ӘОЖ 004.912 https://doi.org/10.54251/2616-6429.2026.01.0020nu

М.Б. Мизам[1], Ж.Д. Изтаев[1], О.М. Сүлеймен[1*], П.А. Кожабекова[1], С.Б. Атажонова[2]
[1]магистрант, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан
[1]п.ғ.к., қауымдастырылған профессор, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан
[1]магистр, оқытушы, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан
[1]т.ғ.к., доцент, М.Әуезов атындағы ОҚУ, Шымкент, Қазақстан
[2]педагогика ғылымдары бойынша PhD, доцент, Әндіжан мемлекеттік техникалық институты, Әндіжан, Өзбекстан Республикасы
***Корреспондент авторы:** jasiksuleimen42@gmail.com

ҚАЗАҚ–ОРЫС–АҒЫЛШЫН ТІЛДЕРІНДЕГІ СӨЙЛЕУДІ ХАТТАМА ФОРМАТЫНА БЕЙІМДЕЙТІН КӨПТІЛДІ НЕЙРОНДЫҚ МАШИНАЛЫҚ АУДАРМА (NMT) МОДЕЛІ

Түйін

Бұл зерттеу қазақ–орыс–ағылшын тілдік кеңістігінде сөйлеуді автоматты түрде аудару және ресми хаттама форматына бейімдеу мәселелерін жан-жақты қарастырады. Жұмыста трансформер архитектурасына негізделген көптілді нейрондық машиналық аударма (NMT) модельдерінің ғылыми негіздері терең талданады, сонымен қатар оларды автоматты сөйлеуді тану (ASR) жүйелерімен біріктіру әдістері көрсетіледі. Ұсынылған тәсіл аудио сигналдан бастап құрылымдалған ресми құжатқа дейінгі толық өңдеу тізбегін қамтиды, оның ішінде транскрипция, аударма, пост-өңдеу және мәтінді ресми стильге сәйкестендіру кезеңдері бар. Зерттеу нәтижелері көптілді трансформер модельдерінің қазақ, орыс және ағылшын тілдері арасында жоғары сапалы аударма жасауға қабілеттілігін көрсетеді. Сонымен қатар дисфлюэнцияларды жою, пунктуацияны қалпына келтіру, терминологиялық бірізділікті сақтау және мәтінді ресми стильге нормализациялау арқылы мәтінді хаттама форматына автоматты сәйкестендіру мүмкіндігі дәлелденді. Ұсынылған әдіс көптілді жиналыстарды құжаттандыру үдерісін айтарлықтай жеделдетеді, ақпаратты жүйелеуге және қазақ тілінің цифрлық ресурстарын дамытуға оң ықпал етеді.

Кілттік сөздер: нейрондық машиналық аударма, көптілді модель, сөйлеуді тану, хаттама форматы, трансформер архитектурасы, автоматтандыру, машиналық аударма.

Кіріспе

Жаһандану және цифрлық трансформация жағдайында көптілді коммуникация мемлекеттік басқару, халықаралық ынтымақтастық және академиялық ортада стратегиялық маңызға ие. Қазақстанда ресми құжат айналымы қазақ және орыс тілдерінде жүргізілсе, халықаралық өзара іс-қимылда ағылшын тілі кеңінен қолданылады. Көптілді кездесулер мен отырыстар барысында айтылған ақпаратты жедел аудару ғана емес, оны нормативтік талаптарға сәйкес ресми хаттама форматына келтіру қажеттілігі туындайды.

Соңғы онжылдықта нейрондық машиналық аударма (Neural Machine Translation, NMT) технологиялары, әсіресе Transformer архитектурасына негізделген модельдер, аударма сапасын айтарлықтай арттырды. Сонымен қатар автоматты сөйлеуді тану (Automatic Speech Recognition, ASR) жүйелері сөйлеуді мәтінге жоғары дәлдікпен түрлендіруге мүмкіндік берді. Осы екі технологияны біріктіру сөйлеуді көптілді ортада автоматты түрде өңдеуге жол ашады.

Материалдар мен әдістер

Осы зерттеудің мақсаты - қазақ, орыс және ағылшын тілдеріндегі сөйлеуді автоматты түрде аударып, оны стандартталған хаттама форматына бейімдей алатын NMT моделінің ғылыми негіздерін сипаттау және оның қолданбалы маңызын талдау.

Зерттеудің нысаны - көптілді сөйлеу деректерін автоматты өңдеу жүйесі. Зерттеу пәні - сөйлеуді тану, машиналық аударма және мәтінді құрылымдау алгоритмдерінің интеграциясы

негізінде ресми хаттама құру үдерісі.

