

O‘ZBEK TILI LEKSIK BIRLIKLARI RAVISH SO‘Z TURKUMINI POS TEGLASH ALGORITMLARI VA MODELLARI

Xudayberganov Nizomaddin Uktambay o‘g‘li,

Alisher Navoiy nomidagi Toshkent davlat

o‘zbek tili va adabiyoti universiteti o‘qituvchisi

nizomaddin@navoiy-uni.uz

ORCID: 0000-0002-6213-3015

Annotatsiya. Ushbu maqola o‘zbek tilidagi ravish so‘z turkumini aniqlash va grammatik POS (Part-of-Speech) teglash algoritmlari hamda modellarining nazariy asoslari, amaliy tatbiq usullari va eksperimental natijalarini chuqur tahlil qilishga bag‘ishlangan. Tadqiqotda qoidalarga asoslangan yondashuv (rule-based), statistik modellar (Hidden Markov Model – HMM, Conditional Random Fields – CRF) va zamonaviy chuqur o‘rganish (Deep Learning) modellarining ravish so‘z turkumini teglashdagi samaradorligi o‘rganiladi. Ravishlarning morfologik tuzilishi, ma‘no turlari va sintaktik vazifalaridan kelib chiqqan holda, ularni avtomatik aniqlash algoritmlari ishlab chiqilgan. Tadqiqot natijalariga ko‘ra, gibrid yondashuv qoidalarga asoslangan va statistik modellarining uyg‘unligi 94.2% aniqlik darajasiga erishgan bo‘lib, bu qoidalarga asoslangan (89.7%) va HMM (91.3%) modellaridan yuqori ko‘rsatkich hisoblanadi. Maqolada algoritm strukturalari, teglar to‘plami (tagset) va samaradorlik ko‘rsatkichlari jadval va diagrammalar orqali vizual tarzda ifodalangan.

Kalit so‘zlar: ravish so‘z turkumi, POS teglash, HMM, CRF, qoidaviy algoritmlar, morfologik tahlil, o‘zbek tili korpusi, Viterbi algoritmi, teglar to‘plami.

Annotation. This article is devoted to an in-depth analysis of the theoretical foundations, practical implementation methods, and experimental results of algorithms and models for identifying and grammatically POS (Part-of-Speech) tagging adverbs in the Uzbek language. The study examines the effectiveness of rule-based approaches, statistical models (Hidden Markov Model – HMM, Conditional Random Fields – CRF), and modern deep learning models in tagging the adverb word class. Based on the morphological structure, semantic types, and syntactic functions of adverbs, algorithms for their automatic identification have been developed. According to the research results, the hybrid approach-combining rule-based and statistical models-achieved an accuracy of 94.2%, which is higher than that of the rule-based (89.7%) and HMM (91.3%) models. The article presents the algorithm structures, tagset, and performance metrics in a visual form using tables and diagrams.

Keywords: adverb word class, POS tagging, HMM, CRF, rule-based algorithms, morphological analysis, Uzbek language corpus, Viterbi algorithm, tagset.

Аннотация. Данная статья посвящена глубокому анализу теоретических основ, методов практической реализации и экспериментальных результатов алгоритмов и моделей для определения и грамматического POS (Part-of-Speech) теггирования наречий в узбекском языке. В исследовании рассматривается эффективность правилowego подхода (rule-based), статистических моделей (Hidden Markov Model – HMM, Conditional Random Fields – CRF) и современных моделей глубокого обучения (Deep Learning) при теггировании наречий. На основе морфологической структуры, семантических типов и синтаксических функций наречий разработаны алгоритмы их автоматического определения. Согласно результатам исследования, гибридный подход, объединяющий правилые и статистические модели, достиг точности 94,2%, что выше по сравнению с правиловой моделью (89,7%) и моделью HMM (91,3%). В статье структуры алгоритмов, набор тегов (tagset) и показатели эффективности представлены в наглядной форме с использованием таблиц и диаграмм.

