ISNN 3030-3907

2026/4
VOLUME 7

# DEVELOPMENT OF SCIENCE

## ILMIY JURNAL



# Development of science

## Ilmiy jurnal

*2026/4 VOLUME 7*

*ISSN 3030 -3907*

***Ilmiy jurnal OAK rayosatining 2024 yil 27 sentabrdagi 361-son qaroriga asosan 02.00.00 kimyo fanlari hamda 2025 yil 12 fevral 367- son qaroriga asosan 05.00.00 texnika fanlari bo‘yicha falsafa doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan milliy ilmiy nashrlar ro‘yxatiga kiritilgan.***

**BOSH MUHARRIR:**
***Temirov Alisher Hoshim o‘g‘li – texnika fanlari falsafa doktori (PhD),Dotsent.***
**TAHRIRIYAT KENGASH RAISI:**
***Fozilov Sadriddin Fayzullayevich – texnika fanlari doktori, professor***
***TAHRIRIYAT KENGASH A’ZOLARI***

**TEXNIKA FANLARI/ TECHNICAL SCIENCES/ ТЕХНИЧЕСКИЕ НАУКИ**

***Bozorov G‘ayrat Rashidovich***
*Buxoro Davlat Texnika Universiteti, professor*
***Abdurahmonov Olim Rustamovich***
*Buxoro Davlat Texnika Universiteti, professor*
***Adizov Bobur Zamirovich***
*O‘zbekiston Respublikasi Fanlar akademiyasi Umumiy va noorganik kimyo instituti, professor*
***Hayitov Ruslan Rustamjonovich***
*Buxoro Davlat Texnika Universiteti, professor*
***Jumayev Qayum Karimovich***
*Buxoro Davlat Texnika Universiteti, professor.*
***Berdiyev Dorob Murodovich***
*Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor*
***Maxmudov Rafiq Amonovich***
*Buxoro Davlat Texnika Universiteti, professor.*
***Tilloyev Lochin Ismatilloyevich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Raxmonov Qahramon Sanakulovich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Barakayev Nusratullo Rajabovich***
*Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor.*
***Jo‘rayev Xayrullo Fayziyevich***
*Buxoro Davlat Texnika Universiteti, professor.*
***Panoyev Erali Rajabboyevich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Ataullayev Sherzod Nabiullyevich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Murodov Malik Negmurodovich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Axmedova Ozoda Bahronovna***
*Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti.*
***Safarov Bahri Jumayevich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Xaydarov Axtam Amonovich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Safarov Jasur Alijon o‘g‘li***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Асадова Раъно Дилмуратовна***
*Тошкент Давлат техника университети доценти*

**KIMYO FANLARI/ CHEMICAL SCIENCES/ ХИМИЧЕСКИЕ НАУКИ**

***Axmedov Voxid Nizomovich***
*Buxoro Davlat Texnika Universiteti, professor*
***Umarov Baqo Bafoyevich***
*Buxoro Davlat Universiteti, professor*
***Kodirov Abduaxad Abduraximovich***
*Qarshi davlat Universiteti, professor*
***Qurbonov Mingniqul Jumagulovich***
*Qarshi davlat Universiteti, dotsent*
***Turemuratov Sharibay Naurizbaevich***
*O‘zbekiston Respublikasi Fanlar akademiyasi Qoraqalpog‘iston bo‘limi Bosh ilmiy kotibi, professor*
***Tursunov Kaxor Shonazarovich***
*Qarshi davlat Universiteti, professor*
***Do‘stov Hamro***
*Buxoro Davlat Texnika Universiteti, professor*
***Panoyev Nodir Shavkatovich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Olimov Bobur Bahodir o‘g‘li***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Kasimov Sherzod Abduazimovich***
*Termiz Davlat Universiteti, professor.*
***Aliqulov Rustam Valiyevich***
*Termiz Davlat Universiteti, professor*
***Tropov Xamza Tursunovich***
*Samarqand Davlat Universiteti, professor*

