

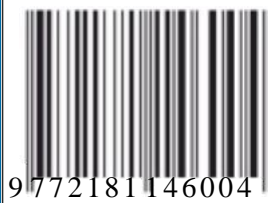
BUXORO DAVLAT UNIVERSITETI ILMIY AXBOROTI



Научный вестник Бухарского государственного университета
Scientific reports of Bukhara State University

5/2025

E-ISSN 2181-1466



9 772181 146004

ISSN 2181-6875



9 772181 687004



@buxdu_uz



@buxdu1



@buxdu1



www.buxdu.uz

5/2025

BUXORO DAVLAT UNIVERSITETI ILMIY AXBOROTI
SCIENTIFIC REPORTS OF BUKHARA STATE UNIVERSITY
НАУЧНЫЙ ВЕСТНИК БУХАРСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Ilmiy-nazariy jurnal
2025, № 5, may

Jurnal 2003-yildan boshlab **filologiya** fanlari bo'yicha, 2015-yildan boshlab **fizika-matematika** fanlari bo'yicha, 2018-yildan boshlab **siyosiy** fanlar bo'yicha, **tarix** fanlari bo'yicha 2023-yil 29-avgustdan boshlab O'zbekiston Respublikasi Oliy ta'lim, fan va innovatsiyalar Vazirligi huzuridagi Oliy attestatsiya komissiyasining dissertatsiya ishlari natijalari yuzasidan ilmiy maqolalar chop etilishi lozim bo'lgan zaruriy nashrlar ro'yxatiga kiritilgan.

Jurnal 2000-yilda tashkil etilgan.

Jurnal 1 yilda 12 marta chiqadi.

Jurnal O'zbekiston matbuot va axborot agentligi Buxoro viloyat matbuot va axborot boshqarmasi tomonidan 2020-yil 24-avgust № 1103-sonli guvohnoma bilan ro'yxatga olingan.

Muassis: Buxoro davlat universiteti

Tahririyat manzili: 200117, O'zbekiston Respublikasi, Buxoro shahri Muhammad Iqbol ko'chasi, 11-uy.

Elektron manzil: nashriyot_buxdu@buxdu.uz

TAHRIR HAY'ATI:

Bosh muharrir: Xamidov Obidjon Xafizovich, iqtisodiyot fanlari doktori, professor

Bosh muharrir o'rinbosari: Rasulov To'liqin Husenovich, fizika-matematika fanlari doktori (DSc), professor

Mas'ul kotib: Shirinova Mexrigiyo Shokirovna, filologiya fanlari bo'yicha falsafa doktori (PhD), dotsent

Kuzmichev Nikolay Dmitriyevich, fizika-matematika fanlari doktori (DSc), professor (N.P. Ogaryov nomidagi Mordova milliy tadqiqot davlat universiteti, Rossiya)

Danova M., filologiya fanlari doktori, professor (Bolgariya)

Margianti S.E., iqtisodiyot fanlari doktori, professor (Indoneziya)

Minin V.V., kimyo fanlari doktori (Rossiya)

Tashqarayev R.A., texnika fanlari doktori (Qozog'iston)

Mo'minov M.E., fizika-matematika fanlari nomzodi (Malayziya)

Mengliyev Baxtiyor Rajabovich, filologiya fanlari doktori, professor

Adizov Baxtiyor Rahmonovich, pedagogika fanlari doktori, professor

Abuzalova Mexriniso Kadirovna, filologiya fanlari doktori, professor

Amonov Muxtor Raxmatovich, texnika fanlari doktori, professor

Barotov Sharif Ramazonovich, psixologiya fanlari doktori, professor, xalqaro psixologiya fanlari akademiyasining haqiqiy a'zosi (akademigi)

