8D06101 – «Ақпараттық жүйелер» білім беру бағдарламасы бойынша PhD дәрежесін алу үшін ұсынылған Сарсенбаева Асия Калисеитовнаның «Қазақстан Республикасының заңнамасы саласындағы мемлекеттік тілге арналған семантикалық тәсілдер негізінде модельдерді зерттеу және интеллектуалды сұрақ-жауап жүйесін әзірлеу» тақырыбындағы диссертациялық жұмысына ғылыми жетекшінің

ПІКІРІ

Диссертациялық жұмыс Қазақстан Республикасының заңнамасы саласындағы мемлекеттік тілге арналған интеллектуалды сұрақ-жауап жүйелерін әзірлеуге арналған және жасанды интеллект, табиғи тілдерді өңдеу (NLP), үлкен тілдік модельдер (LLM) және Retrieval-Augmented Generation (RAG) технологиялары саласындағы өзекті ғылыми бағыттардың бірі болып табылады.

Зерттеу тақырыбының өзектілігі келесі факторлармен анықталады:

— қазақ тіліндегі құқықтық NLP және Legal AI бағыттарының жеткілікті деңгейде дамымауы;

— мемлекеттік тілдегі құқықтық ақпаратқа қолжетімділікті арттыру қажеттілігі;

— құқықтық мәтіндердің жоғары терминологиялық және семантикалық күрделілігі;

— қазақ тілінің агглютинативті құрылымына байланысты semantic retrieval жүйелерін құрудағы қиындықтар;

— құқықтық сұрақ-жауап жүйелерінде LLM модельдерінің hallucination мәселесінің болуы.

Диссертациялық зерттеудің мақсаты – қазақ тілінің морфологиялық ерекшеліктерін ескеретін, semantic preprocessing, retrieval және үлкен тілдік модельдерді біріктіру негізінде Қазақстан Республикасының заңнамасы бойынша интеллектуалды сұрақ-жауап жүйесін әзірлеу.

Қойылған мақсатқа жету үшін келесі міндеттер шешілді:

— Қазақстан Республикасының заңнамалық құжаттары негізінде мамандандырылған құқықтық деректер жиынын қалыптастыру;

— құқықтық мәтіндерді жинау, алдын ала өңдеу және аннотациялау әдістерін зерттеу;

— үлкен тілдік модельдерді құқықтық доменге fine-tuning арқылы бейімдеу;

— пайдаланушы сұраныстарын құқықтық терминологиямен сәйкестендіруге арналған SemanticPrep семантикалық препроцессинг тәсілін әзірлеу;

— retrieval-based құқықтық сұрақ-жауап жүйесінің архитектурасын құру;

— ұсынылған әдістердің тиімділігін эксперименттік бағалау.

Диссертациялық жұмыста келесі ғылыми жаңалығы бар нәтижелер алынған:

— алғаш рет қазақ тіліндегі заңнамалық сұрақ-жауап міндеттеріне арналған 8000-нан астам сұрақ-жауап жұбынан және 5.3 ГБ құқықтық мәтіндерден тұратын мамандандырылған деректер жиыны қалыптастырылды;

— mT5 трансформері мен SVD/LSA әдістерін біріктіретін екі кезенді SemanticPrep семантикалық препроцессинг тәсілі ұсынылды;

— құқықтық доменге бейімделген fine-tuned LLM модельдері әзірленді;

— semantic preprocessing, retrieval және генеративті модельдерді біріктіретін Law-AI интеллектуалды сұрақ-жауап жүйесінің пайплайны құрылды;

— ұсынылған тәсілдің тиімділігі эксперименттік түрде дәлелденіп, ROUGE-L көрсеткішінің 6–21%-ға, BERTScore нәтижесінің 3–7%-ға, Recall@5 мәнінің 17–28%-ға және MRR метрикасының 30–40%-ға жақсарғаны анықталды.

Диссертациялық жұмыстың практикалық маңыздылығы мемлекеттік тілдегі құқықтық ақпаратқа қолжетімділікті қамтамасыз етуге арналған интеллектуалды Law-AI жүйесін әзірлеумен анықталады. Ұсынылған әдістер құқықтық кеңес беру платформаларында, электрондық мемлекеттік сервистерде және заңнамалық ақпараттық жүйелерде қолдануға бағытталған. Сонымен қатар зерттеу нәтижелері low-resource language жағдайындағы Legal AI бағытындағы болашақ зерттеулер үшін ғылыми негіз қалыптастырады.

Диссертациялық зерттеудің негізгі нәтижелері халықаралық ғылыми журналдар мен конференцияларда жарияланған. Атап айтқанда, зерттеу нәтижелері Web of Science және Scopus базаларында индекстелетін басылымдарда, соның ішінде Computers (процентилі 84), халықраралық конференциялар ACIIDS 2025, UBMK 2025 және Springer CCIS жинақтарында жарияланған. Сонымен қатар, ҚР ҒЖБМ Ғылым және жоғары білім саласындағы сапаны қамтамасыз ету комитеті ұсынған басылымдарда 3 мақалалар жарық көрген. Бұл алынған нәтижелердің ғылыми маңыздылығын және халықаралық деңгейдегі апробациядан өткендігін көрсетеді.