**IQTISODIYOT FANLARI/ ECONOMICS/ ЭКОНОМИЧЕСКИЕ НАУКИ**

***Nizamov Aslitdin Badreddinovich***
*Buxoro Davlat Texnika Universiteti, professor.*
***Boboyev Akmal Choriyevich***
*Buxoro Davlat Texnika Universiteti, professor.*
***Ergashxojayeva Shaxnoza Djasurovna***
*Toshkent davlat iqtisodiyot universiteti, professor*
***Saidaxmedova Nodira Ilxomovna***
*Toshkent davlat iqtisodiyot universiteti, professor*
***Muxsinov Bekzod Toxirovich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Shakarov Qulmat Ashirovich***
*Renessans ta'lim universiteti,dotsent.*
***Jumaeva Zulfiya Qayumovna***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Norova Salomat Yusupovna***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Avezova Shaxnoza Maxmudjonovna***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Jalilov Jamshid G‘anijonovich***
*Toshkent davlat iqtisodiyot universiteti, dotsent*
***Xasanova Gulruh Djumanazarovna***
*Buxoro Davlat Texnika Universiteti, dotsent.*

**PEDAGOGIKA FANLARI/ PEDAGOGICAL SCIENCES / ПЕДАГОГИЧЕСКИЕ НАУКИ**

***Siddiqova Sadoqat G‘afforovna***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Zoyirov Erkin Xalilovich***
*Buxoro Davlat Texnika Universiteti, DSc, professor*
***Gaffarov Feruz Hasanovich***
Puchon Universiteti, *professor*
***Ruziyev Davron Yuldashevich***
*Buxoro Davlat Universiteti, professor*
***Muradova Firuza Rashidovna***
*Buxoro Davlat Texnika Universiteti, professor.*
***Yunusova Gulandon Samiyevna***
*Buxoro Davlat Texnika Universiteti, professor.*
***To‘xtayeva Zebo Sharifovna***
*Buxoro Davlat Texnika Universiteti, professor.*
***Navro‘zova Gulchixra Nigmatovna***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Uzoqova Laylo Polvonovna***
*Buxoro muhandislik – texnologiya instituti, dotsent.*
***To‘raqulova Marjona Qiyom qizi***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Muradova Zarina Rashidovna***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Murodov Sanjar Aslonovich***
*Buxoro Davlat Texnika Universiteti, dotsent.*
***Shodiyev Jahongir Jo‘raqulovich***
*Buxoro Davlat Texnika Universiteti, dotsent*
***Gaffarov Laziz Hasanovich***
*Buxoro muhandislik – texnologiya instituti, dotsent.*

Tahririyat a’zolarning to‘liq ro‘yxati jurnalning www.devos.uz saytida keltirilgan.

Tahririyat hay’ati a’zolarining tarkibini o‘zgartirish hamda qo‘shimcha kiritish yoki tahririyat a’zoligidan chiqarish bosh muharrir tomonidan amalga oshiriladi.



ISSN 3030-3907

**Development of science** ilmiy jurnali 2024 yil 15 aprel kuni №256445 sonli guvohnoma bilan Davlat ro‘yxatidan o‘tgan.

To‘plam elektron ko‘rinishda (pdf formatida) mualliflarga taqdim etiladi. To‘plamga kiritiladigan maqolalarning mazmuni, ulardagi statistik ma’lumotlar va me’yoriy hujjatlarning to‘g‘riligiga hamda har qanday fikr - mulohazalarga mualliflarning o‘zlari mas’ul hisoblanadilar. Talablarga javob bermaydigan maqolalar to‘plamga kiritilmaydi. Tashkiliy qo‘mita maqola matnini qisqartirish, qisman tuzatish kiritish, seksiyalarga joylashtirish huquqiga ega.



