IEEE TÜRKİYE SECTION

UBMK'24

Bildiriler Kitabı
Proceedings

Editor Eşref ADALI

9. Uluslararası Bilgisayar Bilimleri ve Mühendisliği Konferansı

9th International Conference on Computer Science and Engineering

26-27-28 Ekim (October) 2024 Antalya - Türkiye

9. Uluslararası Bilgisayar Bilimleri ve Mühendisliği Konferansı (UBMK'2024)

9th International Conference on Computer Science and Engineering

26-28 Ekim 2024 Akdeniz Üniversitesi Antalya Türkiye
26-28 October 2024 Akdeniz University Antalya Türkiye

Telif Hakkı

Bu elektronik kitabın içinde yer alan tüm bildirilerin telif hakları IEEE'ye devredilmiştir. Bu kitabın tamamı veya herhangi bir kısmı yayımcının izni olmaksızın yayımlanamaz, basılı veya elektronik biçimde çoğaltılamaz. Tersi davranışta bulunanlara ABD Telif Hakkı Yasalarına göre ceza uygulanır.

Copyright and Reprint Permission

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. Copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid throught Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For reprint or republication permission, email to IEEE Copyright Manager at pubs-permission@ieee.org

All right reserved. Copyright C 2024

IEEE Catalog Number : CFP24L97-CDR

ISBN : 979-8-3503-6587-0

Additional copies may be ordered from:
Curran Associates, Inc
57 Morehouse Lane Red Hook, NY 12571 USA
Phone: (845) 758 0400
Fax: (845) 758 2633
E-mail: curran@proceeding.com

	ID	Title	Authors	Pages
NLP	1850	HTV-News: A New Dataset with High Novelty Rate for Turkish Text Summarization	Nihal Zuhal Kayalı Sevinç İlhan Omurca	1 - 6
	1853	Sentiment Analysis of Social Network Comments in Uzbek Language	Saboxat Allanazarova	7 - 10
	1861	An Online Platform for Uzbek-Russian and Russian-Uzbek Parallel Corpora	Nigmatova Lolakhon Khamidovna Saidova Mokhira Rasulevna Djuraeva Zulkhumor Radzhabovna Sharipov Sokhib Salimovich Avezov Sukhrob Sobirovich	11 - 16
	1863	Yapım Ekleri ile Özbekçe Sözcük Türetme Deriving Uzbek Words with Derivational Suffixes	Eşref Adalı Khamroeva Shahlo Mirdjonovna	17 - 21
	1864	Database of Grammatical Form and Meaning Correspondence for Uzbek-English Translation Software	Jamoliddinova Odina Rasulovna Muhammedova Saodat Xudoyberdiyevna Khamroeva Shahlo Mirdjanovna Ergasheva Guli Ismoil Qizi Qodirova Madinabonu Murodjon qizi Kholova Muyassar	22 - 26
	1868	Words Speak Louder Than Actions: Decoding Emotions Through NLP	Melda Paksoy Gokhan Bakal	27 - 31
	1872	Özbek Dilinin Pankronik Derleminin Oluşturulması (Kutadgu bilig' Nüshalarına Göre)	Xamroyeva Şahlo Mircanovna Avezova Nigora Safarovna Toştemirova Sitora Bahodir kızı Qodirova Fazilat Şukurovna Atanazarova Shoxida Baxramovna	32 - 37
	1873	Preserving Semantic Integrity in Paraphrasing Texts: Rule-Based Paraphrasing	Zarnigor M. Khayatova Jurakuziyev Nodirbek Kholova Muyassar	38 - 42
	1975	NLLB-Based Uzbek NMT: Leveraging Multisource Data	Nilufar Abdurakhmonova Adkham Zokhirov Mohirdev Mukhammadali Salokhiddinov Anvar Narzullayev Ayrat Gatiatullin	43 - 47
	1976	The Problem of the Archaic Words' Semantic Description in the Alisher Navoi Authorship Corpus	Shuhrat Sirojiddinov Manzura Abjalova Sulton Normamatov Nargiza Gulomova	48 - 53
	1979	Multi-Label Emotion Classification with Fine-Tuned BERT and Contrastive Learning	Ammara Naseem Khan Mehmet Nafiz Aydın	54 - 59
	1885	Algorithm for Aligning Paragraphs and Sentences in Aligner Tool	Elov Botir Boltayevich Khamroeva Shahlo Mirdjonovna Dauletov Adilbek Yusupbayevich Matyakubova Noila Shakirjanovna	60 - 63
	1888	Semantic and Grammatical Issues in Translating Idioms with Automatic Translation Systems	Manzura Abjalova Sarvinoz Sharipova	64 - 69
	1889	Improving Methodology for Selecting Fiction Works for Grades 5-9	Khabibulla Madatov Sapura Sattarova	70 - 73
	1897	Innovating SQL Automation: Evaluating Open-Source Large Language Models with a Dual-Stage Approach for Corporate Data Solutions	Erhan Arslan Eugen Harinda	74 - 79
	1898	Development of a Method for Automatic Extraction of Ontology Entity Names from Natural Language Texts	Sadirmekova Zhanna Elmira Daiyrbayeva Karbozova Indira Bekbolatov Samat	80 - 85
	1899	UZBEKCORPORA.UZ: Creation of Author Corpora on the Platform	Karshiev Abduvali Berkinovich Karimov Suyun Amirovich Tursunov Mukhammadsolikh Sadin ugli Kudratov Iskandar Makhmidov Shirinboy B otirovich	86 - 90
	1901	A Technique for Automatic Extraction of Basis Words: a Case Study on "Uzbek Primary School Corpus"	Khabibulla Madatov	91 - 94

		Shukurla Bekchanov Surayyo Khajibaeva	
1908	Özbek Dili Metinlerindeki Eşgönderge Çözümlemesi Algoritması Algorithm of Coreference Resolution in Uzbek Texts	Elov Botir Boltayeviç Abdisalomova Şahlo Abdimurod kızı Karimova Xurriyat Şaripovna	95 - 100
1909	The Process of Lemmatization and Stemming in the Automatic Morphological Analysis of Uzbek Texts	Manzura Abjalova Eşref Adalı Munojot Adilova	101 - 106
1913	Turkish Question - Answer Dataset Evaluated with Deep Learning	Kadir Tutar Olcay Taner Yıldız	107 - 111
1914	Exploring the Semantic Complexity of Adjective-noun Collocations Between Uzbek and English for Improved Machine Translation	Nilufar Abdurakhmonova Nargiza Shamieva Eşref Adalı	112 - 115
1915	Linguistic Support and Algorithm of Phrase Alignment in Aligner Tool	Xoshimovna Shohida Shahabitdinova Matyakubova Noila Shakirjanovna	116 - 120
1946	Evaluating the Performance of PEFT and LoRA Approaches in Transformer-based Architecture for Handwritten Character Recognition	Pinar Savci Bihter Das Resul Das	121 - 126
1947	Evaluating the Performance of Turkish Automatic Speech Recognition Using the Generative AI-based Whisper Model	Tunahan Gokcimen Bihter Das Resul Das	127 - 131
1961	Özbek Dilinde Kelimelerin Kök ve Köklerini Ayırma Algoritmaları Algorithms for Parsing Roots and Stems of Words in Uzbek Language	Elov Botir Boltayeviç Xusainova Zilola Yuldashevna Umirova Svetlana Mamurjonovna Normamatov Sultonbek A.Şahlo Abdimurod kızı Mahmadiev Shavkatjon	132 - 136
1970	Özbekçe Deyimlerin Dilsel Modellemesinin DDİ Açısından Önemi The Importance of Linguistic Modeling of Uzbek Idioms in NLP	Gülyamova Shahnoza Kahramonovna Nurboyeva Maftuna Vahobcon kızı	137 - 140
1981	IT Service Desk Ticket Classification via Large Language Models	Ezgi Paket Göksu Şenerkek Fatma Betül Akyol Furkan Salman	141 - 146
1998	Enhancing the Quality of Clustering Job Skills Through the Fine-tuning of the Sentence Transformer Model	V. S. Ramazanova A. S. Yerimbetova M. A. Sambetbaeva Zh. B. Sadirmekova S. K. Serikbayeva	147 - 152
2004	Named Entity Recognition from Kazakh Speech	Bauyrzhan Kairatuly Madina Mansurova	153 -156
2005	Sentiment Analysis of Reviews for E-Commerce Applications	Elif Hanife Aydoğan Feyza Yıldırım Okay	157 - 162
2011	LLM and RAG-Based Question Answering Assistant for Enterprise Knowledge Management	Gürkan Şahin Karya Varol Burcu Kuleli Pak	163 - 168
2018	Named Entities Recognition in Kazakh Text by SpaCy NER Models	Nurzhan Mukazhanov Aigerim Yerimbetova Mussa Turdalyuly Bakzhan Sakenov	169 - 174
2022	Efficient Unsupervised Domain Adaptation with	M. Rawhani D. Karaboğa U. Nalbantoğlu A. Baştürk B. Akay	175 - 180
2032	Aligning Sentences for Kazakh-Turkish Parallel Corpora	Diana Rakhimova Eşref Adalı Aidana Karibayeva	181 - 186
2048	Automated-Computer Based Assessment of Free-Text Exam Answers By Transformers	Burak Keskin Melih Günay	187 - 192

