

**8D06103 – “Management information systems” мамандығы бойынша PhD
докторант**

Бекарыстанкызы Акбаянның диссертациялық жұмысына

АҢДАТПА

**тақырыбы: “Агглютинативті тілдердегі сөйлеуді автоматты түрде
тануға арналған end-to-end жүйе”**

Зерттеу тақырыбының өзектілігі. Сөйлеуді автоматты тану жүйелері қазіргі уақытта адам өмірінің әртүрлі салаларында, адамның компьютерлік жүйелермен және әртүрлі қолданбалармен өзара әрекеттесуін жеңілдету үшін өте кең қолданылады. Мысалы, смарт көмекшілер, смарт үй жүйелері, автоматты субтитр жасау жүйелері компьютерлік жүйелер мен құрылғыларды тіпті алыстан-ақ басқаруға мүмкіндік береді. Сонымен қатар, сөйлеуді автоматты түрде тану жүйелері денсаулығында ақаулары бар адамдарға электронды құрылғыларды пайдалануды жеңілдетеді. Мысалы, бір топ ғалымдар дизартриямен ауыратын адамдар үшін автоматты түрде сөйлеуді тану жүйесін жасау жолдарын зерттеп жатыр. Дизартрия - бұл артикуляцияға жауап беретін бұлшықеттердегі ақаулардың бір түрі. Мұндай адамдар үшін сөйлеуді автоматты түрде тану жүйелерін әзірлеу оларға сандық жүйелермен өзара әрекеттесіп қана қоймай, басқа адамдармен де қарым-қатынас жасауға көмектеседі. Сөйлеуді автоматты түрде тану жүйелерін енгізудің тағы бір пайдалы мысалы есту қабілетінің жоғалу дәрежесін бағалау болып табылады. Жүйенің бұл түрі қойылған сұрақтарға жауаптарды бағалау арқылы есту қабілетінің бұзылу деңгейін бағалай алады. Бірақ мұндай мүмкіндіктер ағылшын, қытай, орыс және т.б. сияқты жиі қолданылатын тілдерде сөйлейтін адамдарға ғана қол жетімді. Ресурсы аз тілдерге арналған сөйлеуді автоматты түрде тану жүйелерін әзірлеу әлі де көп күшті талап етеді, мысалы, деректерді жинау және оларды өңдеу, өте таныс архитектураларды бір тілдер үшін немесе олардың белгілі бір тобы үшін сынау және баптау. Агглютинативті тілдердің түркі тілдер тобына қазақ тілінен басқа аз ресурсты көптеген тілдер кіреді.

Агглютинативті тілдер үшін сөйлеуді автоматты түрде танудың сенімді жүйесін алудың нақты тәсілдері мен үлгілерін әзірлеуге арналған зерттеулердің үлкен көлемі бар. Кейбір авторлар морфемаларға негізделген тіл үлгісін ұсынады, мұнда морфемалар сөздегі кез келген префикс, түбір немесе жұрнақ болып табылады. Нәтижесінде авторлар сөздік қоры мол автоматты түрде сөйлеуді тану жүйесін алды. Бір зерттеу контекстке сезімтал және оқыту бірлігі ретінде алынған сөз фрагменттерін пайдалану арқылы оқытылатын трансформерге негізделген CTC жүйесінің өнімділігін зерттейді. Авторлар

олардың әдісінің тек ағылшын және неміс тілдері үшін ғана емес, сонымен қатар агглютинативті тілдердің бірі - түрік тілі үшін де тиімділігін атап өтеді. Келесі зерттеу түрік тілін қолданып, морфологиялық дисамбигуация үшін трансформер архитектурасын қолдануды қарастырды. Бұл мағынаны ажырату құралы кез келген NLP есептерінде пайдаланылуы мүмкін және сөйлеуді тану жүйесі де олардың қатарына жатады. Трансформер архитектурасы тағы бір агглютинативті тіл – хинди тілінде де өзін жақсы дәлелдеді. Мұнда түрлендіргіш архитектурасы, Қосылымның уақытша классификациясы (СТС) және тіл үлгісі үшеуі біріккен кезде хинди тілі үшін ең төменгі қателік деңгейін көрсетті: 3,2%. Агглютинативті тіл үшін сөйлеуді автоматты түрде тану жүйесін әзірлеу үшін трансформер архитектурасын қолданудың тағы бір мысалы фин тілі болып табылады. Мұнда автор Transformer-XL архитектурасының өнімділігін LSTM-мен салыстырып, күрделілік мәнінің жақсаруы 29%-ға жетті және фин тілі үшін сөз қателігінің деңгейі 3%-ға дейін төмендеді деген қорытындыға келді.