Baqoyeva Muhabbat Qayumovna, filologiya fanlari doktori, professor

Bo'riyev Sulaymon Bo'riyevich, biologiya fanlari doktori, professor

Jumayev Rustam G'aniyevich, siyosiy fanlar nomzodi, dotsent

Djurayev Davron Raxmonovich, fizika-matematika fanlari doktori, professor

Durdiyev Durdimurod Qalandarovich, fizika-matematika fanlari doktori, professor

Olimov Shirinboy Sharofovich, pedagogika fanlari doktori, professor

Qahhorov Siddiq Qahhorovich, pedagogika fanlari doktori, professor

Umarov Baqo Bafoyevich, kimyo fanlari doktori, professor

Murodov G'ayrat Nekovich, filologiya fanlari doktori, professor

O'rayeva Darmonoy Saidjonovna, filologiya fanlari doktori, professor

Navro'z-zoda Baxtiyor Nigmatovich, iqtisodiyot fanlari doktori, professor

Hayitov Shodmon Ahmadovich, tarix fanlari doktori, professor

To'rayev Halim Hojiyevich, tarix fanlari doktori, professor

Rasulov Baxtiyor Mamajonovich, tarix fanlari doktori, professor

Eshtayev Alisher Abdug'aniyevich, iqtisodiyot fanlari doktori, professor

Quvvatova Dilrabo Habibovna, filologiya fanlari doktori, professor

Axmedova Shoira Nematovna, filologiya fanlari doktori, professor

Bekova Nazora Jo'rayevna, filologiya fanlari doktori (DSc), professor

Amonova Zilola Qodirovna, filologiya fanlari doktori (DSc), dotsent

Hamroyeva Shahlo Mirjonovna, filologiya fanlari doktori (DSc), dotsent

Nigmatova Lola Xamidovna, filologiya fanlari doktori (DSc), dotsent

Boboyev Feruz Sayfullayevich, tarix fanlari doktori

Jo'rayev Narzulla Qosimovich, siyosiy fanlar doktori, professor

Xolliyev Askar Ergashovich, biologiya fanlari doktori, professor

Artikova Hafiza To'ymurodovna, biologiya fanlari doktori, professor

Hayitov Shavkat Ahmadovich, filologiya fanlari doktori, professor

Qurbonova Gulnoz Negmatovna, pedagogika fanlari doktori (DSc), professor

Ixtiyarova Gulnora Akmalovna, kimyo fanlari doktori, professor

Rasulov Zubaydullo Izomovich, filologiya fanlari doktori (DSc), dotsent

Mirzayev Shavkat Mustaqimovich, texnika fanlari doktori, professor

Samiyev Kamoliddin A'zamovich, texnika fanlari doktori, dotsent

Esanov Husniddin Qurbonovich, biologiya fanlari doktori, dotsent

Zaripov Gulmurot Toxirovich, texnika fanlari nomzodi, professor

Jumayev Jura, fizika-matematika fanlari nomzodi, dotsent

Klichev Qybek Abdurasulovich, tarix fanlari doktori, dotsent

G'aybulayeva Nafisa Izattullayevna, filologiya fanlari doktori (DSc), dotsent

MUNDARIJA *** СОДЕРЖАНИЕ *** CONTENTS		
TILSHUNOSLIK *** LINGUISTICS *** ЯЗЫКОЗНАНИЕ		
Гудзина В.А.	Русские мужские календарные имена в антропонимическом пространстве басен и.а.крылова	4
Шермухамедова Н.А.	Моделирование морфологического анализа на основе окончания слова для узбекского языка	9
Хушманова Ш.И.	Структура концепта: когнитивные и лингвистические аспекты	14
