

MASHINA TARJIMASI TIZIMLARIDA O‘ZBEK TILINING LINGVISTIK XUSUSIYATLARINI HISOBGA OLISH MASALALARI

Abu Rayhon Beruniy nomidagi Urganch Davlat universiteti

Bekchanov Hikmat Mansurbek o‘g‘li

e-mail: bekchanovhikmat209@gmail.com

ANNOTATSIYA

Ushbu maqolada mashina tarjimasi tizimlarida o‘zbek tilining lingvistik xususiyatlarini hisobga olish masalalari ko‘rib chiqiladi. O‘zbek tilining agglutinatív tuzilishi, murakkab morfologiyasi va sintaktik o‘ziga xosliklari tarjima jarayonida qator qiyinchiliklar keltirib chiqaradi. Maqolada Google Translate, DeepL va boshqa zamonaviy tarjima tizimlarining o‘zbek tili bilan ishlash samaradorligi tahlil qilinadi. Shuningdek, neyron tarmoqlar asosidagi tarjima modellarining imkoniyatlari va tarjima sifatini oshirishga qaratilgan amaliy takliflar bayon etiladi.

Kalit so‘zlar: mashina tarjimasi, neyron tarjima, o‘zbek tili, agglutinatsiya, transformer modeli, Google Translate, tarjima sifati, NMT.

1. KIRISH

Globalashuv davrida til to‘siqlari iqtisodiy, ilmiy va madaniy almashinuvning asosiy to‘sig‘iga aylanmoqda. Mashina tarjimasi (Machine Translation — MT) texnologiyasi ushbu to‘siqni bartaraf etishda hal qiluvchi vosita sifatida qaralib, so‘nggi o‘n yil ichida misli ko‘rilmagan sur‘atda rivojlandi. Bugungi kunda Google Translate, DeepL, Yandex.Translate kabi tizimlar kuniga milliard jumladan ortiq tarjima amalga oshirib, barcha qit‘a foydalanuvchilariga xizmat ko‘rsatmoqda.

O‘zbek tili — 35 million nafardan ortiq inson so‘zlaydigan, Turkiy til oilasiga mansub, agglutinatív tuzilishga ega til bo‘lib, uning mashina tarjimasi tizimlariga integratsiya qilinishi o‘ziga xos muammolar majmuasini keltirib chiqaradi. Bir tomondan, o‘zbek tilining boy morfologiyasi va erkin so‘z tartibi tarjima algoritmlarini chalg‘itishi mumkin; ikkinchi tomondan, o‘zbek tili uchun mavjud parallel korpuslarning kamligi (low-resource problem) neyron tarjima modellarini o‘qitishni murakkablashtiradi.

Ushbu maqolaning asosiy maqsadi — mashina tarjimasi tizimlarida o‘zbek tilining lingvistik xususiyatlarini hisobga olish masalalarini tizimli o‘rganish, mavjud yechimlarni tahlil qilish va tarjima sifatini oshirishga qaratilgan ilmiy asoslangan takliflar ishlab chiqishdan iborat. Tadqiqot natijalaridan o‘zbek tilidagi NLP tizimlari, tarjima platformalari va ta‘lim texnologiyalarini ishlab chiquvchi mutaxassislar foydalanishi mumkin.

2. MASHINA TARJIMASI TUSHUNCHASI VA TARIXI

Mashina tarjimasi (MT) — kompyuter dasturlari yordamida bir tabiiy tildan boshqa tabiiy tilga avtomatik tarjima qilish jarayonidir. MT tarixi 1949-yilda Warren Weaver'ning memorandumidan boshlanadi va bugungi katta til modellarigacha bo'lgan ulkan evolutsiyani qamrab oladi.

Mashina tarjimasining asosiy rivojlanish bosqichlari:

- Qoidaviy mashina tarjimasi (Rule-Based MT, RBMT, 1950–1990) — lingvistlar tomonidan qo'lda yozilgan grammatik qoidalar va ikki tilli lug'atlar asosida ishlaydi. Yuqori aniqlik, lekin katta mehnat talab qiladi va yangi domenga moslashish qiyin;
- Statistik mashina tarjimasi (Statistical MT, SMT, 1990–2014) — katta hajmdagi parallel korpuslardan statistik naqshlarni o'renadb, n-gram modellari va IBM modellariga asoslanadi. Google Translate 2006–2016-yillarda ushbu yondashuvdan foydalangan;
- Neyron mashina tarjimasi (Neural MT, NMT, 2014-hozir) — chuqur o'reanish va encoder-decoder arxitekturasida ishlaydi; 2017-yildan transformer modeli ustunlik qilmoqda. Google, DeepL, Yandex hozirda shu yondashuvdan foydalanadi;
- Katta til modellari asosidagi tarjima (LLM-based MT, 2022-hozir) — GPT-4, Claude, Gemini kabi modellar maxsus tarjima o'reatishsiz ham yuqori sifatli tarjima bera oladi; few-shot va zero-shot tarjima imkoniyati.

