

KONFERENSIYALAR.COM

ANJUMANLAR PLATFORMASI

O'ZBEKISTON – 2030: INNOVATSIYA, FAN VA TA'LIM ISTIQBOLLARI

**II RESPUBLIKA ILMIY-AMALIY
KONFERENSIYA MATERIALLARI**

OKTYABR, 2026-YIL



ISBN 978-9910-8337-7-9

O‘ZBEKISTON – 2030: INNOVATSIYA, FAN VA TA’LIM ISTIQBOLLARI. II Respublika ilmiy-amaliy konferensiyasi materiallari. – Toshkent: Scienceproblems team, 2026. – 200 bet.

Elektron nashr: <https://konferensiyalar.com>

Konferensiya tashkilotchisi: “Scienceproblems Team” MChJ

Konferensiya o‘tkazilgan sana: 2026-yil, 2-oktyabr

Mas’ul muharrir:

Isanova Feruza Tulqinovna

Annotatsiya

Mazkur nashrda “O‘zbekiston — 2030: innovatsiya, fan va ta’lim istiqbollari” nomli II Respublika ilmiy-amaliy konferensiyasi doirasida taqdim etilgan ilmiy maqolalar to‘plami jamlangan. Unda O‘zbekistonning turli oliy ta’lim va ilmiy-tadqiqot muassasalari, tarmoq tashkilotlari, mustaqil tadqiqotchilar tomonidan taqdim etilgan ijtimoiy-gumanitar, iqtisodiyot, huquq, biologiya, tibbiyot va boshqa sohalarga oid maqolalar kiritilgan. Maqolalarda ilm-fanning zamonaviy yo‘nalishlari, innovatsion texnologiyalar, ta’lim islohotlari hamda barqaror taraqqiyotga oid masalalar muhokama qilingan. To‘plam akademik izlanishlar, amaliy tajribalar va ilmiy xulosalarni birlashtirgan holda, fanlararo integratsiyani chuqurlashtirish va ilmiy hamkorlikni kuchaytirishga xizmat qiladi.

Kalit so‘zlar: ilmiy-amaliy konferensiya, innovatsiya, fan va ta’lim, O‘zbekiston 2030, barqaror rivojlanish, ilmiy izlanishlar, fanlararo integratsiya, ilmiy hamkorlik, texnologik taraqqiyot, zamonaviy ta’lim.

ISBN 978-9910-8337-7-9

Barcha huquqlar himoyalangan.

© Scienceproblems team, 2026-yil

© Mualliflar jamoasi, 2026-yil

O'ZBEK TILIDA FE'L VAZIFA SHAKLLARINI AVTOMATIK TASNIFLASHNING NLP USULLARI

To'xtayev Shuhrat Furqatovich

Buxoro shahar Kogon shahar IM

ona tili va adabiyot fani o'qituvchisi

Email: shuhrattoxtayev22@gmail.com

Annotatsiya. Mazkur maqolada o'zbek tilidagi fe'l vazifa shakllarini — harakat nomi, sifatdosh va ravishdoshni — avtomatik tasniflash masalasi tabiiy tilni qayta ishlash (NLP) nuqtayi nazaridan tadqiq etiladi. O'zbek tilining agglutinativ tabiati, morfologik shakllarning ko'p funksionaliligi va grammatik ma'nolarning kontekstga bog'liqligi avtomatik tasniflash jarayonida muayyan qiyinchiliklarni yuzaga keltiradi. Tadqiqotda qoidaviy, statistik, mashinali o'rganish va neyron tarmoqlarga asoslangan yondashuvlar tahlil qilinadi. Universal Dependencies (UD) tizimida o'zbek fe'l shakllarini Fin, Part, Conv va Vnoun qiymatlari orqali ifodalash imkoniyati ko'rib chiqiladi. O'zbek tili uchun yaratilgan Uzbek-UT, Uzbek-TueCL, UzUDT va UzbekPOS kabi korpus resurslari tahlil qilinib, ular asosida fe'l vazifa shakllarini avtomatik tasniflashning gibrid modeli taklif etiladi. Model morfologik belgilar, affiksalar tuzilma va kontekstual ma'lumotlarni birlashtirishga asoslanadi.

