O‘ZBEK TILIDAGI NUTQNI QAYTA ISHLASH MODELLARINING SAMARADORLIGINI NEYROLINGVISTIK MEZONLAR ASOSIDA BAHOLASH

Dushamova Nodirabegim – mustaqil tadqiqotchi

Annotatsiya. Mazkur maqolada o‘zbek tilidagi nutqni qayta ishlash modellarining samaradorligini neyrolingvistik mezonlar asosida baholash masalalari yoritiladi. An’anaviy tabiiy tilni qayta ishlash tizimlarida grammatik aniqlik, leksik moslik va semantik natijadorlik asosiy baholash ko‘rsatkichlari hisoblanadi. Biroq bunday ko‘rsatkichlar modelning inson nutqni anglash jarayoniga qanchalik mos kelishini to‘liq ifodalamaydi. Shu sababli maqolada nutqni anglash vaqti, kontekstual bashorat, semantik nomuvofiqlikka sezgirlik, grammatik murakkablikni qayta ishlash hamda inson bahosi kabi mezonlarni birgalikda qo‘llash imkoniyatlari ko‘rib chiqiladi. O‘zbek tilining agglutinativ tuzilishi, morfologik boyligi va kontekstga bog‘liq semantik xususiyatlari hisobga olinib, modellarni baholash uchun ko‘p mezonli metodologik yondashuv taklif etiladi. Mazkur yondashuv o‘zbek tilidagi sun’iy intellekt tizimlarining lingvistik aniqligi va inson nutqini tushunish jarayoniga funksional mosligini chuqurroq o‘rganishga xizmat qilishi mumkin.

Kalit so‘zlar: sun’iy intellekt, neyrolingvistika, tabiiy tilni qayta ishlash, o‘zbek tili, nutqni anglash, semantik qayta ishlash, kognitiv mezonlar, til modeli, EEG, ko‘z harakati tahlili.

Sun’iy intellekt va tabiiy tilni qayta ishlash texnologiyalarining rivojlanishi inson tilini avtomatik tahlil qilish, nutq mazmunini aniqlash va kontekst asosida javob yaratish imkoniyatlarini kengaytirmoqda. Zamonaviy til modellari matnni qayta ishlashda yuqori natijalarga erishayotgan bo‘lsa-da, ularning inson kabi ma’noni anglash mexanizmlariga ega ekanligi masalasi ilmiy munozaralarning muhim yo‘nalishlaridan biri bo‘lib qolmoqda. An’anaviy baholash usullarida modelning natijalari aniqlik, qamrov, F1 ko‘rsatkichi, perplexity yoki yaratilgan matnning etalon javobga mosligi orqali o‘lchanadi.

Ushbu ko‘rsatkichlar muayyan vazifalardagi samaradorlikni baholash uchun foydali, biroq insonning nutqni qayta ishlash jarayonidagi kognitiv xususiyatlarni bevosita ifodalamaydi. Neyrolingvistika tilni idrok etish va tushunish jarayonlarining miya faoliyati, kognitiv mexanizmlar hamda vaqt omili bilan bog‘liq jihatlarini o‘rganadi. Shu sababli neyrolingvistik mezonlarni kompyuter lingvistikasi bilan uyg‘unlashtirish sun’iy intellekt modellarini yanada kengroq baholash imkonini beradi. O‘zbek tilida ushbu masala alohida ahamiyatga ega. Tilning agglutinativ xususiyati, qo‘shimchalar orqali grammatik munosabatlarning ifodalanishi, so‘z tartibining kommunikativ maqsadga qarab o‘zgarishi va ko‘p ma’noli birliklarning keng qo‘llanishi nutqni avtomatik tahlil qilishni murakkablashtiradi. Shuning uchun o‘zbek tilidagi modellarni faqat umumiy texnik ko‘rsatkichlar asosida emas, balki lingvistik va kognitiv mezonlar asosida ham baholash maqsadga muvofiqdir

Tabiiy tilni qayta ishlash modellari til birliklarini aniqlash, grammatik munosabatlarni tahlil qilish, matn mazmunini ifodalash va yangi matn yaratish kabi vazifalarni bajaradi. Ushbu vazifalarning har biri uchun alohida baholash mezonlari talab qilinadi. Masalan, so‘z turkumlarini aniqlash modelining samaradorligi belgilangan grammatik teglarning to‘g‘ri topilishi orqali o‘lchanishi mumkin. Semantik tahlil tizimida esa so‘zlar va gaplar o‘rtasidagi ma’no munosabatlarini to‘g‘ri aniqlash muhim hisoblanadi. Dialog modelida foydalanuvchi maqsadini anglash, kontekstni saqlash va mazmunan mos javob yaratish kabi ko‘rsatkichlar ham hisobga olinadi. Neyrolingvistik yondashuv ushbu natijalarni insonning nutqni qayta ishlash jarayonlari bilan qiyoslash imkonini beradi. Bunda ikki jihatni farqlash lozim:

Lingvistik samaradorlik — modelning grammatik, leksik va semantik vazifalarni to‘g‘ri bajarish darajasi.

