

## МАШИНАЛЫҚ АУДАРМАНЫҢ ДАМУ ТАРИХЫ ЖӘНЕ КЕЗЕҢДЕРІ

**Әбубәкір Эльнура Маратқызы**

6B06118 - «Ақпараттық жүйелер және технологиялар»  
Білім беру бағдарламасының 2 курс студенті

**Аяз Мирас Ерланұлы**

6B06118 - «Ақпараттық жүйелер және технологиялар»  
Білім беру бағдарламасының 4 курс студенті

М.Х.Дуллати атындағы Тараз университеті, Тараз қ., Қазақстан Республикасы

Ғылыми жетекшілері: **Адилова А.К., Балгинбаева К.А.**

[ak.adilova@dulaty.kz](mailto:ak.adilova@dulaty.kz), [ka.balginbaeva@dulaty.kz](mailto:ka.balginbaeva@dulaty.kz)

Интернеттің кең таралуымен, жаһанданудың жеделдеуімен және елдер арасында жиі ақпарат алмасумен адамдардың аудармадағы сұранысы артып келеді және дәстүрлі қолмен аудару аудармаға тез өсіп келе жатқан қажеттілікті қанағаттандырмауда. Деректерді машинамен өңдеу қабілеті үлкен көлемдегі ақпаратты жылдам аударуға мүмкіндік береді. Машиналық аударма дамудың жаңа кезеңін ашты және уақыт талабына сәйкес машиналық аударманың әртүрлі бағдарламалық қамтамасыз етуі пайда болды.

Соңғы жылдары терең оқытудың жылдам дамуымен машиналық аударма нейрондық желілерді (neural machine translation, NMT) зерттеуде үлкен прогресс байқалды. Көптеген тілдерде нейрондық желілерді пайдалану арқылы машиналық аударманың өнімділігі айтарлықтай жақсарып, дәстүрлі статистикалық машиналық аударманың (statistical machine translation, SMT) тиімділігінен әлдеқайда асып түсті және ол ғылыми ортада үлкен ықпалға ие. Бұл осы зерттеудің маңыздылығын түсіндіреді. Таңдалған тақырыптың өзектілігі нейрожелілік технологиялардың жылдам дамуына және олардың аударма сапасын арттыру және қолданушылық тәжірибені жақсарту үшін мүмкіндіктеріне негізделген. Жұмыста қазіргі заманғы машиналық аударма жүйелері қаралады және оларды нейрондық желілердің көмегімен дамыту, қателіктерін түзету тәсілдері ұсынылады.

«Аударма» сөзі латын тіліндегі «translatio» сөзінен шыққан, бұл «көшіру, ауыстыру немесе беру» дегенді білдіреді. Алғашқысы адамдарға географиялық айырмашылықтар мен жалпы сауда-саттық үшін кедергілерден туындаған қиындықтарды еңсеруге мүмкіндік

беруден тұрды. Жаһанданудың тереңдей дамуына қарай қазіргі заманғы әлемдік саяси, экономикалық, ғылыми және мәдени байланыстар барған сайын тығыз бола түсуде және тіл барған сайын маңызды ақпарат алмасуда, бұл кеңінен танылады. Сондықтан шет тілдерін үйренуге және аударма қызметін дамытуға үлкен мән берілуде. Соңғы жылдары машиналық көшіру үлкен маңызға ие болды. Бірақ әлемде тілдер өте көп, адамдардың барлығын игеруге энергиясы мен уақыты жетпейді, адамдардың аударма қызметтеріне деген сұранысы ұзаққа созылады, сондықтан адам аудармасы адамдардың қажеттіліктерін қанағаттандыра алмайды, сондықтан машина аудармасы пайда болды. Машина аудармасының тарихы 1930-1940 жылдардан басталады. Ұшып-көтерілу мен құлдыраудың 60 жылынан кейін машина аудармасы әдетте ғылыми топтар келесі төрт кезеңге бөлетін айналмалы жол дамуынан өтті:

Басталуы (1947-1964). 1950-жылдардың басынан бастап 1960-жылдардың бірінші жартысына дейін машиналық аударма саласындағы жылдық зерттеулер тенденцияның өсуін көрсетті. Америка Құрама Штаттары мен бұрынғы Кеңес Одағы, екі аса әскери, саяси және экономикалық мақсаттар машиналық аударма жобаларына үлкен қаржылай қолдау көрсетті, ал Еуропаның геосаясат пен экономика қажеттіліктерімен де машиналық аударма саласындағы зерттеулерге айтарлықтай көңіл бөлді. Машиналық аударманы зерттеу қарқын алады.