Kalit so'zlar: o'zbek tili, tabiiy tilni qayta ishlash, NLP, fe'l vazifa shakllari, harakat nomi, sifatdosh, ravishdosh, morfologik tahlil, POS-tagging, Universal Dependencies, avtomatik tasniflash.

NLP METHODS FOR AUTOMATIC CLASSIFICATION OF VERBAL NON-FINITE FORMS IN THE UZBEK LANGUAGE

To'xtayev Shuhrat Furqatovich

Teacher of Uzbek Language and Literature,

Kogon City IM, Bukhara, Uzbekistan

Annotation. This article investigates the automatic classification of non-finite verb forms in the Uzbek language, namely the verbal noun, participle, and converb, from the perspective of natural language processing (NLP). The agglutinative nature of Uzbek, the multifunctionality of morphological forms, and the context-dependent nature of grammatical meanings present specific challenges for automatic classification. The study analyzes rule-based, statistical, machine learning, and neural network approaches. It examines the possibility of representing Uzbek verb forms using the Fin, Part, Conv, and Vnoun values within the Universal Dependencies (UD) framework. Corpus resources developed for Uzbek, including Uzbek-UT, Uzbek-TueCL, UzUDT, and UzbekPOS, are analyzed, and a hybrid model for the automatic classification of non-finite verb forms is proposed based on these resources. The model integrates morphological features, affixal structures, and contextual information.

Key words: Uzbek language, natural language processing, NLP, non-finite verb forms, verbal noun, participle, converb, morphological analysis, POS tagging, Universal Dependencies, automatic classification.

Kirish. So'nggi yillarda o'zbek tilini kompyuter yordamida qayta ishlash, xususan, morfologik tahlil, so'z turkumlarini avtomatik belgilash va sintaktik parsing yo'nalishlarida muhim tadqiqotlar amalga oshirilmoqda. O'zbek tili agglutinativ til bo'lgan sababli grammatik ma'nolar ko'pincha o'zakka ketma-ket qo'shiladigan affikslar orqali ifodalanadi. Natijada bitta so'z shaklida bir nechta morfologik belgi mujassamlashishi mumkin. Elov va hammualliflar o'zbek tilining morfologik teglash modelini Universal Dependencies (UD) freymvorki asosida ishlab chiqishda aynan morfotaktik qoidalar, affikslar ketma-ketligi va ularning UD FEATS tizimiga mosligini hisobga olganlar [1]. Fe'l kategoriyasi ushbu jarayonda alohida murakkablikka ega. Chunki o'zbek tilida fe'lning tuslangan va tuslanmagan shakllari turli grammatik vazifalarni bajaradi. Universal Dependencies tizimida o'zbek fe'llari VERB so'z

turkumi ostida berilib, VerbForm xususiyati orqali Fin (finite), Part (participle), Conv (converb) va Vnoun (verbal noun) shakllariga ajratiladi [2]. Mazkur tasnif an'anaviy o'zbek tilshunosligidagi sof fe'l, sifatdosh, ravishdosh va harakat nomi tushunchalarini xalqaro NLP standartlari bilan bog'lash imkonini beradi. Maqolaning maqsadi — o'zbek tilidagi fe'l vazifa shakllarini avtomatik tasniflashda qo'llanadigan NLP usullarini tahlil qilish, mavjud korpus va modellarni o'rganish hamda morfologik va kontekstual belgilarni birlashtiruvchi gibril tasniflash modelini taklif etish.

Tadqiqot metodologiyasi. Tadqiqotda quyidagi usullardan foydalanildi:

ilmiy adabiyotlarni tahlil qilish;

mavjud o'zbek NLP korpuslarini qiyoslash;

Universal Dependencies standartlarini o'rganish;

morfologik va kontekstual belgilarni modellashtirish;

qoidaviy va mashinali o'rganish yondashuvlarini taqqoslash;

avtomatik tasniflash uchun gibril modelni ishlab chiqish.

Tadqiqotning muhim metodologik asosi — tasniflashda so'zning faqat tashqi shakliga emas, balki morfologik tuzilishi va gapdagi kontekstiga birgalikda tayanishdir.

Fe'l vazifa shakllarining NLP tizimidagi ifodasi. An'anaviy o'zbek tilshunosligida harakat nomi, sifatdosh va ravishdosh fe'lining vazifa shakllari sifatida qaraladi. NLP tizimida esa ularning xalqaro standartdagi ifodasi quyidagicha ko'rinadi.