	2073	Knowledge-Augmented Large Language Model	Başak Buluz Kömeçoğlu Burcu Yılmaz	193 - 198
	2092	LLMs for Document-level Text Simplification in Turkish Foreign Language Learning	Fatih Bektaş Kutay Arda Dinç Gülşen Eryiğit	199 - 203
	2105	Mining Software Requirements From Turkish Texts:Techniques and Challenges	Serhat Uzunbayir Senem Kumova Metin	204 - 209
	2108	Using LLMs for Annotation and ML Methods for Comparative Analysis of Large-Scale Turkish Sentiment Datasets	Ercan Ezin Rukiye Savran Kiziltepe Murat Karakus	210 - 215
	2122	Sentiment Analysis for Hotel Reviews in Turkish by using LLMs	Ata Onur Özdemir Efe Batur Giritli Yekta Said Can	216 - 220
	2123	Semantic Knowledge Base for the Emotional Coloring Analysis of Kazakh Texts	Banu Yergesh Dinara Kabdylova Tilekbergen Mukhamet	221 - 224
	2140	Adaptive Batch Budget for LLM Inference	Cağrı Yeşil Berhan Turku Ay Funda Ay Ak Öykü Berfin Mercan Oğuzhan Nefesoğlu	225 - 229
	2154	Open Data Processing for Analysis of Behavior of Users in Digital Environment	Nikolai Prokopyev Pavel Ustin Fail Gafarov	230 - 232
	2164	Evaluating Turkish BERT-based Language Models for Effective Customer Feedback Interpretation in CRM	Can İşcan Muhammet Furkan Özara Ahmet Erkan Çelik Akhan Akbulut	233 - 238
	2175	Does Prompt Engineering Help Turkish Named Entity Recognition?	Arda Serdar Pektezol Ahmet Batuhan Ulugergerli Volkan Öztoklu Seniz Demir	239 - 243
	2188	Araneum Uzbecicum: A Gigaword Web-Crawled Uzbek Corpus	Vladimi-r Benko Radovan Garabik Shahlo Khamroyeva	244 - 2149
	2189	Automatic Topic Detection in Large Text Data in Uzbek using Clustering Methods	Umirova Svetlana Kholmuhamedov Bakhtiyor Karimov Suyun Narzieva Mamura	250 - 255
	2193	Machine Learning-Based Synonymous Word	Diana Rakhimova Madina Mansurova Nurakhmet Matanov Amira Yerik	256 - 260
	2194	Enhancing Text-to-SQL Conversion in Turkish: An Analysis of LLMs with Schema Context	Ferhat D emirkıran Ali Kemal Coşkun Yavuz Kömeçoğlu Başak Buluz Kömeçoğlu Ramazan Güven	261 - 265
	2203	Linguistic Knowledge Graph "Turklang" as Universal Model for Linguistic Resources and Tools in Turkic Languages	Ayrat Gatiatullin Nikolai Prokopyev Lenara Kubedinova Nilufar Abdurakhmonova Rustam Burnashev	266 - 270
	2216	Requirements for the Development of a Website Builder with Adaptive Design	Gulmira Bekmanova Banu Yergesh Assel Omarbekova Laura Orynbay Aiganym Bessembayeva Dinara Kabdylova Altanbek Zulkhazhav Beibarys Sultan	271 - 276
AI	**1870**	A Modified Marine Predators Algorithm for Numerical Function Optimization	Resul Özdemir Murat Taşyürek	277 - 281

		Veysel Aslantaş	
1884	Document Classification and Key Information Extraction Using Multimodal Transformers	Mehmet Selman Baysan Furkan Kızılay Ayşe İrem Özmen Gökhan İnce	282 - 287
1926	Employee Turnover Prediction on Synthetic and Real Datasets	Azhar Murzaeva Volkan Illik Kadir Yunus Koc	288 - 292
1937	Data Augmentation in Remote Sensing Image Change Captioning	Orkhan Karimli Ilyas Mustafazade Ali Can Karaca Fatih Amasyalı	293 - 298
1968	Comparative Evaluation of Word2Vec and Node2Vec for Frequently Bought Together Recommendations in E-commerce	Mustafa Keskin Enis Teper Alya Kurt	299 - 303
1971	Disrupting the Retrieval Phase in RAG-Based LLM Chatbots through Input Manipulation	Felix Pitterling Georgii Koshelev Oliver Haddad Rodrigo Teran	304 - 308
1976	Mutation Testing Reinvented: How Artificial Intelligence Complements Classic Methods	Serhat Uzunbayir Kaan Kurtel	309 - 314
1997	Prompting Large Language Models for Aerial Navigation	Emirhan Balcı Mehmet Sarıgül Barış Ata	315 - 320
2006	Açıklanabilir Yapay Zeka Üzerine Bir İnceleme A Review of Explainable Artificial Intelligence	Samed Al Şeref Sağıroğlu	321 - 326
2020	Investigating the Effects of Pre-trained Deep Learning Models and Fusion Techniques on Fruit Segmentation Performance	Esma İbiş Aybars Uğur	327 - 332
2034	Beyin Tümörü Ameliyatlarında Beyin Kayması Probleminin Telafisi: Bilgisayar Destekli Yöntemler Uzerine Bir Derleme Compensating for the Brain Shift Problem in Brain Tumor Surgeries: A Review of Computer-Assisted Methods	Ayşe Gül Eker Meltem Kurt Pehlivanoğlu Nur Banu Albayrak Nevcihan Duru Tolga Turan Dündar	333 - 338
2039	AI-Enhanced Endometrial Cancer Diagnosis System	Raziye Aslıhan Kürkcü Işın Yeşim Yeşilkaya Baylan Erkut Attar Burcu Selçuk Tacha Serif	339 - 343
2040	El Titreme Sinyalinden Transformer Temelli Parkinson Hastalığı Sınıflandırması Transformer Based Parkinson's Disease Classification from Hand Tremor Signal	Murat Atçeken Lütfü Hanoğlu Mehmet Ersin Bitirgen Bahadır K. Güntürk İbrahim Özşeker	344 - 347
2042	Üretken Yapay Zeka Fırsatlar ve Tehditler: Bibliyometrik Analiz Generative Artificial Intelligence Opportunities and Threats: Bibliometric Analysis	Betül Ersöz Şeref Sağıroğlu Halil İbrahim Bülbül	348 - 353
2053	Users Acceptance of Generative Artificial Intelligence Based on ChatBot Use	Ayşe Yeşim Mutlu	354 - 359
2054	MMBAttn: Max-Mean and Bit-wise Attention for CTR Prediction	Hasan Saribas Cagri Yesil Serdarcan Dilbaz Halit Orenbas	360 - 365
2065	Türkçe Otomatik Konuşma Tanıma Sistemleri için Sentetik Veri Üretme Yöntemi Synthetic Data Generation Method for Turkish Automatic Speech Recognition Systems	Hilal Tekgöz Harun Uz Muhammed Murat Özbek Tolga Büyüktanır	366 - 370
2066	A Federated Learning Framework for Classifying the Images in Ultrasonic Nondestructive Testing	Abdulkadir Gulsen Hilal Hacılar	371 - 375