Құлдырау (1964-1975). 1966 жылғы қарашада ALPAC комиссиясы «Language and Machines: Computers in Translation and Linguistics» есебін жариялады, машиналық аударманың мақсаттылығы екі жақты қабылданбады, машиналық аударма жобаларын қаржылық қолдауды тоқтатуға мүмкіндік берді. Бұл баяндаманы жариялау дамып келе жатқан машина аудармасына айтарлықтай соққы берді және машина аудармасын зерттеу тұйыққа тірелді.

Қалпына келтіру (1975-1989). 1970-жылдары кіргеннен кейін ғылымды және техниканы дамыту және елдер арасында ғылыми-техникалық ақпаратпен жиі алмасу, елдер арасындағы тілдік кедергі, дәстүрлі адами аударма аударма жұмысына қатысу үшін компьютерлердегі қажеттіліктерді қанағаттандырудан алыс болды. Уақыттың ұзақтығы, компьютерлік ғылымдар мен лингвистикалық зерттеулерді дамыту, әсіресе компьютерлік аппараттық технологияларды айтарлықтай жетілдіру және өңдеуде жасанды интеллектті қолдану техникалық деңгейден машиналық аударманы қалпына келтіруге ықпал етті. Машина ауыстыру жобалары дами бастады және дәйекті түрде әртүрлі практикалық және эксперименттік жүйелер іске қосылды.

Жаңа кезең (1990-шы жылдар аралығында): Интернетті кеңінен қолдану, әлемдік экономикалық интеграция процесін жеделдету және халықаралық қарым-қатынас жасау арқылы дәстүрлі адами ауыспалы аудармаға деген бұрын-соңды болмаған өсіп келе жатқан сұранысты қанағаттандыруға қабілетті. Бұл машина аудармасын дамыту үшін жаңа мүмкіндіктер жасады.

1933 жылы кеңестік ғалым Петр Смирнов Троянскийге «бір тілді басқа тілге аудару кезінде сөздерді таңдауға және басып шығаруға арналған машина» деген патенттік баяндама алды. Оның өнертабысы өте қарапайым, соның ішінде карта 4 тілде, жазу машинкасы мен ескі пленкалы фотоаппарат. халықаралық қарым-қатынаспен аударды. Бұл машина аудармасын дамыту үшін жаңа мүмкіндіктер ашты.

1946 жылы әлемдегі алғашқы электрондық компьютер пайда болды. 1947 жылы американдық математик Уоррен Уивер (Warren Weaver) жаратылыстану тілін аудару үшін компьютерлерді пайдалануды ұсынды. Сондықтан осы жыл зерттеушілік облыс ретінде машиналық аударманың туған күні болып саналады. Екі жылдан кейін Уоррен Уивер машиналық аударма тұжырымдамасын қалыптастырды. Ол машина аударма туралы бірқатар мақалалар жазып, «Уивер меморандумын» (Weaver-Memorandum) жариялады. 1954 жылы 7 қаңтарда Нью-Йорктегі IBM штаб-пәтерінде Джорджтаун эксперименті іске қосылды - автоматты түрде ағылшын тіліндегі 60 орыс сөйлеміне IBM701 компьютер, бұл тарихтағы алғашқы демонстрациялық машина аудармасы болды.

1964 жылы облыстағы ғылыми зерттеулердің ілгерілеуін бағалау үшін Американдық ғылым академиясы сөзді автоматты өңдеу жөніндегі комитетті (Automatic Language Processing Advisory Committee) құрды, ол екі жылдық талдау мен тестілеуді бастады. 1966 жылы Сөзді автоматты өңдеу проблемалары жөніндегі американдық комитет (ALPAC) «Language and Machines: Computers in Translation and Linguistics» есебін жариялады, ол осы саладағы мемлекеттік қаржыландыруды тоқтатуға алып келеді.

1980 жылдардың басында машина аудармасын зерттеу компьютерлік техниканы дамытуды жандандырды. 1980-жылдардың соңы мен 1990-жылдардың басында машинамен аудару жүйелері, соның ішінде дауыспен аудару, статистикалық машинамен аудару және басқалары нарыққа шығарылды. 1990-шы жылдардың соңында Интернет пайда болды және автоматтандырылған аударма мен машиналық аударма негізінде мысалдар жылдам дамыды.