Диссертациялық жұмыс аяқталған ғылыми-біліктілік зерттеу болып табылады. Жұмыста құқықтық сұрақ-жауап жүйелері мен жасанды интеллект саласындағы маңызды ғылыми-техникалық мәселені шешуге бағытталған жаңа ғылыми нәтижелер алынған. Зерттеу жұмысының мазмұны, ғылыми деңгейі және алынған нәтижелері Қазақстан Республикасы Ғылым және жоғары білім министрлігінің ғылыми дәрежелер беру туралы талаптарына толық сәйкес келеді.

Сарсенбаева Асия Калисеитовна 8D06101 – «Ақпараттық жүйелер» білім беру бағдарламасы бойынша философия докторы (PhD) дәрежесін алуға лайық деп есептеймін. Диссертация философия докторы (PhD) дәрежесін алуға ұсынылады.

Отандық ғылыми кеңесші:
PhD, қауымдастырылған профессор
әл-Фараби атындағы Қазақ ұлттық университеті
«Ақпараттық жүйелер» кафедрасының
доцент-зерттеушісі
Д.Р. Рахимова

БАСТАЙМЫН
әл-Фараби атындағы ҚазҰУ Ғылыми кадрларды
даярлау және аттестаттау басқармасының басшысы
ЗАВЕРЯЮ
Начальник управления подготовки и аттестации
научных кадров КазНУ им. аль-Фараби
«18» 05 2026 ж./г.

— mT5 трансформері мен SVD/LSA әдістерін біріктіретін екі кезенді SemanticPrep семантикалық препроцессинг тәсілі ұсынылды;

— құқықтық доменге бейімделген fine-tuned LLM модельдері әзірленді;

— semantic preprocessing, retrieval және генеративті модельдерді біріктіретін Law-AI интеллектуалды сұрақ-жауап жүйесінің пайплайны құрылды;

— ұсынылған тәсілдің тиімділігі эксперименттік түрде дәлелденіп, ROUGE-L көрсеткішінің 6–21%-ға, BERTScore нәтижесінің 3–7%-ға, Recall@5 мәнінің 17–28%-ға және MRR метрикасының 30–40%-ға жақсарғаны анықталды.

Диссертациялық жұмыстың практикалық маңыздылығы мемлекеттік тілдегі құқықтық ақпаратқа қолжетімділікті қамтамасыз етуге арналған интеллектуалды Law-AI жүйесін әзірлеумен анықталады. Ұсынылған әдістер құқықтық кеңес беру платформаларында, электрондық мемлекеттік сервистерде және заңнамалық ақпараттық жүйелерде қолдануға бағытталған. Сонымен қатар зерттеу нәтижелері low-resource language жағдайындағы Legal AI бағытындағы болашақ зерттеулер үшін ғылыми негіз қалыптастырады.

Диссертациялық зерттеудің негізгі нәтижелері халықаралық ғылыми журналдар мен конференцияларда жарияланған. Атап айтқанда, зерттеу нәтижелері Web of Science және Scopus базаларында индекстелетін басылымдарда, соның ішінде Computers (процентилі 84), халықараралық конференциялар ACIIDS 2025, UBMK 2025 және Springer CCIS жинақтарында жарияланған. Сонымен қатар, ҚР ҒЖБМ Ғылым және жоғары білім саласындағы сапаны қамтамасыз ету комитеті ұсынған басылымдарда 3 мақалалар жарық көрген. Бұл алынған нәтижелердің ғылыми маңыздылығын және халықаралық деңгейдегі апробациядан өткендігін көрсетеді.

Диссертациялық жұмыс аяқталған ғылыми-біліктілік зерттеу болып табылады. Жұмыста құқықтық сұрақ-жауап жүйелері мен жасанды интеллект саласындағы маңызды ғылыми-техникалық мәселені шешуге бағытталған жаңа ғылыми нәтижелер алынған. Зерттеу жұмысының мазмұны, ғылыми деңгейі және алынған нәтижелері Қазақстан Республикасы Ғылым және жоғары білім министрлігінің ғылыми дәрежелер беру туралы талаптарына толық сәйкес келеді.

Сарсенбаева Асия Калисеитовна 8D06101 – «Ақпараттық жүйелер» білім беру бағдарламасы бойынша философия докторы (PhD) дәрежесін алуға лайық деп есептеймін. Диссертация философия докторы (PhD) дәрежесін алуға ұсынылады.

Отандық ғылыми кеңесші:
PhD, қауымдастырылған профессор
әл-Фараби атындағы Қазақ ұлттық университеті
«Ақпараттық жүйелер» кафедрасының
доцент-зерттеушісі
Д.Р. Рахимова