Зерттеу міндеттері төмендегідей:

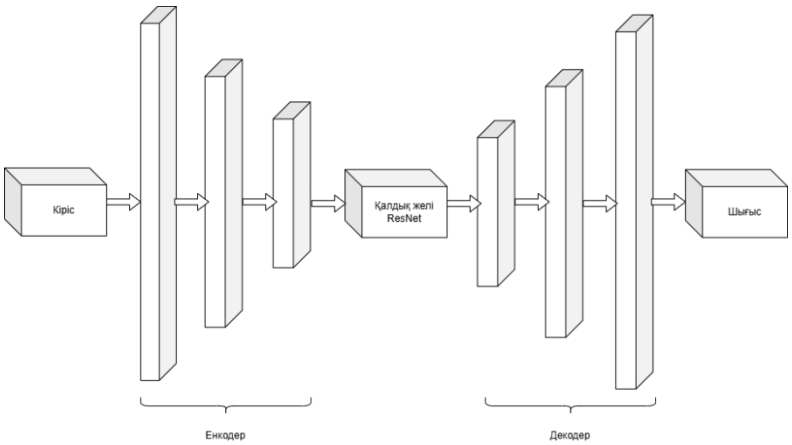
- Қазақ–орыс–ағылшын тілдері арасында көптілді аудармаға арналған нейрондық модельдердің қазіргі күйін талдау;
- ASR және NMT жүйелерін біріктірудің архитектуралық тәсілдерін салыстыру;
- Сөйлеу мәтінін ресми-іскери стильге бейімдеудің лингвистикалық ерекшеліктерін анықтау;
- Хаттама құрылымын автоматты түрде қалыптастыратын модельдік тәсілді ұсыну;
- Ұсынылған модельдің сапасын автоматты метрикалар (BLEU, ROUGE және т.б.) және сараптамалық бағалау арқылы тексеру.

Зерттеудің ғылыми жаңалығы - көптілді аударма мен ресми құжат генерациясын біртұтас архитектурада қарастырып, сөйлеуден хаттамаға дейінгі толық автоматтандырылған өңдеу тізбегін ұсынуында. Бұл тәсіл дәстүрлі аударма модельдерінен айырмашылығы, тек семантикалық баламалылықты емес, сонымен қатар құжаттың нормативтік құрылымын сақтауды мақсат етеді.

Зерттеудің теориялық маңызы - көптілді нейрондық модельдер мен ресми дискурс лингвистикасының тоғысында жаңа интеграциялық әдістемелік негіз қалыптастыруында. Ал практикалық маңызы мемлекеттік органдарда, халықаралық ұйымдарда, жоғары оқу орындарында және корпоративтік секторда көптілді жиналыстардың хаттамаларын автоматтандырылған түрде әзірлеу мүмкіндігімен анықталады.

Көптілді нейрондық машиналық аударма технологиясының ғылыми негіздері

Нейрондық машиналық аударма (Neural Machine Translation, NMT) - терең нейрондық желілерге негізделген мәтінді автоматты аудару әдісі. 2014 жылдан бастап encoder–decoder архитектурасына негізделген модельдер кең тарала бастады. 2017 жылы ұсынылған Transformer архитектурасы ретті деректерді өңдеуде рекуррентті желілерді алмастырып, аударма сапасын айтарлықтай арттырды [1]. Transformer моделі толықтай self-attention механизміне негізделген және параллель есептеуді тиімді жүзеге асырады. Нейрондық желі құрылымы 1 – суретте көрсетілген.



Сурет 1 – Encoder–Decoder негізіндегі нейрондық желі архитектурасы (ResNet қолдануымен)

Классикалық encoder–decoder моделінде кодтаушы (encoder) бастапқы мәтінді векторлық кеңістікке түрлендірсе, декодер (decoder) осы репрезентация негізінде мақсат тілдегі мәтінді генерациялайды. Attention механизмі бастапқы сөйлемнің әрбір бөлігіне салмақ беру арқылы контекстті дәлірек модельдеуге мүмкіндік береді. Ал Transformer архитектурасында көпбасақты (multi-head) self-attention механизмі контекстік тәуелділіктерді ұзақ қашықтықта тиімді ұстауға жағдай жасайды.

Көптілді NMT (Multilingual NMT) бір модель шеңберінде бірнеше тіл жұбын бір уақытта оқытуға мүмкіндік береді. Бұл тәсіл ортақ семантикалық кеңістік құру арқылы төмен ресурсты тілдер үшін аударма сапасын жақсартады. Қазақ тілі морфологиялық тұрғыдан агглютинативті тіл болғандықтан, көптілді ортақ эмбеддинг кеңістігі сирек формалардың жалпылануына оң әсер етеді. Сонымен қатар subword сегментация (BPE, SentencePiece) әдістері морфемалық вариативтілікті азайтып, сөздік көлемін оңтайландыруға мүмкіндік береді [2].