			Burak Kolukısa Burcu Bakir-Gungor	
	2071	AI Assisted Customer Review Sentiment Analysis and Department Classification Tool	Burcu Seçuk Berat Ilgaz Dursun Tacha Serif	376 - 381
	2075	Histopatolojik Görüntülerde Doğru Mitoz Tespiti için Geliştirilmiş Renk Normalleştirme Yöntemi Enhanced Stain Normalization Method for Accurate Mitosis Detection in Histopathological Images	Refik Samet Nooshin Nemati Emrah Hançer Serpil Sak Bilge Ayça Kırmızı	382 - 387
	2110	ATGRUVAE: Reducing Noise and Improving Forecasting Performance in Stock Data	Hüseyin Akkas Burak Kolukisa Burcu Bakir-Gungor	388 - 392
	2113	Evaluating the Impact of Sentiment Analysis on Deep Reinforcement Learning-Based Trading Strategies	Mustafa Etcil Burak Kolukisa Burcu Bakir-Gungor	393 - 398
	2118	PsychSynth: Advancing Mental Health AI Through Synthetic Data Generation and Curriculum Training	Vedanta S P Madhav Rao	399 - 404
	2121	Covid-19 Bilgisayarlı Tomografi Görüntülerinin Segmentasyonu için Topluluk Öğrenmesiyle Birleştirilmiş U-Net Modeli U-Net Model Combined with Ensemble Learning for Segmentation of Covid-19 Computed Tomography Images	Furkan Atlan İhsan Pençe	405 - 410
	2129	Turkish Sign Language Video Generation from Pose Sequences Using a Conditional GAN Model	Feyza Özkan Hatice Kübra Tekin Hacer Yalim Keles	411 - 416
	2131	Statistical Attention Layer for Neural Network Training	Ergun Biçici Hasan Saribaş	417 - 420
	2134	Dudak Şekillerinden Kişilik Analizi: Marifetname ve Açıklanabilir Yapay Zeka Yaklaşımı Personality analysis from lip shapes: Marifetnâme and explainable artificial intelligence approach	Semra Çelebi İbrahim Türkoğlu	421 - 426
	2148	Breast Cancer Detection Using Deep Learning Models	Erkan Akkaş Mustafa Yapar Şeref Sağıroğlu Ali Öter Betül Ersöz	427 - 430
	2191	CGAN Mimarisi ile Kısıtlı Veri Ortamında Kestirimci Bakım Performansının İyileştirilmesi Improving Predictive Maintenance Performance in Limited Data Environment with CGAN	Mehmet Musa Özcan Uğurhan Kutbay	433 - 438
	2207	Integrating the Focusing Neuron Model with N-BEATS and N-HiTS	Şuayb Talha Özçelik Faik Boray Tek	439 - 442
	2217	Advanced Convolutional Neural Networks for Plastic Classification in Recycling Systems	Nazym Alimbekova Ainur Zhumadillayeva Sunggat Aiymbay	443 - 447
ML	**1846**	Semi-Supervised Learning for Sensor-Based Flash Point Prediction in Oil Industry	Mert Sülük Şule Gündüz Öğüdücü	448 - 452
	1847	Robust Stacked Ensemble Model for Lung Cancer Diagnosis	Ayhan Akbas Gonca Buyrukoglu Selim Buyrukoglu	453 - 457
	1856	Kod Kusurlarının Tespitinde Makine Öğrenmesi Tekniklerinin Başarım Analizi The Performance Analysis of Machine Learning Techniques in Code Smell Detection	Arman Yavuz Oya Kalıpsız	458 - 463
	1860	Enhancing Financial Time-Series Analysis with TimeGAN: A Novel Approach	Cemal Öztürk	464 - 467
	1869	From Traditional to Deep: Evaluating Sentiment Analysis Models on a Large-Scale Tweet Dataset	Alisahib Mammadov Gokhan Bakal	468 - 473

1871	Comparison of Feature Selection Methods for Mechanical Properties of Cold Rolled Products in Flat Steel Manufacturing	Didem Bakiler İlme Merve Öper E. Fatih Yetkin	474 - 478
1883	A Machine Learning Approach to Steel Sheet Production Surface Quality	Asena Öztürk Mehmet Nafiz Aydın	479 - 484
1938	NLP Tabanlı Özellik Çıkarımına Dayalı Makine Öğrenme Algoritmaları Kullanılarak SQL Enjeksiyonu Ataklarının Tespit Edilmesi Detection of SQL Injection Attacks Using Machine Learning Algorithms Based on NLP-Based Feature Extraction	Hakan Can Altunay	485 - 489
1940	Enhancing Driver Injury Severity Prediction Using Optimized Oversampling and Feature Selection Techniques	Vahide Nida Kılıç Gizen Mutlu Esra Saraç Çiğdem Acı Murat Özen	490 - 495
1972	Predicting Electric Vehicle Adoption in the EU: Analyzing Classification Performance and Influencing Attributes Across Countries, Gender, and Education Level	Mert Kumbasar Gül Tokdemir Thouraya Gherissi Labben Gurdal Ertek	496- 499
1977	Arabam.com Kurumsal Üyelerinin Satın Alma Davranışlarının Analizi	Dilan Oragaz İsmail Duru	500 - 505
1983	Makine Öğrenmesi ve Derin Öğrenme Modelleri Kullanılarak MR Görüntülerinden Demans Sınıflandırma	Meltem Kurt Pehlivanoğlu Onur Varol Osman Aldemir Ata Emir Uncu Nevcihan Duru	506 - 511
1995	Indoor Localization with GravNetConv and Dynamic Graphs	Mert Bayraktar Alper Ozcan Umit Deniz Ulusar	512 - 516
2023	GPS Spoofing Detection on Autonomous Vehicles with XGBoost	Emre İşleyen Şerif Bahtiyar	517 - 522
2025	SmartSniffer: Predicting Food Spoilage Time with an Electronic Nose-based Gas Monitoring Apparatus Utilizing a Two-Stage Pipeline Model	Pandey Shourya Prasad Sreyas Janamanchi Barath S Narayan Madhav Rao	523 - 528
2064	GNF: Generative Natural Faces Dataset for Face Expression Recognition Task	Cansu Özer Volkan Dağlı Mustafa Kağan Gürkan	529 - 533
2081	Incorporating Knowledge Graph Embeddings into Graph Neural Networks for Sequential Recommender Systems	Kazım Emre Yüksel Susan Üsküdarlı	534 - 539
2095	The Classification of Human and Cognitive Architecture Time Estimation Using Machine Learning Methods	Behiye Şahin Sonay Duman	540 - 543
2101	Optimization of Roadside Assistance Tow Truck Services Provided to Insured Individuals with Traffic and Comprehensive Insurance Policies	Serkan Kırca Fatma Yağmur Erbaş Sedat Cebesoy Buse Özkanat Aysun Yıldırım Ceren Yıldırım	544 - 549
2104	Optimizing Recommendation Systems By Fusion of KNN, Singular Value Decomposition, and XGBoost for Enhanced Performance	Mohammed Basim Mohammed Mohammed Erkut Arıcan	550 - 555
2119	Spoken Accent Detection in English using Audio-Based Transformer Models	Oguzhan Ozturk Hasan Kilimci H.Hakan Kilinc Zeynep Hilal Kilimci	556 - 561
2136	Analysis of Adversarial Training for Resilient Image Recognition Against Different Attacks	Emir Can Karakuş Fırat Kutluhan İslim Ece Gelal Soyak	562 - 567