Машина аудармасы 21 ғасырда кәсіби емес аудармашылар мен аударма бюросы үшін алмастырылмайтын көмекші болды. 2016 жылдың қыркүйегінде Google негізгі нейрондық желілерге (GNMT) аудару жүйесін шығарды. Соңғы бірнеше жылда біз үздік машинаны көшіруді қамтамасыз ету үшін нейрондық желілердің машиналық оқыту сияқты жасанды интеллект технологияларын пайдаланатын аударма компанияларындағы елеулі өзгерістердің куәсі болдық.

Машиналық аударма мәтінді немесе сөзді бір тілден екінші тілге автоматты түрде аудару үшін компьютерлік жүйелерді пайдалануды білдіреді. Есептеу қуаттары мен алгоритмдердің дамуымен, сондай-ақ қол жетімді деректер көлемінің өсуімен машинамен аудару әдістерінде айтарлықтай өзгерістер орын алды. Бүгінде статистикалық аударманы, нейрондық желілер мен гибридік үлгілерді қоса алғанда, бірнеше негізгі тәсілдерді бөліп көрсетуге болады.

Статистикалық машиналық аударма әдісі (Statistical Machine Translation, SMT) нейрондық желілер пайда болғанға дейін ең танымал және кеңінен қолданылатын әдістердің бірі болды. Статистикалық аударманың негізгі идеясы параллельді мәтіндердің үлкен көлемін статистикалық талдау негізінде аудару ықтималдығын бағалайтын математикалық модельді жасау болып табылады. Бұл тәсілдегі негізгі міндет аударманың фрагменттелген үлгілерін жасау болып табылады, онда әрбір фрагмент (немесе сөз тіркесі) ықтималдық үлгілерін пайдалана отырып аударылады.

Фразалық аударма (Phrase-based Translation, PBT) - бұл жеке сөздер емес, сөз тіркестері немесе мәтін фрагменттері деңгейінде аударылатын машиналық аударма әдісі. Жекелеген сөздермен жұмыс істейтін ауызша аударма сияқты неғұрлым қарапайым әдістерден айырмашылығы, фразалық аударма фразалар, сөздер топтары немесе тіпті тұтас сөйлемдер сияқты мәтіннің неғұрлым ұзын бірліктерін өңдеуге бағытталған. Фразалық аударма модельдері үлкен параллель корпустарда, яғни бастапқы тілдегі әрбір фрагменттің нысаналы тілге тиісті аудармасы бар мәтіндерде оқытылады. Оқыту процесінде модельдер бастапқы және мақсатты тілдердегі сөз тіркестері арасындағы ықтималдық байланыстарды талдайды, бұл жүйеге неғұрлым кең контекстерді ескеруге және аударманың дәлдігін жақсартуға мүмкіндік береді.

Фразалық аударма модельдерін оқыту үшін мәтіндердің параллель корпусы қолданылады - бастапқы тілдегі сөйлемдер жиынтығы және олардың мақсатты тілге тиісті аудармалары. Оқу процесі бірнеше негізгі кезеңдерді қамтиды:

- Сөз тіркестерін бөлектеу: бастапқы және мақсатты тілдердің мәтіндері сөз тіркестеріне бөлінеді. Ол үшін, мысалы, статистикалық немесе синтаксистік принциптерге негізделген сөз тіркестерін оқшаулай алатын алгоритм қолданылады;

- Фразалық сәйкестіктерді іздеу: әрі қарай, жүйе екі тілде де бір-біріне сәйкес келетін сөз тіркестерін іздейді. Ол үшін максималды ықтималдықты іздеу әдісі қолданылады, мұнда оқыту деректерінің үлкен көлеміне сүйене отырып, бастапқы тілден бір фразаның осылай аударылу ықтималдығы есептеледі;

– Тіл модельдерін қолдану: аударманың сапасын жақсарту үшін мақсатты тіл контекстінде мақсатты сөз тіркестерінің ықтималдығын бағалайтын қосымша тіл моделі жиі қолданылады. Мысалы, мақсатты тілде сөйлемнің қаншалықты табиғи естілетінін бағалау үшін n-g модельдерін қолдануға болады.

Фразалық аударманы нейрондық желілермен салыстыру. Seq2Seq және Transformer сияқты нейрондық желілер мен модельдердің пайда болуымен фразалық аударма біртіндеп икемді әдістермен ауыстырылды, олар контексті сөйлем деңгейінде және тіпті мәтіндерде ескеруге қабілетті.