Сулайманов Б.Қ.	Терминология государственного управления восточного происхождения в русском и узбекском языках	18
Каримова В.В.	Влияние гендерных стереотипов на социальные нормы и роли в современном лингвистическом обществе	24
Искаков М.С.	Зоонимы как носители лингвокультурной информации	29
Gadoyeva M.I., Hayrulloeva A.T.	Ommaviy axborot vositalari va badiiy matnlardagi kiyim-kechak obrazlari	33
Xamroyeva Sh.I.	Muallif mushohadasining milliy-madaniy o‘ziga xosligi	37
Saidova N.A.	Global reklama bozorida tarjimaning o‘rni: ispan va o‘zbek tillari misolida	42
Sharipova M.B.	Literonimlarning leksik tizim tarkibidagi o‘rni	47
Saidov A.A.	Roald dal ijodida til o‘yini fenomeni (“the bfg” ertagi misolida)	51
Safarova D.B.	O‘zbek badiiy matnida gaplararo aloqaning ifodalanishi: nazar eshonqulning “oq alanga” hikoyasi misolida	59
Ro‘ziyeva N.K.	Badiiy diskurs va unda rad etish kommunikativ strategiyalarining qo‘llanilishi	64
Abdulkhakova D.Sh.	The evolution and development of cognitive models in cognitive linguistics	69
Asadov A.B.	Preventiv nutqiy aktining sotsiopragmatik xususiyatlari	74
Azimova A.O’., Asadov T.H.	Muallif nutqi neologizmlarining toshpo‘lat ahmad ijodidagi ifodasi	78

Barnoyeva Z.S.	Imom G`azzoliy asarlarida so`filik terminlarining zamonaviy interpretatsiyasi	82
Esirgapov M.N.	The impact of the internet and digital technologies on english	86
Eshto'xtarova B.B.	Linguo-pragmatic analysis of speech acts in english and uzbek languages	91
Burkhanova D.I.	Comparative study of cognitive linguistics in three languages	95
Egamnazarova D.Sh.	Zamonaviy ingliz va o'zbek she'riyatida sintaktik deviatziyalarning badiiy-funksional tahlili	101
Halimova N.H.	The role of informal speech in reflecting cultural changes	105
Hojiyeva N.H.	O'z qatlamga oid meva nomlari xususida	111
Ibraximova D.T.	Ingliz va o'zbek tillariga xos so'z guruhlarining o'rganilishi	115
Isakova Z.Z.	Bilish va baholashning kognitiv bosqichlari	119
Jamoldinova G.N., Sobirova D.A.	Lingua-stylistics features of the texts of english and uzbek news	124
Karimjonova Sh.R.	Vizual kognitiv fe'llarning morfologik va sintaktik xususiyatlari	128
Kendjayeva G.F., Ochilova G.A.	Ingliz va o'zbek maqollaridagi metaforalarning ma'naviy va tarbiyaviy roli	133
Khametov I.R.	Digital dynamics: studying the formation of neologisms in the era of online globalisation	137
Koziva I.K., Faxriddinova M.	A.Qodiriyning "O'tkan kunlar" asaridagi antroponimlarning o'rganilishi	142
Maxsudova N.I.	Antroponim asosli regionimlar haqida ba'zi mulohazalar	149
Najmiddinova S.N.	Siyosiy diskursning turlari va ularning zamonaviy ijtimoiy-siyosiy kontekstdagi ahamiyati	153
Ortiqova H.M.	Shaxs nutqiga yosh xususiyatlarining ta'siri masalasi	157
Navruzova N.X.	Comparative study of connotative meanings in english and uzbek languages	161
Ochilova M.R.	Semantic and stylistic analysis of art terminology in	166