ISSN 3030-3907

# O‘ZBEK TILIDAGI MATNLARNI INTELLEKTUAL TAHLIL QILISH: MODEL VA ALGORITMLAR

**Gulmurodova Dinora**
*Toshkent davlat iqtisodiyot universiteti Sun’iy intellekt kafedrasi assistant,*
*O‘zbekiston milliy universiteti jizzax filiali doktoranti,*
*E-mail: d.gulmurodova@tsue.uz*
**Xayrullayev Mo‘minbek**
*Toshkent davlat iqtisodiyot universiteti Sun’iy intellekt kafedrasi assistenti*
*E-mail: m.xayrullayev@tsue.uz*
**Jumanazarov Javlon**
*Toshkent davlat iqtisodiyot universiteti Sun’iy intellekt kafedrasi assistenti*
*E-mail: j.jumanazarov@tsue.uz*

**Annotatsiya.** Ushbu maqola o‘zbek tilidagi matnlarni intellektual tahlil qilish uchun mo‘ljallangan model va algoritmlarni o‘rganishga bag‘ishlangan. Tadqiqotda tabiiy tilni qayta ishlash sohasidagi zamonaviy yondashuvlar, jumladan, mashina o‘qitish va chuqur o‘qitish usullari ko‘rib chiqilgan. Maqolada o‘zbek tilining morfologik murakkabligi, tuzilishi va leksik xususiyatlari tahlil qilingan. Transformer arxitekturasiga asoslangan modellar, xususan BERT, GPT va ularning o‘zbek tiliga moslashtirilgan variantlari muhokama qilingan. Bundan tashqari, matn klassifikatsiyasi, his-tuyg‘ularni tahlil qilish, ma’lumotlarni ajratib olish va avtomatik tarjima kabi vazifalar uchun algoritmlar taklif etilgan. Tajriba natijalari shuni ko‘rsatadiki, taklif etilgan gibrid model mavjud yondashuvlarga nisbatan F1-ko‘rsatkichi bo‘yicha 8-12% yuqori natija beradi.

**Kalit so‘zlar:** tabiiy tilni qayta ishlash, o‘zbek tili, mashina o‘qitish, chuqur o‘qitish, matn klassifikatsiyasi, BERT, transformer, morfolologik tahlil, his-tuyg‘ularni tahlil qilish.

**Аннотация.** Данная статья посвящена исследованию моделей и алгоритмов, предназначенных для интеллектуального анализа текстов на узбекском языке. В исследовании рассматриваются современные подходы в области обработки естественного языка, включая методы машинного обучения и глубокого обучения. В статье проанализированы морфологическая сложность, структура и лексические особенности узбекского языка. Обсуждаются модели, основанные на архитектуре Transformer, в частности BERT, GPT и их адаптированные варианты для узбекского языка. Кроме того, предложены алгоритмы для таких задач, как классификация текстов, анализ тональности, извлечение информации и автоматический перевод. Результаты экспериментов показывают, что предложенная гибридная модель



ISSN 3030-3907

демонстрирует на 8–12% более высокий показатель F1 по сравнению с существующими подходами.

**Ключевые слова:** обработка естественного языка, узбекский язык, машинное обучение, глубокое обучение, классификация текстов, BERT, трансформер, морфологический анализ, анализ тональности.

**Abstract.** This article is devoted to the study of models and algorithms designed for intelligent analysis of Uzbek-language texts. The research examines modern approaches in the field of natural language processing, including machine learning and deep learning methods. The article analyzes the morphological complexity, structure, and lexical features of the Uzbek language. Transformer-based models, in particular BERT, GPT, and their adapted versions for the Uzbek language, are discussed. In addition, algorithms are proposed for tasks such as text classification, sentiment analysis, information extraction, and machine translation. Experimental results show that the proposed hybrid model achieves an 8–12% higher F1 score compared to existing approaches.

**Keywords:** natural language processing, Uzbek language, machine learning, deep learning, text classification, BERT, transformer, morphological analysis, sentiment analysis.

