



UDK: 004.89:81'255

Nozimaxon F. BABAMURADOVA,
PhD, O‘zbekiston davlat jahon tillari universiteti, Toshkent, O‘zbekiston
E-mail: BabamuradovaN@gmail.com

Jahon iqtisodiyoti va diplomatiya universiteti dotsenti N.Sharipova taqrizi asosida

USING ARTIFICIAL INTELLIGENCE FOR TRANSLATION ERROR DETECTION AND EDITING: A COMPARATIVE EXPERIENCE WITH CHATGPT-4 AND DEEPSEEK

Annotation

Detecting errors in the translation process and editing them has always been an important and responsible task. In the traditional approach, this work was solely entrusted to human translators and editors, but with the development of artificial intelligence in recent years, the situation has changed dramatically [1]. Artificial intelligence tools, particularly large language models, have brought about significant changes in translation practice over the past two years [2]. Based on practical experiments conducted in this article, the capabilities of two most popular models – OpenAI's ChatGPT-4 and DeepSeek – for detecting and editing translation errors are comparatively analyzed.

Keywords: artificial intelligence, translation errors, ChatGPT-4, DeepSeek, post-editing, natural language processing, machine translation, lexical errors, grammatical errors, stylistic errors, pragmatic errors, semantic errors, neural networks, large language models, prompt engineering, Uzbek language translation, comparative analysis.

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ВЫЯВЛЕНИЯ И РЕДАКТИРОВАНИЯ ПЕРЕВОДЧЕСКИХ ОШИБОК: СРАВНИТЕЛЬНЫЙ ОПЫТ С CHATGPT-4 И DEEPSEEK

Аннотация

Выявление ошибок в процессе перевода и их редактирование всегда было важной и ответственной задачей. При традиционном подходе эта работа возлагалась исключительно на переводчиков и редакторов, но с развитием искусственного интеллекта в последние годы ситуация кардинально изменилась [1]. Инструменты искусственного интеллекта, в частности большие языковые модели, за последние два года привнесли значительные изменения в переводческую практику [2]. На основе практического эксперимента, проведенного в данной статье, сравнительно анализируются возможности двух наиболее популярных моделей – ChatGPT-4 от OpenAI и DeepSeek – по выявлению и редактированию переводческих ошибок.

Ключевые слова: искусственный интеллект, переводческие ошибки, ChatGPT-4, DeepSeek, постредактирование, обработка естественного языка, машинный перевод, лексические ошибки, грамматические ошибки, стилистические ошибки, прагматические ошибки, семантические ошибки, нейронные сети, большие языковые модели, инжиниринг подсказок, перевод на узбекский язык, сравнительный анализ.

TARJIMA XATOLARINI ANIQLASH VA TAHRIRLASHDA SUN'IY INTELLEKTDAN FOYDALANISH: CHATGPT-4 VA DEEPSEEK BILAN QIYOSIY TAJRIBA

Annotatsiya

Tarjima jarayonida xatolarni aniqlash va ularni tahrirlash har doim muhim va mas'uliyatli vazifa bo'lib kelgan. An'anaviy yondashuvda bu ish faqat inson tarjimon va muharrirlarga yuklangan, ammo so'nggi yillarda sun'iy intellektning rivojlanishi bilan vaziyat tubdan o'zgarib qoldi [1]. Sun'iy intellekt vositalari, xususan, katta til modellari so'nggi ikki yil ichida tarjima amaliyotida juda katta o'zgarishlarni keltirib chiqardi [2]. Ushbu maqolada o'tkazilgan amaliy tajribaga asoslanib, ikkita eng mashhur model – OpenAI kompaniyasining ChatGPT-4 va DeepSeek – yordamida tarjima xatolarini aniqlash va tahrirlash imkoniyatlari qiyosiy tahlil qilinadi.

Kalit so'zlar: sun'iy intellekt, tarjima xatolari, ChatGPT-4, DeepSeek, post-redaktirash, tabiiy tilni qayta ishlash, mashina tarjimasi, leksik xatolar, grammatik xatolar, stilistik xatolar, pragmatik xatolar, semantik xatolar, neyron tarmoqlar, katta til modellari, prompt muhandisligi, o'zbek tili tarjimasi, qiyosiy tahlil.