	modern uzbek and english	
Ortiqova H.M.	Verbal assotsiatsiyalarning matn yaratishdagi oʻrni	171
Raxmonova G.N.	Soʻz birikmalari semantik tuzilishida freym andozalaridan foydalanish	175
Inomova D.U.	Sources of the appearance of neologisms and their functions in the process of globalization (english and russian languages)	180
Kilichev B.E., Kozieva I.K.	National names of russia	184
Nazarov S.Sh.	A comparative study of the structural-semantic features of english and uzbek word combinations	194
Sarimsokov X.A.	Sharh va sharhlashga doir nazariy yondashuvlar	198
Baxtiyorova M.B.	Badiiy onimlarning stilistik va konnotativlik xususiyatlari	204
Ataboyev N.B., Abdullayeva N.A.	Corpus analysis of 263 islamic supplications in english and uzbek	208
Isakova N.A., Nafasova G.Sh.	Key concepts in literary translation: theoretical overview	215
Mardiyev B.X.	Shaxs nomlarining oʻrganilish tarixi va u bilan bogʻliq nazariy qarashlar	223
Saidov Z.Sh.	Causes of language interfrence in second language learning	227
Isakova N.A., Abdurahimova M.U.	Loanwords in language evolution: classification, adaptation, and sociolinguistic impact	231
Sodiqova O.Sh.	Belgini anglatuvchi metaforalarning tavsifiy oʻrganiliishi	236
Tagaeva S.U., Urazkulova A.A.	Mutterschaft / onalik borligʻi subfreymlarining frazeologik birliklar vositasida ifodalanishi	241
Xurramova S.X.	Yel-yogʻin otlari tushunchasi va tizimi	246
Saidov Z.Sh.	Explanatory dimensions of language mismatch in the process of second language acquisition	251

УДК:

МОДЕЛИРОВАНИЕ МОРФОЛОГИЧЕСКОГО АНАЛИЗА НА ОСНОВЕ ОКОНЧАНИЯ СЛОВА ДЛЯ УЗБЕКСКОГО ЯЗЫКА

Шермухамедова Насиба Алмасовна,

*Бухарский государственный университет, филологический факультет,
кафедра русского языка и литературы, старший преподаватель
Узбекистан. Бухара.*

Аннотация. Узбекский язык, обладающий агглютинативным характером, образует слова путем сочетания аффиксов с корнями, используя флективные окончания для различных морфологических признаков. Это свойство создает большое количество комбинаций окончаний слов и значительно увеличивает словарный запас и объем данных проблемы разреженности статистических моделей. В этой статье рассматривается модель морфологического анализа, которая включает в себя выделение корней, лемматизацию и извлечение морфологической информации с учетом морфофонетических исключений. Суть модели заключается в разработке полного набора окончаний слов с соответствующей морфологической информацией и дополнительных наборов данных для морфологического анализа. Предложенная модель была протестирована с использованием специально разработанного набора тестов, включающего 5,3 тыс. слов. Точность на уровне слов превысила 91%, что было определено с помощью ручной проверки основы, леммы и исправлений морфологических признаков, проведенных экспертами-лингвистами. Созданный на основе предложенной методологии инструмент доступен в виде пакета Питон с открытым исходным кодом, а также веб-приложения, включающего общедоступный API.

Ключевые слова: узбекский язык, морфологический анализ, морфологическая сегментация, стемминг, лемматизация, связанная морфема, флективное окончание.

MODELING MORPHOLOGICAL ANALYSIS AT THE END OF A WORD FOR THE UZBEK LANGUAGE

Abstract. Uzbek, an agglutinative language, forms words by combining affixes with roots, utilizing inflectional endings for various morphological features. This property makes a large number of combinations of word ending, and greatly increases the word-vocabulary size, and data sparseness problems for statistical models. This paper discusses a morphological analyzing model which includes stemming, lemmatizing and extraction of morphological information considering morpho - phonetic exceptions. A main point of the model involves developing a complete set of word-ending with assign morphological information, and additional datasets for morphological analysis. The proposed model was evaluated using a curated test set comprising 5.3K words. It achieved a word-level accuracy over 91%, as determined through manual verification of stem, lemma, and morphological feature corrections conducted by linguistic experts. The created tool based on the proposed methodology is available as an open-source Python package, as well as a web-based application including a public API.

Keywords: Uzbek language, morphological analyzing, morphological segmentation, stemming, lemmatizing, bound morpheme, inflectional ending.

