

## SUN'IY INTELEKT VA ZAMONAVIY TILSHUNOSLIK

**Umarxodjaev M.I.**

filologiya fanlari doktori, professor  
Andijon davlat chet tillari instituti

XX asrning so'ngi yillarida buyuk kashfiyotlardan biri bo'lgan sun'iy intellekt (sun'iy tafakkur, sun'iy ong) algoritmlarining kashf etilishi insoniyatning fan-texnika, iqtisodiy taraqqiyotini misli ko'rilmagan yangi bosqichga olib chiqdi. Hozirgi kunda taraqqiy etgan va taraqqiyot sari intilayotgan mamlakatlarda biron-bir intellektial, iqtisodiy soha yo'qki sun'iy tafakkur algoritmlaridan samarali foydalanmayotgan bo'lsa. Ta'lim, tibbiyot, qishloq xo'jaligi, harbiy soha, davlat boshqaruvidagi raqamlashtirish kabi yo'nalishlar sun'iy ong yordamida o'z samaradorligini oshirib bormoqda.

Jahon tilshunosligining bir qator yo'nalishlarida sun'iy intellekt imkoniyatlaridan keng foydalanib kelinmoqda. Bulardan eng mashhurlari sifatida Google translate, Grammarly, ChatGPTlarni ko'rsatish mumkin. Tilning fonetika va nutq texnologiyalarini, ko'p miqdordagi til ma'lumotlarini qayta ishlash, tarjima sifatini yaxshilash, multimodal NLP, ya'ni tilni rasmlar, audio va video ma'lumotlar bilan ta'minlash, tilning leksik va frazeologik birliklarini aniqlash va tahlil qilish kabi boshqa bir qator yo'nalishlarda keng qo'llanilmoqda.

Alohida ta'kidlash lozimki, sun'iy intellekt algoritmlari zamonaviy lug'atshunoslikda ham juda ahamiyatlidir: avtomatik ma'lumot yig'ish va tahlil qilish, til korpuslarini yaratish, ko'p tilli, audio va vizual lug'atlar tuzish, yoki animatsiyalar va ovozli materiallar orqali interaktiv lug'atlarga tartib berish va boshqalar. Yuorida aytilgan fikrlarga misol tariqasida Andijon davlat chet tillari instituti qoshida faoliyat yuritayotgan ALTAIIM (Andijon lug'atshunoslik-tarjimashunoslik amaliy ilmiy innovatsion markaz) hodimlari tomonidan yaratilayotgan "*O'zbek xalqining til xazinasi*" (tezaurus) nomli lug'atga tartib berish jarayonida sun'iy intellekt algoritmlarining ahamiyatini ko'rsatish lozim.

Ma'lumki, lug'at tuzishning eng mashaqqatli yo'llaridan biri bu lug'atning so'zlar kartotekasini tuzish hisoblanadi. Mashhur lug'atshunoslar (N. Vebster, S.

Djonson, V. Dal) bu ish uchun butun umrlarini sarflaganlar. Hozirgi kunda esa kompyuter texnologiyalari yordamida bu masala tez va keng ko‘lamda o‘z yechimini topmoqda. Markazning lug‘atshunoslari hamda dasturchi-muhandislari hamkorligida bir yil ichida 1 milliarddan ortiq til birliklarini o‘z ichiga olgan “xazina” korpus yaratildi. Dastlabki ma’lumotlar tahlili xazinaga 1 millionga yaqin so‘zlar, terminlar, frazemalar va paremiyalar tanlab olish mumkinligini ko‘rsatmoqda. Ushbu til birliklari 2300 dan ortiq turli lug‘atlar, ilmiy-ommabop, badiiy, diniy adabiyotlar hamda 700 dan ortiq jurnal va gazeta materiallari asosida shakllantirildi. Tadqiqotning dastlabki (to‘liq bo‘lmagan) natijalari asosida so‘zlarning faollik darajasi (chastotasi) ochiqlanmoqda.

Masalan, “*agar*” so‘zi “1550011” marotaba;

“*bu*” so‘zi “8558562” marotaba;

“*edi*” so‘zi “10886248” marotaba;

“*men*” so‘zi “1374432” marotaba;

“*va*” so‘zi “17757757” marotaba;

“*u*” so‘zi “19303266” marotaba.

Ushbu yo‘l bilan aniqlangan so‘zlarning faollik (chastota) ko‘rsatkichlari nafaqat ilmiy ahamiyatga, balki turli xil darsliklar, qo‘llanmalar, hamda bir tilli, ko‘p tilli o‘quv lug‘atlari tuzish uchun ham katta amaliy ahamiyatga egadir.

Birgina qiziq ma’lumotga e’tiborimizni qaratsak. Dastlab savol qo‘ysak: O‘zbek tilida “*siz*” so‘zi ko‘p qo‘llaniladimi yoki “*sen*” so‘zimi? Tadqiqotimiz ma’lumotlari “*siz*” so‘zini 3040090 marotaba qo‘llanilganini va bu so‘z yordamida yasalgan so‘zlarning soni:

“*sizni*” - 929663;

“*sizga*” - 764093;

“*sizning*” - 659308 marotaba ekanligini ko‘rsatdi.

“*Sen*” so‘zining chastotasi esa 438896 tani tashkil etdi:

“*seni*” - 319240;

“*senga*” - 175861.

Demak, tilimizda “*siz*” so‘zi va u bilan bog‘liq bo‘lgan so‘zlar “*sen*” va u bilan bog‘liq so‘zlardan bir necha marotaba faolroq ekan. Bunday ma’lumotlar zamonaviy lug‘atshunoslik uchun o‘ta ahamiyatli hisoblanadi.