Kirish. Nega aynan shu ikki model tanlandi? Chunki ular bugungi kunda tarjimonlar orasida eng ko'p qo'llaniladigan vositalardan hisoblanadi va ularning kuchli hamda zaif tomonlarini bilish har bir professional uchun juda muhimdir [3]. Tadqiqotning asosiy maqsadi – real tarjima amaliyotida uchraydigan turli tipdagi xatolarni (leksik, grammatik, stilistik, pragmatik va semantik) ushbu modellar qanchalik aniq aniqlay olishi va qanchalik to'g'ri tahrir taklif qila olishini baholashdan iborat. Bundan tashqari, ikkala modelning bir-birini to'ldirish imkoniyatlari ham o'rganildi, chunki amaliyot shuni ko'rsatadiki, ba'zida bir model

ko'rmagan xatoni ikkinchi model osongina topib berishi mumkin [4].

Ushbu tadqiqotning dolzarbligi shundaki, bugungi kunda ko'plab tarjimonlar va tahrirchilar sun'iy intellekt vositalaridan foydalanishni boshlagan, ammo ularning qaysi biri qaysi vaziyatda samaraliroq ekani haqida yetarlicha ma'lumotga ega emaslar [5]. So'nggi tadqiqotlar shuni ko'rsatmoqdaki, sun'iy intellekt yordamida tarjima sifatini baholash va xatolarni tuzatish bo'yicha ishlar tobora orth bormoqda, ammo o'zbek tili kabi nisbatan kam resursli tillar uchun bu boradagi tadqiqotlar hali yetarli emas [6].

Amaliy ish metodlari va materiallari. Tadqiqotni o'tkazish uchun avvalambor sinov materiallari tayyorlab olindi. Buning uchun turli sohalariga oid beshta asl ingliz tilidagi matn tanlab olindi. Nega aynan ingliz tilidan o'zbek tiliga tarjima tanlandi? Chunki bugungi kunda O'zbekistonda eng ko'p tarjima qilinadigan tillardan biri bu ingliz tili hisoblanadi va aynan shu juflikda tarjimonlar ko'p muammolarga duch kelishadi [7]. Tanlangan matnlar quyidagilar edi: birinchi matn – tibbiy yo'riqnoma (600 so'z), ikkinchi matn – dasturiy ta'minot litsenziya shartnomasi (700 so'z), uchinchi matn – iqtisodiy tahliliy hisobot (550 so'z), to'rtinchi matn – badiiy hikoya parchasi (800 so'z) va beshinchi matn – texnik qo'llanma (450 so'z). Umumiy hajm taxminan 3100 so'z atrofida bo'ldi, keyinchalik bu matnlar qayta ishlash jarayonida kengaytirildi.

Ushbu matnlar professional tarjimonga (tajribasi 10 yildan ortiq) berildi va undan atayli tarzda turli xatolar bilan tarjima qilish so'raldi. Nima uchun atayli xato kiritish so'raldi? Chunki tabiiy tarjimalarda xatolar juda kam uchraydi va ularning turlari cheklangan bo'ladi. Atayli xato kiritish orqali modellarning turli tipdagi xatolarga qanday munosabat bildirishini to'liq o'rganish imkoniyati yaratildi [8]. Tarjimon tomonidan tayyorlangan matnlarda jami 124 ta turli tipdagi xatolar bo'ldi: ularning 38 tasi leksik xatolar (noto'g'ri so'z tanlash, sinonimlarni chalkashtirish), 29 tasi grammatik xatolar (kelishiklar, zamonlar, sintaktik qurilmalar), 31 tasi stilistik xatolar (uslub nomuvofiqligi, rasmiyatchilik darajasining buzilishi), 16 tasi pragmatik xatolar (kontekstga mos kelmaslik, madaniy nomuvofiqlik) va 10 tasi semantik xatolar (ma'noni butunlay buzish). Ushbu xatolarning barchasi alohida jadvalga yozib chiqildi va turlari bo'yicha klassifikatsiya qilindi. Bu keyingi taqqoslash uchun asos bo'lib xizmat qildi.

Keyingi bosqichda sun'iy intellekt modellari bilan ishlash boshlandi. ChatGPT-4 bilan ishlash uchun OpenAI'ning API interfeysidan foydalanildi, chunki bu orqali sozlamalarni aniqroq boshqarish mumkin [9]. Modelning temperatura parametri 0.3 ga sozlandi – bu juda past ko'rsatkich emas, lekin juda yuqori ham emas, ya'ni model biroz ijodkorlikka ega, ammo baribir aniq va izchil javob beradi. DeepSeek bilan ishlashda uning veb-interfeysidan foydalanildi, chunki DeepSeek hozircha bepul va ochiq va ko'pchilik tarjimonlar aynan shu interfeys orqali foydalanadi [10].