O'ZBEK TILI UCHUN SO'ZNING ASOS OXIRIDA MORFOLOGIK TAHLILNI MODELLASHTIRISH

Annotatsiya. Agglyutinativ xususiyatga ega bo'lgan o'zbek tilida affikslarni o'zak bilan bog'lash, turli morfologik belgilarga ko'ra qo'shimchalar qo'llash orqali so'z hosil qiladi. Bu xususiyat juda ko'p sonli so'z birikmalarini yaratadi va statistik modellarning siyraklik muammosining so'z boyligi va ma'lumotlar hajmini sezilarli darajada oshiradi. Ushbu maqolada morfofonetik istisnolarni hisobga olgan holda ildizni ajratib olish, lemmatizatsiya va morfologik ma'lumotni ajratib olishni o'z ichiga olgan morfologik tahlil modeli muhokama qilinadi. Modelning mohiyati mos keladigan morfologik ma'lumotlar va morfologik tahlil uchun qo'shimcha ma'lumotlar to'plami bilan so'zlarning to'liq to'plamini ishlab chiqishdir. Taklif etilayotgan model maxsus ishlab chiqilgan testlar to'plami, jumladan 5,3 ming so'z yordamida sinovdan o'tkazildi. So'z darajasidagi aniqlik 91% dan oshdi, bu tilshunoslar tomonidan qo'lda o'zak, lemma va morfologik

xususiyatlarni tuzatish orqali aniqlandi. Taklif etilgan metodologiya asosida yaratilgan vosita ochiq manba Python paketi, shuningdek, umumiy API ni o'z ichiga olgan veb-ilova sifatida mavjud.

Kalit so'zlar: o'zbek tili, morfologik tahlil, morfologik segmentatsiya, o'zak, lemmatizatsiya, bog'langan morfema, flektiv son.

Вступление. Компьютерная лингвистика интегрирует моделирование естественного языка человека с помощью основанного на правилах подхода, основанные на статистических моделях, моделях машинного и глубокого обучения. Тот факт, что в некоторых языках широко используются суффиксы и префиксы для передачи грамматического значения (например, согласование подлежащего и глагола), представляет собой проблему для большинства современных языковых технологий. Суффиксы и префиксы в таких языках в более общем смысле можно назвать морфемами, которые определяются как значимые составные части слов. Правила, которые языки используют для сочетания морфем, а также сами используемые ими морфемы называются морфологией языка. Языки, которые широко используют морфемы для построения слов, считаются морфологически богатыми. Узбекский язык, как и другие тюркские языки, имеет богатую морфологию и множество исключений в морфологической структуре слов. Морфология фокусируется на том, как части слова, такие как основы, префиксы и суффиксы, организованы или изменены для передачи различных значений. Существует две основные категории морфологии: словоизменяющая морфология и словообразовательная морфология. Каждая из этих двух категорий имеет свое значение в различных аспектах НЛП. Словоизменяющие морфемы, как правило, суффиксы, добавляются до конца слова, чтобы передать грамматические нюансы. Эти словоизменяющие окончания служат для изменения значения слова, сохраняя при этом его фундаментальное значение. В этом проекте мы исследуем модель морфологического анализа, основанную на полном наборе окончаний слов, которые образовались только из словоизменяющих морфем. Модель морфологического анализа включает в себя методы стемминга, лемматизации и морфологической сегментации наряду с извлечением морфологической информации. Морфологический анализ узбекского языка предполагает разбор сложных структур слов, включающих префиксы, суффиксы и корневые формы. Высокая флексивность языка требует тщательного подхода к выделению признаков и пониманию контекстуальной семантики. Модели нейронных сетей, обладающие способностью улавливать сложные паттерны и контекстуальные свойства, обладают потенциалом для улучшения работы в этом сложном лингвистическом ландшафте.