Tarjima sifatini o'lchovchi asosiy mezonlar: BLEU (Bilingual Evaluation Understudy) — n-gram mos kelishiga asoslangan avtomatik baholash; METEOR, ChrF, TER — BLEU'ning kamchiliklarini to'ldiruvchi qo'shimcha o'lchovlar; MQM (Multidimensional Quality Metrics) — professional tarjimonlar tomonidan ko'p o'lchamli baholash; COMET — neyron tarmoq asosidagi, inson baholashiga yaqin avtomatik metrika.

O'zbek tili uchun mashina tarjimasi 2000-yillar boshidan e'tiborga olinib boshlangan. Dastlab qoidaviy tizimlar (o'zbek-rus juftligi uchun) ishlab chiqilgan bo'lsa, hozirda neyron tarjima modellari ustuvorlik qilmoqda. Biroq o'zbek tili hali ham "past resursli" (low-resource) tillar toifasida qolmoqda.

3. GOOGLE TRANSLATE VA BOSHQA TARJIMA TIZIMLARI

Bugungi kunda o'zbek tilini qo'llab-quvvatlovchi bir nechta yirik mashina tarjimasi platformalari mavjud bo'lib, ularning har biri o'ziga xos texnologiya va yondashuvdan foydalanadi.

Google Translate — 2006-yilda ishga tushirilgan, hozirda 133 dan ortiq tilni qo'llaydigan eng ommabop tarjima xizmati. O'zbek tilini 2010-yildan qo'llab-quvvatlaydi. Google'ning Transformer asosidagi GNMT (Google Neural Machine Translation) tizimi 2016-yilda statistik tarjimani almashtirgach, o'zbek tili uchun

tarjima sifati sezilarli yaxshilandi. Biroq o'zbek tilida yuqori sifatli parallel korpuslar etishmasligi tufayli ingliz, nemis, fransuz kabi tillarga nisbatan aniqlik past bo'lib qolmoqda.

Boshqa yirik tarjima platformalarining o'zbek tili bilan ishlashi:

- Yandex.Translate — o'zbek tilini 1990-yillardan qo'llab-quvvatlaydi; rus-o'zbek juftligi uchun nisbatan ko'proq o'ratish ma'lumotlariga ega bo'lib, bu yo'lda yuqoriroq natija ko'rsatadi;
- DeepL — Yevropa tillarida eng yuqori sifat ko'rsatuvchi tizim hisoblanadi; o'zbek tilini hozircha rasmiy qo'llamaydi, lekin API orqali cheklangan ko'rinishda ishlash mumkin;
- Microsoft Translator (Azure) — 100 dan ortiq tilni qo'llab-quvvatlaydi, o'zbek tili ham kiritilgan; Office va Teams mahsulotlarga integratsiyalashgan;
- Amazon Translate — bulutli tarjima xizmati; o'zbek tilini 2022-yildan qo'llab-quvvatlaydi; korporativ foydalanuvchilarga yo'naltirilgan;
- Nllb (No Language Left Behind, Meta AI, 2022) — 200 dan ortiq tilni, jumladan o'zbek tilini (lotin va kiril alifbolarida) qo'llab-quvvatlovchi ochiq manbali model; past resursli tillar uchun alohida optimallashtirilgan.

Amaliy sinovlar shuni ko'rsatadiki, ushbu tizimlarning barchasi o'zbek tilidagi murakkab morfologik shakllar, qo'shma fe'llar va maqol-iboralarda xatolarga yo'l qo'yadi. Xususan, Google Translate o'zbek tilidagi kelishik qo'shimchalarini ingliz tiliga tarjima qilishda 15–20% xato ko'rsatadi (Mirzayev va boshq., 2023). Yandex.Translate esa kirildan lotinga konversiyada ba'zan xatoga yo'l qo'yadiki, bu o'zak muammo hisoblanadi.