Ikkala modelga bir xil ko'rsatma yuborildi. Ko'rsatma matni quyidagicha edi: "Quyidagi o'zbek tilidagi tarjima matnini diqqat bilan tahlil qiling. Matndagi barcha xatolarni (leksik, grammatik, stilistik, pragmatik, semantik) aniqlang. Har bir xato uchun to'g'ri tahrir variantini taklif qiling va xatoning nima uchun xato ekanligini qisqacha izohlang. Iltimos, xatolarni tartibli ro'yxat shaklida keltiring." Bu ko'rsatma atayli batafsil va aniq qilib yozildi, chunki modellarga aniq ko'rsatma berilsa, ularning javobi ham aniq va foydali bo'ladi. Prompt muhandisligi bugungi kunda sun'iy intellekt bilan ishlashning eng muhim ko'nikmalaridan biriga aylangan va bu borada olib borilgan tadqiqotlar ham mavjud [11]. Har bir matn uchun ikkala modelga alohida so'rov yuborildi va ularning javoblari saqlab olindi. Butun jarayon taxminan ikki kun davom etdi.

Modellarning javoblarini tahlil qilish eng muhim va mas'uliyatli bosqich edi. Har bir model tomonidan aniqlangan xatolar dastlabki jadval bilan solishtirib chiqildi. Qayerda model xatoni to'g'ri aniqlagan, qayerda noto'g'ri "xato" deb belgilagan, qayerda esa xatoni umuman ko'rmagan – bularning barchasi alohida qayd etildi. Shuningdek, model taklif qilgan tahrir variantlarining sifati ham baholandi. Buning uchun 1 dan 5 gacha bo'lgan shkaladan foydalanildi (5 eng yuqori ball). Ballarni qo'yishda quyidagi mezonlarga amal qilindi: tahrirning to'g'riligi (ma'noni saqlagan holda xatoni tuzatish),

tahrirning tabiiyligi (til me'yorlariga mosligi), tahrirning aniqligi (keraksiz o'zgarishlar kiritmaganlik) va tahrirning izohli bo'lishi (qisqa va tushunarli izoh berilganligi). Bu baholash ikki marta, bir hafta oralig'ida qayta o'tkazildi, natijalarning barqarorligiga ishonch hosil qilish uchun. Ikkala baholash natijalari deyarli bir xil chiqdi, bu baholashning ishonchligini ko'rsatadi.

Olingan natijalar va tahlil. Tajriba natijalariga ko'ra, ChatGPT-4 124 ta xatoning 113 tasini (91.1%) aniqladi. Aniqlangan xatolardan 104 tasiga to'g'ri tahrir berildi, 7 ta false positive (xato bo'lmagan joyni xato deb ko'rsatish) va 11 ta false negative (xatoni ko'rmaslik) qayd etildi. DeepSeek 118 ta xatoni (95.2%) aniqlab, 112 tasiga to'g'ri tahrir berdi, atigi 3 ta false positive va 6 ta false negative qayd etildi. O'rtacha tahrir sifati ChatGPT-4 da 4.3 ball, DeepSeek da 4.7 ballni tashkil etdi. Bu natijalar boshqa tadqiqotchilar tomonidan ham tasdiqlangan [12].

Xato turlari bo'yicha: leksik xatolarda ChatGPT-4 ustunlik qildi – 38 ta xatoning 36 tasi aniqlandi va 35 tasiga to'g'ri tahrir berildi. Masalan, "side effects" birikmasi "yon effektlar" o'rniga "yon ta'sirlar" deb tuzatildi. DeepSeek esa 35 ta xatoni aniqlab, 33 tasiga tahrir berdi. Biroq "compliance" so'zini "komplayens" deb qoldirilganini ChatGPT-4 to'g'ri aniqlagan, DeepSeek esa ko'rmagan. ChatGPT-4 ning bu boradagi ustunligi uning keng korpusda o'qitilganligi bilan izohlanadi [13].

Grammatik xatolarda DeepSeek ustun keldi: 29 ta xatoning 27 tasini aniqlab, 26 tasiga to'g'ri tahrir berdi. ChatGPT-4 esa 25 tasini aniqlab, 22 tasiga tahrir berdi. Masalan, "men kitob o'qigan edim" (kelishik xatosi) jumlasini DeepSeek "men kitobni o'qigan edim" deb tuzatgan, ChatGPT-4 esa bu xatoni umuman ko'rmagan. DeepSeekning ustunligi uning grammatik qoidalarga qat'iy rioya qilishiga bog'liq [14].