An'anaviy shakl	UD VerbForm	Misol	Asosiy funksiyasi
Sof fe'l	Fin	keldi, boradi	predikativ
Harakat nomi	V noun	o'qish, kelish, kelmoq	nominal
Sifatdosh	Part	kelgan, boradigan	atributiv/predikativ
Ravishdosh	Conv	kelib, o'qib, borgach	Adverbial

O'zbek tili uchun mavjud NLP resurslari. O'zbek tili uchun korpus resurslarining rivojlanishi avtomatik tasniflagichlarni yaratish imkoniyatlarini kengaytirmoqda. 2025-yilda Akhundjanova va Talamo tomonidan o'zbek tili uchun birinchi Universal Dependencies treebanki taqdim etilgan. Uzbek-UT deb nomlangan ushbu resurs 500 ta gap va 5850 ta tokenni o'z ichiga oladi hamda lemma, POS, morfologik xususiyatlar va dependency munosabatlari bilan annotatsiyalangan [4]. 2026-yilda Matlatipov va Aripov tomonidan taqdim etilgan UzUDT esa 684 ta gap va 7582 ta tokenni o'z ichiga oladi. Mualliflar ushbu resurs to'liq qo'lda annotatsiyalanganini, ko'p annotatorli tekshiruvdan o'tkazilganini va annotatorlararo kelishuv koeffitsiyentlari 0,90 dan yuqori bo'lganini bildiradi [5]. Bundan tashqari, Sharipov, Kuriyozov va Vičič tomonidan 2026-yilda taqdim etilgan UzbekPOS korpusi qariyb 4500 ta gap va 53 mingdan ortiq token/teg juftligini o'z ichiga oladi. Korpus adabiyot, yangiliklar, fan, ta'lim va ommaviy nutq kabi turli domenlardan tuzilgan bo'lib, annotatsiyalar kamida ikki annotator tomonidan tekshirilgan [6].

O'zbek NLP korpuslarining qiyosiy tavsifi:

Resurs	Yil	Hajmi	Annotatsiya	Asosiy vazifa
--------	-----	-------	-------------	---------------

Uzbek-UT	2025	500 gap; 5850 token	POS, morfologiya, dependency	UD parsing
Uzbek-TueCL	—	148 gap; 939 token	POS, morfologiya, dependency	grammatik konstruksiyalar
UzUDT	2026	684 gap; 7582 token	to'liq UD	parsing va tagging
UzbekPOS	2026	~4500 gap; >53 ming token	POS	POS tagging

Uzbek-TueCL korpusida ham Conv, Fin, Part va Vnoun qiymatlari amaliy annotatsiyada qo'llangan.

Qoidaviy tasniflash usuli. Fe'l vazifa shakllarini avtomatik tasniflashning dastlabki bosqichi qoidaviy yondashuv bo'lishi mumkin. Bunda so'z oxiridagi affikslar aniqlanadi va dastlabki sinf belgilanadi.

Masalan:

-ish, -moq, -uv → Vnoun;

-gan, -adigan, -ayotgan, -ar → Part;

-ib, -a/y, -gach, -gani, -gancha → Conv.

Biroq bunday qoidalar yakuniy tasniflash vositasi bo'la olmaydi. Chunki bir xil shakl turli grammatik vazifalarda kelishi mumkin.

Masalan:

Kelgan bola kitob o'qidi.

Bu yerda kelgan otni aniqlab kelgani sababli sifatdosh funksiyasida.

U kecha kelgan.

Bu yerda kelgan predikativ kontekstda keladi.

Kitob o'qish foydalgapida o'qish harakat nomi sifatida qo'llangan. Demak, avtomatik tizim affiksni aniqlashdan keyin kontekstni ham hisobga olishi kerak.

Statistik va mashinali o'rganish usullari. Statistik tasniflashda so'zning o'zi bilan bir qatorda uning atrofidagi birliklar ham modelga beriladi. Bunda quyidagi belgilar muhim:

1. tokenning o'zi;
2. lemma;
3. so'z oxiridagi 2–5 belgidan iborat suffiks;
4. oldingi token;
5. keyingi token;
6. oldingi va keyingi POS teglari;
7. morfologik segmentlar;
8. gapdagi sintaktik pozitsiya.