Обзор литературы. Работа [1] была посвящена разработке морфологических анализаторов, которые разделяют слова на последовательности морфем или слогов в уйгурском языке. Процесс сегментации морфем является центральным в создании всеобъемлющего корпуса уйгурского языка. В то время как контролируемый подход в сочетании с правилами и алгоритмами статистического обучения эффективно справляется с сегментацией морфем, значение лингвистических морфем в расширении охвата слов и сокращении размера словарного запаса. Кроме того, эти морфологические анализаторы предназначены для обработки как стандартных, так и поверхностных форм, учитывающих фонетические изменения и сложные морфологические изменения.

Хранение всех поверхностных слов в словаре для морфологически богатого тюркского языка уже давно является сложной задачей и накладывает ограничения на широкое применение нейронного машинного перевода (NMT). В работах [2], [3] устраняется это ограничение путем внедрения инновационного подхода к морфологической сегментации тюркских языков, основанного на концепции полного набора окончаний (CSE). Основанный на CSE метод сегментации, как показано на примере казахского, кыргызского и узбекского языков, эффективно уменьшает объем словарного запаса в исходных текстах. Вычислительные эксперименты NMT, в частности, с акцентом на казахском языке, показали убедительные результаты. По сравнению с сегментацией, основанной на парном байтовом кодировании (BPE), подход, основанный на CSE based, повышает балл двуязычного дублера для казахско-английских и англо-казахских пар. В контексте морфологического анализа узбекского языка было предложено несколько моделей с использованием конечных автоматов (FSM). В [4] авторы используют предопределенные грамматические правила для извлечения морфологической информации из заданных слов с помощью автоматов. Однако важно отметить, что эта методология использует словарь словоформ в качестве фундаментальной базы данных, в отличие от использования грамматических, морфотаксических и морфологических правил. В работе [5], которая является одной из работ по морфологическому анализу слов узбекского языка на основе модели конечного автомата (FSM), исследуется задача определения корня с извлечением соответствующей морфологической информации. Тем временем для узбекского языка были разработаны морфотаксические правила [6] и

модели морфологического анализа [7], [8]. Считается, что узбекские слова имеют довольно сложную структуру, и существует множество исключительных случаев. Узбекский язык представляет собой сложную задачу из-за наличия омонимии и синонимии аффиксов, гармонии гласных, что может привести к ошибкам в морфологической сегментации. Это также может привести к снижению точности моделей морфологического анализа, основанных на FSM. Методология. В этом разделе мы представляем подробное описание узбекской модели морфологического анализа. Используемая модель дополняет набор данных об окончаниях слов и дополнительных наборах данных и включает в себя методы определения корней, лемматизации и морфологического анализа на основе правил.

Начальный этап предлагаемой методологии включает в себя разработку полного набора языковых данных по окончаниям (CSE). Морфемная часть слова будет разделена на две части: словообразовательную и словоизменяющую морфологию. Словоизменяющая морфология включает в себя последовательности словосочетаний, лексические и синтаксические суффиксы. Словообразовательные морфемы изменяют часть речи слова. В этой работе мы сохраняем словоизменяющую часть слова. При формировании набора данных мы собрали все возможные окончания слов; была предоставлена морфологическая информация, связанная с окончаниями каждого слова. Эта морфологическая информация включает различные лингвистические атрибуты, такие как часть речи (POS), время, принадлежность, связка, форма единственного или множественного числа, форма вопроса, падеж и многое другое. Набор данных окончаний слов создается экспертами-лингвистами для обеспечения точности морфологической информации. На основе анализа морфологической структуры слов была разработана структура базы данных для хранения всей необходимой информации для модели морфологического анализа.

Из-за гармонии гласных и правил аффиксации суффикс в узбекском языке может иметь несколько алломорфов. Итак, в наборе данных окончаний слов отмечены конкретные обозначения для учета вариаций и алломорфов. Для описания алломорфных суффиксов мы используем следующие обозначения из предыдущей работы [5]:

G: {g,k,q}; K: {g,k}; Q: {g,g',k,q}; Y: {a,y}; T: {t,d}; (): буква, заключенная в круглые скобки, может быть опущена.

