әл-Фараби атындағы Қазақ Ұлттық университеті Ақпараттық-технологиялар факультетінің «6D075100 – Информатика, есептеу техникасы және басқару» мамандығының докторанты Елеусинов Арман Инабатұлының «Қазақ тілінде қолжазба мәтіндерді тану және машиналық оқыту алгоритмдері мен ақпараттық моделін құру» тақырыбына жазған философия докторы (PhD) дәрежесін алу үшін дайындалған диссертациясына ғылыми жетекшісі PhD, доцент м.а Бурибаев Жолдас Алладиновичтің

ПІКІРІ

Елеусинов Арманның диссертациялық жұмысы қазақ тіліндегі қолжазба мәтіндерін тануға арналған. Диссертациялық жұмыстың өзектілігі цифрлық технологиялардың қарқынды дамуына байланысты мәтіндік ақпаратты автоматты түрде өңдеу және тану жүйелерінің қажеттілігінде. Әсіресе, тарихи, архивтік және әкімшілік құжаттарда сақталған қолжазбаларды цифрландыру қажеттілігі күн тәртібінде тұр. Алайда, қазіргі қолжетімді жүйелердің көпшілігі қазақ тілінің графикалық, морфологиялық және синтаксистік ерекшеліктерін ескермейді. Осы орайда, диссертациялық жұмыс қазақ тіліндегі қолжазба мәтіндерін тану мәселесін терең қарастырып, заманауи машиналық оқыту әдістері негізінде тиімді шешу жолдарын ұсынуы – ғылыми тұрғыда да, практикалық мағынада да аса өзекті болып табылады.

Диссертациялық жұмыстың негізгі ғылыми жаңалығы – қазақ тілінде жазылған қолжазба мәтіндерге арналған толыққанды өңдеу жүйесінің құрылуында және отандық деректер жиынтығы негізінде бейімделген нейрондық желілердің тиімділігін дәлелдейтін нақты эксперименттік нәтижелердің ұсынылуында. Арман қазақ тіліндегі қолжазбалар үшін арнайы деректер жинағын құрып, оның құрылымдық ерекшеліктеріне негізделген аннотациялау әдістерін қолданды. Бұл деректер жиыны – зерттеудің сапалы негізі ретінде танылып, әрі қарайғы модельдерді оқытуда басты рөл атқарды. Сондай-ақ, жұмыста қолжазба мәтінді тану үшін гибридті нейрондық архитектуралар (CNN+Transformer) қолданылған. Сонымен қатар, деректерді жасанды ұлғайту үшін GAN (Generative Adversarial Network) үлгілерін қолдану – қазақ тіліндегі мәтіндерге бейімделген алғашқылардың бірі. Бұл тәсілдер дәстүрлі әдістермен салыстырғанда жоғары нәтижелер көрсеткен.

Зерттеу нәтижелері теориялық тұрғыдан машиналық оқыту алгоритмдерін бейімдеу, мәтіндік деректерді алдын ала өңдеу, сегменттеу және тану үдерістерін жүйелеу бағытында маңызы зор. Жұмыста ұсынылған архитектуралар мен алгоритмдер тек қазақ тілі үшін ғана емес, басқа түркі тілдерінде жазылған қолжазбаларды тану үшін де негіз бола алады.

Практикалық тұрғыдан бұл зерттеу архив құжаттарын автоматты тану, білім беру және мемлекеттік құжат айналымы салаларында қолдануға бейімделген нақты ақпараттық жүйе жасауға мүмкіндік береді. Жобада әзірленген модельдер NVIDIA Triton Inference Server платформасында интеграцияланып, тәжірибелік тұрғыдан жоғары деңгейде жүзеге асырылған.

Диссертациялық жұмыста барлық ұсынылған тәсілдер мен нәтижелер бірнеше дербес эксперименттер арқылы тексеріліп, сандық метрикалармен бағаланған. Әсіресе,

жол, сөз және таңба деңгейінде тану дәлдігінің артуы, сондай-ақ WER (Word Error Rate), CER (Character Error Rate) сияқты көрсеткіштермен есептелген нәтижелердің нақты сипаттамалары жұмыстың сенімділігін көрсетеді. Жүйенің тиімділігі бірнеше модельмен салыстырмалы түрде зерттеліп, олардың артықшылықтары мен шектеулері көрсетілген.

Елеусинов Арманның зерттеу үдерісінің барлық кезеңдерін – бастапқы деректерді жинау, оларды өңдеу және аннотациялау, модельдерді әзірлеу мен оқыту, нәтижелерді бағалау және тестілеу жұмыстарын – өз бетімен орындаған. Оның машиналық оқыту және нейрондық желілер саласындағы терең білімдері мен техникалық дағдылары зерттеу сапасына тікелей әсер еткен.

Диссертациялық жұмыста ұсынылған барлық негізгі нәтижелер нақты деректер жиынтықтарында тестіленіп, заманауи әдістермен салыстырылған. Автор бірнеше модельдердің нәтижелерін бір-бірімен салыстырып, өзінің ұсынған тәсілінің артықшылықтарын нақты дәлелдей алған. Бұл қорғауға ұсынылған ғылыми нәтижелердің жеткілікті дәрежеде негізделгенін көрсетеді.

Сонымен қатар, зерттеу барысында алынған нәтижелер халықаралық және отандық ғылыми басылымдарда жарияланған. Автордың бірнеше мақаласы жоғары импакт-факторлы журналдарда және беделді конференцияларда ұсынылған, бұл оның зерттеуінің ғылыми қоғамдастық тарапынан мойындалғанын білдіреді. Бұл диссертациялық жұмыстың жариялылық дәрежесі мен академиялық құндылығын айқындайды.

Жалпы алғанда, Елеусинов Арманның диссертациялық жұмысы ғылыми тұрғыдан терең, құрылымы жағынан жүйелі, мазмұны жағынан өзекті және тәжірибелік қолдануға бейімделген еңбек болып табылады. Диссертациялық жұмыс ҚР ҒжЖБМ «Ғылыми дәрежелерді беру тәртібі туралы ережеңің» талаптарына сәйкес келеді, ал диссертациялық жұмыстың авторы Елеусинов Арман Инабатұлы «6D075100 - Информатика, есептеу техникасы және басқару» мамандығы бойынша философия докторы (PhD) дәрежесін алуға лайықты.

Ғылыми жетекшісі:
PhD, доцент м.а,
әл-Фараби атындағы ҚазҰУ,
Ақпараттық технологиялар факультеті,
Компьютерлік ғылымдар кафедрасы Бурибаев Ж.А.

әл-Фараби атындағы Қазақ ұлттық университеті ғылыми және педагогикалық кадрларды даярлау және аттестаттау басқармасының басшысы
ЗАВЕРЯЮ
Начальник управления подготовки и аттестации научных кадров КазНУ им. аль-Фараби
«12» 08 2025 ж./г.

«ӘЛ-ФАРАБИ АТЫНДАҒЫ ҚАЗАҚ ҰЛТТЫҚ УНИВЕРСИТЕТІ» КОММЕРЦИЯЛЫҚ ЕМЕС АКЦИОНЕРЛІК ҚОҒАМЫ
НЕКОММЕРЧЕСКОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО «КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ»
NON-PROFIT JOINT-STOCK COMPANY «AL-FARABI KAZAKH NATIONAL UNIVERSITY»
FARABI KAZAKH NATIONAL UNIVERSITY 1934
БИН 990140001154