

AI VOSITALARI ORQALI GRAMMATIK XATOLARNI AVTOMATIK TAHLIL QILISHNING LINGVISTIK ASOSLARI

Raximova Shahnoza Yorqin qizi

O'zDJTU, Sharq filologiyasi fakulteti Turk filologiyasi 3-bosqich talabasi
rakhimova0109@gmail.com

Annotasiya. Mazkur maqolada sun'iy intellekt texnologiyalarining zamonaviy tilshunoslik, xususan, grammatik xatolarni avtomatik aniqlash va tuzatish (GED/GEC) sohasidagi o'rni tahlil qilinadi. Muallif tabiiy tilni qayta ishlash (NLP) jarayonining asosiy bosqichlari sifatida morfologik, sintaktik va semantik tahlil usullarini batafsil yoritib bergan. Ayniqsa, o'zbek tili kabi agglyutinativ tillarda morfema tahlilining o'ziga xos murakkabliklari va ularning SI tizimlari tomonidan qayta ishlanishi masalalariga alohida e'tibor qaratilgan. Maqolada an'anaviy usullardan farqli o'laroq, zamonaviy transformer arxitekturalari — BERT, GPT va T5 modellarining kontekstni anglashdagi ustunliklari ochib beriladi.

Kalit so'zlar: Sun'iy intellekt (SI); tabiiy tilni qayta ishlash (NLP); kompyuter lingvistikasi; morfologiya; sintaksis; pragmatika; agglyutinativ tillar; BERT.

Аннотация. В данной статье анализируется роль технологий искусственного интеллекта в современной лингвистике, в частности в области автоматического выявления и исправления грамматических ошибок (GED/GEC). Автор подробно освещает основные этапы обработки естественного языка (NLP), включая морфологический, синтаксический и семантический анализ. Особое внимание уделено специфическим трудностям морфемного анализа в агглютинативных языках, таких как узбекский, а также тому, как эти особенности обрабатываются системами на основе ИИ. В отличие от традиционных методов, в статье раскрываются преимущества современных трансформерных архитектур — BERT, GPT и T5 — в задаче понимания контекста.

Ключевые слова: Искусственный интеллект (ИИ); обработка естественного языка (NLP); компьютерная лингвистика; морфология; синтаксис; прагматика; агглютинативные языки; модель BERT.

Abstract. This article analyzes the role of artificial intelligence technologies in modern linguistics, particularly in the field of automatic grammatical error detection and correction (GED/GEC). The author provides a detailed explanation of the main stages of natural language processing (NLP), including morphological, syntactic, and semantic analysis. Special attention is given to the unique complexities of morphemic analysis in agglutinative languages such as Uzbek and the challenges they pose for AI-based systems. Unlike traditional approaches, the article highlights the advantages of modern transformer architectures—BERT, GPT, and T5—in understanding contextual meaning.

Keywords: Artificial intelligence (AI); natural language processing (NLP); computational linguistics; morphology; syntax; pragmatics; agglutinative languages; BERT model.

Kirish. Hozirgi globallashtirish jarayonida sun'iy intellekt texnologiyalari inson faoliyatining deyarli barcha sohalariga chuqur kirib borib, tilni qayta ishlash, ma'lumotni tahlil qilish va axborot almashinuvini kuchaytirishda keng foydalanilmoqda. Shuni ham ta'kidlash joizki, u inson faoliyatidagi ko'plab jarayonlarni avtomatlashtirish, optimallashtirish va tezkor hamda aniq qarorlar qabul qilish

imkoniyatini yaratib beruvchi eng muhim texnologiyalardan biri sifatida 8maydonga chiqmoqda. Shu bilan birga, SI modellarining rivojlanishi tilning murakkab tuzilmalarini o'rganish va kontekstual ma'nolarni tushunish imkonini beradi. Mazkur tadqiqot sun'iy intellektning tilshunoslikdagi o'rni va uning amaliy qo'llanilishi jihatlarini yoritishga qaratilgan.