Ключевые слова: наречие, POS-теггирование, HMM, CRF, правилые алгоритмы, морфологический анализ, корпус узбекского языка, алгоритм Витерби, набор тегов.

Kirish. Tabiiy tilni qayta ishlash (Natural Language Processing, NLP) sohasining jadal rivojlanishi bilan birga, kam resursli tillar uchun lingvistik vositalarni yaratish dolzarb masalalardan biriga aylandi. O‘zbek tili agglyutinativ tillar guruhiga mansub bo‘lib, so‘z yasalihi va so‘z o‘zgarishida affiksatsiyaning muhim rol o‘ynashi tilni avtomatik tahlil qilish jarayonini murakkablashtiradi. Xususan, so‘z turkumlarini aniqlash (POS tagging) matndagi har bir so‘zning grammatik turkumini teglash vazifasi keyingi barcha tahlillar (sintaktik tahlil, semantik tahlil, mashina tarjimai) uchun asos bo‘lib xizmat qiladi.

Ravish so‘z turkumi o‘zbek tilining muhim mustaqil so‘z turkumlaridan biri hisoblanadi. Ravishlar harakat va holatning belgisini, shuningdek, belgining belgisini bildiradi. Ularning morfologik xususiyatlari (daraja kategoriyasi, maxsus yasovchi qo‘shimchalar), ma’no turlarining xilma-xilligi (holat, o‘rin, payt, daraja-miqdor, maqsad, sabab) va sintaktik vazifalarining (hol, aniqlovchi, kesim) ko‘p qirraliligi ularni avtomatik teglashda qator qiyinchiliklarni keltirib chiqaradi.

O‘zbek tilida ravish so‘z turkumining xususiyatlari. Ravishlar quyidagi asosiy xususiyatlar quyidagicha tavsiflanadi:

1. Ma’no turlari:

- Holat (tarz) ravishlari: *tez, sekin, piyoda, birdan*
- O‘rin ravishlari: *uzoqda, yaqindan, pastda, narida*
- Payt ravishlari: *hozir, kecha, bugun, erta*
- Daraja-miqdor ravishlari: *ancha, sal, kam, juda*
- Maqsad ravishlari: *ataylab, jo‘rttaga, qasddan*
- Sabab ravishlari: *noiloj, ilojsiz, chorasizlikdan*

2. Tuzilishiga ko‘ra turlari:

- Sodda ravishlar: *tez, kech, oz*
- Qo‘shma ravishlar: *har dam, bir yo‘la, har qachon*
- Juft ravishlar: *kechakunduz, qishin-yozin*
- Takroriy ravishlar: *oz-oz, tez-tez, ko‘p-ko‘p*

3. Morfologik xususiyatlari:

- Daraja kategoriyasi: oddiy (*tez*), qiyosiy (*tezroq*), orttirma (*eng tez, juda tez*)
- Maxsus yasovchi qo‘shimchalar: *-cha, -ona, -larcha, -lab, -an*
- O‘zgarimlilik (otlashish hollaridan tashqari)

Ushbu tadqiqotning asosiy maqsadi o‘zbek tilidagi ravish so‘z turkumini avtomatik POS teglashning samarali algoritmlari va modellarini ishlab chiqish hamda ularni qiyosiy tahlil qilishdan iborat.

Korpus va teglar to‘plami. Tadqiqotni amalga oshirish uchun 50000 so‘zdan iborat maxsus korpus shakllantirildi. Korpus tarkibiga badiiy, publitsistik, ilmiy va so‘zlashuv uslubidagi matnlar kiritilgan bo‘lib, ularning har biri mutaxassislar tomonidan qo‘lda teglangan. Korpusdagi ravish so‘zlarining umumiy soni 4320 tani tashkil etadi.