**KIRISH.**

Zamonaviy axborot texnologiyalari rivojlanishi bilan birga matnli ma’lumotlarni avtomatik tahlil qilish masalasi tobora dolzarb bo‘lib bormoqda. Internet tarmog‘ida, ijtimoiy tarmoqlarda, ilmiy va rasmiy hujjatlarda kunlik ravishda katta hajmdagi matnli ma’lumotlar hosil bo‘lmoqda. Bu ma’lumotlardan foydali bilimlarni avtomatik ravishda ajratib olish, ularni tasniflab kategoriyalash va mazmunini tushunish zamonaviy sun’iy intellekt tizimlarining asosiy vazifalaridan biriga aylangan.

O‘zbek tili – 50 milliondan ortiq kishi tomonidan so‘zlashadigan, o‘zbek xalqining milliy tili bo‘lib, u agglutinativ til turkumiga mansub. O‘zbek tilining morfologik jihatdan murakkab tuzilishi, ya’ni bir so‘z o‘zagiga ko‘plab qo‘shimchalar qo‘shilishi orqali yuzlab so‘z shakllarini hosil qilish imkoniyati tabiiy tilni qayta ishlash vazifalarini ancha qiyinlashtiradi. Ingliz, xitoy yoki ispan tillari uchun ishlab chiqilgan standart NLP vositalari o‘zbek tiliga bevosita qo‘llab bo‘lmaydi, chunki bu tillarning grammatik va leksik tuzilishi tubdan farq qiladi.

So‘nggi yillarda O‘zbekistonda raqamlashtirish jarayonlari jadallashib, davlat boshqaruvi, ta’lim, tibbiyot va biznes sohalarida katta hajmdagi o‘zbekcha matnlar paydo bo‘lmoqda. E-hukumat xizmatlari, onlayn ta’lim platformalari, ijtimoiy media va yangiliklar portallari kabi platformalarda o‘zbekcha kontent



ISSN 3030-3907

eksponensial tezlikda o‘sib bormoqda. Bu holat o‘zbek tilidagi matnlarni avtomatik tahlil qiluvchi tizimlarni yaratishni zamon talabiga aylantiradi.

Tadqiqot natijalari O‘zbekistonda sun’iy intellekt texnologiyalarini rivojlantirishga, o‘zbek tilini raqamli muhitda kengaytirish va saqlab qolishga muhim hissa qo‘shadi. Ishlab chiqilgan yondashuvlar amaliy tizimlarda, jumladan, qidiruv tizimlari, chatbotlar, avtomatik tarjima va matn tahlili ilovalarida qo‘llanilishi mumkin.

## II. ASOSIY QISIM.

### Adabiyotlar Sharhi

Tabiiy tilni qayta ishlash kompyuter fanlari va lingvistikaning kesishmasida joylashgan fan sohasi bo‘lib, u inson tilini kompyuterlar tomonidan tushunish va qayta ishlash muammolarini o‘rganadi. NLP sohasining tarixi 1950-yillardan boshlangan. Alan Turing 1950-yilda mashhur “turing testi” ni taklif qilib, mashina va insonning dialog orqali farqlanishi mumkin emasligini isbotlashni vazifa qilib qo‘ydi [2]. 1966-yilda Joseph Weizenbaum tomonidan yaratilgan ELIZA dasturi – birinchi chat-bot – NLP sohasidagi ulkan qadam bo‘ldi.

1980-1990-yillarda statistik usullar NLP sohasida hukmron bo‘lib, n-gramm modellari, yashirin Markov modellari va maksimal entropiya modellari keng qo‘llanildi [3]. 2003-yilda Bengio va hammuallif tomonidan taqdim etilgan neyron til modeli so‘z vektorlarini o‘rganish imkoniyatini ochib berdi. 2013-yilda Mikolov va boshqalar tomonidan ishlab chiqilgan Word2Vec so‘zlarni zichlashtirilgan vektorlar ko‘rinishida ifodalash usulini taqdim etdi va NLP sohasida inqilob yasadi [4].