Kognitiv moslik — model natijalari va ichki hisoblash ko‘rsatkichlarining inson nutqni qayta ishlashidagi ayrim kuzatiladigan xususiyatlar bilan mosligi.

Bu ikki tushuncha o‘zaro bog‘liq bo‘lishi mumkin, biroq ular aynan bir hodisani anglatmaydi. Modelning grammatik jihatdan to‘g‘ri javob berishi uning inson kabi fikrlashi yoki ma’noni aynan inson kabi anglashini isbotlamaydi.

Shuning uchun neyrolingvistik baholash modelni inson bilan tenglashtirishga emas, balki inson tilni qayta ishlashining muayyan xususiyatlarini miqdoriy va eksperimental jihatdan qiyoslashga qaratilishi kerak. O‘zbek tilidagi nutqni qayta ishlash modellarini baholashda bir-birini to‘ldiruvchi bir nechta mezonlardan foydalanish mumkin. Inson nutqni tushunish jarayonida oldingi so‘zlar va gap mazmunidan foydalanib, keyingi birliklarni taxmin qilishi mumkin. Til modellari ham kontekst asosida keyingi so‘z yoki tokenning ehtimolini hisoblaydi.

Masalan:

“Talaba imtihonga puxta tayyorlandi, shuning uchun savollarga ...”

Ushbu kontekstda “javob berdi” kabi davomlar mazmunan kutiladigan variantlardir. Modelning kontekstga mos davomni tanlashi uning bashorat qilish qobiliyatini baholashga yordam beradi. Biroq yuqori bashorat aniqligi insonning ma’noni to‘liq tushunishini o‘z-o‘zidan tasdiqlamaydi. Shu sababli mazkur mezon semantik, grammatik va kognitiv ko‘rsatkichlar bilan birgalikda qo‘llanishi kerak. Inson nutqni qayta ishlashda kutilmagan yoki mazmunan nomuvofiq birliklarga turlicha munosabat bildiradi. Neyrolingvistik tadqiqotlarda semantik kutilmaganlikni o‘rganish uchun elektroensefalografiya (EEG) orqali qayd etiladigan N400 komponenti keng qo‘llanadi.

Masalan, quyidagi gaplarni solishtirish mumkin:

“U nonushtada issiq non yedi.”

“U nonushtada issiq tosh yedi.”

Ikkinchi gap odatiy vaziyatda semantik jihatdan kutilmaganroq. Biroq u muayyan badiiy yoki noodatiy kontekstda mazmunli bo‘lishi mumkin. Shuning uchun semantik

nomuvofiqlikni aniqlashda faqat alohida gap emas, balki uning konteksti ham hisobga olinadi.

Til modelining semantik kutilmaganlikni qanday baholashi va insonlarda kuzatiladigan N400 javobi bilan qay darajada bog‘liqligi eksperimental tekshirilishi mumkin. Bunda modelning token ehtimollari yoki surprisal ko‘rsatkichi bilan insonning EEG natijalari o‘rtasidagi munosabat tahlil qilinadi. Inson matnni o‘qish yoki og‘zaki nutqni tinglash jarayonida ayrim birliklarga ko‘proq vaqt sarflaydi. Kam uchraydigan so‘zlar, murakkab grammatik tuzilmalar va kontekstga mos kelmaydigan birliklar qayta ishlash vaqtiga ta’sir ko‘rsatishi mumkin. Ko‘z harakati kuzatuvi orqali birinchi qarash davomiyligi, so‘zga qayta murojaat qilish va umumiy o‘qish vaqti kabi ko‘rsatkichlar o‘lchanadi. Ushbu ma’lumotlar modelning kontekstual bashoratlari bilan qiyoslanishi mumkin. Masalan, o‘zbek tilidagi murakkab qo‘shimchali so‘zlar va grammatik jihatdan sodda shakllarni o‘z ichiga olgan gaplar taqqoslanadi. Agar insonlar muayyan shakllarni qayta ishlashda ko‘proq vaqt sarflasa, modelning ehtimollik ko‘rsatkichlari bilan bu farq o‘rtasidagi bog‘liqlik tekshiriladi. Shuni hisobga olish kerakki, o‘qish vaqti faqat tilni tushunish bilan belgilanmaydi. Diqqat, so‘zning uzunligi, o‘quvchining tajribasi va matnning murakkabligi ham natijalarga ta’sir qiladi.