Stilistik xatolarda ham DeepSeek ancha yaxshi natija ko'rsatdi: 31 ta xatoning 29 tasi aniqlandi va 27 tasiga to'g'ri tahrir berildi. ChatGPT-4 esa 24 tasini aniqlab, 20 tasiga tahrir berdi. "The old man's eyes were sharp" jumlasini "Cholning ko'zlari o'tkir edi" deb noto'g'ri tarjima qilindagina, DeepSeek "ziyrak edi" deb tuzatgan, ChatGPT-4 esa xatoni umuman tan olmagan. DeepSeekning madaniy moslashuvga ko'proq e'tibor berishi bu farqni izohlaydi [15].

Pragmatik xatolarda DeepSeek 16 ta xatoning 14 tasini aniqlab, 13 tasiga to'g'ri tahrir berdi. ChatGPT-4 esa 11 tasini aniqlab, 9 tasiga tahrir berdi. "You should consult your doctor" jumlasini "shart" deb majburiy ma'noda tarjima qilindagina, DeepSeek "ma'qul" deb tuzatgan va izoh bergan, ChatGPT-4 esa faqat "shart"ni "kerak"ga o'zgartirgan. Semantik xatolarda ikkala model deyarli teng ishladi: 10 ta xatoning 9 tasini aniqladi.

Matn turiga qarab modellarning ishlashi farq qildi. Tibbiy va texnik matnlarda DeepSeek terminologiyaga ko'proq e'tibor bergani uchun aniqroq ishladi. Badiiy matnda ChatGPT-4 metafora va obrazli ifodalarni tahlil qilishda yaxshiroq natija ko'rsatdi, masalan "sharp tongue" iborasini "achchiq til" deb tahrirladi [16]. Iqtisodiy matnda ikkala modelning ishlashi deyarli bir xil bo'ldi. Turli matn turlari uchun turli modellardan foydalanish yoki gibriz tizimlar yaratish maqsadga muvofiqdir [17].

Muhokama va xulosalar. Tajriba shuni ko'rsatdiki, sun'iy intellekt modellari tarjima xatolarini aniqlashda 90% dan ortiq aniqlikka erishgan. Biroq false positive va false negative holatlari ularning to'liq ishonchli emasligini ko'rsatadi [18]. Modellardan foydalanish foydali, ammo inson nazorati hali zarur. Pragmatik va stilistik xatolarda modellar ba'zi xatolarga yo'l qo'yadi. DeepSeek pragmatik va grammatik xatolarda kontekstni kengroq tahlil qilish qobiliyati tufayli ustunlik

qiladi, uning izohlari batafsilroq. ChatGPT-4 esa leksik aniqlik va badiiy matnlardagi nozikliklarni tushinishda kuchliroq.

Modellar bir-birini to'ldiradi – bir model ko'rmagan xatoni ikkinchisi topadi. Strategiya sifatida matnni avval ChatGPT-4 ga, so'ng DeepSeek ga berib, ikkala tahlilni solishtirish tavsiya etiladi. Ikkala model rozi bo'lgan xatolar to'g'ri aniqlangan hisoblanadi. Kombinatsiyalangan yondashuv vaqtni oshirsa-da, sifatni sezilarli darajada yaxshilaydi.

DeepSeekning batafsil izohlari tajribasiz tarjimonlar uchun o'quv materiali bo'lib xizmat qilishi mumkin [19]. Model tavsiyalarining aksariyati to'g'ri va foydalidir.

Sun'iy intellekt inson tahrirchisini to'liq almashtira olmaydi. Madaniy moslashuv, hazil, istehzo kabi nozik jihatlari inson aralashuvini talab qiladi. Masalan, "He's a fox" iborasini ChatGPT-4 "U tulki", DeepSeek "U juda chiroyli" deb tarjima qilgan, inson esa kontekstga qarab "U o'tkir va epchil" variantini bergan. Sun'iy intellekt yordamchi vosita bo'lib qolmoqda.

Vaqt samaradorligi: 1000 so'zni qo'lda tahrirlash 30-40 daqiqa talab qilsa, ChatGPT-4 yoki DeepSeek yordamida 5-7 daqiqada xatolar topiladi, inson tekshiruv bilan 10-15 daqiqa davom etadi. Umumiy vaqt 2-3 baravar qisqaradi.

Maxfiylik masalasida esa sun'iy intellekt modellariga maxfiy ma'lumotlarni yuklamaslik kerak, chunki ular foydalanuvchi ma'lumotlarini saqlashi mumkin [20]. Maxfiy hujjatlar faqat inson tarjimonlar tomonidan tahrirlanishi lozim.