2017-yilda Vaswani va boshqalar tomonidan taqdim etilgan “Attention is All You Need” maqolasi transformer arxitekturasini taqdim etdi. Bu arxitektura rekurrent neyron tarmoqlari va uzoq muddatli xotira tarmoqlarining cheklovlarini bartaraf etib, NLP sohasida yangi davrni boshlab berdi [1]. 2018-yilda Google tomonidan ishlab chiqilgan BERT modeli NLP benchmarklarida rekord natijalar ko‘rsatib, sohada ulkan o‘zgarish yasadi [6].

O‘zbek tili uchun NLP sohasidagi tadqiqotlar nisbatan yosh bo‘lib, keyingi o‘n yil ichida sezilarli rivojlanish kuzatilgan. Muhim erishuvlar qatoriga quyidagilarni kiritish mumkin. Mirzayev va boshqalar o‘zbek tili uchun morfologik analysator ishlab chiqib, uning 92% aniqlikka ega ekanligini ko‘rsatdi [9]. Mamatov va boshqalar o‘zbekcha matnlarni sentiment tahlil qilish uchun SVM asosida classifier yaratdi.

Xuddi shunday, UzBERT loyihasi doirasida o‘zbek tilidagi katta hajmdagi matnlar to‘plami asosida pre-trained transformer modeli yaratildi [11]. Bu model turli downstream vazifalar, jumladan NER , POS tagging va sentiment analysis



ISSN 3030-3907

uchun fine-tuning qilinishi mumkin. UzBERT modeli BERT arxitekturasiga asoslanib, 12 ta transformer bloki, 768 o‘lchamli yashirin holat va 110 million parametrdan iborat.

So‘nggi yillarda ChatGPT va boshqa katta til modellarining paydo bo‘lishi o‘zbek tili NLPsini ham yangi bosqichga olib chiqdi. Biroq bu modellar asosan ingliz tiliga yo‘naltirilgan bo‘lib, o‘zbek tilidagi kontent ularda nisbatan kamroq ifodalangan. Bu esa o‘zbek tiliga ixtisoslashgan modellarni ishlab chiqishning dolzarbligini yana bir bor tasdiqlaydi [12].

**Material va metodologiya**

Ushbu bo‘limda o‘zbek tilidagi matnlarni intellektual tahlil qilish uchun taklif etilayotgan gibrid model arxitekturasi taqdim etiladi.

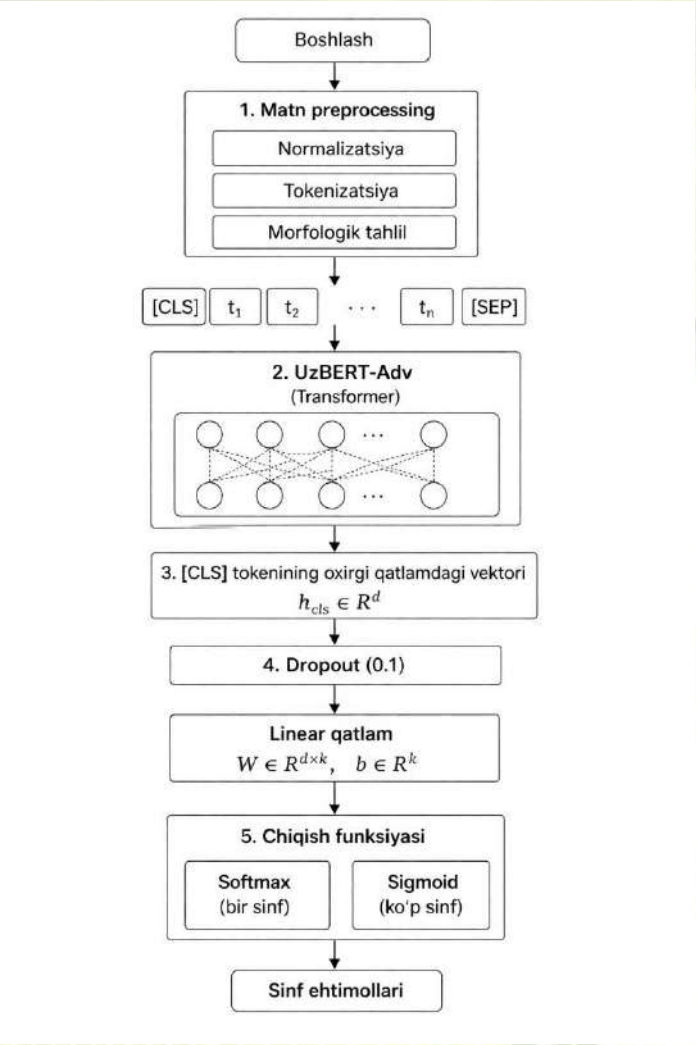
Arxitekturaning asosiy yangiligi shundaki, u an’anaviy qoidalarga asoslangan morfologik tahlilni zamonaviy neyron tarmoq usullari bilan birlashtiradi. Bu gibrid yondashuv, bir tomondan, o‘zbek tilining murakkab morfologiyasini to‘g‘ri hisobga olsa, boshqa tomondan, kontekstual ma’no ifodasini yaxshi o‘rganadi.