Для процесса ввода мы разработали веб-приложение, в котором каждому эксперту-лингвисту присваиваются уникальные имя пользователя и пароль, с помощью которых он может получить доступ к веб-сайту и ввести новое слово, заканчивающееся соответствующей морфологической информацией. Значительная часть словоизменяющих окончаний правильно соответствует глаголу и существительному как грамматической функции слова. В наборе данных 1205 и 150 записей являются глаголами и существительными соответственно. Кроме того, набор данных содержит 22 записи числительных, 10 записей прилагательных, 20 записей местоимений и 10 записей наречий.

Кроме того, мы сформировали пять наборов данных для повышения точности производительности методологии. Исключительные корни (доллар, ташкари, ...), список корней, для которых в нем есть какой-либо дополнительный конец. Неаффиксированные основы (va, lekin, ya na, ...), список основ, к которым никогда не было никаких прикреплений.

Числовые основы (bir, ikki, ellik, ...).

Основы короткой длины (длина \leq 2) (u, bu, ot, un, ...).

Исключительная лемма (bitta, bir, ta / single, singil, i).

На следующем этапе методологии для каждого элемента будут сгенерированы алломорфы из столбца affix в наборе данных CSE. Затем в процесс моделирования будут включены спецификации алфавита, которые используются для проверки морфологических правил. После этого вводимые слова будут перейти к процессу очистки, который включает в себя преобразование всех символов в нижний регистр и замену определенных специальных символов по мере необходимости. Это гарантирует, что введенные слова будут надлежащим образом подготовлены для последующей обработки и проверки правил.

Прежде чем определить окончание введенного слова, мы проверяем предопределенные основы в дополнительных созданных наборах данных, что гарантирует быструю работу модели и низкий уровень ошибок. Извлекается окончание слова, что предоставляет дополнительный уровень морфемной информации. Для повышения точности морфологического анализа при извлечении окончаний слов применяются правила аффиксации. Затем извлеченное окончание слова сравнивается с набором данных CSE для выявления совпадений. При обнаружении совпадения результирующему набору присваивается соответствующая морфологическая информация. Когда прогнозируемая длина стебля составляет менее двух знаков, выполняется дополнительный этап проверки. Для подтверждения правильности определения стебля используется набор данных о короткой длине стебля. Этот процесс

проверки обеспечивает точность определения коротких стеблей, что повышает общую надежность морфологического анализа. Результаты этих этапов обработки возвращают список, включающий основу, лемму и морфологическую информацию для данного слова и POS-тега.

Результаты и обсуждения. В качестве цели оценки предложенная модель была проанализирована с использованием данных построенного корпуса для метода морфологической сегментации. Для тестирования был использован построенный корпус, состоящий из 40 документов, полученных с новостной платформы (daquoz.uz), которые были разделены на четыре группы по десять документов в каждой. Этот корпус содержит в общей сложности 11 952 слова, из которых 5 288 уникальных составляют 44,24% от общего числа слов. Слова были аннотированы экспертом-лингвистом с правильной сегментацией. В процессе оценки время работы модели составляет около 1000 слов в секунду. Выходные данные модели сравниваются для каждого слова с фактическими ожидаемыми выходными данными. Выходные данные модели были разделены на пять различных вариантов, один из которых указывал на правильные прогнозы, в то время как остальные четыре указывали на различные типы ошибок, обнаруженных во время оценки. В таблице 1 результаты оценки модели на основе собранного корпуса демонстрируют разбивку результатов по категориям в случаях. В таблице поверхностное слово определено как образец китобимдан (из моей книги), оно разделено на три морфемы: китоб [основа], -(i)м [принадлежность к первому лицу] и -дан [аблятивный падеж].