	2150	Business Process Management Anomaly Detection through Semantic Embedding-Integrated Graph Neural Networks	Teoman Berkay Ayaz Ege Gülce Stanley Hsu Alper Özcan Akhan Akbulut	568 - 573
	2155	Multi-Aspect Anomaly Detection with Graph Neural Networks and Kolmogorov-Arnold Networks in Business Process Management	Teoman Berkay Ayaz Ege Gülce Stanley Hsu Alper Özcan Akhan Akbulut	574 - 579
	2156	Comparison of a Deep Learning and a Hybrid Model for Classification of an Unbalanced Urgent Cases Dataset for Human Faces	Faruk Özgür Neslihan Arıkan Özge Öztimur Karadağ	580 - 585
	2171	Estimating the Manufacturing Cost of a Metal Part from Textual and Geometric Features	Talha Rehman Abid Mert Daloğlu Cem Yıldız Ali Erman Erten Kamer Kaya	586 - 591
	2176	Object Detection in Hyperspectral Images with Unsupervised Domain Adaptation	Sinem Aybüke Şakacı Alp Ertürk Erchan Aptoula	592 - 596
	2182	Machine Learning Approaches to Predict Thyroid Cancer Recurrence: A Comparative Study	Candide Ozturk Ozgur Sagir Ulas Vural	597 - 602
	2197	Yangın Söndürme Süreçlerinde Su Tüketiminin Makine Öğrenmesi ile Tahmini Prediction of Water Consumption in Fire Extinguishing Processes Using Machine Learning Approaches	Emin Ölmez Ahsen Usta Orhan Akbulut	603 - 606
	2202	Synthetic Vibration Data Generation and Fault Classification in CNC Machines Using Transformer GANs and ConvLSTM Networks	Özlem Erbay Batıray Erbay	607 - 612
IoT	**1862**	Design and Implementation of a Management System for Wireless Electronic Combination Locks	Batuhan Kol Metin Bilgin	613 - 618
	1887	Solar IoT: Monitoring the Orientation and Electrical Parameters of the Solar Panel	Hakan Dalkılıç Oğuz Gora	619 - 624
	1973	Multifunctional Smart Spoon for Parkinson's: Stability Enhancement and Diagnostic Tools	Divyansh Singhal Sasi Snigdha Yadavalli Nupur Patil Siddharth Chauhan Madhav Rao	625 - 629
	1988	An AI-Assisted Autonomous IoRT Agent for Smart Spaces	Yakup Kayataş Sanem Kabadayı	630 - 635
	1991	Advanced Detection and Prevention of Sinkhole Attacks in 6TiSCH Networks	Burak Aydın Hakan Aydın Sedat Görmüş	636 - 641
	2000	Comparison of Different Weather Data Acquisition Methods	Emre Evcin Yusuf Murat Erten	642 - 647
	2057	Employing Digital Twin to Forest Fire Management Systems	Bugra Aydin Sema Fatma Oktug	648 - 653
	2085	Distributed Key Value Store for IoT Edge Devices	Burak Aslantaş Elif Nurdan Pektaş Şebnem Baydere	654 - 659
EMB	**2096**	Soft Error Reliability Assessment of TinyML Algorithms on STM32 Microcontroller	Ahmet Selim Karakuş Osman Buğra Göktaş Sadık Akgedik Sanem Arslan	660 - 664
HUM	**1999**	Development of An Algorithm for Converting Json Formats to Xml by Forming its File Data Structure	Aigul Mukhitova Aigerim Yerimbetova Vladimir Barakhnin	665 - 670

			Elmira Daiyrbayeva	
	2130	Facial Emotion Recognition for Imitation in Human-Robot Interaction	İbrahim Yanç Aykan İpek Selma Yılmazyıldız Kayaarma	671 - 676
HPER	**2030**	A Glimpse to Scalable LLM Architectures: Building Real-Time Sentiment Analyser with Kafka and FastAPI	Uzay Çetin Yunus Emre Gündoğmuş	677 - 681
	2165	Parallelization of BitColor Algorithm via Multithreading and GPU for Graph Coloring	Burak Kocausta Gizem Sungu Terci Alp Arslan Bayrakci	682 - 687
BCHN	**1922**	Security Dynamics of Blockchain-Enabled SDN Systems: A Taxonomic Approach	Deniz Dudukcu Murat Karakus	688 - 693
	2056	Comparative Performance Analysis of Ethereum and Optimism Smart Contracts in Health Insurance	Beyhan Adanur Dedetürk Bilge Kagan Dedetürk	694 - 699
BIO	**1907**	Disease Prediction From Human Microbiome by Utilizing Machine Learning	Bahadir Emin Temel Bora Kocapinar Zerrin Isik	700 - 705
	1923	Medikal Görüntülerin Etkili Bir Şekilde SınıflandırılmasıIçin Dikkat Tabanlı Temsil Öğrenimi	Selim Arslan	706 - 711
		Attention Based Representation Learning for Effective Classification of Medical Images	Kemal Polat	
	1933	Utilizing Tree-Based Algorithms for Genetic Variant Interpretation	Rumeysa Aslıhan Ertürk Mehmet Baysan	712 - 717
	1978	A Test Bench for Replicating Human Breathing: Evaluating Thermal Effects of N95 Filtering Facepiece Respirator Leaks ' Preliminary Findings	Geoffrey Marchais Barthelemy Topilko Mohamed Arbane Jean Brousseau Clothilde Brochot Yacine Yaddaden Ali Bahloul Xavier Maldague	718 - 723
	1979	Development of Multiphysics Models for the Study of Airflow and Thermal Effects During the Use of Filtering Facepiece Respirators	Barthelemy Topilko Geoffrey Marchais Mohamed Arbane Jean Brousseau Ali Bahloul Xavier Maldague Clothilde Brochot Yacine Yaddaden	724 - 730
	1994	Awareness of Gamification as a Non-Pharmacological Intervention in Sleep Health Research: A Systematic Literature Review	Arife Gülah Erol Murat Y ılmaz Paul M. Clarke	731 - 736
	2003	Optik Çukur Bölütlemesi Segmentation of the Optic Cup	Saadet Aytaç Arpacı Songül Varlı	737 - 741
	2090	Clinically Significant Prostate Cancer Detection and Diagnosis in Bi-parametric MRI with Deep Learning Models	Clinton Binda Asoh-Itambi Mecit Yüzkat Songül Varlı	742 - 744
	1993	Malaria Incidence Prediction Using Climate Factors with Machine Learning Models	Ayoade Adeyemi Cyrille Mesue Njume Goodnews Akindele Nana Aisha Umar Ayodele James Oyejide	748 - 753
CDSec	**1857**	Dual-layered Approach for Malicious Domain Detection	Nadide Bilge Doğan Alp Barış, Beydemir Şerif Bahtiyar Umutcan Doğan	754 - 759
	1953	Analysis of the Zero-Day Detection of Metamorphic Malware	Sibel Gulmez Arzu Gorgulu Kakisim Ibrahim Sogukpinar	760 - 765
	1960	White-Box Style Intrusion Detection System Integration into Industrial Metaverses	Yasir Kılıç	766 - 771