Asosiy qism. Grammatik xatolarni avtomatik aniqlash va tuzatish (Automatic Grammar Error Detection and Correction – GED/GEC) - bugungi kunga kelib keng ko'lamda foydalanilayotgan yo'nalishlardan biri hisoblanadi. Ushbu texnologiya tabiiy tilni qayta ishlash (NLP - Natural Language Processing), kompyuter lingvistikasi, korpus lingvistikasi va mashinaviy o'qitish kabi bir qancha fanlarning o'zaro birlashuvidan shakllanadi. Ushbu jarayonning asosini tilning morfologik, sintaktik, semantik va pragmatik xususiyatlari tashkil qiladi. Zamonaviy SI modellari shunchaki xatoni topibgina qolmay, tilning semantik xususiyatlarini hisobga olib gapning mazmun - mohiyatidan kelib chiqqan holda tuzatish funksiyasini amalga oshirmoqda. Morfologiya grammatik xatolarni aniqlashning birinchi va eng muhim bosqichi hisoblanadi. Sun'iy intellekt ham aynan morfologiya qoidalariga tayangan holda xatolarni aniqlaydi. U har bir so'zning negizini (o'zagini), grammatik shakllarini, qo'shimchalarini, so'z turkumini ajratish orqali tahlil qiladi. Misol uchun, o'zbek tilidagi "borishmayotgandirlar" so'zini oladigan bo'lsak, SI bu kabi qo'shimchalari ketma-ket qo'shilgan so'zlarni birgina -dir (tahmin) yoki -ma (inkor) qo'shimchasini noto'g'ri tushunsa, butun so'zning ma'nosi o'zgarib ketishi mumkin. Shu sababli, SI uchun "Agglyutinativ" (so'z o'zagi yoki negiziga qo'shimchalar qo'shilishi natijasida yangi so'zlar hosil bo'luvchi tillar) bu biroz murakkabroq hisoblanadi va har bir belgi (morfema) ustida alohida ishlashi, mantiqiy tomonlama ham kelib chiqqan ma'noga alohida e'tibor qaratish lozim.

Grammatik xatolarni aniqlashning ikkinchi bosqichi sintaksis hisoblanadi. Sintaktik tahlil yordamida gap bo'laklari, ularning o'zaro bog'lanishi, gapda fe'lning o'rni, bosh va ergash gaplar, so'z tartibi kabi jihatlarni aniqlab, xatolar ustida ishlaydi. SI uchun eng qiyin bosqich bu semantik tahlildir. Bunga quyidagi qiyinchiliklarni sabab qilib keltirishimiz mumkin:

1. Ma'no mos kelmasligi (Semantic mismatch). Misol uchun: "Men kecha keldim" — "Men kecha kelaman". Bu yerda zamon qo'shimchasi noto'g'ri tarjima qilinganligi sababli ma'no buzilgan.

2. Ko'p ma'noli so'zlar - bitta so'zning bir nechta ma'nodoshlari bo'lishi mumkin va bu SI uchun tarjima asnosida qiyinchilik tug'dirishi mumkin. Misol uchun, "yuz"- son, "yuz" - yuz qismi. SI bu holatda konteks ma'nosidan kelib chiqib to'g'ri tarjimani qo'yishi lozim.

Bu qiyinchiliklar SI ustida yanada chuqurroq izlanish olib borishga, har bir tilning agglyutinativ bo'lishini hisobga olgan holda ichki algoritm yaratishga turtki bo'ladi.