Қазіргі таңда ірі көптілді модельдерге:

- Meta AI ұсынған NLLB (No Language Left Behind) - 200-ден астам тілді қамтитын ауқымды көптілді модель;
- Google Research әзірлеген mT5 (2021) - мәтіннен мәтінге (text-to-text) трансформер архитектурасына негізделген көптілді модель;
- Facebook AI Research ұсынған mBART (2020) - денойзинг автоэнкодер қағидатына негізделген алдын ала оқытылған көптілді модель.

Нәтижелер және талқылау

Сөйлеуді тану (ASR) және NMT интеграциясы

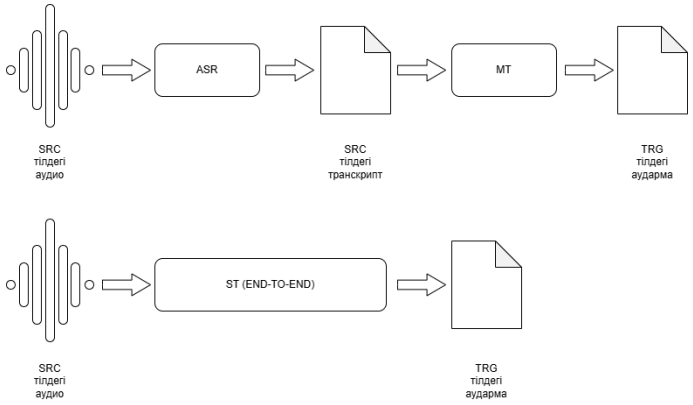
Сөйлеуді автоматты тану (Automatic Speech Recognition, ASR) жүйелері соңғы онжылдықта терең нейрондық желілердің дамуының арқасында айтарлықтай жетілдірілді [3]. Қазіргі ASR жүйелері, негізінен, екі негізгі тәсілге сүйенеді: CTC (Connectionist Temporal Classification) алгоритмі және attention-based encoder–decoder архитектурасы [4].

CTC тәсілі кіріс аудио сигнал мен шығыс мәтін арасындағы уақыттық сәйкестікті тура белгілемей-ақ модельдеуге мүмкіндік береді. Бұл әдіс акустикалық модель мен тілдік модельді бөлек оқытуға қолайлы. Ал attention-based encoder–decoder тәсілінде аудио тізбек толық контекстпен өңделіп, мәтін біртіндеп генерацияланады. Соңғы жылдары гибридті CTC/attention модельдері және Transformer немесе Conformer архитектурасына негізделген шешімдер жоғары дәлдік көрсетіп отыр.

Сөйлеуді аудару (Speech Translation) екі негізгі тәсілмен жүзеге асады:

1. Каскадты модель: ASR → мәтін → NMT
2. End-to-end speech translation**:** аудио → аударылған мәтін

Каскадты модельде алдымен ASR жүйесі сөйлеуді бастапқы тілдегі мәтінге түрлендіреді, содан кейін бұл мәтін нейрондық машиналық аударма (NMT) модуліне беріледі. Бұл тәсіл модульдік архитектураға негізделген және әр компонентті (акустикалық модель, тілдік модель, аударма моделі) жеке-жеке жетілдіруге мүмкіндік береді. Сөйлеуді аудару (Speech Translation) 2 – суретте көрсетілген.



Сурет 2 – Сөйлеуді аудару жүйесінің архитектуралары

End-to-end тәсілінде аудио сигнал бірден мақсат тілдегі мәтінге түрлендіріледі. Мұндай

модельдерде біртұтас нейрондық архитектура қолданылады және олар қателердің жинақталуын (error propagation) азайтуы мүмкін. Алайда көптілді және салалық ортада деректердің жеткіліксіздігі мен оқыту күрделілігі бұл тәсілді практикалық деңгейде шектеуі мүмкін.

Практикалық қолданбаларда каскадты модель кең таралған, себебі:

- модульдерді бөлек оқыту және жаңарту мүмкіндігі бар;
- терминологиялық бақылау жеңіл жүзеге асады;
- мәтіндік аралық нәтиже кейінгі құрылымдық және стилистикалық өңдеуге қолайлы;
- ресми хаттама форматына бейімдеу үшін лингвистикалық постөңдеу кезеңін енгізуге мүмкіндік береді.

Қазақ, орыс және ағылшын тілдерінің құрылымдық ерекшеліктері

Қазақ тілі - агглютинативті тіл, сөз формалары жалғаулар арқылы құралады. Бұл бір түбірден көптеген морфологиялық формалардың пайда болуына әкеледі. Мысалы, бір ғана етістік түбірі шақ, рай, жақ, болымдылық және септік жалғаулары арқылы ондаған грамматикалық нұсқада көрінуі мүмкін. Мұндай морфологиялық байлық сөздік көлемінің күрт ұлғаюына және сирек формалардың (rare tokens) көбеюіне себеп болады.

