

AVTOMATIK TARJIMA MASALALARI

Mirzakamolova Madina

Andijon davlat chet tillari instituti magistranti

mirzakamolovam@gmail.com

Ibragimova Gulchehra

Andijon davlat chet tillari institutida f.f.f. (PhD), dotsent

iguli2496@gmail.com

Annotatsiya. Zamonaviy tilshunoslikda avtomatik tarjima texnologiyalari muhim o‘rin tutmoqda. Ushbu maqolada biz avtomatik tarjima tizimlarining rivojlanish tarixi, hozirgi holati va kelajakdagi istiqbollarini tahlil qilamiz. Maqolada neyron tarmoqlarga asoslangan tarjima modellari (masalan, Transformer arxitekturasini) va ularning til kontekstini tushunishdagi muvaffaqiyatlari alohida e‘tibor markazida bo‘ladi. Shuningdek, avtomatik tarjima tizimlarining cheklovlari, xususan, madaniy nuanslar va kam resursli tillarni qayta ishlashdagi muammolar ko‘rib chiqiladi. Ushbu maqola tilshunoslar, sun‘iy intellekt tadqiqotchilari va texnologiya ishlab chiqaruvchilari uchun avtomatik tarjima sohasidagi muammolar va imkoniyatlarni tushunishda muhim manba bo‘lib xizmat qiladi.

Kalit so‘zlar: avtomatik tarjima, sun‘iy intellekt, neyron tarmoqlar, Transformer, tilshunoslik, sifat baholash, machine translation (MT).

Ayni paytda zamonaviy dunyoda texnologik olam jadallik bilan rivojlanmoqda va bu bilan tillararo muloqotning ahamiyati ortishiga olib kelmoqda. Jamiyatning turli sohalarida, jumladan, ilmiy tadqiqotlar, biznes, ta‘lim va madaniy almashinuvda ko‘p tilli muhitda samarali kommunikatsiya qilish zarurati avtomatik tarjima (machine translation-MT) ni o‘rni muhim ahamiyatga ega. Mashina tarjimasi yoki avtomatik tarjima-inson ishtirokisiz muayyan so‘z yoki tayyor matnlar yoki nutqning kompyuter

dasturlari yordamida bir tildan boshqa tilga qilingan tarjima jarayonidir¹. Hozirda kompyuter lingvistikasining asosiy yo'nalishlaridan biri avtomatik tarjima bo'lib, avtomatik tarjima tizimi prinsip va metodlari haqidagi bilimlarni hosil qilish, tilshunoslik ta'minotini yaratish, mavjud tizimlar va lug'at yaratish tarixi, til haqidagi bilimlar, ko'nikmalarni shakllantirish, mashina tarjimasi tizimining lingvistik tahlilini algoritmlash tizimini amalga oshiradi. Ushbu texnologiya so'nggi yillarda sun'iy intellect (AI) va neyron tarmoqlar (neural networks) yutuqlari tufayli keskin rivojlanmoqda, biroq hali ham bir qator metodologik, lingvistik va texnik muammolarga ega.

Avtomatik tarjima tizimlari global miqyosda informatsiya almashinuvini osonlashtirish, tillararo tushunishni yaxshilash va turli madaniyatlar o'rtasidagi bog'lanishni mustahkamlashda muhim rol o'ynaydi. Xususan, ilmiy hamjamiyatda xorijiy tadqiqotlarni tezda o'qish va tushunish imkoniyatini beradi. Hamda ta'lim sohasida chet tillarini o'rganishda yordamchi vosita sifatida va turizm va madaniy almashinuvda sayyohlar va mahalliy aholi o'rtasidagi kommunikatsiyani yaxshilashga xizmat qiladi. Aytish joizki, jamiyatning har sohasida avtomatik tarjimadan faol foydalanmoqdamiz. Lekin avtomatik tarjima tizimlari tarixan bir necha bosqichlardan o'tganini kuzatishimiz mumkin. Dastlabki tizimlar lug'at va Grammatik qoidalariga asoslangan (Rule-Based Machine Translation-RBMT) qoida bo'yicha shug'ullangan. Bu usul aniq grammatik tuzilmalarni talab qiladi, 1 ekin murakkab iboralarda samaradorligi pastligini aytib o'tishimiz lozim. Katta hajmdagi parallel korpuslar asosida statistika yordamida tarjima qiluvchi (Statistical Machine Translation-SMT)-Statistikdir. Google translate dasturining dastlabki versiyalari shu usuldan foydalangan. Hozirda esa eng ilg'or tizimlar (DeepL, Google NMT)-Neyron (Neural Machine Translation-NMT) Sun'iy neyron tarmoqlar yordamida kontekstni hisobga