Klassifikator quyidagi ko'rinishda tasavvur qilinishi mumkin:

$Y=f(X_{\text{morph}}, X_{\text{context}}, X_{\text{syntax}})$ $Y=f(X_{\{\text{morph}\}}, X_{\{\text{context}\}}, X_{\{\text{syntax}\}})$

bu yerda:

YY — yakuniy fe'l shakli;

XmorphX_{morph} — morfologik belgilar;

XcontextX_{context} — kontekstual belgilar;

XsyntaxX_{syntax} — sintaktik belgilar.

Bunday yondashuvda CRF, SVM, Random Forest yoki boshqa klassifikatorlardan foydalanish mumkin. Biroq o'zbek tilidagi mavjud resurslarning hajmi hali katta bo'lmagani sababli model tanlashda o'quv korpusining hajmi va sifati hisobga olinishi zarur.

Taklif etilayotgan gibrid model. Yuqoridagi tadqiqotlar asosida o'zbek tilidagi fe'l vazifa shakllarini avtomatik tasniflash uchun uch komponentli gibrid model taklif etiladi. Taklif etilayotgan model bosqichlari:

Bosqich	Amal	Natija
1	Morfologik analiz	lemma va affikslar
2	Kontekstual tahlil	kontekst belgilar
3	Klassifikatsiya	Fin / Part / Conv / Vnoun
4	UD mapping	standart morfologik teg
5	Validatsiya	yakuniy natija

Tasniflashda uchraydigan asosiy muammolar: Affikslarning ko'p funksiyaliligi: Bir affiksning turli grammatik vazifalarda qo'llanishi avtomatik tasniflashni qiyinlashtiradi. Shu sababli -gan, -ish, -a, -ib kabi shakllarni faqat morfologik qoidalar bilan aniqlash yetarli emas.

Otlashish: Harakat nomlari otga xos kelishik va egalik shakllarini qabul qilishi mumkin: o'qishni, kelishimiz, borishga.UD korpuslarida Vnoun shakllarining Case va boshqa morfologik belgilar bilan birikishi amalda kuzatiladi.

Analitik va murakkab shakllar: O'zbek tilida fe'lning yordamchi fe'llar bilan birikishi natijasida murakkab konstruksiyalar yuzaga keladi. Bunday holatlarda token darajasidagi tasnif bilan sintaktik konstruksiya darajasidagi tahlilni bir-biridan farqlash zarur.

Kam resurslilik: O'zbek tili uchun yuqori sifatli annotatsiyalangan korpuslar soni hali cheklangan. Shu sababli kichik korpuslarda o'qitilgan model boshqa janr va domenlarda bir xil natija bermasligi mumkin. UzbekPOSning ko'p domenli tuzilishi aynan shu muammoni qisman kamaytirishga xizmat qiladi [6].

Baholash mezonlari. Tasniflagichni baholashda Accuracy bilan bir qatorda Precision, Recall va F1-score ko'rsatkichlaridan foydalanish maqsadga muvofiq.

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Har bir sinf bo'yicha alohida natija berilishi kerak:

Sinf	Precision	Recall	F1
Vnoun	—	—	—
Part	—	—	—
Conv	—	—	—
Fin	—	—	—

Izoh: jadvaldagi natijalar model amaliy tajribada o'qitilib, test qilinmaguncha to'ldirilmasligi kerak. Shu sababli maqolada sun'iy ravishda eksperimental raqamlar keltirilmagan.

Natijalar va muhokama. Mavjud tadqiqotlar o'zbek tilida avtomatik morfologik va POS tasniflash uchun bir nechta muhim yo'nalish shakllanganini ko'rsatadi.

Birinchidan, UD tizimi o'zbek fe'l shakllarini yagona standart asosida ifodalash imkonini beradi. Fin, Part, Conv va Vnoun qiymatlarining mavjudligi an'anaviy grammatik tasnifni NLP tizimlariga moslashtirish uchun qulay asos yaratadi [2; 3].

Ikkinchidan, o'zbek tili uchun mavjud treebanklar hajmi oshib bormoqda. Uzbek-UTdan keyin UzUDTning yaratilishi to'liq qo'lda annotatsiyalangan yangi resursni taqdim etdi [4; 5].

Uchinchidan, morfologik analizatorning rivojlanishi neyron tasniflagich uchun qo'shimcha morfologik belgilar olish imkonini beradi.