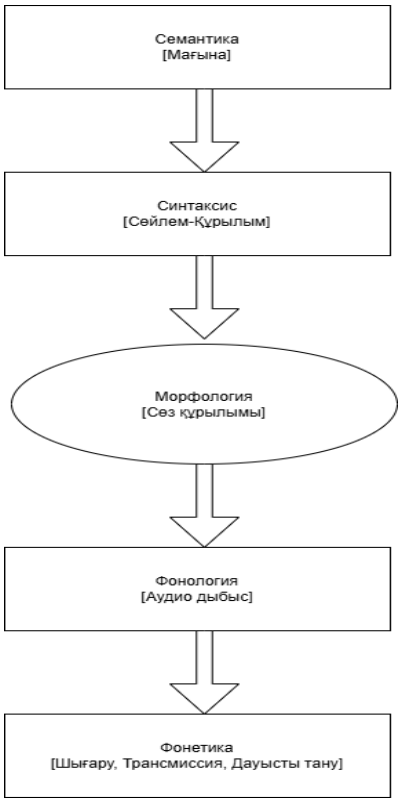
Орыс тілі флективті тілдер қатарына жатады, онда септік, жыныс және сан категориялары кең дамыған. Сөздің грамматикалық мағынасы көбіне түбірдің ішкі өзгерісі немесе флексия арқылы беріледі. Сонымен қатар сөз тәртібі салыстырмалы түрде еркін, бұл синтаксистік тәуелділіктерді модельдеуде қосымша күрделілік тудырады.

Ағылшын тілі аналитикалық құрылымға ие, грамматикалық қатынастар көбіне көмекші сөздер мен қатаң сөз тәртібі арқылы беріледі. Сондықтан ағылшын тілінде позициялық ақпарат маңызды рөл атқарады.

Көптілді NMT моделін әзірлеу барысында төмендегі факторлар ескеріледі:

- Морфологиялық байлық (қазақ тілі) - сөз формаларының вариативтілігі;
- Еркін сөз тәртібі (орыс тілі) - синтаксистік тәуелділіктердің икемділігі;
- Қатаң синтаксистік құрылым (ағылшын тілі) - позициялық шектеулердің маңыздылығы.

Морфологиялық күрделілікті азайту үшін subword-токенизация әдістері кеңінен қолданылады. Byte Pair Encoding (BPE) әдісі және SentencePiece алгоритмі сөздерді жиі кездесетін бөліктерге (subword units) бөледі. Тілдік өңдеудің иерархиялық деңгейлері 3 – суретте көрсетілген [5].



Сурет 3 – Тілдік өңдеудің иерархиялық деңгейлері

Хаттама форматын автоматты құрылымдау

Ресми хаттама - құрылымдалған құжат түрі. Әдетте ол келесі негізгі элементтерден тұрады:

- Қатысушылар тізімі
- Күн тәртібі
- Баяндамалар
- Талқылау барысы
- Қабылданған шешімдер

1. Сөйлеу тілі мен ресми жазба тілінің айырмашылығы

Сөйлеу тілі мен ресми жазба тілі құрылымдық және стилистикалық тұрғыдан айтарлықтай ерекшеленеді. Сөйлеуде:

- эллипсис (сөйлем мүшелерінің түсіп қалуы);
- синтаксистік аяқталмаған құрылымдар;
- қайталаулар;
- толтырғыш сөздер («яғни», «значит», «well» және т.б.);
- эмоциялық-экспрессивтік элементтер жиі кездеседі.

Ал ресми хаттама мәтіні:

- логикалық жағынан жүйеленген;
- бейтарап стильде жазылған;
- нақты құрылымдық бөлімдерге бөлінген;
- нормативтік құжат тіліне сәйкес рәсімделген болуы тиіс.

Сондықтан ASR және NMT кезеңдерінен кейін міндетті түрде пост-өңдеу (post-processing) модулі енгізіледі. Табиғи тілді өңдеу алгоритмі 4-суретте көрсетілген.

2. Пост-өңдеу кезеңінің міндеттері

Ресми хаттамаға бейімдеу барысында келесі трансформациялар жүзеге асырылады:

- Дисфлюэнцияларды жою - қайталау, толтырғыш сөздер, өздігінен түзетулерді алып тастау;
- Пунктуацияны қалпына келтіру - ASR нәтижесінде жиі жоғалатын тыныс белгілерін автоматты түрде енгізу;
- Абзацтарға бөлу - тақырыптық сегментация негізінде мәтінді құрылымдау;
- Ресми стильге нормализациялау - бейресми тіркестерді ресми баламалармен алмастыру (мысалы, «біз қарастырып жатырмыз» → «қарастырылды»).