1-jadval. Ravish soʻz turkumi uchun ishlab chiqilgan teglar toʻplami (tagset)

Teg	Maʼnosi	Misol	Korpusdagi chastotasi
ADV-MNR	Holat (tarz) ravishi	tez, sekin, yaxshi	38.5%
ADV-LOC	Oʻrin ravishi	uzoqda, yaqin, pastda	12.3%
ADV-TMP	Payt ravishi	hozir, kecha, erta	24.7%
ADV-DEG	Daraja-miqdor ravishi	juda, ancha, sal	15.2%
ADV-PRP	Maqsad ravishi	ataylab, qasddan	5.1%
ADV-CAU	Sabab ravishi	noiloj, chorasiz	4.2%

Metodlar.

I. Qoidalarga asoslangan (Rule-based) algoritim

Qoidalarga asoslangan yondashuvda ravishlarni aniqlash uchun quyidagi bosqichlardan iborat algoritim ishlab chiqildi:

1-bosqich. Tokenlash va morfologik tahlil

- Matnni soʻzlarga (tokenlarga) ajratish
- Soʻzlarning asos va qoʻshimchalarini ajratish

2-bosqich. Lugʻat asosida tekshirish

- Tub ravishlar bazasi bilan solishtirish (1200 ta tub ravish)
- Agar soʻz bazada mavjud boʻlsa, teg berish

3-bosqich. Affiks qoidalari

Ravish yasovchi qoʻshimchalar asosida teglash:

Agar soʻz quyidagi qoʻshimchalardan biriga ega boʻlsa: *-cha, -ona, -larcha, -lab, -an, -day, -dek, -siz, -chan* ravish (ADV) deb belgilanadi.

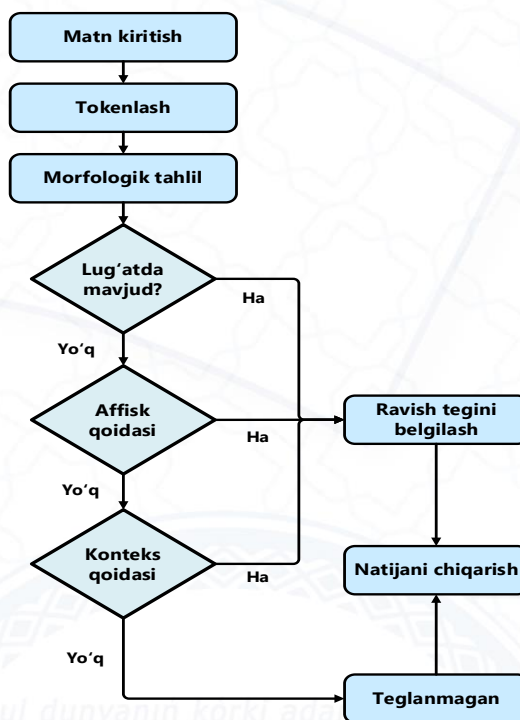
4-bosqich: Kontekstual qoidalar

Ravishlarning boshqa soʻz turkumlaridan (ayniqsa, sifat va otdan) farqlash uchun kontekstual qoidalar qoʻllanadi:

Agar soʻz feʼlga bogʻlangan boʻlsa va soʻz ot vazifasida kelmagan boʻlsa u holda soʻz ADV deb belgilanadi.

5-bosqich: Aniqlanmagan holatlar

Yuqoridagi qoidalar bilan aniqlanmagan soʻzlar uchun tegi qoʻyilmaydi.



1-rasm. Qoidalarga asoslangan yondashuv asosida teglash algoritmining blok-sxemasi

II. Statistika modellar

Yashirin Markov modeli (HMM)

HMM asosidagi POS teglash tizimi quyidagi ehtimollik modeliga asoslanadi:

Berilgan soʻzlar ketma-ketligi $W = w_1, w_2, \dots, w_n$ uchun eng ehtimolli teglar ketma-ketligi $T = t_1, t_2, \dots, t_n$ quyidagi formula bilan topiladi:

$$T^* = \operatorname{argmax}_T P(T | W) = \operatorname{argmax}_T P(W | T)P(T)$$