Примерный пример кейсов %

1. Правильное предсказание китобимни > китоб 4821 91,2%

2. Ни один аффикс не был вычеркнут, хотя аффикс(ы) существует китобимни > китобимни 24 0,5%

3. Аффикс(ы) был удален из основы, в то время как основа имеет аффикс(ы) китобимни > кито 131 2,5%

4. Частично полосатые аффиксы китобимни > китобим 158 3,0%

Большинство ошибок возникает из-за наличия коротких окончаний слов в основах (3-й, 5-й падежи), и все варианты еще не охвачены в наборе данных CSE (4-й случай).

Инструмент Питон, созданный для этой работы, находится в открытом доступе, а также может быть легко установлен с помощью следующей команды, популярной в сообществе Python: pip устанавливает морфологический анализатор.

Существует также общедоступная система API 2 для интеграции модели в другое программное обеспечение, и более подробную информацию о ней можно найти в репозитории проекта на гитхаб 3.

Вывод. В этом исследовании проводится всесторонний морфологический анализ узбекского языка, охватывающий различные задачи, связанные со структурой слов и морфологическими свойствами. Методология включает в себя разработку набора данных о словоизменительных окончаниях, который лингвистические эксперты дополняют морфологической информацией. помощью существительных и глаголов, которые играют определенную роль в изменении грамматических свойств слов. Процесс оценки включает в себя определение основы, лемматизацию и морфологический анализ, что приводит к впечатляющему показателю точности – более 91% для проверки основы и леммы. Кроме того, мы представили код на Питон, веб-инструмент и API, созданные для использования предложенной модели.

ЛИТЕРАТУРА:

1. М. Аблимит, Т. Кавахара, А. Паттар и А. Хамдулла, “Уйгурский язык, основанный на стволовых аффиксах” *Morphological Analyzer*, “Международный журнал коммуникации будущего поколения и Сетевое взаимодействие, том 9, № 2, 2016, doi: 10.14257/ijfgcn.2016.9.2.07.

2. У. Тулеев, А. Карибаева и З. Х. Жуманов, “Метод морфологической сегментации Нейронный машинный перевод на тюркский язык”, *Cogent Eng*, том 7, № 1, стр. 1856500, 2020, doi: 10.1080/233111916.2020.1856500.

3. Шермухамедова Н. А. Графика и транскрипция заимствованных имен // *Barqaror Taraqqiyot va Rivojlanish Tamoyillari*. – 2023. – Т. 1. – №. 1. – С. 34-36.

4. Шермухамедова Н. А. Лингвистические проблемы перевода: текст как объект переводческой деятельности // *Polish science journal*. – 2020. – С. 335.

5. Шермухамедова Н. А. Ходжи на Насреддина как источник духовно-нравственно воспитания молодёжи // *Тошкент-2031*- С.31.

6. Мукимова Г. А. оптимальные пути обучения иностранному языку: Янги Узбекистон тараккиотида таджикотлар урни ва ривожланиши омиллари // Янги Узбекистон тараккиотида таджикотларни о'рни ва ривойланиши омиллари. – 2024. – Т. 4. – №. 1. – С. 159-163.
7. Шермухамедова Н.А. словообразование аспект компьютерного слена и интернет-дискурса.
8. Туйлиева Л. А. Методы. Методы исследования глаголов состояния: //Образование и инновационные исследования международный научно-методический журнал. – 2023. – №. 3. – С. 65-73.
9. Истамовна Д. С. Категории оппозиции и контраста как объекты лингвистического анализа // EPRA International Journal of Multidisciplinary Research (IJMR). – 2023. – Т. 9. – №. 11. – С. 407-409.
10. Авезова Д. литературные псевдонимы на примере русских и узбекских писателей //Наука и инновации. – 2024. – Т. 3. – №. С3. – С. 90-95.