграждане на автоматизирана система за решаване на задълбочени бизнес проблеми върху информация от база данни, както и 3)

симулации и експерименти с автоматизирани системи за решения от бази данни.

В първата глава се изясняват теоретичните основи на същността на големите езикови модели (LLMs) и тяхното приложение в бизнеса. Особено внимание се отделя на тяхната мултимодалност, системите за генерация с обогатено извличане (RAG), както и на специализираното обучение на големите езикови модели. Обхванати са и проблемите на обучението на моделите, както и техните способности за аргументация. Дефинирани са и опорните моменти за внедряването на LLM в бизнеса. В тази част е поместена и концептуалната постановка на изследването, в която се прави мотивация за фокуса върху SQL. Втората глава е посветена на методиката на изграждане на автоматизирани системи за решаване на задълбочени бизнес проблеми върху информация от база данни. В нея фокусът е върху методите за прогнозиране с големи езикови модели, както и върху процесите на вграждане (embeding) на данни. В главата се изяснява и архитектурата на RAGSQL, както и на различни възможности за реализация на тази система. В последната глава се провеждат симулации и експерименти с разработената система за решения. В нея се представят принципните приложения на системата, както и изследване на критичните компоненти.

Разработката постига целта си и достига до фундаментите на разработването на многоагентни автоматизирани системи, базирани на големи езикови модели и интегрирани с бази данни. В работата проличава практическият опит на автора и търсенето на силно приложни резултати. Общото впечатление от труда е за много доброто познаване на материята и за направата на задълбочен анализ. След запознаване с материалите, считам текстовете в труда за авторски.

Определено могат да бъдат посочени и редица приноси в труда. В основата е разработената архитектура на система за автоматизирано решаване на задачи за анализ, базирана на многослойната координация между отделни LLM агенти и анализ на получените справки (RAGSQL). Предложен е нов подход за специализиращо обучение на модели със семантично вграждане, в контекста на BIRD. Също така е предложено практическо приложение на системи, базирани на ReAct. Представени са различни принципни постановки за функционирането на системата, като се използва задачата SQL Murder Mystery. За нейното решаване се започва със

симплифициран модел, след което се надгражда до използването на LLM супервайзър и няколко LLM анализатори на базата данни. По този начин се цели представяне на функционирането на системата в различни условия и да се проследи нейното поведение.

Към труда са приложени три самостоятелни публикации, както и четвърта в съавторство с научния ръководител. Като цяло те популяризират разработката, като някои от тях правят и допълнително надграждане на темата в посока логистичен сектор. Авторефератът също може да бъде оценен положително.

Към разработката могат да бъдат отправени и критични бележки, като ще маркирам част от тях, които имат по-фундаментален характер. Като цяло връзката с бизнеса и управлението може да бъде засилена, като се навлезе във функционалните направления на организацията. По този начин могат да бъдат дефинирани по-ясно областите на приложение в бизнес среда, като конкретни процеси в организацията. Темата може да се разшири и в условията на интердисциплинарност и връзка между функционалните направления. Отделно от това трудът би спечелил, ако мотивацията за избора на SQL базите данни се разшири и се обосновава от конкретните потребности на бизнес-организациите.

Въпреки посочените бележки, дисертационният труд на тема „Приложение на методите на машинното самообучение за решаване на бизнес задачи - Многоагентни автоматизирани системи за решаване на сложни задачи в бази данни, с помощта на големи езикови модели“, изготвен от Сергей Сергеев Вичев, представлява актуално научно изследване, притежаващо научни и научно-приложни приноси за теорията и практиката. Въз основа на това давам своята положителна оценка и предлагам на уважаемите членове на научното жури да присъдят на Сергей Сергеев Вичев образователната и научна степен „доктор“ по професионално направление 3.8 „Икономика“, ДП-Аналитични изследвания върху данни (Data Science).

29 май 2025 г. *доц. д-р Николай Драгомиров Иванов*

гр. София /……../