¹ Martínez Lorena Guerra. Human Translation versus Machine Translation and Full Post-Editing of Raw Machine Translation Output. Dissertatin of MA. -Dublin, 2003.

olgan holda yuqori aniqlikdagi tarjimalardir². Bulardan DeepL yevropa tillarida yuqori aniqlik bilan ishlaydi.

Zamonaviy avtomatik tarjima tizimlari bugungi kunda Google translate, DeepL, Yandex translate kabi tizimlar foydalanuvchilar tomonidan keng qo'llanilmoqda. Ayniqsa, transformer modellariga asoslangan DeepL tarjima sifatida yuqori baholanmoqda. Ushbu tizimlar kontekstni hisobga olib, tildagi ifoda shakllarini aniqroq tarjima qila oladi.

Muhokama va natijalar. Avtomatik tarjimaning o'ziga hos afzalliklari va ba'zi bir kamchiliklarni kuzatishimiz mumkin. Aytish joizki avtomatik tarjimaning asosiy afzalligi - tezkorlik va ko'p tilda ishlash imkoniyatidir. U nafaqat kundalik muloqotda, balki ta'lim, ilmiy tadqiqotlar, diplomatik aloqalarda, texnik hujjat tarjimalarida ham muhim ahamiyatga egadir. O'zbekistonda ham davlat tilining rivoji bilan birga AT tizimlarining milliy variantlarini ishlab chiqishga ehtiyoj ortib bormoqda. Afzalliklariga qaramay uning ba'zi grammatik murakkabliklar, metafora va madaniy kontekstlarni to'liq anglay olmasligi kabi cheklovlar mavjud. Ayniqsa, o'zbek tilidagi so'zlar ko'p ma'nolilik kasb etgani uchun tarjimada aniqlik yo'qolishi mumkin.

Transformer tarmoqlari o'z-o'ziga e'tibor (self-attention) mexanizmi orqali so'zlar o'rtasidagi bog'liklarni aniqlaydi. Misol qilib olishimiz mumkinki, "bank" so'zi "daryo sohili" va "moliyaviy bank" kontekstida har xil tarjima qilinadi³. NMT tizimlari kirish va chiqish matnlarini to'g'ridan to'g'ri xaritalashni o'rganadi. Bu yondashuv SMT ga nisbatan kam xatoliklar va yaxshi oqimga ega. Bundan tashqari yana madaniy va kontekstual nuanslarni kuzatishimiz mumkin. Masalan, rus tilidagi "рука" so'zi ingliz tiliga "hand"(qo'l) yoki "arm"(qo'l/oyoq)tarzida tarjima qilinishi mumkin. AT tizimlari kontekstni noto'g'ri talqin qilsa, ma'no buzilishi yuzaga keladi. Avtomatik

² Нелюбин Л.Л. Компьютерная лингвистика и машинный перевод. - М.: ВЦП, 1991.

