

ZAMON QO'SHIMCHALARINI AVTOMATIK ANIQLASH MODELI

Erkinjonova Go'zal Baxtiyor qizi,

Komputer lingvistikasi mutaxassisligi magistranti

uzbekistan-24@mail.ru

ToshDO'TAU

Annotatsiya. Tilshunoslikda fe'l so'z turkumiga mansub so'zlarning zamonini aniqlash muhim ahamiyatga ega. Fe'llarning zamonlarini aniqlash orqali matn mazmunini to'g'ri tushunish, tarjima qilish va tabiiy tilni qayta ishlash (NLP) sohasida ko'plab muammolarni hal qilish mumkin. Zamon qo'shimchalarini aniqlashda lingvistik jihatdan o'zbek tilidagi fe'l zamon kategoriyasining morfologik xususiyatlari tahlil qilinadi. Amaliy jihatdan esa qoidalarga asoslangan metod yordamida modelni qurish jarayoni tavsiflanadi.

Abstract. In linguistics, it is important to determine the tense of words belonging to the verb group. Determining verb tenses can solve many problems in text comprehension, translation, and natural language processing (NLP). Morphological characteristics of the verb tense category in the Uzbek language are analyzed from a linguistic point of view when determining the tenses. In practical terms, the process of building a model using the rule-based method is described.

Аннотация. В лингвистике важно определять время слов, принадлежащих к группе глаголов. Определение времен глаголов может решить многие проблемы понимания текста, перевода и обработки естественного языка (NLP). Морфологические характеристики категории времени глагола в узбекском языке анализируются с лингвистической точки зрения при определении времен. В практическом плане описывается процесс построения модели с использованием метода, основанного на правилах.

Kalit so'zlar: *Zamon qo'shimchalari, qoidalarga asoslangan metod, NLP, Python.*

Kirish

Dunyo tilshunosligida XX asr oxirlarida til korpuslarini tuzish ularni lingvistik annotatsiyalash dolzarblasha boshladi. XXI asrda ko'plab so'zlarni o'z ichiga olgan korpuslar paydo bo'la boshladi. Natijada fe'l zamonlarining avtomatik tahlil qilish muhim masalaga aylana boshladi. Fe'l so'z turkumida zamon qo'shimchalarini aniqlash biroz murakkab jarayon hisoblanadi, sababi fe'l zamonlarining ko'chishi, kontekstda boshqa bir zamonga ishora qilishi avtomatik tahlilda qiyinchiliklar keltiradi.

Metodologiya

Jahon korpus lingvistikasida zamon shakllarini ko'plab olimlar tadqiq etishgan sirasiga M.Verhagen, G.Marsik, N. UzZaman, L.Bierkandt, V. Gast kabi olimlar faqat muayyan grammatik shakllar orqali ifodalangan zamon ma'nosi emas, balki turli vaqt ma'noli lug'aviy birliklar, shuningdek, birikmalar va gaplarning o'zaro munosabatidan kelib chiqadigan zamon ma'nosini modellashtirish masalalari ham o'rganiladi. O'z tadqiqotlarida korpus va uning tuzilishi, razmetka standartlari, matnning avtomatik morfologik va sintaktik tahlilini amalga oshirish kabi masalalarga e'tibor bergan olimlar V.P.Zaxarov, S.Y.Bogdanova, A.B.Kutuzov, Y.V.Nedoshivina, D.Y.Gruzdev, I.O.Kuznetsov, I.M.Nojoy kabilardir.

O'zbek tilshunosligida kompyuter lingvistikasi, matnga leksikografik ishlov berish va lingvostatistik tahlil etish borasida ko'plab tadqiqotlar olib borilgan. Jumladan, O'zbek korpus lingvistikasida ham zamon shakllarini annotatsiyalash yuzasidan fikrlar morfologik teglash, morfoanalizator ishlab chiqish masalalari doirasidagi tadqiqotlarda bildirilgan. Sh.Hamroyeva “O'zbek tili morfologik analizatorining lingvistik ta'minoti” mavzusidagi dissertatsiyasida o'zbek tili morfoanalizatori uchun grammatik kategoriyalarni o'zbek tili qonuniyatlari asosida qat'iy belgilash lozimligini ta'kidlab, o'zbek tilidagi 11 ta grammatik kategoriyani, jumladan, zamon kategoriyasini ham ko'rsatib o'tadi [1]. M.Abjalova o'zbek tili matnlarining avtomatik tahlili va tahriri bo'yicha amalga oshirgan tadqiqotida tahlilning morfologik bosqichi doirasida zamon shakllariga ham to'xtaladi [2]. Sh. Hamroyeva, O'.Xoliyorov, G.Abdualimovalar tomonidan tuzilgan “Fe'l shakllari lug'ati”da zamonning o'zida 49ta, variantlari bilan esa 66ta shakl mavjudligi qayd etiladi [3]. Zamon kategoriyasi shakllari to'liqsiz fe'llar bilan hosil qilingan analitik shakllar hisobiga shunday miqdorni ko'rsatishi mumkin. A.Q.Po'latov, T.P.Mo'minova, I.O.Po'latovalarning “Dunyoviy o'zbek tili” kitobida ingliz tilida bir fe'l taxminan 100ta shaklga ega bo'lgani holda o'zbek tilida uning shakllari 100000dan oshishi haqida ma'lumot beriladi [4].

Asosiy qism

O'zbek tilida zamon fe'lining muhim grammatik kategoriyalaridan biri bo'lib, uch asosiy shaklga ega: Hozirgi zamon, o'tgan zamon, kelasi zamon.