4. O'ZBEK TILINING AGGLUTINATIV XUSUSIYATLARI

Agglutinatsiya — so'z o'zagiga ketma-ket qo'shimchalar (affiks) qo'shish orqali yangi so'z shakllari hosil qilish usuli bo'lib, turkiy tillarga xos asosiy morfologik xususiyat hisoblanadi. Ingliz, rus yoki fransuz tillariga nisbatan o'zbek tilida bitta so'zning o'n va undan ortiq grammatik shakli bo'lishi mumkin.

O'zbek tilining mashina tarjimasi uchun muhim lingvistik xususiyatlari:

- Morfologik boylik — "o'qituvchilarimizning" so'zida: o'qit (o'zak) + uvchi (kasb yasovchi) + lar (ko'plik) + imiz (egalik, 1-shaxs ko'plik) + ning (qaratqich kelishigi) = 5 morfema bitta so'zda. Inglizda bu "of our teachers" — 3 ta alohida so'z bilan ifodalanadi;
- Erkin so'z tartibi — o'zbek tilida SOV (Ega–To'ldiruvchi–Kesim) asosiy tartib bo'lib, gapning ma'nosi o'zgarmasdan so'zlar joyi almashishi mumkin. Bu tarjima modellarining so'z tartibini o'rganishini murakkablashtiradi;
- Kesimning oxirida kelishi — o'zbek tilida kesim (fe'l) doim gapning oxirida turadi, holbuki ingliz tilida ega va kesim qo'shni bo'ladi. Encoder-decoder modellari bu uzoq masofali bog'likni o'rganishda qiyinchilikka duch keladi;

- Predikativ qo‘shimchalar — egalik, zamon, shaxs va son qo‘shimchalari bitta fe‘lga qo‘shiladi: "yozayotibmiz" = yoz + ayot + ib + miz (hozirgi zamon, 1-shaxs, ko‘plik);
- Ikki alifbo — lotin (1993-yildan rasmiy) va kiril yozuvi parallel qo‘llanilishi; tarjima tizimi har ikki variantni tanib olishi shart, aks holda tarjima butunlay xato bo‘lib qolishi mumkin;
- Qo‘shimchalar alomorfizmi — bir qo‘shimchani fonetik muhitga qarab turlicha yozilishi: ko‘plik qo‘shimchasi -lar/-ler, -da/-de/-ta/-te kabi kelishik variantlari;
- Postpozitsiyalar — o‘zbek tilida predloglar emas, postpozitsiyalar ishlatiladi: "uy oldida" (inglizcha "in front of the house"); bu tarjimada grammatik rol belgilashni qiyinlashtiradi.

Qiyosiy misol: "Men universitetimizda o‘reanayapman" jumlasini tarjima qilganda Google Translate "I am learning at our university" deb to‘g‘ri tarjima qiladi. Lekin "Universitetimizda o‘reanayotgan talabalarning referat topshirish muddati uzaytirildi" kabi murakkab gaplarda tarjima sifati keskin pasayadi, chunki bir nechta morfologik zanjir va sintaktik bog‘liklar bir vaqtda hal etilishi kerak bo‘lib qoladi.

5. TARJIMA JARAYONIDAGI MUAMMOLAR

O‘zbek tilining mashina tarjimasida uchraydigan muammolarni bir necha asosiy toifaga ajratish mumkin. Ushbu muammolar nafaqat til strukturasidan, balki resurslar etishmasligidan ham kelib chiqadi.

Morfologik darajadagi muammolar:

- Qo‘shimchalarni noto‘g‘ri segmentlash — tarjima modeli "kitoblar" so‘zini ba‘zan "kitob" + "lar" emas, balki bitta leksema sifatida ko‘rishi va uni lug‘atda topa olmasligi;
- Noaniq kelishik — o‘zbek tilida bosh kelishik va tushum kelishigi ba‘zan shakllanmasdan qoladi, tarjima modeli ularni farqlay olmasligi mumkin;
- Fe‘lning murakkab zamon shakllari — "yozayotgan edi", "yozib bo‘lgan edi" kabi analitik shakllarni ingliz zamonlariga to‘g‘ri moslashtirish qiyin.