Shu asosda fe'l vazifa shakllari uchun faqat qoidaviy tizimdan foydalanish o'rniga morfologik analiz + kontekstual model + UD mapping kombinatsiyasi tadqiqot uchun asosli yo'nalish sifatida taklif etiladi.

Xulosa. O'zbek tilidagi fe'l vazifa shakllarini avtomatik tasniflash o'zbek NLP tizimlarining muhim vazifalaridan biridir. Tadqiqot shuni ko'rsatadiki, o'zbek tilining agglyutinativ tabiati va morfologik shakllarning kontekstga bog'liq xususiyatlari sababli faqat affiksga asoslangan qoidaviy tasniflash yetarli emas. Universal Dependencies tizimida Fin, Part, Conv va Vnoun qiymatlarining qo'llanishi fe'l shakllarini standartlashtirish uchun aniq metodologik asos yaratadi [2]. O'zbek-UT, UzUDT va UzbekPOS kabi yangi resurslar esa avtomatik tasniflagichlarni o'qitish va sinash uchun zarur ma'lumotlar bazasini kengaytirmoqda [4–6]. Maqolada o'zbek tilidagi fe'l vazifa shakllarini avtomatik tasniflash uchun morfologik analizator va kontekstual neyron modelni birlashtiruvchi gibrid yondashuv taklif qilindi. Bunda morfologik analizator affiks va grammatik belgilarni, kontekstual model esa shaklning gapdagi funksiyasini aniqlashda foydalaniladi. Kelgusidagi tajriba-sinovlarda Vnoun, Part, Conv va Fin sinflari bo'yicha alohida Precision, Recall va F1 ko'rsatkichlarini hisoblash, turli janrdagi matnlarda modelni tekshirish hamda qoidaviy, klassik mashinali o'rganish va BERT asosidagi modellarni bir xil test to'plamida qiyoslash maqsadga muvofiqdir.

Adabiyotlar/Literatura/References:

1. Elov, B. B., Khamroeva, Sh. M., Kurbanova, M. M., Khusainova, Z. Y., Ganiyeva, A. N. (2025). A Morphological Tagging Model of the Uzbek Language in the Universal Dependencies Format. 2025 10th International Conference on Computer Science and Engineering (UBMK), 380–384. DOI: 10.1109/UBMK67458.2025.11207015.
2. Universal Dependencies. UD for Uzbek. Universal Dependencies v2.
3. Universal Dependencies. Statistics of VERB in UD_Uzbek-UT.
4. Akhundjanova, A., Talamo, L. (2025). Universal Dependencies Treebank for Uzbek. Proceedings of the Third Workshop on Resources and Representations for Under-Resourced Languages and Domains (RESOURCEFUL-2025), 1–6.
5. Matlatipov, S. G., Aripov, M. (2026). UzUDT: Uzbek Universal Dependencies Treebank. Proceedings of the Fifteenth Language Resources and Evaluation Conference, 11642 – 11649. DOI: 10.63317/2uedjqezjxn5.
6. Sharipov, M., Kuriyozov, E., Vičič, J. (2026). UzbekPOS: A multi-domain dataset for Uzbek part-of-speech tagging. Data in Brief, 66, 112640. DOI: 10.1016/j.dib.2026.112640.

O‘ZBEKISTON — 2030: INNOVATSIYA, FAN VA TA’LIM ISTIQBOLLARI

II RESPUBLIKA ILMIY-AMALIY KONFERENSIYASI MATERIALLARI

2026-yil, 6-oktyabr

Mas’ul muharrir: *F.T.Isanova*
Texnik muharrir: *N.Bahodirova*
Diszayner: *I.Abdihakimov*

O‘ZBEKISTON — 2030: INNOVATSIYA, FAN VA TA’LIM
ISTIQBOLLARI. II Respublika ilmiy-amaliy konferensiyasi
materiallari. – Toshkent: Scienceproblems team, 2026. – 200 bet.

Elektron nashr: <https://konferensiyalar.com>

Konferensiya tashkilotchisi: Scienceproblems Team

Konferensiya o‘tkazilgan sana: 2026-yil, 2-oktyabr

ISBN 978-9910-8337-7-9

Barcha huquqlar himoyalangan.
© Scienceproblems team, 2026-yil.
© Mualliflar jamoasi, 2026-yil.