			Vahide Nida Kılıç Ali İnan	
	2037	A Practical Investigation of Spear Phishing Spam Emails: Comparative Analysis and Evaluation	Kendrick Kurt Günter Bollens	772 - 777
	2091	Multipurpose Malware Detection System	Mert Gursimsir Cem Ayar Ibrahim Sogukpinar	778 - 782
CRYP	**1867**	A New Method to Detect Malicious DNS over HTTPS via Feature Reduction	Ali K. Bozkurt Halil E. Aköz Ataberk Taşpınar Şerif Bahtiyar	783 - 788
	1939	Generative Adversarial Networks for Synthetic Jamming Attacks on UAVs	Burcu Sönmez Sarıkaya Şerif Bahtiyar	789 - 794
	1958	Detecting Corruptive Noise Rounds for Statistical Disclosure Attacks	Alperen Aksoy Doğan Kesdoğan	795 - 800
	1996	Future Directions of Cybersecurity in Industrial Internet of Things Through Edge Computing	Tamara Zhukabayeva Lazzat Zholshiyeva Nurdaulet Karabayev	801 - 806
	2061	Resource-Efficient Ensemble Learning for Edge IIoT Network Security against OSINT-based Attacks	Mert İlhan Ecevit Zakire Çukur Muhammed Ali İzgün Noor Ul Ain Hasan Dağ	807 - 812
	2074	Transfer Learning for Phishing Detection: Screenshot-Based Website Classification	Furkan Çolhak Mert İlhan Ecevit Hasan Dağ	813 - 818
	2145	Blok Şifrelerin Karıştırma ve Yayılım Tabakaları için Yeni Bir Analiz Aracı A New Analysis Tool for Confusion and Diffusion Layers of Block Ciphers	Mehmet Ali Demir Meltem Kurt Pehlivanoğlu Pınar Savaştürk Emir Öztürk Muharrem Tolga Sakallı Sedat Akleylek	819 - 824
CVIS	**1851**	Impact of Image Augmentation on Deep Learning-Based Classification of Granite Tiles	Gaye Ediboglu Bartos Sibel Ünaldı Nesibe Yalçın	825 - 828
	1877	Advanced Facial Expression Classification with CNN-Transformer Integration for Human-Computer Interaction	Ali Azmoudeh Cigdem Altin Gumussoy Hazım Kemal Ekenel	829 - 834
	1936	Word Image Representation at Local and Global Levels Based on Vision Transformers	Baha Edine Harrath Mohamed Mhiri Mohamed Cheriet	835 - 840
	1949	Detecting Duplicate Products in E-Commerce Images Using Siamese Networks	Enis Teper Furkan Eseoğlu Mustafa Keskin	841- 846
	1963	Comparative Analysis of Visual Attribute Tagging Models for Upper-Body Clothing Products	Engin Kaya Mert Yanık	846 - 850
	1986	Development of A Model of Kazakh Sign Language Recognition Based on Deep Learning Method	Aigerim Yerimbetova Bakzhan Sakenov Ulmeken Berzhanova Nurzhan Mukazhanov Elmira Daiyrbayeva Mohamed Othman	851 - 856
	1989	Recognising Kazakh Sign Language with Mediapipe	Aigerim Yerimbetova Diana Kaidina Bakzhan Sakenov Elmira Daiyrbayeva Mussa Turdalyuly Ulmeken Berzhanova	857 - 862
	1990	Ultrason Görüntülerinden Meme Kanseri Teşhisi için Lezyon Tespitli Hibrit Derin Öğrenme Modelleri	Osman Doğuş Gülgün	863 - 868

		Hybrid Deep Learning Models with Lesion Detection for Breast Cancer Diagnosis from Ultrasound Images	Hamza Erol	869 - 874
	1992	Olumsuz Hava Koşullarında Gemi Tespiti ve Sınıflandırılması Ship Detection and Classification in Adverse Weather Conditions	Yahya İzala Yaşar Becerikli	
	2010	Removing Background from Noisy Handwritten Signatures on Banking Documents using GANs	Ege Dinçer Sacide Kalaycı Emre Yurdakul Bilge Köroğlu	875 - 879
	2013	A Faster R-CNN Model for Multi-class Classification and Detection of Land, Air, and Sea Vehicles	Enes Güvelioğlu Çiğdem İnan Acı	880 - 885
	2098	Çelik Hurdasının Sınıflandırılmasında ResNet ve Görüntü Dönüştürücü Tabanlı Modellerin Başarımı	Sefa Temur Levent Karacan	886 - 891
	2117	Advanced Computer Vision Techniques for Reliable Gender Determination in Budgerigars (Melopsittacus undulatus)	Atalay Denknalbant Efe Ilhan Cemalcılar Majid Ahangari Abdussamat Saidburkhan Alireza Zirak Ghazani Erkut Arıcan	892 - 897
	2151	Teslimat Süresi Tahminlerinde Makine Öğrenmesi Modellerinin Yorumlanabilirliği Interpretability of Machine Learning Models in Delivery Time Predictions	Serhat Agit Satıcı Habil Kalkan	898 - 903
	2169	Deep Learning based Order Form Recognition	Enes Alperen Buğaz Orhan Akbulut Aysun Taşyapı Çelebi Uğur Yıldız	904 - 908
	2170	Learning Based Photo Management on Smartphones	Beyza Nur Şenay Orhan Akbulut Aysun Taşyapı Çelebi Uğur Yıldız	909 - 912
	2177	Retinal Disease Classification Using Optical Coherence Tomography Angiography Images	Omer Faruk Aydın Muhammet Serdar Nazlı F. Boray Tek Yasemin Turkan	913 - 918
	2178	Segmentation Based Classification of Retinal Diseases in OCT Images	Öykü Eren F. Boray Tek Yasemin Turkan	919 - 924
	2183	Unsupervised Translation from Shortwave Infrared Images to RGB Images: A Comparative Evaluation	Duygu Tasbas Hacer Yalim Keles	925 - 930
DSCI	**1882**	Automatic Segmentation of Time Series Data with PELT Algorithm for Predictive Maintenance in the Flat Steel Industry	Saygın Kaçar Tuğçe Ballı E. Fatih Yetkin	931 - 936
	1921	Otomobil Kredilerinde Temerrüt Tahmini ve Araç Geri Kazanım Olasılığı Analizi - Bir Segmentasyon Çalışması Default Prediction and Vehicle Recovery Probability Analysis in Auto Loans - A Segmentation Study	Sahin Nicat Anıl Ferdi Kaya	937 - 942
	1955	E-Ticaret Sadakat Programı Müşteri Eğilim Tahmini Customer Propensity Prediction in E-Commerce Loyalty Program	Yunus Emre Gündoğmuş Sinan Keçeci Ege Erdem Emre Rençberoğlu	943 - 946
	1967	Yapay Zeka ve Makroekonomik Göstergeler ile Tüzel Kredilerin Değerlendirilmesi Evaluation of Corporate Loans with Artificial Intelligence and Macroeconomic Indicators	Burak Yüksel Hakkı Berkay Çiçek	947 - 951
	1974	Enhanced Bot Detection on TwiBot-20 Dataset	Mehmet Ali Osman Atik Şevket Umut Çakır Alper Özcan	952 - 956
	1975	Drug-Drug and Drug-Protein Link Prediction on DTINet dataset	Mehmet Ali Osman Atik Yusuf Çelik Alper Özcan	957 - 960
	2029	Perakende Verilerinde Anomali Tespiti ve Döviz Kuru ÄİlişkisiÜzerine ChatGPT Destekli Yorumlama	Şadi Evren Şeker	961 - 966