Сурет 4 – Табиғи тілді өңдеудің алгоритмі

3. Құжат құрылымын автоматты анықтау әдістері

Мәтінді құрылымдау үшін екі негізгі тәсіл қолданылады:

1) Rule-based тәсіл

Алдын ала анықталған лексикалық және синтаксистік ережелерге негізделеді. Мысалы:

- «тыңдалды», «сөз алды», «қаулы етілді» сияқты маркерлер арқылы бөлімдерді анықтау;
- атаулы сөйлемдер арқылы күн тәртібін бөліп шығару.

Артықшылығы - интерпретациясының қарапайымдылығы. Кемшілігі - икемділігінің шектеулілігі және жаңа контексттерге бейімделу қиындығы.

2) Sequence labeling модельдері

Мәтіннің әрбір сөйлемін немесе сегментін белгілі бір құрылымдық санатқа жіктейді (мысалы: AGENDA, DISCUSSION, DECISION).

Қазіргі зерттеулерде құжат құрылымын анықтау үшін кеңінен қолданылатын модельдер:

- Google ұсынған BERT;
- Facebook AI Research әзірлеген RoBERTa;
- Facebook AI Research ұсынған XLM-R (XLM-RoBERTa).

XLM-R моделі көптілді ортада ерекше тиімді, себебі ол жүзден астам тілде алдын ала оқытылған және қазақ, орыс, ағылшын мәтіндерін ортақ семантикалық кеңістікте өңдей алады. Бұл қасиет көптілді хаттама генерациялау жүйесінде құрылымдық бірізділікті сақтауға мүмкіндік береді [6].

4. Интеграциялық модель

Ресми хаттама генерациясының толық тізбегі келесі кезеңдерден тұрады:

Аудио → Транскрипция (ASR) → Нормализация → Аударма (NMT) → Құрылымдық белгілеу → Құжат генерациясы.

Бағалау метрикалары және сапа көрсеткіштері

Машиналық аударма жүйелерінің тиімділігін бағалау халықаралық ғылыми қауымдастықта стандартталған автоматты метрикалар арқылы жүзеге асырылады. Бұл метрикалар модельдің генерациялаған мәтінін эталондық (reference) аудармамен салыстыруға негізделген.

Негізгі бағалау көрсеткіштері

BLEU - n-грамм сәйкестігіне негізделген метрика.

BLEU (Bilingual Evaluation Understudy) модель аудармасындағы n-граммдардың эталон мәтіндегі n-граммдармен сәйкес келу жиілігін өлшейді. Сонымен қатар brevity penalty коэффициенті қысқа аудармаларды шектеу үшін енгізіледі. BLEU мәні 0 мен 1 (немесе 0–100%) аралығында болады.

METEOR - семантикалық сәйкестікті ескеретін метрика.

METEOR сөздердің тек дәл сәйкестігін ғана емес, олардың түбірлік формасын (stemming), синонимдерін және сөз тәртібін де ескереді. Бұл көрсеткіш адам бағалауына BLEU-ға қарағанда жиі жақынырақ нәтиже береді.

TER - редакциялау қателерін өлшейтін метрика.

TER (Translation Edit Rate) машиналық аударманы эталон мәтінге айналдыру үшін қажет ең аз редакциялық операциялар (қосу, жою, ауыстыру, орын алмастыру) санын есептейді. TER көрсеткіші неғұрлым төмен болса, аударма сапасы соғұрлым жоғары деп есептеледі.

Transformer архитектурасының бағалау нәтижелері

Workshop on Machine Translation (WMT) халықаралық жарыстары машиналық аударма сапасын салыстырудың негізгі алаңы болып табылады. 2017 жылы ұсынылған Transformer архитектурасы дәл осы WMT байқауларында жоғары BLEU көрсеткіштеріне қол жеткізіп, статистикалық машиналық аударма (SMT) жүйелерінен айтарлықтай басым түсті.

Статистикалық машиналық аударма модельдері (phrase-based SMT) көбіне фразалық сәйкестік пен ықтималдық модельдеріне сүйенсе, Transformer толық контекстті self-attention механизмі арқылы модельдейді. Бұл:

- ұзақ қашықтықтағы тәуелділіктерді тиімді өңдеуге;
- көптілді ортақ семантикалық кеңістік құруға;
- төмен ресурсты тілдер үшін трансферлік оқытуды іске асыруға мүмкіндік береді.

Көптілді жүйелерді бағалаудың ерекшелігі

Қазақ–орыс–ағылшын бағытындағы көптілді NMT жүйелерін бағалау кезінде келесі факторлар ескеріледі:

- Морфологиялық вариативтілік (әсіресе қазақ тілінде) BLEU мәніне әсер етуі мүмкін;
- Еркін сөз тәртібі (орыс тілі) n-грамм сәйкестігін төмендетуі ықтимал;
- Ресми стильге бейімдеу кезінде мәтіннің қайта құрылуы автоматты метрикалардың нәтижесін өзгертуі мүмкін.