Келесі мақалада авторлар сөзді автоматты түрде танудың ең кең таралған және тиімді архитектуралары бар екенін айтады: олар – байланысты уақытша жіктеу және назар аудару механизмдері. Бұл жұмыста сондай-ақ сөйлеуді автоматты түрде танудың сенімді жүйелерін жасау үшін және оқыту үшін агглютинативті тілдерге арналған транскрипцияланған аудио-мәтін жұптарының ресурстарының жеткіліксіздігі туралы айтылады.

Агглютинативті тілдерге арналған жоғарыда аталған зерттеулерге сүйене отырып, сөздік қорын кеңейту және трансформер архитектурасы агглютинативті тілдер үшін сөйлеуді автоматты түрде тану жүйелерін дамытудың ең тиімді тәсілдері болып табылатыны атап өтілді. Сонымен қатар, агглютинативті тілдердің түркі отбасының тілдері арасында оқыту деректерінің аздығы, біркелкі морфологиялық ережелер мен дыбыстық ұқсастықтардың болуы осы тілдерді трансферлік оқыту және көптілді оқыту сияқты біріктіру эксперименттеріне негіз болды.

Диссертациялық жұмыстың мақсаты. Бұл диссертациялық жұмыс түркі тілдестер отбасына жататын тілдер мысалы негізінде агглютинативті тілдер үшін сөйлеуді автоматты түрде тану жүйелерін жетілдіру жолдарын зерттеу мақсатында жазылған.

Зерттеу тапсырмалары.

- 1) Сөйлеуді тануға арналған зерттеулердің жалпы жағдайын және агглютинативті тілдерге арналған жағдайларын сараптау.
- 2) Түркі тілдестер отбасына жататын агглютинативті тілдер үшін мәліметтер корпусын құру және үлкейту.
- 3) Агглютинативті тілдерде сөйлеуді тануға арналған модельдер мен әдістерді дайындау.

4) Агглютинативті тілдерде сөйлеуді тануға арналған жүйе құру.

Зерттеу объектілері. Сөйлеуді автоматты түрде танудың заманауи әдістері мен жолдары, әсіресе көптілдік және трансферлік оқыту, кеңейтілген тілдік модель.

Зерттеу нысаны. Түркі тілдестер отбасына жататын агглютинативті тілдер, машиналық оқыту әдістері, нақтырақ айтқанда сөйлеуді автоматты түрде тануға арналған нейрондық желілер, конволюциялық нейродық желілер, мәліметтер тапшылығы өте жоғары деңгейдегі тілдерде тануды жақсартуға арналған әдістер, сөздер кіріктірілуі, сөздерді сипаттауыш векторлар, демонстративті телебот және әртүрлі форматтағы аудиофайлдардың мазмұнын мәтінге айналдыратын бағдарламалық қосымша.

Зерттеу әдістері. Машиналық оқыту әдістері, сөйлеуді автоматты түрде тану әдістері мен технологиялары, табиғи тілді өңдеу әдістері, математикалық статистика және ықтималдықтар теориясы.

Ғылыми жаңалығы. Диссертациялық жұмыста әсіресе қазақ тіліне назар аударып, агглютинативті тілдерге арналған интеграцияланған автоматты түрде сөйлеуді тану жүйелерін жетілдіруге арналған ғылыми және практикалық жаңалықтар алға қойылған. Ұсынылған тәсілдер басқа тілдер үшін де пайдалы болып табылады. Агглютинативтік тілдер корпусы құрылды, осы жұмыстар аясында қазақ тіліне арналған таңбаланған деректердің көлемін арттыруға да айтарлықтай үлес қосылды. Зерттеу нәтижесінде алынған негізгі оң нәтижелер төменде келтірілген:

- 1) Агглютинативті тілдерде сөйлеуді тануға арналған мәліметтер корпусы құрылды.
- 2) Түркі тілдестер отбасына жататын аз ресурсты агглютинативті тілдерде сөйлеуді тануды жақсартуға арналған тиімді модельдер құрылды: трансферлік әдіс, көптілдік оқыту, тілдік модельді кеңейту.
- 3) Агглютинативтік тілдерде сөйлеуді автоматты түрде тануға арналған жүйе.

Жұмыстың теориялық және практикалық маңызы. Зерттеудің теориялық маңыздылығы мынада: ол сыртқы «үлкен мәтіні» бар тіл үлгісін жақсарту арқылы ғана арқылы сөйлеуді автоматты түрде танитын жүйенің өнімділігін жақсарту мүмкіндігін ұсынады және бір тілдер тобына жататын барлық тілдер үшін өнімділікті жақсарту мүмкіндігін көрсетеді және тілдер аралық трансферлік оқытуың тиімділігін анықтайды. Түркі тілдестер отбасына жататын агглютинативті тілдер үшін барлық аталған теориялық принциптерді қолдану мүмкіндігі осы диссертацияның практикалық маңыздылығын көрсетеді. Сонымен қатар, мәтінді өңдеудің алдын ала өңдеу алгоритмі мәтінді өңдеу есептерінің кең ауқымын шешу үшін пайдаланылуы мүмкін. Зерттеу

барысында жиналған аудио-мәтіндік жұптық деректер сөйлеуді өндеудің әртүрлі есептерінде қолданылуы мүмкін.