		ChatGPT Supported Interpretation on Anomaly Detection in Retail Data and Exchange Rate Relationship	Hatice Nizam-Özoğur	
	2033	Text to SQL Transformation Using LLM: a Comparative Research of T5, Seq2Seq, and SQLNet Models	Zhazira Shaikhiyeva Madina Mansurova Gulshat Amirkhanova	967 - 972
	2076	Sağlık Sigortası Sahiplerinin Davranışsal Analizi ve Kümelenmesi Clustering and Behavioral Analysis of Health Insurance Owners	Omer Sezer Koyuncu Seçil Arslan	973 - 978
	2087	On symbolic Prediction of Time Series for Predictive Maintenance Based on SAX-LSTM	Aykut Güler Tuğçe Ballı E. Fatih Yetkin	979 - 983
	2135	Profiling Driver Behaviors Using AI-Based Methods and Deep Learning Techniques for Improving Road Safety: A Comparative Study of Algorithms	Volkan Oban Mustafa Kaya Güzide Safi İrem Nur Çimen Tubanur Çatak Bulut Karadağ Gökhan Gümüş Aslıhan Çandır Fatih Alagöz	984 - 989
IR	**1896**	ReRag: A New Architecture for Reducing the Hallucination by Retrieval-Augmented Generation	Robin Koç Mustafa Kağan Gürkan Fatoş T. Yarman Vural	990 -994
	1941	Enhancing Object Detection in Aerial Images Using Transformer-Based Super-Resolution	Aslan Ahmet Haykır İlkay Öksüz	995 - 1000
NET	**1985**	Proof of Concept Implementation for RSVP TSN Control Plane	Necip Gozuacik	1001 - 1004
	2100	Integrating Blockchain and SDN for Centrality-Aware Virtual Multicast Tree Embedding	Furkan Ayaz Evrim Guler Murat Karakus Davut Hanbay	1005 - 1010
	1969	QoS Aware Routing Approaches in Software Defined Smart Grids	Sedef Demirci	1011 - 1016
	2008	Deep Reinforcement Learning Routing in Mobile Networks	Arif Burak Dikmen Hasari Çelebi	1017 - 1022
RBOT	**1942**	Endüstriyel Robotik Sistemlerin Güvenlik Doğrulaması Safety Verification of Industrial Robotic Systems	Fatih Furkan Arslan Metin Özkan	1023 - 1028
	2077	EKF Based Localization: Integrating IMU and LiDAR Data in the Hilti SLAM Challenge	Behice Bakır Havvanur Bozömeroğlu Ebu Yusuf Güven	1029 - 1034
SING	**1965**	Communication (Educational) Kit (HaKi)	Murat Sever Utku Bilgin	1035 - 1038
	2089	Manyetik Parçacık Görüntülemede Sistem Matrisi için Farklı Dalgacık Dönüşümlerinin Seyreklik Seviyesi Karşılaştırması Sparsity Level Comparison of Different Wavelet Transforms for the System Matrix in Magnetic Particle Imaging	Vildan Atalay Aydın	1039 - 1043
	2097	Sparse Channel Estimation For M-QAM-Based Underwater Acoustic Communication Systems	Mhd Tahssin Altabbaa Berkay Tekat Emin Tarik Iseri	1044 - 1048
OTH	**1858**	The 80/20 Principle in Morphemics-Morphology in the Educational Corpus of the Uzbek Language	Shahlo Khamroeva Bakhtiyor Mengliyev Muyassar Kholova	1049 - 1052
	1904	Gamification as a Tool for Personalized Learning in Inclusive Education	Dilaram Baumuratova Tamara Zhukabayeva Mira Rakhimzhanova	1053 - 1058
	1918	A Metaheuristic Algorithm for the Fixed Charge Transportation Problem	Nermin Kartli	1059 - 1062
	2027	Eğitimde Sürükleyici Teknolojilerin Kullanılması Fırsatlar ve Beklentiler	Atamuratov Rasuljon Kadırjanovich Majıdova Gulhayo Abdırazzoq qızı Bayjonov Furqat Baxramovıch Ongarov Mansurbek Bayrambekovıch	1063 - 1068

			Saydullayev Zafar Erkınovıch	
	2103	Bilgisayar Mühendisliği Öğrencilerinin Perspektifinden Bilişim Hukukunun Güncel Sorunları ve Çözüm Önerileri Current Challenges and Solution Proposals in IT Law from the Perspective of Computer Engineering Students	Sevda Bora Çınar	1069 - 1075
	2200	A Comparison of shcU-Net Based GAN and U-net Based GAN in Adult Dental Segmentation	Gürdal Altundağ Hakan Öcal	1075 - 1080
	1932	Leveraging Quantum Computing and Optimization to Estimate Financial Crashes in Small and Medium-Sized Enterprises	Ege Dincer Berkay Coskuner Ege Bilaloglu Bilge Koroglu	1081 - 1086
SW	**1859**	Investigating The Adoption of International Software Quality Standards in Turkey: A Comprehensive Analysis	Sevgi Koyuncu Tunç	1087 - 1093
	1886	Development of the Functional Structure of the Science and Education Information System	Dauletov Adilbek Yusupbayevich Matyakubova Noila Shakirjanovna	1094 - 1098
	1892	React ve Preact Javascript Çerçevelerinde Karşılaştırmalı Analiz Comparative Analysis on React and Preact Javascript Frameworks	Muhammed Furkan Uygur Nesibe Yalçın	1099 - 1104
	1917	CAGE: A Tool for Code Assessment and Grading	Ümit Kanoğlu Oğuz Kerem Yıldız Hasan Sözer Olcay Taner Yıldız	1115 - 1110
	1957	Extracting Driving Styles from Automotive Sensor Data to Develop Personas	M. Cagri Kaya Tayssir Bouraffa Krzysztof Wnuk	1111 - 1114
	1962	Lojistik Sipariş Dağıtım Entegrasyonu Sürecinde Sipariş Geri Çağırma Süreci Tasarımı ve Yazılım Geliştirmesi Design and Software Development of The Order Recall Progress in The Logistics Order Distribution Integration Process	İklim Barman Ersin Şengül	1115 - 1120
	2009	The Dimension of Green Coding in Software Quality Control Processes	Volkan Abur	1121 - 1126
	2055	Are We Asking the Right Questions to ChatGPT for Learning Software Design Patterns?	Çağdaş Evren Gerede	1127 - 1132
	2060	Optimizing LLVM IR: Transforming Multiplication to Addition for Enhanced Execution Efficiency	Huseyin Karacalı Efecan Cebel Nevzat Donum	1133 - 1138
	2080	Estimation of Software Integration Test Duration via UML Statecharts	Fehim Göler Tolga Ovatman	1139 - 1144
	2093	DIA4M: A Tool to Streamline DevOps Processes of Distributed Cloud-Native Systems	Eren Tarak H. Hakan Kilinc	1145 - 1150
	2111	Software Industry Perception of Academic Collaboration	Deniz Akdur	1151 - 1156
	2139	Görüntü İşlemeyle Doğrulamalı Robotik Test Otomasyon Kullanımı: POS Cihazları Üzerine Uygulama	Miraç Emektar Harun Kadıoğlu Ahmet Efendioğlu Fatih Mehmet Harmancı	1157 - 1161
	2141	VoIP Sistemlerinde Zihin Haritası Tabanlı Test Stratejiler : SIP Pbx Ürünü Üzerine Bir İnceleme Mind Map-Based Testing Strategies in VoIP Systems: A Case Study on SIP Pbx Products	Miraç Emektar Furkan Günaydın Fatih Mehmet Harmancı	1162 - 1167
	2173	A Robust Microservices Framework for Indoor Tracking System Development	Gafur Hayytbayev Kerem Küçük Mahmut Çavur	1168 - 1172
DM	**1927**	Unsupervised Pattern Extraction of Time Series Data for Energy Disaggregation	Şirin Azazi Deveci Melih Günay	1173 - 1178
	1944	Topic Modeling Enhanced Tripartite Graph for Recommendation using Metapaths	Yaren Yılmaz Irem İşlek Şule Gündüz Öğüdücü	1179 - 1184
	1948	Community Detection on Software Library Dependency Graphs using Graph Neural Networks	Şevket Umut Çakır	1185 - 1190