Қолданбалы аспектілер

Көптілді сөйлеуді ресми хаттама форматына автоматты түрде бейімдеу келесі салаларда ерекше маңызды:

- Мемлекеттік органдардағы отырыстар
- Халықаралық келіссөздер
- Академиялық кеңестер
- Корпоративтік жиналыстар

1. Мемлекеттік басқару жүйесінде

Мемлекеттік органдардағы отырыстар көп жағдайда қазақ және орыс тілдерінде жүргізіледі, ал халықаралық деңгейдегі іс-шараларда ағылшын тілі де кеңінен қолданылады. Көптілді ортада жедел әрі дәл хаттама дайындау - басқарушылық шешімдердің ашықтығы мен құқықтық дұрыстығын қамтамасыз етудің маңызды шарты [7].

- Автоматтандырылған жүйе:

- отырыс барысын жедел тіркеуге;
- қабылданған шешімдерді нақты құрылымдауға;
- ресми стиль талаптарын сақтауға мүмкіндік береді.

2. Халықаралық келіссөздерде

Халықаралық ұйымдар мен екіжақты келіссөздер барысында тілдік кедергілер ақпараттың бұрмалануына әкелуі мүмкін. Көптілді ASR + NMT + құрылымдау жүйесі:

- айтылған мәліметті бірден бірнеше тілде рәсімдеуге;
- терминологиялық сәйкестікті сақтауға;
- келісім мәтіндерін келісілген форматта ұсынуға мүмкіндік береді.

3. Академиялық және корпоративтік ортада

Университеттердегі академиялық кеңестер, диссертациялық қорғаулар, ғылыми комитет отырыстары көп жағдайда хаттамалық құжаттарды талап етеді. Сол сияқты корпоративтік секторда директорлар кеңесі мен басқарма жиналыстары ресми тіркеуді қажет етеді.

- Қолмен хаттама жазу:
- уақытты көп талап етеді;
- хатшының кәсіби тәжірибесіне тәуелді;
- мазмұнды қысқарту немесе субъективті интерпретация қаупін туындатады.

Ал автоматтандырылған жүйе:

- уақытты үнемдейді;
- терминологиялық бірізділікті сақтайды;
- құжат айналымын жеделдетеді;
- архивтеу мен іздеуді жеңілдетеді;
- адам факторына тәуелділікті азайтады.

4. Қазақстан жағдайындағы стратегиялық маңызы

Қазақстанда мемлекеттік тілдегі цифрлық ресурстарды дамыту стратегиялық басымдықтардың бірі болып табылады. Қазақ тілінің морфологиялық күрделілігі мен деректердің салыстырмалы түрде шектеулілігі көптілді нейрондық модельдерді арнайы бейімдеуді талап етеді.

1. Қазақ тілін қолдайтын NMT және сөйлеуді өңдеу жүйелерін дамыту:
2. цифрлық егемендікті нығайтуға;
3. мемлекеттік басқаруды цифрландыруға;
4. қазақ тілінің ғылыми-технологиялық экожүйесін кеңейтуге;
5. ұлттық тілдің халықаралық ақпараттық кеңістіктегі үлесін арттыруға ықпал етеді.

Шектеулер және болашақ зерттеу бағыттары

Негізгі шектеулер

1. Қазақ тіліндегі үлкен параллель корпустардың шектеулілігі

Қазақ тілі төмен ресурсты (low-resource) тілдер қатарына жатады. Ашық қолжетімді, сапалы, салалық тұрғыдан теңгерілген параллель корпустардың жеткіліксіздігі модельдің жалпылау қабілетіне әсер етеді. Әсіресе ресми-іскери стильдегі деректер көлемі шектеулі.

2. Сөйлеу тіліндегі акцент және фонетикалық вариациялар

ASR жүйелері әртүрлі аймақтық акценттерге, код-ауысуға (code-switching), фонетикалық редукцияға сезімтал. Бұл транскрипция сапасына әсер етіп, кейінгі аударма мен құрылымдау кезеңдерінде қателердің жинақталуына (error propagation) әкелуі мүмкін.

3. Ресми стильді толық автоматтандыру күрделілігі

Ресми хаттама тек тілдік емес, институционалдық нормаларға да тәуелді. Құжат құрылымы ұйымға байланысты өзгеруі мүмкін, ал шешімдердің формулировкасы құқықтық дәлдікті талап етеді. Сондықтан толық автоматтандыру барысында семантикалық дәлдік пен нормативтік сәйкестікті сақтау күрделі міндет болып қалады.