³. Абдурахмонова Н. Машина таржимаси. -Тошкент,2008

tarjima tizimlari, resursli tillarda (masalan, katta matnlar korpusi, parallel ma'lumotlar, lingvistik vositalar mavjud bo'lgan tillar) yuqori sifatli tarjima qilish uchun katta ma'lumotlar bazasiga tayanadi⁴. Ammo, resursi kam tillar (masalan, kichik jamoatchilikka ega tillar yoki raqamli resurslar ,kam rivijlangan tillar) uchun bu jarayon qiyinroq kechadi. Bunda avtomatik tarjima bir nechta usullarda ishlashi mumkin. Misol tariqasida Transfer Learning usulini olsak, u ko'p tilli modellar (masalan, Google Translate, Facebook/Meta ning M2M-100, OpenAI modellari) bir nechta tillarni o'rganadi. Kam resursli tillar uchun yuqori resursli tillarning ma'lumotlari va grammatik qoidalaridan foydalaniladi. Bu shuni anglatadiki agar tilda parallel korpus kam bo'lsa, ingliz yoki rus tilidagi ma'lumotlar orqali o'qitishni boshlashi mumkin. Yana bir keng tarqalgan yo'llaridan biri bu bir modelni bir nechta tillarga moslashtiriladi. Masalan, BERT, mT5, XLM-R kabi modellar 100 dan ortiq tillarni o'z ichiga oladi⁵. Ushbu modellar tillar o'rtasidagi umumiy lingvistik xususiyatlarini o'rganadi, bu esa kam resursli tillar uchun tarjimani yaxshilaydi. Qo'shni tillardan foydalanib tarjima qilish ham mumkin. Bunda til oilalariga asoslanib (masalan, turkiy tillar, hind-evropa tillari) tarjima qilishlik⁶. Agar maqsad tili boshqa yuqori resursli tilga yaqin bo'lsa, ularning lingvistik o'shashliklaridan foydalaniladi. Pivot til usuli esa to'g'ridan-to'g'ri parallel korpus yo'q bo'lganda, uchinchi til orqali tarjima qiladi⁷. Bunda masalan, o'zbekchadan qirg'izchaga tarjima uchun avval rus tiliga, so'ngra qirg'iz tiliga tarjima qilishi mumkin. Bu uning samarali ishlashiga ta'sir qilishi mumkin. Bundan tashqari ko'plab kam resursli tillar (masalan, turkey, fin-ugor yoki afrika tillari) morfologik jihatdan murakkab bo'lib, so'z shakllari ko'p variantda o'zgaradi. Bu modellarning so'zlarni to'g'ri kontekstda tarjima qilishni qiyinlashtiradi. O'zbek tilidagi fe'l shakllari (kelinglar, keladilar, kelganlar) tarjimada

⁴ Пўлатов А., Мухамедова С. Компьютер лингвистикаси (ўқув3.қўлланма). –Тошкент, 2009

⁵ Пулатов А.К. Тексты лекций по математической и компьютерной лингвистике (электронный вариант)

⁶ Уринбоев Д.Б. Компьютер лингвистикаси (Маърузалар матни).

⁷ Раҳимов А. Компьютер лингвистикаси асослари. Тошкент, Академнашр, 2011

aniq va murakkab talqinni talab qiladi⁸. Shuning uchun ham umumiy modellar (Google Translate, DeepL) sifatli ishlamasligi mumkin, chunki ular asosan yuqori resursli tillar (ingliz, xitoy) uchun optimizatsiya qilingan⁹.

Yuqorida sanab o'tilgan muammolarning yechimlarini taklif etmoqchiman. Birinchidan, yuqori resursli tillardan o'rgangan modellarni kam resursli tillarga moslashtirish, ma'lumotlarni ko'paytirish usullari yordamida sun'iy parallael korpuslar yaratish, vikipediya, kitoblar yoki ijtimoiy tarmoqlardan ma'lumotlar yig'ish. Bunda maxsus modellar: Meta AI'ning NLL (No Language Left Behind) kabi loyihalari kam resursli tillar uchun tarjima sifatini oshirishga xizmat qiladi. Endi aynan o'zbek tiliga to'htaladigan bo'lsak, o'zbek tilining aglutinativ tabiati tufayli avtomatik tarjima tizimlari morfologik tahlilda qiyinchiliklarga duch kelmoqda. Masalan, "uyg'onmoq" so'zining turli shakllari (uyg'ondi, uyg'onypati) kontekstga qarab turlicha tarjima tlab qiladi. Bu muammoni hal qilish uchun tizimlarga o'zbek tiliga xos korpuslar integratsiyasi zarur.