			Mehmet Ali Osman Atik Ümit Deniz Uluşar	
	2190	Enhancing Mesh and Point Cloud Similarity Detection through Geometric Features and ICP	Talha Rehman Abid Mehtap Öklü Cem Yıldız Ali Erman Erten Kamer Kaya	1191 - 1196
	2214	Comparative Analysis and Practical Implementation of Machine Learning Algorithms for Phishing Website Detection	Samad Najjar-Ghabel Shamim Yousefi Payam Habibi	1197 - 1202
	2215	A Technical Analysis and Practical Implementation of Machine Learning Algorithms for Predicting Survival in Breast Cancer Patients	Shamim Yousefi Samad Najjar-Ghabel Hamidreza Shafaei	1203 - 1208
BIG	**1881**	Comparison Between Time Series and Relational Databases	Alpar Türkoğlu Onurcan Ersen Ibrahim Onuralp Yiğit Dincer Unal Hatice Golcuk	1209 - 1212
	1930	A Performance Evaluation Study on a Data Analytics Platform for Emergency Calls	Engin Yakar H. Hakan Kilinc	1213 - 1218
	2079	Adaptive Composite Market Volatility Index (CMVI) for Enhanced Stock Price Forecasting	Rabia Çevik Uğur Barış Özyürek Ali Kanal Vael Kokach Büşra Kocaçınar Oznur Şengel Fatma Patlar Akbulut	1219 - 1223
	2142	Hybrid Deep Learning Framework for Stock Price Prediction Incorporating Technical and Macroeconomic Indicators	Ali Can Turan Vael Kokach Büşra Kocaçınar Oznur Şengel Fatma Patlar Akbulut	1224 - 1228
	2125	Emotion-Aware Multimodal Biometric Identification by using Biosignals	Yekta Said Can Beyzanur Bektan Fatih Alagöz	1229 - 1235
	1854	Özbekçe-Türkçe Otomatik Çeviri Yazılımı için Deyimlerin Veritabanını Teşkil Etmede Karşılaşılan Güçlükler Automatic Translation Software Difficulties in Organizing the Database of Idioms for Uzbek and Turkish	Manzura Abjalova Umida Raşidova Eşref Adalı	1236 - 1240
	2028	Reversible Steganographic System for the Transmission of Personal Medical Data	Elmira Daiyrbayeva Ekaterina Merzlyakova Aigerim Yerimbetova Aigul Mukhitova	1241 - 1246

Özbek Dilinde Kelimelerin Kök ve Köklerini Ayırma Algoritmaları

Algorithms for Parsing Roots and Stems of Words in Uzbek Language

Elov Botir Boltayeviç
Bilgisayar Dilbilimi ve Dijital Teknolojiler,
Taşkent Devlet Özbek Dili ve Edebiyatı Üniversitesi
Taşkent, Özbekistan
elov@navoiy-uni.uz

Xusainova Zilola Yuldashevna
Bilgisayar Dilbilimi ve Dijital Teknolojiler,
Taşkent Devlet Özbek Dili ve Edebiyatı Üniversitesi
Taşkent, Özbekistan
xusainovazilola@navoiy-uni.uz

Umirova Svetlana Mamurjonovna
Özbek Dilbilimi Bölümü,
Semerkant Devlet Üniversitesi
Semerkant, Özbekistan
umirova.s.m06@mail.ru

Normamatov Sultonbek Ermamatovich
Taşkent Devlet Özbek Dili ve Edebiyatı Üniversitesi Taşkent, Özbekistan
normatov@navoiy-uni.uz

A.Şahlo Abdimurod kızı
Bilgisayar Dilbilimi ve Dijital Teknolojiler, Taşkent Devlet Özbek Dili ve Edebiyatı Üniversitesi
Taşkent, Özbekistan
abdisalomovshahlo@gmail.com

Mahmadiev Shavkatjon Suyunkulovich *Özbek Dilbilimi Bölümü,*
Semerkant Devlet Üniversitesi
Semerkant, Özbekistan
shavkatmahmadyev1975@gmail.com

***ÖZ.*—Özbek dilinde kelime köklerini belirlerken şu sorunlar ortaya çıkabilir: Kök ve ekin aynı kök ile eşsesli olması; Bir kelimenin ses değişikliğiyle buluşması; yeni sözcük ve NER(Named entity recognition)leri stemlemek. Bu makalede sunulan performans göstergesi %97,5'e eşit UzbStemming algoritma yukarıdaki sorunları çözer ve onu bir yazım denetleyici, makine çevirisi, soru-cevap sistemleri, sözdizimsel ve anlamsal analizörler gibi NLP uygulamaların geliştirilmesinde kullanılabilir. Lemmalamak – metindeki kelimelerin kanonik biçimlerinin belirlenmesi işlemi olarak, her hangi NLP sistemin önemli bir bileşenidir ve doğal dil anlayışına dayanan çoğu uygulama için önemli bir işlem adımıdır. Lemming işlemi, bir kelimenin anlamının kökten ayırma işleminden daha kesin bir temsilini sağlar. Bu yazıda çeşitli Özbekçe kelimeler, deyimler, NER (ünlü isim), neologizm, lemming kısaltmalarının yöntemleri ve algoritmaları sunulmaktadır.**

Anahtar Sözcükler—component, formatlama, dserlem, kökü bulma, NLP algoritmaları,

***Abstract.*-When determining the roots of words in the Uzbek language, the following problems may arise: homonymity of the root and the suffix with the same root; the meeting of a word with a sound change; stemming new words and NER (Named entity recognition). The performance indicator presented in this article is equal to 97.5% UzbStemming algorithm solves the above problems and can be used in the development of NLP applications such as a spelling checker, machine translation, question-answer systems, syntactic and semantic analyzers. Lemming - as the process of determining the canonical forms of words in the text, is an important component of any NLP system and an important processing step for most applications based on natural language understanding. Lemming operation provides a more accurate representation of the meaning of a word than the process of rooting. This article presents methods and algorithms for the processing of various Uzbek words, idioms, NER (famous name), neologisms, lemming abbreviations.**

Keywords- component, formatting, corpus, lemming

I. GIRIŞ

Kökten çıkarma, DDİ'de metin işlemenin ilk adımıdır ve bu süreç birçok DDİ görevi için gereklidir. Kök çıkarma işleminin temel amacı kelimenin çekimli hallerini kök haline getirmektir. Lemming'den farklı olarak kök çıkarmada sözlük veya biçim bilimsel çözümleme yapılmaz [1]. Kökten ayrılmayla oluşturulan kökün, asıl kelimeyle veya onun biçim bilimsel köküyle aynı olması gerekmez. Köklenme süreci kelimelerin oluşumuyla ilgilidir: Kökün değişmesine ve oluşumuna dayanır. Sonuç olarak, kelime biçimleri birleştirmeye getirilir.

Köklenme süreci kelimelerin oluşumuyla ilgilidir: Kökün değişmesine ve oluşumuna dayanır. Sonuç olarak, kelime biçimleri birleştirmeye getirilir. Kelime formları önek ve son ekleri içerebilir. Kelime formlarında yapım ve çekim ekleri olarak ikiye ayrılır.

II. BILIMSEL ARAŞTIR ANALIZI

J.B.Lovins 1968'de metinlerdeki köklerden biri hakkında ayrıntılı bilgileri yayımlayan ilk kişi oldu [2]. Lovins önerdiğii algoritmaya dayanarak malzeme bilimi ve mühendisliği alanında10.000'den fazla belge işlendi. Lovins algoritması iki aşamadan oluşur:

1. kökü tanımlamak için mümkün olduğu kadar uzun bir son ekin silinmesi;
2. yazım kuralının istisnalarını gözden geçirmek.