Нейрондық желілер:

– Сөздер, сөз тіркестері мен сөйлемдер арасындағы күрделі контексттер мен байланыстарды қарастыруға болады;

– Ұзақ тәуелділіктер мен контексттер тиімді өңделеді, бұған фразалық аударма қол жеткізе алмайды;

– Олар алдын-ала дайындалған фразалық сөздіктерді қажет етпейді және үлкен көлемде оқи алады.

Гибридті статистикалық модельдер. Машиналық аудармадағы гибридті статистикалық модельдер - бұл аударма сапасын жақсарту мақсатында ережелер (лингвистикаға негізделген) және статистикалық әдістер сияқты әртүрлі әдістерді біріктіретін тәсіл. Бұл тәсіл осы әдістердің әрқайсысының кемшіліктерін жою мақсатында статистикалық машиналық аударманың (SMTS) және лингвистикалық ережелерге негізделген аударма жүйелерінің жақсы жақтарын біріктіреді.

Гибридті статистикалық модельдердің артықшылықтары: грамматикалық нормаларды жақсы сақтау. Жақсартылған дәлдік:

– Лингвистикалық ережелерді қолдану аударманың дәлдігін жақсартуға мүмкіндік береді, әсіресе статистикалық әдістер қате болуы мүмкін жағдайларда, мысалы, көп мағыналы сөздерді немесе стандартты емес грамматикалық құрылымдарды аудару кезінде;

– Ережелер мақсатты тілдің грамматикалық құрылымы мен синтаксистік тәртібін сақтауға көмектеседі, бұл әртүрлі құрылымдары бар тілдер үшін өте маңызды.

Сөз тіркестеріне негізделген статистикалық аударманы қолданатын, бірақ аударманы түзету үшін грамматикалық және семантикалық ережелерді одан әрі қолданатын модель келесідей болуы мүмкін:

– Статистикалық компонент кіріс мәтінін талдайды және сөз тіркестері мен сөздердің ықтималдық сәйкестігіне негізделген алдын ала аударманы орындайды;

– Содан кейін жүйе аударманы нақтылау үшін лингвистикалық ережелерді қолданады, мысалы, сөздерді дұрыс ауыстыру немесе контекстке байланысты сөздің дұрыс түрін таңдау;

– Нәтижені мақсатты тілде табиғи және грамматикалық тұрғыдан дұрыс формаға келтіру үшін жұмыс істейтін алгоритмдерден кейінгі арнайы түзетулер енгізуге болады.

Гибридті статистикалық модельдер машиналық аударма саласындағы қуатты құрал болып табылады, өйткені олар статистикалық әдістер мен лингвистикалық ережелердің ең жақсы жақтарын біріктіреді. Бұл тәсіл әртүрлі тілдік жұптар мен контексттерге икемділік пен бейімделуді сақтай отырып, аударманың дәлдігін айтарлықтай жақсартуға мүмкіндік береді. Есептеу шығындарының жоғарылауына және сапалы лингвистикалық ережелердің қажеттілігіне қарамастан, гибридті модельдер жоғары сапалы аудармаға қол жеткізудің тиімді әдістерінің бірі болып қала береді.

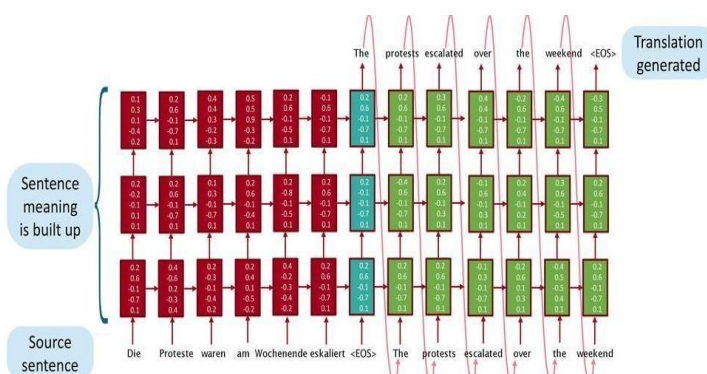
Машиналық аудармадағы нейрондық желілер. Соңғы жылдары нейрондық желілер үлкен көлемдегі деректермен жұмыс істеудегі артықшылықтары мен күрделі тәуелділіктер мен контексттерді модельдеу қабілетінің арқасында машиналық аударма саласында жетекші рөл атқарды. Нейрондық желілер, статистикалық әдістерден айырмашылығы,

мәтінді фразаларға немесе жеке сөздерге алдын-ала бөлуді қажет етпейді, бірақ деректердің үздіксіз тізбегі ретінде бастапқы мәтінмен жұмыс істей алады.