Leksik-semantik darajadagi muammolar:

- Ko‘p ma‘onolilik — "qo‘l" so‘zi qo‘l (raqam), qo‘l (a‘yzo) ma‘onosida kelishi mumkin; kontekstsiz tarjima xato bo‘lib chiqadi;
- Ekvivalentsiz leksika — o‘zbek tilidagi "hashar", "gap" (to‘plashuv ma‘nosida), "mahalla" kabi so‘zlarning boshqa tillarda to‘liq ekvivalenti yo‘q;
- Frazologizmlar va maqollar — "Ko‘zining oqi-qarasi" iborasini to‘g‘ridan-to‘g‘ri tarjima qilish mazmunsiz natija beradi;
- Terminologiya nostandartligi — o‘zbek tilidagi ilmiy-texnik terminlar hali to‘liq standartlashtirilmagan, bir tushuncha uchun bir necha variant mavjud.

Sintaktik darajadagi muammolar:

- Uzoq masofali bogʻliklik — oʻzbek tilida ega va kesim oʻrtasida koʻplab iboralar joylashishi mumkin; transformer modeli bu uzoqlikni doim toʻgʻri qayta tuzavermaydi;

- Egasiz gaplar — oʻzbek tilida ega tushib qolishi odatiy hol ("Ketdi" — Kim ketdi? — feʻldan maʼlum boʻladi); inglizcha tarjimada ega qoʻshilishi shart;

- Sifatdosh va ravishdosh aylanmalari — oʻzbek tilida keng qoʻllaniladigan ushbu shakllar ingliz tiliga tarjimada koʻpincha ergash gap bilan beriladi va tarjima hajmi oʻzgaradi.

Resurs va texnik muammolar:

- Parallel korpuslar kamligi — oʻzbek-ingliz juftligi uchun sifatli parallel matnlar hajmi 10 million jumladan oshmaydi; bunga nisbatan nemis-ingliz juftligi uchun 4 milliarddan ortiq jumla mavjud;

- Ikki alifbo muammosi — bir xil matn lotin yoki kirildan kiritilishi mumkin; bu esa lugʻat hajmini ikki barobar oshirib, modelni chalkashtirib yuboradi;

- Dialekt va lahja farqlari — tarjima modeli asosan adabiy til asosida oʻreatiladi, xalq tili va lahjadagi soʻzlarni taniy olmaydi.

6. NEYRON TARMOQLAR ASOSIDAGI TARJIMA MODELLARI

Neyron mashina tarjimasi (Neural Machine Translation — NMT) — chuqur oʻreanish texnologiyasi yordamida parallel matnlardagi naqshlarni oʻreanib, mazmunni boshqa tilda qayta ifodalash usuli. NMT 2014-yildan boshlab statistik tarjimani siqib chiqara boshladi va hozirda sohada mutlaq ustunlik qiladi.

NMT arxitekturalarining asosiy turlari:

- Encoder-Decoder (RNN asosida, 2014–2017) — manba jumlani vektorga kodlaydi (encoder), soʻngra maqsad tilda dekodlaydi (decoder); LSTM va GRU qatlamlaridan foydalanadi. Uzoq jumalalarda "unutish" muammosi mavjud edi;

- Attention mexanizmi (Bahdanau, 2015) — dekoder har bir chiqish soʻzini yaratishda manba jumlaning barcha pozitsiyalariga eʼtibor qaratishi mumkin; bu uzoq masofali bogʻliklarni oʻreanishni yaxshiladi;

- Transformer arxitekturasini (Vaswani va boshq., 2017) — toʻliq attention asosida qurilgan, parallel hisoblash imkonini beradi; BERT, GPT, T5 kabi modellarning asosi; bugungi kunda barcha yirik tarjima tizimlari ushbu arxitekturalardan foydalanadi;

- Katta til modellari (GPT-4, Claude, Gemini, 2022–2025) — trillion parametrlilik modellar; maxsus tarjima oʻreatishsiz ham (zero-shot) yuqori sifatli tarjima bera oladi; kontekstni uzoq saqlash xususiyati mavjud.

Oʻzbek tili uchun ishlab chiqilgan NMT modellari:

- UzTranslate — Oʻzbekistonda ishlab chiqilgan, oʻzbek-ingliz va oʻzbek-rus yoʻnalishlari uchun maxsus oʻreatilgan NMT modeli; 5 million jumlati parallel korpusda oʻreatilgan;

- NLLB-200 (Meta AI, 2022) — 200 tilni qo'llab-quvvatlovchi; o'zbek tili uchun FLORES-200 sinovida BLEU ko'rsatkichi 28–32 ga teng; past resursli tillar uchun maxsus optimallashtirish qo'llangan;

- mBART-50 (Facebook AI) — 50 tilli ko'p tilli model; o'zbek tilini lotin alifbosida qo'llab-quvvatlaydi; fine-tuning orqali muayyan domen uchun yanada yaxshilanadi;

- OpusMT — Helsinki NLP guruhining ochiq manbali modellari; o'zbek-ingliz juftligi uchun samarali; Hugging Face platformasida erkin yuklab olish mumkin.