Markov modelida bu quyidagicha hisoblanadi:

$$T^* \approx \operatorname{argmax}_T \prod_{i=1}^n P(w_i | t_i) \cdot P(t_i | t_{i-1})$$

Bu yerda:

- $P(t_i | t_{i-1})$ – oʻtish ehtimoli (transition probability)
- $P(w_i | t_i)$ – emissiya ehtimoli (emission probability)

2-jadval. HMM modeli uchun oʻtish ehtimollari matritsasi (qisqartirilgan)

Teg	ADV-MNR	ADV-LOC	ADV-TMP	ADV-DEG	Ot	Sifat
ADV-MNR	0.25	0.10	0.15	0.20	0.15	0.15
ADV-LOC	0.20	0.15	0.25	0.10	0.20	0.10
ADV-TMP	0.30	0.10	0.20	0.15	0.15	0.10
Ot	0.35	0.15	0.20	0.05	0.15	0.10

Viterbi algoritmi yordamida eng ehtimolli teglar ketma-ketligi topiladi. Algoritmning asosiy bosqichlari:

1. **Initializatsiya:** $\delta_1(t) = \pi_t \cdot P(w_1 | t)$
2. **Rekursiya:** $\delta_i(t) = \max_s [\delta_{i-1}(s) \cdot P(t | s)] \cdot P(w_i | t)$

3. Tugatish: $T_n^* = \operatorname{argmax}_t \delta_n(t)$

4. Orqaga qaytish: $T^*i = \psi_i + 1(T_{i+1}^*)$

Conditional Random Fields (CRF)

HMM dan farqli o'laroq, CRF butun so'zlar ketma-ketligi uchun shartli ehtimollikni modellashtiradi. CRF modelida quyidagi xususiyat funksiyalaridan foydalanildi:

- So'zning o'zi va uning n-gramlari
- So'zning qo'shimchalari (oxirgi 2-4 harf)
- Oldingi va keyingi so'zlar
- So'zning kichik/katta harf bilan boshlanishi
- So'zda raqam mavjudligi

III. Gibrid model

Aralash modelda qoidalarga asoslangan va statistik yondashuvlarning afzalliklari birlashtirildi. Model quyidagi bosqichlardan iborat:

1-bosqich: Qoidalarga asoslangan tahlil – aniq qoidalar asosida ravishlarni teglash

2-bosqich: HMM/CRF tahlili – qoidaviy tahlilda aniqlanmagan yoki shubhali holatlarni statistik usulda teglash

3-bosqich: Post-processing – qoidalar asosida natijalarni tekshirish va tuzatish

Natijalar. Modellarning qiyosiy samaradorligi. Ishlab chiqilgan modellar 10000 so'zdan iborat test korpusida sinovdan o'tkazildi. Quyidagi jadvalda har bir modelling umumiy aniqlik darajasi (accuracy) va ravish so'z turkumi uchun alohida aniqlik (precision, recall, F1-measure) ko'rsatkichlari keltirilgan.

3-jadval. Modellarning qiyosiy samaradorlik ko'rsatkichlari

Model	Umumiy aniqlik	Ravish uchun Precision	Ravish uchun Recall	Ravish uchun F1
Qoidalarga asoslangan model	87.3%	89.7%	85.2%	87.4%
HMM modeli	90.1%	91.3%	89.8%	90.5%
CRF modeli	91.8%	92.5%	91.2%	91.8%
Gibrid model	93.5%	94.2%	92.9%	93.5%

Ravish turlari bo'yicha tahlil aralash (Gibrid) modelda ravishlarning ma'naviy turlarini aniqlashdagi samaradorligi quyidagi jadvalda keltirilgan.

