

«6D070300 – Ақпараттық жүйелер» мамандығы бойынша философия докторы (PhD) дәрежесіне іздену үшін ұсынған Кәрібаева Айдана Сейілғазықызының «Нейронды машиналық аударма үшін қазақ тіліндегі мәтіндерді морфологиялық сегменттеу модельдері мен әдістерін құру және зерттеу» атты диссертациялық жұмысына ғылыми кеңесшінің

ПІКІРІ

А.С. Кәрібаева диссертациялық жұмысы қазақ тілінің нейронды машиналық аудармасында морфологиялық сегменттеуді зерттеуге арналған. Қазіргі уақытта нейронды машиналық аудармада морфологиялық сегменттеу өзекті есептердің бірі болып табылады.

Диссертациялық зерттеу көлемінде машиналық аудармада сегменттеуге байланысты шетелдік және Отандық зерттеулерге талдау жасалынып, күрделі морфологиясы бар қазақ тілі үшін нейронды машиналық аудармада сегменттеу жасаудың қажеттілігі анықталды, диссертациялық жұмыстың мақсаты мен есептері құрылды.

Сөздіктер нейронды машиналық аудармада (NMT) маңызды рөл атқарады. NMT оқыту кезінде тиісті NMT сөздігінің көлемі тез артады; сондықтан компьютер жадының шамадан тыс қорын қажет етеді. Үлкен сөздікке едәуір жады қажет, бұл NMT қолдану мүмкіндігін шектейді және жады мәселесін тудыруы мүмкін. Бұл шектеуді әр сөзді параллель қайнар көздерде морфемаларға бөлу арқылы шешуге болады.

Бұл диссертациялық жұмыста түркі тілдері үшін жалғаулардың толық жиынтығына (CSE - complete set of endings) негізделген жаңа морфологиялық сегменттеу тәсілі ұсынылған, бұл тәсіл бастапқы корпустың сөздігін (шамамен 2,5 есе), тиісінше, қажетті жады көлемін де азайтқан.

А.С. Кәрібаева зерттеу жұмысында келесі ғылыми жаңалықтарды алды:

1. Тілдің толық жалғаулар жиынын құрумен ерекшеленетін мүмкін болатын жалғауларды қарастырудың негізінде қазақ тілі морфологиясының жетілдірілген есептеуіш моделі әзірленді.

2. Танымал әдістер мен алгоритмдерден толық жалғаулар жиынын шешімдер кестесін ретінде құрумен ерекшеленетін және нейрондық машиналық аударма сөздігінің көлемін қысқартуға мүмкіндік беретін қазақ тілі морфологиясының жетілдірілген есептеу моделіне негізделген морфологиялық сегменттеу әдісі мен алгоритмі әзірленді. Бұл нейронды машиналық аударманы үлкен көлемді кіріс мәліметтерде (сөздіктің үлкен көлемінде) оқытуға мүмкіндік береді.

Зерттеу жұмысының өзектілігі, ғылыми жаңалығы және практикалық маңыздылығы диссертацида толығымен ашылып, көрсетілді. Жұмыстың нәтижелері ғылыми негізделген және тәжірибе негізінде тиімділігі көрсетілген. Алынған ғылыми нәтижелер әртүрлі оқыту конфигурациялары бар эксперименттермен расталады.

Диссертациялық жұмыстың негізгі нәтижелері 20 Қазақстандық және

шетелдік публикация баслымдарында жарық көрген. Морфологиялық сегменттеу программалық құралына авторлық куәлік алынған.

Жоғарыда айтылғандарды негізге ала отырып, «Нейронды машиналық аударма үшін қазақ тіліндегі мәтіндерді морфологиялық сегменттеу модельдері мен әдістерін құру және зерттеу» тақырыбындағы диссертациялық жұмыс толық түрде аяқталған, рейтингтік ғылыми журналдар мен ғылыми конференцияларда апробацияланған, құнды ғылыми және тәжірибелік нәтижелерге ие ғылыми жұмыс болып табылады.

Диссертациялық жұмыс Қазақстан Республикасы Білім және ғылым министрлігінің «Ғылыми дәрежелер беру ережесі» талаптарына толық сай келеді, ал жұмыстың авторы Кәрібаева Айдана Сейілғазықызы «6D070300 – Ақпараттық жүйелер» мамандығы бойынша философия докторы (PhD) ғылыми дәрежесін алуға лайық.

Ғылыми жетекші:

т.ғ.д., әл-Фараби атындағы
Қазақ ұлттық университеті,
«Ақпараттық жүйелер» кафедрасының
профессоры



У.А. Тукеев

РАСТАЙМЫН
әл-Фараби кафедрасы ҚазҰУ Ғылыми кадрларды
даярлау және аттестаттау басқармасының басшысы
ЗАВЕРЯЮ
Начальник управления подготовки и аттестации
научных кадров КазНУ им. аль-Фараби
Р.Е. Кудалбергенова
« _____ » _____ 20__ ж.г.