Transfer o'reanish (Transfer Learning) yondashuvi o'zbek tili uchun alohida muhimdir. Ko'p ma'lumotli tillar (masalan, turk, qozoq, qirg'iz) bilan birga o'reatilgan modelni o'zbek tiliga moslashtirish (fine-tune) orqali parallel korpus hajmini 10 baravar kamaytirgan holda ham qabul qilingan sifatga erishish mumkin. Bu yondashuv Flores va boshqa benchmarklarda tasdiqlangan.

O'zbek tilining morfologik boyligini hisobga olish uchun maxsus usullar taklif etilmoqda: morfemik tokenizatsiya (so'zni qo'shimchalarga bo'lib tokenizatsiya qilish), faktorlangan tarjima modellari (so'zning morfologik belgilari alohida o'reatiladi), va morfologik generatsiya moduli (tarjimadan keyin maqsad tildagi to'g'ri morfologik shaklni tiklash). Ushbu usullarning kombinatsiyasi BLEU ko'rsatkichini 3–7 punktga oshirishi mumkin.

7. TARJIMA SIFATINI OSHIRISH BO'YICHA TAKLIFLAR

O'zbek tilidagi mashina tarjimasi sifatini oshirish uchun texnologik, lingvistik va tashkiliy darajadagi kompleks chora-tadbirlar zarur. Quyida muallif tomonidan ishlab chiqilgan va adabiyotlar tahlili asosida umumlashtirilgan takliflar keltiriladi.

1. Parallel korpuslarni kengaytirish:

- Davlat hujjatlari, qonunchilik bazasi, Oliy Majlis materiallari, Jahon banki va BMT hujjatlarining o'zbek tilidagi versiyalarini parallel korpusga kiritish — bu parallel matn hajmini 50 milliondan ortiq jumlagacha yetkazish imkonini beradi;

- Veb-krauling orqali parallel matnlar to'plash — ikki tilli veb-saytlardan (hukumat, xalqaro tashkilotlar) parallel matnlarni avtomatik yig'ish;

- Inson tarjimasi annotatsiyasi — professional tarjimonlar tomonidan tayyorlangan yuqori sifatli kichik korpus (100–500 ming jumla) neyron tarjimini "yo'lga solish" uchun muhim rol o'ynaydi.

2. Morfologik tahlilni tarjima konveyeriga qo'shish:

- Tarjimadan oldin morfologik normalizatsiya — so'zlarni o'zak shakliga keltirish (lemmatizatsiya) modelni kichikroq lug'at bilan ishlatish imkonini beradi;

- Morfemik tokenizatsiya — BPE o'rniga morfema chegarasidan xabardor bo'gan tokenizatsiya (linguistically motivated tokenization) BLEU ko'rsatkichini 4–6 punktga oshirishi aniqlangan;

- Morfologik belgilar (POS-tagging) ni encoder kirishiga qo'shish — model so'zning grammatik rolini yanada aniqroq o'rganadi.

3. Ikki alifbo muammosini hal etish:

- Tarjimadan oldin kiril-lotin konverter moduli — yagona yozuv standartiga o'tkazish lug'at hajmini ikki barobar kamaytiradi va tarjima aniqligini oshiradi;
- Alifbo-invariant kodlash — modelni ikki alifboda ham ishlashga o'reatish; o'zak birligini alifbodan mustaqil ifodalash.

4. Ko'p tilli transfer o'rganish:

- Turkiy tillar guruhida birgalikda o'reatish — turk, qozoq, ozarbayjon, qirg'iz tillarining katta korpuslari o'zbek tilidagi modelni qo'shimcha o'rgatib beradi;
- Cross-lingual transfer — inglizcha o'reatilgan modelni o'zbek tiliga moslashtirish (zero-shot cross-lingual transfer) parallel korpus talabini kamaytiradi.