Қорғауға ұсынылатын негізгі тұжырымдама.

- 1) Агглютинативті тілдерге арналған мәліметтер корпусы құрылды.
- 2) Аз ресурсты агглютинативтік тілдер үшін сөйлеуді тануды жақсартудың әдістері.
- 3) Агглютинативті тілдерге арналған сөйлеуді тану жүйесі құрылды.

Нәтижелердің сенімділік және апробациялану деңгейі.

Диссертациялық жұмыс тақырыбы аясындағы зерттеулер мен оның нәтижелері әртүрлі конференциялар мен семинарларда ұсынылып, талқыланды, ал кейбіреулері жарияланды. Сондай-ақ, авторға семинар спикері және баяндаманың үздік презентациясы ретінде сертификаттар берілді:

- 1) O. Mamyrbayev, D. Oralbekova, A. Kudyrbekova, T. Turdalykyzy and A. Bekarystankyzy, "End-to-End Model Based on RNN-T for Kazakh Speech Recognition," 3rd International Conference on Computer Communication and the Internet (ICCCI) (25-27 маусым 2021 ж., Токио).
- 2) Семинар спикері сертификаты, тақырыбы: "Improved Speech Recognition for Agglutinative languages", Coimbra Institute of Engineering (ISEC), (21 апреля 2023 г., Коимбра, Португалия).
- 3) Ең үздік баяндама үшін сертификат, тақырыбы: "Improve Automatic Speech Recognition for Kazakh Language using Extended Language Model", "ACeSYRI Young Researchers School" (5-10 маусым 2023 ж., Алматы, Қазақстан).
- 4) A. Bekarystankyzy, O. Mamyrbayev, "IMPROVE AUTOMATIC SPEECH RECOGNITION FOR KAZAKH LANGUAGE USING EXTENDED LANGUAGE MODEL", 21-научная конференция, (20-21 сәуір 2023ж., Рига, Латвия).
- 5) Automatic Speech Recognition Improvement for Kazakh Language with Enhanced Language Model // Recent Challenges in Intelligent Information and Database systems. ACIIDS 2023. Part of The Communications in Computer and Information Science book series. – 2023. - Vol. 1, - P.538-545 (Springer, Cham).

Зерттеушінің жеке үлесі. PhD-студент өзіне қойылған тапсырмаларды өз бетінше орындап, шешіп шықты. Автор қазақ және агглютинативті тілдер үшін сөйлеуді танудың интеграцияланған модельдерін өз бетінше құрды. Ол агглютинативті тілдер мәліметтер қорын жиыстыруға, қазақ тілінің таңбаланған корпусын үлкейтуге, әсіресе әртүрлі орталарда жинақталған деректерді енгізуге өзінің жеке үлесін қосты. Қолданыстағы және жетілдірілген модельдер бойынша эксперименттер және оларды бағалау үрдістерін орындады.

Диссертация тақырыбының ғылыми-зерттеу жұмыстарының жоспарымен байланысы.

Диссертация тақырыбы аясындағы зерттеу жұмыстары келесі гранттық жоба негізінде орындалды: «Агглютинативті тілдердегі сөйлеуді автоматты түрде тануға арналған арналған интегралдық жүйе құру» (2020-2022, мемлекеттік тіркеу нөмірі: 0120PK00344). Зерттеу жұмысы Қазақстан Республикасы Ғылым және Жоғарғы Білім министрлігінің Ғылым Комитетіне қарасты Ақпараттық және Есептеу Технологиялар институтында орындалды.

Диссертациялық зерттеудің негізгі нәтижелерін жариялау. Зерттеу тақырыбы аясында бірнеше мақалалар жарияланды, олардың біреуі импакт-факторы нөлден жоғары мерзімдік журналда жарияланды және Scopus және Web of Science мәліметтер базасында индекстелді, ал үшеуі Қазақстан Республикасы Ғылым және Жоғарғы Білім министрлігінің Білім және ғылым саласындағы бақылау комитеті ұсынған журналдарда жарияланды.