4-jadval. Ravish turlari bo'yicha aniqlik ko'rsatkichlari (aralash model)

Ravish turi	Precision	Recall	F1-measure	Korpusdagi ulushi
Holat (tarz) ravishlari	95.3%	94.8%	95.0%	38.5%
O'rin ravishlari	92.1%	90.5%	91.3%	12.3%
Payt ravishlari	96.2%	95.7%	95.9%	24.7%
Daraja-miqdor ravishlari	94.8%	93.2%	94.0%	15.2%
Maqsad ravishlari	88.5%	85.3%	86.9%	5.1%
Sabab	86.7%	83.1%	84.9%	4.2%

ravishlari

Xatolik tahlili. Aralash (Gibrid) modelda yo‘l qo‘yilgan xatoliklarning tahlili shuni ko‘rsatdiki, eng ko‘p uchraydigan xatolik turlari quyidagilardan iborat:

5-jadval. Xatolik turlari va ularning ulushi

Xatolik turi	Tavsifi	Ulushi
Sifat-ravish almashinuvi	Sifatlarni ravish deb yoki aksincha teglash	42.3%
Ot-ravish almashinuvi	Otlashgan ravishlarni aniqlashdagi xatolar	23.7%
Ko‘p ma’noli so‘zlar	Kontekstga qarab ma’nosi o‘zgaradigan so‘zlar	18.5%
Qo‘shma ravishlar	Qo‘shma va juft ravishlarni ajratishdagi xatolar	10.2%
Boshqa xatolar	Yuqoridagilarga kirmaydigan xatolar	5.3%

HMM va CRF modellarning kontekst uzunligiga (oldindan va keyindan qaraladigan so‘zlar soni) bog‘liqligi o‘rganildi va jadvalda keltirilgan.

6-jadval: Kontekst uzunligining aniqlikka ta’siri

Kontekst uzunligi	HMM aniqlik	CRF aniqlik
1 so‘z (oldingi)	87.5%	88.2%
2 so‘z (oldingi)	89.8%	90.5%
3 so‘z (oldingi)	90.1%	91.4%
2 so‘z (oldingi + keyingi)	90.8%	91.8%
3 so‘z (oldingi + keyingi)	91.0%	92.1%

Natijalar tahlili. Tadqiqot natijalari shuni ko‘rsatdiki, o‘zbek tilidagi ravish so‘z turkumini POS teglashda gibrid yondashuv eng yuqori samaradorlikka erishdi (93.5% F1-measure). Bu ko‘rsatkich sof qoidaviy modeldan (87.4%) va sof HMM modelidan (90.5%) sezilarli darajada yuqori.

Qoidalarga asoslangan modelling nisbatan past ko‘rsatkichi (87.4%) o‘zbek tilidagi ravishlarning morfologik va sintaktik xususiyatlarining murakkabligi bilan izohlanadi. Xususan, sifat va ravish o‘rtasidagi chegaraning noaniqligi, ko‘p ma’noli so‘zlarning mavjudligi qoidalarga asoslangan yondashuvda to‘liq qamrab olinmagan. Shu bilan birga, qoidaviy modelning afzalligi kam resursli sharoitda ham ishlay oladi va tushunarli qoidalarga asoslanadi.

HMM modeli (90.5%) qoidalarga asoslangan modeldan yuqori natija ko‘rsatdi, bu statistik yondashuvning kontekstual ma’lumotlardan samarali foydalana olishi bilan bog‘liq. HMM ayniqsa, ko‘p ma’noli so‘zlarni teglashda yaxshi natija berdi. Biroq HMMning cheklovi faqat oldingi tegga asoslanadi va keyingi kontekstni hisobga olmaydi.

CRF modeli (91.8%) HMMdan yuqori natija ko‘rsatdi, chunki u butun kontekstni (oldingi va keyingi so‘zlar) hisobga oladi hamda murakkab xususiyat funksiyalaridan foydalanish imkonini beradi. Kontekst uzunligi orttirilgan sari aniqlik oshib boradi, ammo 3 so‘zdan keyin bu o‘sish sekinlashadi.