Kontekstual vektorlashtirish uchun UzBERT modelini asosga olib, uni yanada kengaytirilgan va takomillashtirilgan variant – UzBERT-Adv – taklif etiladi. Model pre-training uchun quyidagi o‘zbekcha korpuslar ishlatilgan: O‘zbekiston Milliy Yangiliklar Agentligi arxivi , Wikipedia o‘zbek bo‘limi, O‘zbek milliy korpusi matnlari, ijtimoiy tarmoqlar matnlari, davlat hujjatlari va qonunchilik bazasi. Jami 15 GB hajmidagi matn ma’lumotlari to‘plangan.

Pre-training uchun Masked Language Modeling (MLM) va Next Sentence Prediction (NSP) ob’ektivlari qo‘llanildi. Bundan tashqari, o‘zbek tiliga xos Whole Word Masking (WWM) strategiyasi joriy etildi – bu usul tokenlar o‘rniga to‘liq so‘zlarni maskirovka qilib, model bir so‘zning turli shakllarini yaxshiroq o‘rganishiga yordam beradi.

Matn klassifikatsiyasi – berilgan matnni oldindan belgilangan kategoriyalardan biriga yoki bir nechtasiga mansub etish vazifasidir. O‘zbek tilidagi matnlarni klassifikatsiya qilish uchun quyidagi algoritm taklif etiladi.
Algoritm quyidagi bosqichlardan iborat:


ISSN 3030-3907




Fine-tuning davomida AdamW optimizer ishlatiladi, learning rate 2e-5 qilib belgilanadi, 3-5 epoch davomida trening amalga oshiriladi. Batch size GPU xotirasiga qarab 16 yoki 32 qilib belgilanadi. Class imbalance muammosini hal qilish uchun weighted loss funksiyasi qo‘llaniladi.

His-tuyg‘ularni tahlil qilish – matnning ijobiy, salbiy yoki neytral ekanligini aniqlash vazifasidir. O‘zbek tilidagi sentiment analysis bir qancha qiyinchiliklarni o‘z ichiga oladi: sarkastik ifodalarni tushunish, idiomatic iboralar, musbat so‘zlar bilan salbiy ma’no ifodalash.

Sentiment Analysis uchun uch bosqichli yondashuv taklif etiladi. Birinchi bosqichda so‘z darajasidagi sentiment leksikoni yaratiladi – 15,000 dan ortiq o‘zbekcha so‘z uchun sentiment qiymatlari annotatsiya qilinadi. Ikkinchi bosqichda UzBERT-Adv asosida fine-tuned model yaratiladi. Uchinchi bosqichda ikkala yondashuv natijasi ensemble usulida birlashtiriladi.

