

проф. д-р инж. Милена Карова, преподавател, КНТ
Пламена Атанасова Атанасова, студент, КНТ, председател на студентски клуб “Creative Code;”
Виктор Иванов Иванов, студент, СИТ, член на студентски клуб “Creative Code;”
Атанас Петров Маргаритов, студент, СИТ, член на студентски клуб “Creative Code;”
Данаил Николаев Димитров, студент, СИТ, член на студентски клуб “Creative Code;”

Въведение

Развитието на интелигентни системи за автоматизиран достъп до информация е важен елемент от съвременния научен и технологичен напредък. В сферата на висшето образование тяхното значение е особено голямо, тъй като дигиталните асистенти подобряват ефективността на комуникацията между студенти, преподаватели и администрация. Тези системи се базират на големи езикови модели (Large Language Models – LLM), които могат да разбират контекста и да генерират уместни отговори, съобразени с начина на общуване на потребителя. Основните предизвикателства при създаването на генеративни асистенти са свързани с необходимостта от изграждане на качествени набори от данни и ограничената наличност на ресурси за български език. За да се постигне надеждност на системата, данните трябва да бъдат внимателно етикетирани и да покриват широк спектър от възможни сценарии.

Избор на AI модел

В рамките на проекта е избран езиков модел от типа BERT, който благодарение на двупосочното си контекстуално обучение осигурява висока степен на семантично разбиране на текста. Това го прави особено подходящ за университетска информационна среда, където приоритет са точността, надеждността и коректността на предоставяната информация, а не единствено генерирането на нов текст. Изборът на подходящ модел е от съществено значение, тъй като определя качеството на взаимодействието между потребителя и системата, както и степента, в която дигиталният асистент може да отговори на реалните информационни потребности на студентите и преподавателите.

Заклучение

Проектът успешно реализира генеративен дигитален асистент, който улеснява достъпа до академична информация и оптимизира комуникацията в университетската среда. Разработената система вече доказва своята практическа стойност чрез предоставяне на персонализирани и надеждни отговори на често задавани въпроси. Тя е устойчива, мащабируема и създава основа за бъдещо внедряване на допълнителни функционалности в подкрепа на студентите и преподавателите.

Резултати

В рамките на проекта бе изградена и експериментално валидирана многокомпонентна система за генеративен дигитален асистент, която обединява мобилно приложение, сървърна инфраструктура, релационна база от данни и модул за изкуствен интелект. Особен акцент бе поставен върху интеграцията на модели за обработка на естествен език (NLP), което позволи разработване на система с висока степен на адаптивност към различни сценарии на взаимодействие. Чрез задълбочени теоретични изследвания и практически анализи бяха приложени съвременни трансформаторни архитектури и методи за векторни представяния на текст, които допринесоха за разширяване на познанията в областта на семантичния анализ и управлението на диалогови системи. Експерименталната работа обхвана събиране, етиктиране и структуриране на данни, включително реални информационни запитвания на студенти и преподаватели, като по този начин бе гарантирана представителност и практическа приложимост на обучителните множества. Прилагането на количествени и качествени метрики като точност, прецизност, пълнота и експертна оценка осигури научната достоверност и надеждността на получените резултати.

Разработеният прототип на дигиталния асистент демонстрира успешна интеграция на генеративния езиков модел в реална университетска информационна среда. Асистентът предоставя персонализирани, контекстуално релевантни и своєвременни отговори на въпроси, свързани с академични програми, административни услуги, събития и ежедневието на студентите. Използването на Rasa Framework даде възможност за модулно управление на диалога, което позволява гъвкавост и последващо разширяване с допълнителни модули. Проведените стрес-тестове и симулации на експлоатационни условия доказаха устойчивостта на системата при високо натоварване и многопоточна обработка на заявки, като се гарантира запазване на бързодействието и надеждността на генерираните отговори. Допълнително, системата е проектирана така, че да бъде лесно мащабируема, което позволява интеграция на нови бази данни, функционалности и услуги без необходимост от съществени промени в архитектурата.

Постигнатите резултати имат висока практическа значимост, тъй като доказват възможността системата да бъде внедрена в реална академична среда с цел подобряване на достъпа до информация и оптимизация на комуникацията между студенти, преподаватели и администрация. Създадената технологична основа е мащабируема и подлежи на надграждане с нови функционалности, което гарантира устойчивост и дългосрочна приложимост. По този начин проектът допринася както за научното развитие в областта на изкуствения интелект и обработката на естествен език, така и за утвърждаването на Технически университет – Варна като среда за иновации и практическо приложение на съвременни технологии

Публикации по проекта

1. Атанасова П., Димитров Д., Сравнителен анализ на съвременни AI модели за обработка на естествен език, СНС “Наука в бъдещето 2025”, Технически Университет Варна 2025.

Благодарности

Научните изследвания, резултати, от които са представени в настоящата публикация, са извършени по проекта в рамките на научноизследователската дейност на Технически университет – Варна, финансиран от държавния бюджет.