Aralash modelning yuqori samaradorligi (93.5%) uning ikkala yondashuvning afzalliklarini birlashtira olishi bilan izohlanadi. Qoidalarga asoslangan qism aniq va istisnosiz holatlarni tez va ishonchli teglaydi, statistik qism esa murakkab va noaniq holatlarni kontekst asosida hal qiladi.

Ravish turlari bo'yicha farqlar:

– Ravishlarning ma'no turlari bo'yicha tahlil (4-jadval) qiziqarli natijalarni ko'rsatdi. Eng yuqori aniqlik payt ravishlarida (95.9%) kuzatildi. Buning sababi payt ravishlarining ko'pchiligi (kecha, bugun, erta, hozir) lug'aviy birliklar sifatida oson aniqlanadi va ularning kontekstual qo'llanilishi nisbatan bir xil.

– Holat ravishlari ham yuqori aniqlik (95.0%) ko'rsatdi, ammo bu guruhda sifat-ravish almashinuvi muammosi mavjud. Masalan, "yaxshi" so'zi sifat sifatida ("yaxshi kitob") va ravish sifatida ("yaxshi o'qiydi") qo'llanishi mumkin. Model kontekst orqali bu farqni aniqlay olgan.

– Daraja-miqdor ravishlari (94.0%) ham nisbatan yuqori aniqlik ko'rsatdi, chunki ularning ko'pchiligi (juda, ancha, sal) o'ziga xos leksik birliklar hisoblanadi.

– O'rin ravishlarida aniqlik nisbatan pastroq (91.3%). Bu ularning ba'zilar (uzoqda, yaqinda) o'rin-payt kelishidagi otlar bilan o'xshash shaklga ega ekanligi bilan bog'liq.

– Eng past ko'rsatkich maqsad (86.9%) va sabab (84.9%) ravishlarida kuzatildi. Bu guruhlarning korpusdagi kamchastotaligi (mos ravishda 5.1% va 4.2%) va ularning shakllanishidagi murakkablik (ko'pincha harakat nomi yoki otlardan yasalishi) bilan izohlanadi.

Tadqiqotda quyidagi xatolikni uchratish mumkin. Xatolik tahlili (5-jadval) eng katta muammo sifat-ravish almashinuvi ekanligini ko'rsatdi (42.3%). Bu ikki turkum o'rtasidagi chegara ba'zi hollarda noaniq bo'lishi bilan bog'liq. Masalan, sifatlar ham ravishlar ham belgi bildiradi, farq shundaki, sifatlar predmet belgisini, ravishlar esa harakat belgisini bildiradi. Bu farqni avtomatik aniqlash uchun chuqur sintaktik tahlil talab etiladi.

Ot-ravish almashinuvi (23.7%) asosan otlashgan ravishlar va ravish vazifasida kelgan otlar bilan bog'liq. Masalan, "kechasi" so'zi ot ("kechasi keldi") va ravish ("kechasi ishlaydi") vazifasida kelishi mumkin.

Ko'p ma'noli so'zlar (18.5%) kontekstga qarab turli ma'no va turkumlarda kelishi mumkin. Masalan, "tez" so'zi ravish ("tez yugurdi"), sifat ("tez poyezd") va ot ("tez oqimi") vazifalarida keladi.

Qo'shma va juft ravishlarni aniqlashdagi xatolar (10.2%) ularning yozilishidagi variantlilik (qo'shib, chiziqcha bilan yoki alohida yozilishi) bilan bog'liq.

Tadqiqotning asosiy cheklovlari quyidagilardan iborat:

1. **Korpus hajmi.** 50000 so'zlik korpus o'zbek tilining barcha uslub va janrlarini to'liq qamrab olmaydi. Kengaytirilgan korpus modellarning umumlashtirish qobiliyatini oshiradi.

2. **Teglar to'plami.** Ishlab chiqilgan teglar to'plami (6 teg) ravishlarning barcha o'ziga xos grammatik xususiyatlarini (masalan, daraja kategoriyasi) to'liq aks ettirmaydi.