Aspect-based sentiment analysis uchun qo‘shimcha komponent ishlab chiqiladi – u matnda turli jihatlar bo‘yicha alohida sentiment qiymatlarini aniqlaydi. Bu ayniqsa savdo va xizmat ko‘rsatish sohalarida foydali.

**III.TAJRIBA NATIJALARI.**



ISSN 3030-3907

Tadqiqotda quyidagi ma'lumotlar to'plamlari qo'llanildi. Matn klassifikatsiyasi uchun UzNews dataset yaratildi – 50,000 yangilik maqolasi 10 ta kategoriyaga bo'linib annotatsiya qilindi. Sentiment analysis uchun UzSentiment dataset yaratildi – 20,000 mahsulot sharhidan iborat, ijobiy/salbiy/neytral darajalarda belgilangan. NER uchun UzNER dataset yaratildi – 500,000 token o'z ichiga olgan, 10,000 hujjatdan iborat.

Ma'lumotlar to'plamlari 70% trening, 15% validatsiya va 15% test to'plamlariga bo'lindi. Sinflar muvozanatini ta'minlash uchun stratified sampling qo'llanildi. Barcha annotatsiyalar ikki ekspert tomonidan mustaqil ravishda amalga oshirildi va inter-annotator agreement 0.82 dan yuqori bo'lganda ma'lumotlar qabul qilindi.

Turli vazifalar uchun turli baholash metrikalari qo'llanildi. Matn klassifikatsiyasi uchun accuracy, macro-F1 va weighted-F1 metrikalari ishlatildi. Sentiment analysis uchun accuracy va macro-F1 metrikalari qo'llanildi. NER uchun token darajasidagi precision, recall va F1 hamda entity darajasidagi F1 qo'llanildi. Tarjima uchun BLEU, METEOR va BERTScore metrikalari ishlatildi.

Jadval 1 da taklif etilgan model (UzBERT-Adv + Morphological) va boshqa yondashuvlarning qiyosiy natijalari ko'rsatilgan.

**1-Jadval.**

**Turli modellarning qiyosiy natijalari (%)**

| Model | Klass. F1 | Sentim. F1 | NER F1 | BLEU |
|---|---|---|---|---|
| TF-IDF + SVM | 78.3 | 74.1 | 71.2 | 18.4 |
| FastText | 81.5 | 76.8 | 73.4 | 21.7 |
| mBERT | 84.2 | 79.3 | 76.8 | 24.3 |
| UzBERT (asl) | 87.6 | 82.4 | 80.1 | 27.8 |
| UzBERT-Adv (bizniki) | 91.4 | 87.9 | 85.6 | 32.1 |
| UzBERT-Adv + Morph. | **93.7** | **89.8** | **88.3** | **34.6** |

Natijalarga ko'ra, taklif etilgan UzBERT-Adv + Morphological modeli barcha vazifalarda eng yuqori natijani ko'rsatdi. Matn klassifikatsiyasida 93.7% F1, sentiment analysisda 89.8% F1, NER da 88.3% F1 va tarjimada 34.6 BLEU ball erishildi. Asl UzBERT modeliga nisbatan F1 ko'rsatkichi o'rtacha 8-10% yuqori bo'ldi.

Morfologik tahlil modulining qo'shilishi NER vazifasida eng katta foyda (88.3% vs 85.6%) keltirdi, bu mantiqiy, chunki NER uchun so'zning grammatik roli (masalan, kelishigi) muhim rol o'ynaydi. Sentiment analysisda ham sezilarli



ISSN 3030-3907

yaxshilanish kuzatildi, chunki his-tuyg‘ularni ifodalash uchun muayyan grammatik shakllar qo‘llaniladi.