Hozirgi kunda esa chuqur o'rganish va transformer arxitekturasi asosidagi modellar eng ilg'or yondashuv hisoblanadi. BERT (Devlin va boshq., 2018), RoBERTa, GPT (OpenAI), T5 kabi modellar til birliklarini kontekstda tushunish imkoniyatini berdi. Bu esa grammatik va semantik xatolarni aniqlashda yuqori samaradorlikka olib keldi. Turkiy tillar uchun ham bunday ishlanmalar mavjud. Masalan, turk tilida Zemberek ochiq manbali NLP platformasi ishlab chiqilgan, keyinchalik BERTurk modeli yaratilib, turk tilida chuqur o'rganishga asoslangan dasturlar uchun keng qo'llanilmoqda". (1, 310-b) Chuqur o'rganish (deep learning) va transformer

arxitekturasi asosidagi modellar tilshunoslikda yangi bosqich yaratilishida asos bo'lib xizmat qildi. BERT, RoBERTa, GPT va T5 kabi modellar nafaqat matnni tahlil qilish, balki uni kontekstual jihatdan tushunish imkonini yaratdi. An'anaviy NLP yondashuvlari ko'pincha qat'iy qoidalar bilan cheklangan bo'lib, ular murakkab til birliklari o'rtasidagi bog'lanishlarni aniqlashda ba'zi bir qiyinchiliklarga uchraydi. Transformerlar esa aksincha, matnning har bir so'zini atrofda boshqa so'zlar bilan o'zaro bog'lab tahlil qilgani sababli, tilning sintaktik, semantik hamda pragmatik xususiyatlarini chuqurroq aks ettira oladi. Natijada, ushbu modellar grammatik xatolarni aniqlash va tuzatishda avvalgi tizimlarga qaraganda ancha samarali ishlamoqda. Ayniqsa, aniqlashda qiyinchilik tug'diruvchi xatolar — bog'lovchi muvofiqligi, fe'l zamon mosligi, ot-sifat uyg'unligi hamda matnning ma'no jihatdan mosligi bilan bog'liq xatolarni aniqlashda yuqori aniqlik ko'rsatmoqda. Demak, transformer arxitekturasi grammatika tekshiruv va avtomatik tahrir sohasida ilg'or texnologiyalardan biri sifatida tilshunoslik amaliyotiga chuqur kirib bormoqda.

Xulosa. Yuqoridagi tahlillar shuni ko'rsatadiki, sun'iy intellekt vositalari bugungi kunda grammatik xatolarni avtomatik aniqlash va tuzatishda muhim o'rin egallagan. Morfologik, sintaktik va semantik qatlamlar asosida ishlovchi tizimlar, ayniqsa transformer arxitekturasi ega chuqur o'rganish modellarining paydo bo'lishi bilan yangi bosqichga ko'tarildi. BERT, RoBERTa, GPT va T5 kabi modellar til birliklarini kontekst asosida tahlil qilgani bois nafaqat grammatik xatolarni, balki murakkab semantik nomuvofiqliklarni ham aniqlay oladi. Agglyutinativ tillarda uchraydigan morfologik murakkabliklar ham endilikda yanada aniqroq qayta ishlanmoqda.

Demak, zamonaviy transformer modellariga asoslangan SI tizimlari grammatik tahlilning aniqligi va samaradorligini sezilarli darajada oshirib, tilshunoslik amaliyotiga yangi imkoniyatlar yaratmoqda. Ushbu yondashuv kelajakda yanada takomillashib, tilni chuqur semantik anglashga asoslangan yanada mukammal avtomatik tahlil vositalarini yaratishga zamin yaratadi.

FOYDALANILGAN ADABIYOTLAR:

1. O.Q.Abduvosiqov. Tilda grammatik xatolarni aniqlashda sun'iy intellekt texnologiyalaridan foydalanish, 310-bet 2025.
2. Schweter S. BERTurk: BERT Models for Turkish. 2020.
3. Godwin-Jones, R. (2019). AI tools for language learning. *Language Learning & Technology*, 23(1), 4–11.
4. Abjalova M.A. Tahrir va tahlil dasturlarining lingvistik modullari /monografiya / Toshkent: Nodirabegim, 2020. – 176 b.
5. Peterson, James. *Computer Programs for Detecting and Correcting Spelling Errors*. 1980.
6. Ҳамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа докт. (PhD) ... дисс. автореф. – Қарши, 2018. – 52 б