3. **O'xshash turkumlar.** Sifat va ravish o'rtasidagi farqlashda hali ham muammolar mavjud.

Xulosa. Ushbu tadqiqotda o‘zbek tilidagi ravish so‘z turkumini POS teglashning algoritmlari va modellari ishlab chiqildi va qiyosiy tahlil qilindi. Tadqiqot natijasida quyidagi asosiy xulosalarga kelindi:

1. O‘zbek tilidagi ravishlarni avtomatik teglashda gibrid yondashuv qoidalarga asoslangan va statistik modellarning uyg‘unligi eng yuqori samaradorlikni ko‘rsatadi (93.5% F1-measure).

2. Ravishlarning ma‘no turlari bo‘yicha teglash aniqligi turlicha: payt ravishlari eng yuqori (95.9%), sabab ravishlari esa eng past (84.9%) aniqlikka ega.

3. Asosiy xatolik manbalari sifat-ravish almashinuvi (42.3%), ot-ravish almashinuvi (23.7%) va ko‘p ma‘noli so‘zlar (18.5%) bilan bog‘liq.

4. Kontekst uzunligi model samaradorligiga ta‘sir qiladi: 3 so‘zlik kontekst (oldingi va keyingi so‘zlar) optimal hisoblanadi.

5. CRF modellari HMM modellaridan yuqori aniqlik ko‘rsatadi, chunki ular murakkab xususiyat funksiyalaridan foydalana oladi.

Ishlab chiqilgan modellar va teglar to‘plami o‘zbek tili uchun keyingi NLP vazifalari (sintaktik parsing, semantik tahlil, mashina tarjimai) uchun asos bo‘lib xizmat qilishi mumkin. Shuningdek, ushbu tadqiqot natijalari boshqa agglyutinativ tillar (xususan, turkiy tillar) uchun POS teglash tizimlarini yaratishda ham qo‘llanilishi mumkin.

Foydalanilgan adabiyotlar ro‘yxati:

1. Bobojonova, L., Akhundjanova, A., Ostheimer, P. S., & Fellenz, S. (2025, January). BBPOS: BERT-based part-of-speech tagging for Uzbek. In *Proceedings of the First Workshop on Language Models for Low-Resource Languages* (pp. 287-293).
2. Inamzhanovna, M. M., Mirdjonovna, K. S., & Qizi, E. G. Z. B. (2025, September). Automatic POS Tagging of Verbs in the Uzbek Language and the Algorithm for Identifying Person-Number and Tense Categories. In *2025 10th International Conference on Computer Science and Engineering (UBMK)* (pp. 1639-1643). IEEE.
3. Mansurova, S. (2025). Sifat so‘z turkumini grammatik POS teglash. *Computer linguistics: problems, solutions, prospects, 1*(1).
4. O‘zbek tilining izohli lug‘ati. Ravish // Izoh.uz. – 2024.
5. Sharipov, M., Abjalova, M., & Sobirov, O. (2025, September). Design and Implementation of a Markov Model-Based POS Tagger for the Uzbek Language. In *2025 10th International Conference on Computer Science and Engineering (UBMK)* (pp. 343-347). IEEE.
6. Sharipov, M., & Jumaniyozova, Z. (2023). O‘zbek tilidagi matnlardan sifat, ravish, yordamchi so‘z va oraliq so‘zlar turkumiga oid so‘zlarni aniqlash. *Computer linguistics: problems, solutions, prospects, 1*(1).
7. Sharipov, M., Kuriyozov, E., Yuldashev, O., & Sobirov, O. (2023). UzbekTagger: The rule-based POS tagger for Uzbek language. *arXiv preprint arXiv:2301.12711*.
8. Xudayberganov, N., & Karimova, Z. (2025). Ravish so‘z turkumini grammatik POS teglash algoritmlari: nazariy yondashuv, amaliy tatbiq va kelgusidagi yo‘nalishlar. *Computer linguistics: problems, solutions, prospects, 1*(1).