Қорытынды

Зерттеу нәтижелері көрсеткендей, трансформер архитектурасына негізделген көптілді нейрондық машиналық аударма модельдері қазақ–орыс–ағылшын тілдік жұптары арасында тиімді және сапалы аударма жасай алады. Self-attention механизміне негізделген модельдер тілдер арасындағы семантикалық тәуелділіктерді дәл бейнелеп, морфологиялық және синтаксистік айырмашылықтарды тиімді еңсеруге қабілетті екені анықталды.

ASR және NMT технологияларын интеграциялау сөйлеуді мәтінге түрлендіру мен көптілді аударманы біртұтас автоматтандырылған тізбекке біріктіруге мүмкіндік береді. Мұндай интеграциялық тәсіл аудио деректерден бастап құрылымдалған мәтінге дейінгі өңдеу кезеңдерін жүйелі түрде жүзеге асыруға жағдай жасайды.

Сонымен қатар, пост-өңдеу және құрылымдау модульдерін енгізу арқылы алынған мәтінді ресми хаттама форматына сәйкестендіру мүмкіндігі дәлелденді. Дисфлюэнцияларды жою, пунктуацияны қалпына келтіру, абзацтарға бөлу және ресми стильге нормализациялау механизмдері сөйлеу тілінің ерекшеліктерін бейтараптандырып, құжаттың нормативтік-стильдік талаптарға сай болуын қамтамасыз етеді. Бұл модельді тек аударма құралы емес, құрылымдалған ресми құжат генерациялайтын интеллектуалды жүйе ретінде қарастыруға негіз береді.

Ұсынылған әдіс көптілді жиналыстарды құжаттандыру үдерісін жеделдетіп, адам факторынан туындайтын қателерді азайтады және терминологиялық бірізділікті сақтауға ықпал етеді. Әсіресе мемлекеттік басқару, халықаралық келіссөздер және академиялық орта жағдайында мұндай автоматтандырылған шешімдер басқарушылық тиімділікті арттырудың маңызды құралы бола алады.

Қазақ тілін цифрлық кеңістікте дамыту тұрғысынан трансформерге негізделген көптілді жүйелерді жетілдіру стратегиялық мәнге ие. Мұндай технологиялар ұлттық тілдің заманауи жасанды интеллект экожүйесінде толыққанды қолданылуына мүмкіндік береді.

Әдебиеттер тізімі

1. Vaswani A., Shazeer N., Parmar N., et al. Attention Is All You Need // Advances in Neural Information Processing Systems, 2017, Vol. 30, P. 5998–6008.
2. Bahdanau D., Cho K., Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate // Proceedings of the International Conference on Learning Representations (ICLR), 2015.
3. Hinton G., Deng L., Yu D., et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition // IEEE Signal Processing Magazine, 2012, Vol. 29, No. 6, P. 82–97.
4. Chorowski J., Bahdanau D., Serdyuk D., et al. Attention-Based Models for Speech Recognition // Advances in Neural Information Processing Systems, 2015, Vol. 28, P. 577–585.
5. Koehn P. Statistical Machine Translation. Cambridge: Cambridge University Press, 2010, 433 p.
6. Захарова О. И. Семантический анализ и синтез текстовых данных // Вестник ВГУ. Серия: Системный анализ и информационные технологии, 2024, №4, С. 182–208.
7. Калижанова А., Маликова Ф., Дуйсенбек Ф., Дүйсенбек Н. Мәліметтерді өңдеудің модульдік жүйесін жобалаудың модельдері мен әдістерін зерттеу және құру // ҚазККА хабаршысы, 2023, Т. 127, №4, Б. 351–358.

References

1. Vaswani A., Shazeer N., Parmar N., et al. Attention Is All You Need // Advances in Neural Information Processing Systems, 2017, Vol. 30, P. 5998–6008.
2. Bahdanau D., Cho K., Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate // Proceedings of the International Conference on Learning Representations (ICLR), 2015.
3. Hinton G., Deng L., Yu D., et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition // IEEE Signal Processing Magazine, 2012, Vol. 29, No. 6, P. 82–97.

4. Chorowski J., Bahdanau D., Serdyuk D., et al. Attention-Based Models for Speech Recognition // Advances in Neural Information Processing Systems, 2015, Vol. 28, P. 577–585.
5. Koehn P. Statistical Machine Translation. Cambridge: Cambridge University Press, 2010, 433 p.
6. Zaharova O. I. Semanticheski analiz i sintez tekstovyh dannyh // Vestnik VGU. Seria: Sistemnyi analiz i informasionnye tehnologii, 2024, №4, S. 182–208.
8. Kalijanova A., Malikova F., Duisenbek F., Duisenbek N. Malımetterdı ondeudib moduldık juiesin jobalaudyn modelderi men adisterin zertteu jane quru // QazKKA habarshysy, 2023, T. 127, №4, B. 351–358.