5. Domen moslashuvi (Domain Adaptation):

- Tibbiy, huquqiy, texnik va ta'lim domenlarida alohida fine-tuned modellar yaratish — umumiy tarjima modeli har doim ham ixtisoslashgan terminologiyani to'g'ri tarjima qilavermaydi;

- O'zbek tilidagi terminologik lug'atlar yaratish va ularni tarjima tizimiga integratsiya qilish — bu terminlar tarjimada aniqlikni keskin oshiradi.

6. Tarjima sifatini baholash tizimini takomillashtirish:

- O'zbek tili uchun maxsus BLEU va COMET modellarini o'reatish — mavjud mezonlar ko'pincha morfologik boylikni hisobga olmaydi;
- Professional tarjimonlar bilan hamkorlikda inson tomonidan baholash protokollarini ishlab chiqish va muntazam sinovlar o'tkazish.

8. XULOSA

Mashina tarjimasi tizimlarida o'zbek tilining lingvistik xususiyatlarini hisobga olish masalasi ko'p qirrali bo'lib, texnologik, lingvistik va resurs muammolarini bir vaqtda hal etishni talab qiladi. Ushbu tadqiqot quyidagi muhim xulosalarni shakllantiradi:

- O'zbek tilining agglutinativ morfologiyasi, erkin so'z tartibi va ikki alifbo tizimi mavjud tarjima tizimlariga jiddiy qiyinchiliklar tug'dirib, o'zbek tilini "murakkab" til toifasiga kiritadi;

- Google Translate, Yandex.Translate va NLLB-200 kabi tizimlar o'zbek tilida ishlash imkoniga ega bo'lsa-da, morfologik murakkab konstruktsiyalar, frazeologizmlar va domen terminologiyasida xato ko'rsatkichi yuqori bo'lib qolmoqda;

- Neyron tarjima modellari, ayniqsa transformer arxitekturasini, o'zbek tilini o'rganishda yangi imkoniyatlar yaratmoqda; transfer o'rganish yondashuvi past resursli holatlarda samarali ekanligini tasdiqlamoqda;

- Parallel korpuslarni ko‘paytirish, morfemik tokenizatsiya, kiril-lotin konversiyasi va domen moslashuvi — tarjima sifatini oshirishning asosiy yo‘llari sifatida aniqlangan;
- O‘zbekiston oliy ta‘lim muassasalari, xususan Urganch Davlat universiteti, ushbu muammolarni hal etishda ilmiy-amaliy tadqiqotlar olib borishi milliy NLP ekotizimini mustahkamlashda muhim ahamiyat kasb etadi.

Kelajakda o‘zbek tilida inson darajasidagi tarjima sifatiga erishish real maqsad bo‘lib, buning uchun davlat, akademiya va texnologiya sektori o‘rtasidagi tizimli hamkorlik zarur. Ushbu maqolada taklif etilgan usullar va yondashuvlar o‘zbek tilidagi mashina tarjimasi sifatini yangi bosqichga ko‘tarishga xizmat qiladi deb umid qilinadi.

FOYDALANILGAN ADABIYOTLAR

1. Vaswani, A., Shazeer, N., Parmar, N. et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
2. Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *ICLR 2015*.
3. Sennrich, R., Haddow, B., & Birch, A. (2016). Neural machine translation of rare words with subword units. *ACL 2016*, 1715–1725.
4. Costa-jussà, M. R. et al. (2022). No Language Left Behind: Scaling Human-Centered Machine Translation. *arXiv:2207.04672*.
5. Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: a method for automatic evaluation of machine translation. *ACL 2002*, 311–318.
6. Mirzayev, F. va boshqalar. (2023). O‘zbek-ingliz mashina tarjimasidagi morfologik muammolar tahlili. *O‘zMU axborotnomasi*, 4, 112–125.
7. Koehn, P. (2020). *Neural machine translation*. Cambridge University Press.
8. Bekchanov, H. M. (2024). O‘zbek tilidagi parallel korpuslarni shakllantirish metodikasi. *Urganch Davlat universiteti ilmiy jurnali*, 2(2), 34–45.
9. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL-HLT 2019*.
10. O‘zbekiston Respublikasi Raqamli texnologiyalar vazirligi. (2023). *Sun‘iy intellektni rivojlantirish milliy strategiyasi*. Toshkent.
11. Tiedemann, J. (2012). Parallel Data, Tools and Interfaces in OPUS. *LREC 2012*, 2214–2218.
12. Forcada, M. L., & Ginestí-Rosell, M. (2011). Apertium: a free/open-source platform for rule-based machine translation. *Machine Translation*, 25(2), 127–144.