O‘zbek tilining agglutinativ tuzilishi morfologik shakllarning ko‘pligi bilan tavsiflanadi. Bir so‘z tarkibida bir nechta grammatik ma’no ifodalanishi mumkin. Shu sababli modelning grammatik munosabatlarni qayta ishlash qobiliyatini alohida baholash zarur. Masalan, “Talabalarimizga” shaklida asos, ko‘plik, egalik va jo‘nalish kelishigi ma’nolari birlashgan. Bunday shakllarni tahlil qilishda modelning morfologik teglarni to‘g‘ri aniqlashi va so‘zning gapdagi vazifasini tushunishi tekshiriladi. Sintaktik murakkablikni baholash uchun esa ega-kesim mosligi, kelishik shakllari, ergash gaplar va uzoq masofadagi grammatik bog‘lanishlarni o‘z ichiga olgan namunalar tuzilishi mumkin.

Mazkur ko‘rsatkichlarni bitta umumiy ballga birlashtirish mumkin, ammo buning uchun mezonlarning vazni oldindan asoslanishi kerak. Aks holda, umumiy ball modelning

muhim kamchiliklarini yashirib qo‘yishi mumkin. Shu sababli natijalarni har bir mezon bo‘yicha alohida ko‘rsatish va keyin umumlashtirish tavsiya etiladi.

Amaliy jihatdan ushbu yondashuv o‘zbek tilidagi avtomatik tarjima, matn muharrirlari, dialog tizimlari, ta’lim platformalari va ovozli yordamchilarni baholashda qo‘llanishi mumkin. Masalan, ta’lim tizimida modelning grammatik xatoni topishi bilan birga, o‘quvchi javobidagi semantik nomuvofiqlikni aniqlashi va tushunarli izoh berishi ham baholanishi mumkin.

Taklif etilayotgan metodologiya kelgusida o‘zbek tilidagi nutqni qayta ishlash modellarini qiyoslash uchun standart sinov to‘plamlarini yaratish, inson ishtirokidagi tajribalarni tashkil etish va modellarni lingvistik xususiyatlar bo‘yicha chuqurroq tahlil qilishga yordam beradi.

O‘zbek tilidagi nutqni qayta ishlash modellarining samaradorligini neyrolingvistik mezonlar asosida baholash sun’iy intellekt va inson tilini anglash jarayonlari o‘rtasidagi munosabatni o‘rganish uchun istiqbolli yo‘nalishdir. Bunday baholash an’anaviy texnik ko‘rsatkichlarni kontekstual bashorat, semantik nomuvofiqlikka sezgirlik, nutqni qayta ishlash vaqti va inson bahosi kabi mezonlar bilan to‘ldiradi. O‘zbek tilining morfologik boyligi va grammatik xususiyatlarini hisobga olish baholash metodologiyasining muhim shartidir. Shu sababli sifatli belgilangan lingvistik korpuslar, standartlashtirilgan sinov namunalari va zarur hollarda EEG yoki ko‘z harakati ma’lumotlariga asoslangan tajribalarni birlashtirish maqsadga muvofiq.

Shu bilan birga, texnik samaradorlik va kognitiv moslikni farqlash zarur. Modelning inson nutqini qayta ishlashidagi ayrim ko‘rsatkichlarni yaxshi bashorat qilishi uning inson kabi fikrlashini yoki ma’noni aynan inson kabi anglashini anglatmaydi. Natijalarni ehtiyotkor ilmiy talqin qilish va har bir mezonning chegaralarini hisobga olish tadqiqotning ishonchliligini oshiradi. Umuman olganda, neyrolingvistik va kompyuter-lingvistik yondashuvlarning integratsiyasi o‘zbek tilidagi sun’iy intellekt tizimlarini yanada

chuqurroq baholash, ularning kuchli va zaif tomonlarini aniqlash hamda kelajakdagi til texnologiyalarini takomillashtirish uchun nazariy va amaliy asos yaratadi.

Foydalanilgan adabiyotlar:

1. Abdurahmonova N. Kompyuter lingvistikasi. — Toshkent: Nodirabegim, 2021. — 394 b.

2. Xolmanova Z. Kompyuter lingvistikasi. — Toshkent: Mehr-nuri, 2024. — 234 b.

3. Hojiyev A. Tilshunoslik terminlarining izohli lug‘ati. — Toshkent: O‘zbekiston Milliy ensiklopediyasi, 2002. — 168 b.

4. Ne’matov H., Bozorov O. Til va nutq. — Toshkent: O‘qituvchi, 1993.

5. Ne’matov H., Rasulov R. O‘zbek tili sistem leksikologiyasi asoslari. — Toshkent: O‘qituvchi, 1995.