№	Zamon shakllari	Qo'shimchalari
1.	Hozirgi zamon	-yap,-yotir/ ayotir, -yotib/ ayotib, -moqda

	-di, -gan/ kan/ qan, -b/ ib, -r edi/ ar edi, //rdi/ ardi,
2. O'tgan zamon	-yotgan edi/ ayotgan edi, -yotib edi/ ayotib edi, -guvchi edi, -ydigan edi/ adigan edi, -moqda edi, -moqchi edi.
3. Kelasi zamon	- y/ a, - ar, -r, -moqchi, -y/ ay - y/ a, - ar, -r, -moqchi, -y/ ay

Zamonni ifodalovchi grammatik shakllarning annotatsiyasi morfologik tahlil uchun muhim ahamiyat kasb etadi. Zamon shakllarining morfologik tahlilda ya'ni POS taggingda fe'l so'z turkumiga oid so'zshaklga grammatik xarakteristika berishda muhim hisoblanadi. Zamon qo'shimchalarini avtomatik tahlil qilish uchun formal grammatikadan foydalaniladi. Sababi formal grammatika asosi o'zak va affiks qoidalardan iborat.

Qoidalarga asoslangan model (Rule-based) orqali zamon kategoriyasini avtomatik aniqlash mumkin. Ushbu usul lingvistik qoidalarga asoslangan holda ishlaydi. Qoidalarga asoslangan yondashuv aniq va tushunarlik darajasi yuqori bo'lib, aynan morfologik tahlil talab qilinuvchi tizimlar uchun samarali yondashuv hisoblanadi.

Qoidalarga asoslangan model quyidagicha ishlaydi:

1. O'zbek tilidagi zamonni ifodalovchi grammatik shakllarning tizimli ro'yxatini shakllantirish.
2. Berilgan matndan fe'llar ajratib olinadi
3. Matn tarkibidan zamonni ifodalovchi grammatik shakllarni topishga xizmat qiluvchi lingvistik modellar ishlab chiqilib va ular asosida zamon shakllarini aniqlanadi.
4. Qoidalarga asoslanib qaysi zamon ekanligi aniqlanadi.

Zamon qo'shimchalarini avtomatik aniqlashda python dasturlash tilidan foydalanildi. Bunda asosiy vazifa shaxs-son qo'shimchalarini olib tashlash, fe'lning oxiri o'tgan, hozirgi yoki kelasi zamon qo'shimchalari bilan tugashini tekshirish va aniqlangan fe'lni tegishli zamon ro'yxatiga joylashtirish bilan yakunlandi. Zamon qo'shimchalarini avtomatik aniqlashda quyidagi metodlardan foydalanildi.

Funksiyalar	Vazifasi
.split()	Matnni so'zlarga ajratish uchun ishlatilgan.

.endswith(qoshimcha)	So'z oxiridagi qo'shimchalarni tekshirish uchun ishlatilgan
[:len(qoshimcha)]	So'zdan qo'shimchani olib tashlash uchun ishlatilgan
append()	So'zlarni yangi ro'yxatlarga qo'shish uchun ishlatilgan.
for ... in ...:	Ro'yxatlar ustida aylanish uchun ishlatilgan.
break	So'zdan qo'shimcha topilganda, keraksiz tekshiruvlarni to'xtatish uchun ishlatilgan.

Foydalanilgan metodlar orqali ushbu natijaga erishildi.

Misol (*bordim, bormoqchisan, qilmoqdamiz*)

O'tgan zamon so'zlari:

Asosiy: *bor*, Qo'shimcha: -di

Hozirgi zamon so'zlari:

Asosiy: *qil*, Qo'shimcha: -moqda

Kelasi zamon so'zlari:

Asosiy: *bor*, Qo'shimcha: -moqchi

Xulosa

O'zbek tilshunosligida zamonni ifodalovchi grammatik shakllar va ularning tasnifi munozarali masalalardan hisoblanadi. Korpus tilshunosligida zamon kategoriyasini aniqlash uchun uning grammatik darajasi, semantik darajasi va kontekstual darajasini qamrab olish zarur. Natijada zamon qo'shimchalarini matnda aniqlash birmuncha osonlashadi.

Foydalanilgan adabiyotlar:

1. Хамроева Ш. Ўзбек тили морфологик анализаторининг лингвистик таъминоти: Филол. фан. док. ...дис. – Тошкент, 2021. – Б. 150.
2. Абжалова М. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар таҳрири дастури учун): Филол. фан. бўйича фалсафа доктори (PhD) ...дисс. – Фарғона, 2019. – Б.125–126.
3. Hamroyeva Sh., Xoliyoyorov O., Abdualimova G. O'zbek tilining fe'l shakllari lug'ati. ("kiymoq" fe'li misolida) I jild. – Termiz, 2021. – B. 19.
4. Пўлатов А.Қ., Мўминова Т.П., Пўлатова И.О. Дунёвий ўзбек тили. Ўзбек тилида феъл шакллари ва уларнинг рус, инглиз тилидаги кўринишлари. – Тошкент: Universitet, 2003. – Б. 3.



5. Elova D. Fe'l grammatik shakllarining uslubiy xoslanishiga ko'ra lingvistik annotatsiyalash masalasi xususida.
6. Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Бухоро, 2018. – Б. 3.
7. Elov B., Hamroyeva Sh., Abdullayeva O., Uzakova M. O'zbek tilida POS tagging masalasi: muammo va takliflar / Uzbekistan: Language and Culture. Applied philology. 2022/2(5). – 51-68-b.
8. Ғуломов А. Феъл. – Тошкент: Фан, 1954. – Б. 19–38
9. Сулаймонов А., Ҳожиёв А., Жўраева Ж. Феъл замонлари. – Тошкент, 1962. – 168 б.
10. <https://ucrel.lancs.ac.uk/papers/ClawsWordTaggingSystemRG87.pdf>
11. <https://uznatcorpara>