Kelajakda avtomatik tarjima tizimlarini rivojlantirish uchun quyidagi yo'nalishlar tavsiya etiladi:

1. O'zbek tiliga xos katta hajmdagi parallel korpuslarni yaratish.
2. Transformator modellarini (transformers) o'zbek tilining morfologik xususiyatlariga moslashtirish.
3. Insoniy fikr-mulohazalarni (human feedback) tizimni o'qitishda ko'proq qo'llash.

Kelajakda avtomat tarjima tizimlari yanada takomillashib, kontekstni tushunish va madaniy adptatsiya qobiliyatlari oshadi. Multimodal tarjima, ya'ni matn, tovush va tasvirlarni birlashtirgan tizimlar rivojlanishi kutilmoqda. Shuningdek, kam resursga

⁸ Абдурахмонова Н. Инглизча матнларни узбек тилига таржима қилиш дастурининг лингвистик таъминоти. Фило. фан... (PhD) дисс. - Тошкент, 2018..

⁹ Абдурахмонова Н. Машина таржимаси. -Тошкент, 2008

ega tillar uchun ma'lumotlar bazasini kengaytirish va AI'ning etik muammolari (masalan , tarjimadagi hato yoki noaniqliklar)ustida ishlash muhim vazifalardan biridir. Uning takomillashuvi uchun lingvistik bilimlar bilan texnologik yondashuvni uyg'unlashtirish zarur.

Xulosa. Avtomat tarjima texnologiyalari tillar o'rtasidagi to'siqlarni bartaraf etishda muhim o'rin tutadi. Qoidaga asoslangan tizimlardan nevrал tarjimaga o'tish tarjima sifatini sezilarli oshirdi, ammo madaniy va kontekstual muammolar hali ham davom etmoqda. Keljakda multimodal va inklyuziv tizimlarning rivojlanishi avtomat tarjimani yanada universal vositaga aylantiradi.

Foydalanilgan adabiyotlar

1. Абдурахмонова Н. Машина таржимаси. -Тошкент, 2008
2. Абдурахмонова Н. Инглизча матнларни узбек тилига таржима қилиш дастурининг лингвистик таъминоти. Фило. фан... (PhD) дисс. - Тошкент, 2018.
- 3.Martínez Lorena Guerra. Human Translation versus Machine Translation and Full Post-Editing of Raw Machine Translation Output. Dissertatin of MA. -Dublin, 2003.
4. Нелюбин Л.ЛТ. Компьютерная лингвистика и машинный перевод. - М.: ВЦП, 1991.
5. Нурмонов А., Йўлдошев Б. Тилшунослик ва табиий фанлар. –Тошкент: Ўқитувчи, 2001.
6. Пўлатов А., Мухамедова С. Компьютер лингвистикаси (ўқув.қўлланма). – Тошкент, 2009
7. Пулатов А.К. Тексты лекций по математической и компьютерной лингвистике (электронный вариант)
8. Пўлатов А. Компьютер лингвистикаси. –Тошкент, Академнашр, 2011

9. Рахимов А. Компьютер лингвистикаси асослари. Тошкент, Академнашр, 2011
10. Ibragimova, G. M. (2024). Study of the Parentheses by World Linguistics. Excellencia: International Multi-Disciplinary Journal of Education (2994-9521), 2(3), 5-7. <https://doi.org/10.5281/>