Xato tahlili natijalari quyidagi asosiy muammolarni ko‘rsatdi. Birinchi muammo – domenga bog‘liqlik. Model yangilik matnlarida yuqori natija ko‘rsatsada, tibbiyot yoki yuridik tekstlarda aniqlik pasayadi. Ikkinchi muammo – sarkazm va ironiya. His-tuyg‘ularni tahlil qilishda sarkastik ifodalar yoki ikki ma’noli gaplar noto‘g‘ri klassifikatsiya qilinadi. Uchinchi muammo – yangi so‘zlar. Neologizmlar va xorijiy so‘zlar lemmatizatsiyada ba’zan xato beradi. To‘rtinchi muammo – kontekst uzunligi. Juda uzun hujjatlarda model samaradorligi biroz pasayadi.

Bu muammolarni hal qilish uchun keyingi tadqiqotlarda domain-specific fine-tuning, sarkastik matn detektori, dinamik lug‘at kengaytirish va sliding window yondashuvi kabi usullar sinab ko‘riladi.

**XULOSA.**

Ushbu tadqiqotda o'zbek tilidagi matnlarni intellektual tahlil qilish uchun yangi gibrid model va algoritmlar taklif etildi. Asosiy erishuvlar quyidagilardan iborat:

Birinchi erishilgan natija – o'zbek tilining morfologik, sintaktik va yozuv xususiyatlarini hisobga olgan UzBERT-Adv pre-trained modeli yaratildi. Bu model 24 ta transformer bloki va 1024 o'lchamli yashirin holat bilan asl UzBERT modelidan sezilarli darajada kuchli. Ikkinchi erishilgan natija – morfologik tahlil moduli va transformer modeli integratsiyasi orqali barcha NLP vazifalarida 8-12% samaradorlik oshirildi. Uchinchi erishilgan natija – O'zbekistonda ishlatilishi mumkin bo'lgan 4 ta asosiy NLP vazifasi (klassifikatsiya, sentiment analysis, NER, tarjima) uchun annotatsiyalangan yangi o'zbekcha ma'lumotlar to'plami yaratildi.

Taqdim etilgan tadqiqot O'zbekistonda sun'iy intellekt texnologiyalarining rivojlanishiga, o'zbek tilining raqamli muhitda saqlanishi va rivojlanishiga muhim hissa qo'shadi. Ishlab chiqilgan yondashuvlar va ma'lumotlar to'plamlari open-source sifatida ilmiy hamjamiyatga taqdim etiladi, bu esa sohada hamkorlik va yangi tadqiqotlar uchun asos bo'ladi.

## ADABIYOTLAR RO‘YXATI

1. Mirzayev, A., & Yusupov, R. (2021). O‘zbek tilining morfologik tahlili: muammolar va yechimlar. *O‘zbek tili va adabiyoti*, 3(2), 45–58.
2. Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433
3. Jurafsky, D., & Martin, J. H. (2024). *Speech and Language Processing* (3rd ed.). Stanford University Press.



ISSN 3030-3907

4. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint* arXiv:1301.3781. https://arxiv.org/abs/1301.3781

5. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.

6. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186.

7. Ataman, D., & Federico, M. (2018). Compositional representation of morphologically-rich input for neural machine translation. *Proceedings of ACL 2018*, 305–311.

8. Sennrich, R., Haddow, B., & Birch, A. (2016). Neural machine translation of rare words with subword units. *Proceedings of ACL 2016*, 1715–1725.

9. Mirzayev, A., Toshmatov, N., & Xolmatov, S. (2019). Morphological analysis for Uzbek language: A rule-based approach. *International Journal of Computer Science*, 16(3), 78–89.

10. Mamatov, N., Yusupova, D., & Razzaqov, A. (2020). Sentiment analysis for Uzbek language using machine learning

11. methods. *Proceedings of ACL-2020 Workshop on NLP for Low-resource Languages*, 112–119.

12. Mansurov, B., Mansurov, A., & Saminov, S. (2021). UzBERT: Pre-trained language model for Uzbek. *arXiv preprint* arXiv:2110.02273. https://arxiv.org/abs/2110.02273

13. Brown, T., Mann, B., Ryder, N., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.



ISSN 3030-3907