- 1) M. Orken, A. Keylan, O. Dina, B. Akbayan and Z. Bagashar. (2022). Identifying the influence of transfer learning method in developing an end-to-end automatic speech recognition system with a low data level // Eastern-European Journal of Enterprise Technologies. - Vol. 1, № 115. -P. 84–92 // <https://doi.org/10.15587/1729-4061.2022.252801>(Scopus, percentile 34);
- 2) Bekarystankyzy A. and Mamyrbayev O. (2023). INTEGRATED AUTOMATIC SPEECH RECOGNITION SYSTEM FOR AGGLUTINATIVE LANGUAGES // News of the National academy of sciences of the republic of Kazakhstan. - 2023. - Vol. 1, № 345. -P. 37-49 // <https://doi.org/10.32014/2022.2518-1726.167>
- 3) Bekarystankyzy A., Mamyrbayev O., Oralbekova D., Zhumazhanov B. (2023). Transfer learning for an integrated low-data automatic speech recognition system // Scientific and technical journal "Bulletin of the Almaty University of Power Engineering and Telecommunications". - 2023, -Vol. 1, №. 60. -P. 185-198 // https://doi.org/10.51775/2790-0886_2023_60_1_185
- 4) Bekarystankyzy A. and Mamyrbayev O. (2023). End-to-end speech recognition systems for agglutinative languages // Scientific Journal of Astana IT University. - 2023. -Vol. 13. -P. 86-92 // DOI: 10.37943/13IMII7575
- 5) "UniCodeKaz" бағдарламалық жабдығына авторлық куәлік, № 38545, 21.08.2023, Автор: Бекарыстанқызы А.
- 6) "Система автоматического транскрибирования аудио файлов в текст" бағдарламалық жабдығына авторлық куәлік, № 38833, 31.08.2023, Авторы: Бекарыстанқызы А., Мамырбаев О., Дуйсенхан Б.

Диссертациялық жұмыстың құрамы және көлемі. Диссертациялық жұмыс кіріспеден, 4 бөлімнен, қорытындыдан, 163 атаудан тұратын

әдебиеттер тізімінен және 4 қосымшадан тұрады. Жұмыс 107 бетте баяндалған және 38 суреттен, 16 кестеден тұрады.

Диссертациялық зерттеу жұмысының қысқа сипаттамасы.

Кіріспеде зерттеу тақырыбының өзектілігі, мақсаты, диссертацияның зерттеу сұрақтары мен міндеттері, ғылыми жаңалығы, зерттеудің теориялық және практикалық маңыздылығы және диссертацияда қолданылған зерттеу әдістері туралы мәліметтер берілген.

Бірінші тарауда агглютинативті және агглютинативті емес тілдерде сөйлеуді автоматты түрде тануға арналған жұмыстарға жан-жақты шолу жасалады. Осы тараудағы бастапқы талқылау қазіргі заманғы зерттеулерді шолудың әдістемесіне бағытталған. Келесі кезекте сөйлеуді автоматты түрде танудың алғашқы мысалдары, сөйлеуді өңдеу саласында нейрондық желілерді енгізу және қазіргі заманғы архитектуралық шешімдер үлгілері талқыланады. Бұл тараудың соңғы екі бөлімшелері агглютинативті тілдер мен қазақ тілінің зерттеулері мен жетістіктеріне арналған.

Сөйлеуді танудың теориялық негіздеріне арналған **екінші** тарау математикалық және акустикалық модельдеуден басталады да, ары қарай агглютинативті тілдерді танудың, табиғи тілді өңдеудің, сөйлеуді автоматты түрде тану, өнімділік метрикасының және сөйлеуді тану модельдерінің ең маңызды тұжырымдамалары қарастырылады, мысалы, жасырын Марков үлгілері, Гаусс қоспасының үлгілері, жасырын Марков үлгілері/жасанды нейрондық желілер, RNN, GRU және LSTM сияқты реттілік үлгілеріне арналған соңғы үлгілер және желілік блок түрлері. Соңғы бөлімде назар аудару механизмі мен байланыстырушы уақытша классификацияға негізделген заманауи архитектуралар сипатталады: трансформер, конформер және бранчформер.

Үшінші тарау эксперименттерге және олардың нәтижелеріне арналған. Бұл тарау деректерді жинау процесінен басталады. Одан әрі түркі тектес агглютинативті тілдердің әліпбилері бірдей тобына арналған көптілді экспериментке арналған бөлімдер; трансформер архитектурасын қолданып «Үлкен мәтінді» оқыту арқылы алынған және декодтау кезеңінде қолданылатын кеңейтілген тіл моделі; қазақ және әзірбайжан тілдері бойынша трансферлік оқыту жүргізілді.

Төртінші тарауда диссертацияға енгізілген барлық эксперименттер нәтижелері және олардың нәтижелері талқыланады, бұл нәтижелер жүргізілген зерттеулерге ұқсас алдыңғы зерттеулердің нәтижелерімен салыстырылады.

Қорытындыда диссертацияға қарастырылған зерттеудің негізгі қорытындылары жинақталып, әрі қарай атқарылатын жұмыстар туралы ақпарат беріледі.