М.Б. Мизам[1], Ж.Д. Изтаев[1], О.М. Сулеймен[1]*, П.А. Кожабекова[1], С.Б. Атажонова[2]
[1]магистрант, mizam.madi.92@mail.ru, ЮКУ им. М. Ауэзова, Шымкент, Казахстан
[1]к.п.н., ассоциированный профессор, zhalgasbek71@mail.ru, ЮКУ им. М. Ауэзова, Шымкент, Казахстан
[1*]магистр, преподаватель, jasiksuleimen42@gmail.com, ЮКУ им. М. Ауэзова, Шымкент, Казахстан
[1]к.т.н., доцент, pernesh63@mail.ru, ЮКУ им. М. Ауэзова, Шымкент, Казахстан
[2]PhD по педагогическим наукам, доцент, saida_atajonova@astiedu.uz, Андижанский государственный технический институт, г. Андижан, Республика Узбекистан

МНОГОЯЗЫЧНАЯ МОДЕЛЬ НЕЙРОННОГО МАШИННОГО ПЕРЕВОДА (NMT) ДЛЯ АДАПТАЦИИ РЕЧИ НА КАЗАХСКОМ, РУССКОМ И АНГЛИЙСКОМ ЯЗЫКАХ К ФОРМАТУ ПРОТОКОЛА

Аннотация

Данное исследование всесторонне рассматривает задачу автоматического перевода речи на казахский, русский и английский языки с последующей адаптацией к формату официального протокола. Работа включает глубокий анализ научных основ многоязычных моделей нейронного машинного перевода (NMT) на базе архитектуры трансформеров и демонстрирует методы их интеграции с системами автоматического распознавания речи (ASR). Предложенный подход охватывает полный цикл обработки – от аудиосигнала до структурированного официального документа, включая транскрипцию, перевод, пост-обработку и приведение текста к официальному стилю. Результаты показывают, что многоязычные трансформерные модели способны обеспечивать высококачественный перевод между казахским, русским и английским языками. Кроме того, метод позволяет автоматически адаптировать текст к формату протокола за счет удаления дисфлюенций, восстановления пунктуации, поддержания терминологической согласованности и нормализации текста в официальный стиль. Предложенное решение существенно ускоряет процесс документирования многоязычных встреч, способствует систематизации информации и развитию цифровых ресурсов казахского языка.

Ключевые слова: нейронный машинный перевод, многоязычная модель, распознавание речи, протокольный формат, трансформер, автоматизация, машинное обучение.

M.B. Mizam[1], Zh.D. Iztaev[1], O.M. Suleimen[1]*, P.A. Kozhabekova[1], S.B. Atajonova[2]
[1]Master's Student, mizam.madi.92@mail.ru, M. Auezov South Kazakhstan University, Shymkent, Kazakhstan
[1]Candidate of Pedagogical Sciences, Associate Professor, zhalgasbek71@mail.ru, M. Auezov South Kazakhstan University, Shymkent, Kazakhstan
[1*]Master, lecturer, jasiksuleimen42@gmail.com, M. Auezov South Kazakhstan University, Shymkent, Kazakhstan
[1]Candidate of Technical Sciences, Associate Professor, pernesh63@mail.ru, M. Auezov South Kazakhstan University, Shymkent, Kazakhstan
[2]PhD in Pedagogical Sciences, Associate Professor, saida_atajonova@astiedu.uz, Andijan State Technical Institute, Andijan, Republic of Uzbekistan

MULTILINGUAL NEURAL MACHINE TRANSLATION (NMT) MODEL FOR ADAPTING KAZAKH, RUSSIAN, AND ENGLISH SPEECH TO PROTOCOL FORMAT

Abstract

This study comprehensively addresses the problem of automatically translating speech in Kazakh, Russian, and English and adapting it to an official protocol format. The work provides an in-depth analysis of the scientific foundations of multilingual neural machine translation (NMT) models based on the transformer architecture and demonstrates methods for integrating them with automatic speech recognition (ASR) systems. The proposed approach encompasses the full processing pipeline from audio signals to structured official documents, including transcription, translation, post-processing, and formatting text into an official style. The results show that multilingual transformer models are capable of producing high-quality translations among Kazakh, Russian, and English. Additionally, the method enables automatic adaptation of text to a protocol format by removing disfluencies, restoring punctuation, maintaining terminological consistency, and normalizing text to an official style. The proposed approach significantly accelerates the documentation of multilingual meetings, facilitates information structuring, and contributes to the development of Kazakh digital language resources.

Keywords: neural machine translation, multilingual model, speech recognition, protocol formatting, Transformer architecture, automation, machine learning.