Машиналық аудармадағы нейрондық желілер (Neural Machine Translation, NMT) – мәтінді бір тілден екінші тілге автоматты түрде аудару үшін нейрондық желілерді пайдалануға негізделген машиналық аударма әдісі. Нейрондық желілер, дәстүрлі статистикалық және фразалық аударма әдістерінен айырмашылығы, жүйеге кең контекстті ескеруге және нақты грамматикалық тұрғыдан дұрыс аудармаларды құруға мүмкіндік беретін 1-суретте көрсетілген күрделі тәсіл болып табылады.

Нейрондық желілерге негізделген машиналық аударма немесе терең оқыту машиналық аударма жүйесі (Deep learning machine translation), қазір белгілі болғандай, 90-шы жылдардың басынан бастап сөйлеуді тану бағдарламаларында алғаш рет қолданыла бастады. Машиналық аудармада нейрондық желілерді қолдану туралы алғашқы ғылыми зерттеу жарияланды 2014, мұндай жүйелердің көптеген кейінгі артықшылықтары бар (үлкен сөздік базалары, графикалық ақпаратты анықтау үшін пайдалану мүмкіндігі, субтитрлерді аудару, көптілділік).

Нейрондық желілер бастапқыда адам миының биологиялық нейронынан шабыттан-ды, онда нейрондар негізгі ақпаратты жеткізеді және өндейді, содан кейін ми күрделі ұғымдар мен идеяларды құрастырады. Жасанды нейрондық желілер, адам миы сияқты, иерархиялық құрылымда жиналған ақпараттың әртүрлі бөліктерінен күрделі тұжырымдамалар құра алады. Тереңдетіп оқыту жүйелері, кем дегенде теория жүзінде адамның бөгде араласуынсыз, оның қолындағы деректерге сүйене отырып, оқу процесін толығымен автоматты түрде жасауға мүмкіндік берді.



Сурет 1 – Машиналық аудармадағы нейрондық желілер.

Машиналық аудармадағы нейрондық желілер статистикалық әдістерді алмастырды (мысалы, фразалық аударма) және деректердің үлкен көлемінде оқу қабілеті мен жоғары икемділігінің арқасында осы салада тез жетекші орынға ие болды. Машиналық аудармадағы нейрондық желілер терең оқытуға негізделген модельдерді қолдануға бейім, бұл оларға үлкен көлемдегі деректермен тиімді жұмыс істеуге және тілдер арасындағы күрделі байланыстарды анықтауға мүмкіндік береді. Бұл модельдердің негізгі сипаттамалары:

– Аударма үшін қолмен реттелген ережелер мен статистиканы қолданған дәстүрлі әдістерден айырмашылығы, нейрондық желілер адамның алдын-ала араласуынсыз барлық қажетті белгілерді (мысалы, синтаксистік, морфологиялық, семантикалық) тікелей деректерден ала алады;

– Нейрондық желілерді басынан аяғына дейін, яғни бастапқы мәтінді енгізуден аударманы шығаруға дейін үйретуге болады. Бұл оқу процесін едәуір жеңілдетеді және статистикалық модельдердегідей аралық кезеңдерді қолмен реттеу қажеттілігі керек емес.

Машиналық аудармаға қатысты тереңдетіп оқыту жүйелері алгоритмдерді мүмкін етті, мұнда тек бірнеше параметрлер қолмен реттеледі, ал идея жүйенің өзі - өзінің мәліметтер базасына сүйене отырып, ең жақсы сәйкестікті таңдай алады. Осы тұрғыдан алғанда, бұл идея таза статистикалық аударма модельдерінде ұсынылған, бірақ іс жүзінде көптеген параметрлерді қолмен конфигурациялау қажет болды. Мысалы, біз жоғарыда қарастырған бес IBM жүйесінде әрбір жаңа модель қолмен жазылып, алдыңғысының кемшіліктерін түзетуге арналған. Тереңдетіп оқыту жүйелері, керісінше, кем дегенде теория жүзінде адамның бөгде араласуынсыз, оның қолындағы деректерге сүйене отырып, оқу процесін толығымен автоматты түрде жасауға мүмкіндік жасады.