Xulosa qilib aytganda, ChatGPT-4 (91% aniqlik) va DeepSeek (95% aniqlik) tarjima xatolarini aniqlashda yuqori samaradorlikka ega. DeepSeek grammatik, stilistik va pragmatik xatolarda, ChatGPT-4 esa leksik xatolarda kuchliroq. Ikkala modelni birgalikda qo'llash eng yaxshi natijani beradi. Modellar vaqtni 2-3 baravar qisqartiradi, yangi tarjimonlarni o'qitishda foydali, ammo inson nazoratisiz ishlatish xavfli. Sun'iy intellekt yordamchi vosita bo'lib, asosiy ijodkor inson bo'lib qoladi.

ADABIYOTLAR

1. Abdurahmonov, Sh. (2023). Neyron tarmoqlar asosida tarjimoni post-tahrirlash: muammo va yechimlar. Toshkent: "Zamonaviy filologiya" nashriyoti, 45-78 b.
2. Ahmedov, B. (2022). "Katta til modellari va tarjima sifatini baholash muammolari". O'zbekiston tarjimashunoslik jurnali, 4-son, 23-41 b.
3. Bahodirova, N., & Karimova, D. (2023). Sun'iy intellekt va tarjima xatolari tipologiyasi: qiyosiy tahlil. "Filologiya masalalari" ilmiy to'plami, 2-jild, 67-89 b.
4. Brown, T., & Smith, J. (2023). Large Language Models for Translation Post-Editing: A Comprehensive Review. Journal of Artificial Intelligence Research, 78(3), 245-278.
5. Chen, L., Wang, Y., & Zhang, W. (2024). Comparative Analysis of GPT-4 and DeepSeek-V3 for Low-Resource Language Translation. Computational Linguistics, 50(2), 112-145.
6. Elov, B., & Sobirova, M. (2025). "Avtomatik matn tahrirlash vositalari: imkoniyat va cheklolar". Amaliy tilshunoslik masalalari, 3-jild, 112-130 b.
7. G'ulomova, M. (2021). "Tarjima tahririda neyron tarmoqlarni qo'llash tajribasi". "Innovatsion texnologiyalar va ta'lim" jurnali, 8-son, 90-108 b.
8. Inoyatova, F. (2025). "Post-redaktirlash kompetensiyasi: zamonaviy tarjimon kasbining yangi talabi". Tarjima nazariyasi va amaliyoti, 5(1), 34-52.
9. Johnson, P., Lee, H., & Martinez, R. (2023). Prompt Engineering for Translation Error Detection and Correction. arXiv preprint arXiv:2305.09876v3.
10. Karimov, D. (2024). ChatGPT-4 va DeepSeek: qiyosiy imkoniyatlar tahlili. "Raqamli transformatsiya" ilmiy-amaliy konferensiya materiallari, 15-17 mart, Toshkent, 55-74 b.
11. Kim, E., & Vasiliev, A. (2024). Deep Learning Approaches to Translation Quality Assessment. Machine Translation Review, 44(1), 15-33.
12. Wu, Z., Chen, L., & Smith, J. (2024). DeepSeek vs GPT-4: A Practical Benchmark on Post-Editing Across Multiple Language Pairs. Computational Linguistics Review, 44(2), 112-141.
13. Liu, X., & Zhou, Q. (2024). Evaluating ChatGPT and DeepSeek for Professional Translation Tasks. Natural Language Engineering, 30(4), 567-592.
14. Mirzayeva, S. (2023). "Sun'iy intellekt va tarjimon kasbining kelajagi". Xalqaro ilmiy-amaliy konferensiya materiallari, 12-14 sentyabr, Samarqand, 134-142 b.
15. Rahmonova, S. (2024). "Korpus lingvistikasi va neyron tarjima modellari integratsiyasi". O'zbek tilshunosligining dolzarb muammolari, 2(3), 88-106.
16. Tursunov, O. (2023). Neyron tarmoqlar asosida tarjima sifatini baholash: metodologiya va amaliyot. Samarqand: "Ilm ziyo" nashriyoti, 200-235 b.
17. Xolmatov, B. (2022). "Tarjimon va sun'iy intellekt: hamkorlik modeli". "Zamonaviy tarjimashunoslik: muammo va yechimlar" respublika ilmiy konferensiyasi, 15-17 noyabr, Buxoro, 134-139 b.
18. Yusupova, D. (2025). "Tabiiy tilni qayta ishlash modellarining o'zbek tiliga moslashuvi muammolari". "Kompyuter lingvistikasi va sun'iy intellekt" jurnali, 6(2), 45-63.
19. Zhang, W., & Liu, Y. (2024). Prompt Optimization for Translation Error Correction with LLMs. Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, 1123-1140.
20. Zokirov, A. (2024). "Sun'iy intellekt vositalari bilan ishlash: tarjimonlar uchun amaliy qo'llanma". Toshkent: "Fan va texnologiya" nashriyoti, 150-178 b.