Нейрондық желілерге негізделген машиналық аударма немесе "терең оқыту" машиналық аударма жүйесі (Deep learning machine translation), қазір белгілі болғандай, 90-шы жылдардың басынан бастап сөйлеуді тану бағдарламаларында алғаш рет қолданыла бастады. Машиналық аудармада нейрондық желілерді қолдану туралы алғашқы ғылыми зерттеу жарияланды 2014, мұндай жүйелердің көптеген кейінгі артықшылықтары бар (үлкен сөздік базалары, графикалық ақпаратты анықтау үшін пайдалану мүмкіндігі, субтитрлерді аудару, көптілділік).

Нейрондық желілер бастапқыда адам миының биологиялық нейрондарынан шабыттанды, онда нейрондар негізгі ақпаратты жеткізеді және өндейді, содан кейін ми күрделі ұғымдар мен идеяларды құрастырады. Жасанды нейрондық желілер, адам миы сияқты, иерархиялық құрылымда жиналған ақпараттың әртүрлі бөліктерінен күрделі тұжырымдамалар құра алады. Тереңдетіп оқыту жүйелері, кем дегенде теория жүзінде адамның бөгде араласуынсыз, оның қолындағы деректерге сүйене отырып, оқу процесін толығымен автоматты түрде жасауға мүмкіндік берді.

Нейрондық машиналық аударма адам миының жұмысын имитациялайтын жасанды нейрондық желілерге негізделген. NMP негізгі элементтері-кодерлер мен декодерлер. Кодер бастапқы мәтінді жасырын көрініске (векторға) түрлендіреді, ал декодер бұл векторды мақсатты тілдегі мәтінге түрлендіреді. NMP-де назар аудару механизмі (attention mechanism) ерекше рөл атқарады, бұл модельге аударманы құру кезінде бастапқы мәтіннің әртүрлі бөліктеріне назар аударуға мүмкіндік береді. Бұл контекстті жақсырақ қарастыруға және аударма сапасын жақсартуға көмектеседі. Қарапайым тілмен айтқанда, тіл ережелеріне немесе сөз тіркестеріне негізделген бұрынғы машиналық аударма жүйелерінен айырмашылығы, нейрондық желілер келесі сөзді болжау үшін бүкіл бастапқы сөйлемнің контекстін, сондай-ақ бұрыннан жасалған аударма сөздерінің реттілігін пайдаланады. Бұл тәсіл кәсіпқойдың қалай аударатынын еске салады.

Сонымен қатар, нейрондық желілер тиісті мәліметтерден сабақ ала отырып, белгілі бір пәндік салалар мен стильдерге тез бейімделе алады. Бұл олардың икемділігі мен әртүрлі сценарийлерде қолданылуын айтарлықтай арттырады.

Машиналық аударманы бағалаудың негізгі аспектілерінің бірі-оның нәтижелерін адам аудармасымен салыстыру. Ол үшін мәтіндерді нейрондық желілер де, кәсіби аудармашылар да аударатын стандартталған тесттер қолданылады. Содан кейін нәтижелер дәлдік, мағынаны сақтау және стилистикалық дұрыстықты қоса алғанда, әртүрлі параметрлер бойынша талданады.

### **Қолданылған әдебиеттер тізімі:**

1. И.Н.Дьяченко, Ю.В.Матыченко. Нейронный машинный перевод: преимущества, сложности, перспективы // Языки и литература в поликультурном пространстве. 2020, №6. Стр 28-33.
2. К.А.Дроздова, М.И.Левин. Машинный перевод: история, классификация, методы// слово молодым. Гуманитарные исследования. 2015, № 3 (7). стр 156-158
3. И.В.Карасев, Е.А.Артюшина. Системы машинного перевода// Успехи современного естествознания. 2011, №7, стр 117-118.

4. Н.А.Крашенинникова. Классификация типичных ошибок при использовании систем машинного перевода // Симбирский научный вестник.2011, № 2 (4). стр 190-194.

5. М.С.Легостина. Метрики оценки качества машинного перевода В сборнике ИННОВАТИКА-2019. Сборник материалов XV Международной школы-конференции студентов, аспирантов и молодых ученых. 2019. стр 490-493.

6. А.А.Нанкевич. Оценка качества результатов систем машинного перевода// Системы компьютерной математики и их приложения. 2019, №20-1. стр 215-221.

7. К.И.Фаустова. Нейронные сети: применение сегодня и перспективы развития Территория науки. 2017. № 4. С. 83-87.

8. Лопатин, И.А. Методы и алгоритмы машинного перевода с использованием нейронных сетей. - М.: Наука, 2019. стр 110 - 145.