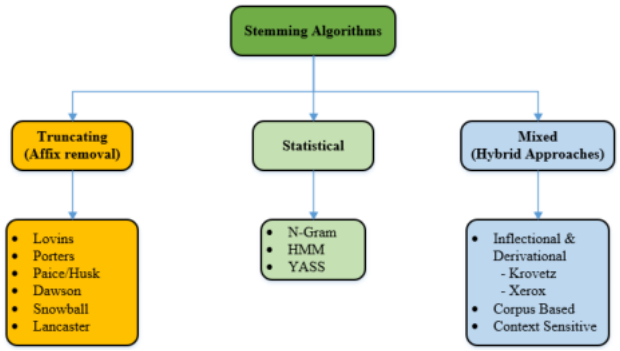
1980 yılında M. Porter ekleri ayırma kurallarına dayanan bir kök ayırıcı algoritması önerdi. Algoritma, beş aşamalı ve 10.000 kelimelik bir sözlüğe sahip kök ayırıcının metin/derlem boyutunu %33 oranında azalttığını buldu. Temel Kombine Programlama Dili (BCPL) derlemindeki birçok kelime, 1. adımda çoğul ve geçmiş zaman kelimeleri kısaltılarak kısaltılmıştır.

1990 yılında Lancaster Üniversitesi'nden Chris Pace, dilbilgisi kurallarına dayalı olarak son ekleri yinelemeli olarak kaldıran yeni bir kök ayırıcı geliştirdi. Bu kök düzeltme algoritması, kural istisnalarının oluşmaması için son ekleri değiştirmeyi amaçlamaktadır. Bu kök saplayıcıya genellikle Paice/Husk kök saplayıcı veya Lancaster kök saplayıcı denir.

2001 yılında M. Porter geliştirdiği kök ayırıcının birçok uygulamasında hatalar olduğunu keşfetti. Aynı zamanda Hollanda dili için bir Kraaij-Pohlmann kökü [3] açık kaynaklı ANSI C yazılımı olarak yayımlandı ancak algoritmaya ilişkin herhangi bir açıklamanın bulunmadığı belirtildi. Bu sorunları çözmek için Porter, kuralların doğal bir şekilde ifade edilebildiği Snowball sistemini önerdi ve C dilinde bir programlama kodu geliştirdi. İngilizce için Snowball uygulamasına Porter2 rootmer da denir. O 2019'a kadar, Snowball C, Java, Python, Go, Rust, C# ve diğer dillerde oluşturulan kodlarla doğal dilleri destekler.

A.Uyar Google, 2009'dan bu yana arama motorunda kullanılan kökten düzeltme mekanizmalarını araştırıyor. 18.000 farklı kelimeyi kullanarak Google'ın sisteminin belge tabanlı bir algoritma kullandığı belirledi [4]. Dokümanların anlamsal olarak güçlü bir şekilde ilişkili kelime formlarına göre indekslendiğini belirtti.

G. Anjali, kök çıkarma algoritmalarını üç gruba ayırır: kesme (truncating), istatistiksel (statistical) ve karma (mixed) [5]. Bu grupların her birine yönelik algoritmalar, belirli yaklaşımlara dayalı olarak kelime varyant köklerini tanımlamaya yardımcı olur. Bu algoritmaların ve yöntemlerin sınıflandırması aşağıdaki Şekil 1'de gösterilmektedir:



Şekil 1. Köklendirme algoritmalarının sınıflandırılması

Lemmalamak, bir sözcük biçiminden veya kökten bir lemmanın (sözcük biriminin) anlamlı biçimini belirleme ve bu birimleri doğal dilde işleme yöntemine verilen addır. Lemmaing birçok DDİ görevi için önemli bir adımdır. Lemmas, bir kelimenin sözcüksel olarak normalleştirilmiş biçimini bulma sürecidir.

Lemmas iyi çalışılmış bir süreçtir. Özellikle A. Bjorkeland, D. Seddah ve A. Fraser, metinlerin söz dizimsel analizinde ve makine çevirisinde lemming kullanımı konusunu incelediler [6]. Kelimeleri sözcük kaynaklarıyla karşılaştırmak ve çekim biçimleri arasında ilişkiler kurmak için Lemma tanımlaması gereklidir. Lemmalar biçim bilimsel açıdan zengin diller için özellikle önemlidir. G. Chrupala, doğal dilden bağımsız olarak jeton tabanlı bir lemma yöntemi önerdi ancak bu yöntemin verimliliğinin düşük olması nedeniyle bu algoritma pratikte nadiren kullanılıyor [7].

Dil derlem metinleri bağlamında sözcük birimlerinin lemması, POS etiketi ve biçim bilimsel nitelikleri arasında bir ilişki vardır. S. Manning, L. Padro ve E. Stanilovski, araştırmalarında niteliklerin lemması kelimesinin ifadesine dikkat çektiler [8].

Kanerva, F. Ginter, T. Salakoski, diziye (sequence-To-sequence) dayalı sinir ağı mimarisine ve morfo söz dizimsel bağlam gösterimine dayanan yeni bir lemma yöntemi sundular [9]. Önerilen yöntemde bağlama duyarlı bir lemmatizer, biçim bilimsel etiketleyiciden ve onun morfo söz dizimsel özelliklerinden elde edilen bilgilere dayanarak bir lemmayı tanımlar. Geliştirilen lemmatizer 52 farklı dil ve 76 farklı treebankalar üzerinde değerlendirildi. Ortaya çıkan sonuçlar, 76 treebankadan 62'sinde ortak temel UDPipe Future'dan daha iyi başarım gösterdi ve hataları ortalama %19 oranında azalttı. Lemmatizer, tüm eğitilmiş modellerle birlikte Turku-neural-parsing-pipelinenin bir parçası olarak Apache 2.0 lisansı altında sağlanmaktadır.

E. Arslan ve U. Orhan, morfolojik benzerliği kullanarak Türkçe kelimeler içinin grafik tabanlı bir lemma yöntemi geliştirdiler [10]. Bu lemma yöntemi, POS etiketleme, yazı denetimi ve belge kümeleme gibi DDİ uygulamalarında kullanılmıştır. O. ÄÖztürkmenoğlu ve A.T Alpkojak Türkçedeki bir dizi metin hakkında bilgi edinmek amacıyla farklı lemma yaklaşımlarını karşılaştırmak için yöntemler geliştirdiler. Çalışma, bir dizi Türkçe metin üzerinde dört farklı lemmatizer ve belirli uzunluk budama yaklaşımlarını karşılaştırdı. Birincisi, Türk dilinin morfolojik çözümleyicisine dayanmaktadır; bir diğeri ise veri yapısını kullanan Dictionary-based Türkçe Lemmatizer'dır (DTL); üçüncüsü basit bir sözlük tabanlı, yukarıdan aşağıya

ayrıştırmadır ve sonuncusu, kelimelerin tam uzunluğa göre kesilmesidir.

Özbek dilinde sözcük birimlerinin Lemming algoritmalarla ilgili çeşitli çalışmalar küçük metinler üzerinde yapılmış ve sınanmıştır. Özellikle, I. Bakayev dil kurallarına dayalı bir yaklaşımla Özbek dilindeki kelimelerin köklerini belirleme algoritmasının geliştirilmesi konularını inceledi [11]. M. Sharipov, O. Sobirov'ler Özbek dili için sonlu durum makinesi (Finite State Machine) ile kural tabanlı bir lemma algoritması oluşturdu [12].

Bu bildirimizde Özbek dilinde eklerin ayrılmasına dayanan UzbStemming algoritmasını ve farklı yapıya sahip kelimelerin lemming algoritmalarını sunuyoruz.

III. EK AYRIMINA DAYALI UZBSTEMMING ALGORITMASI

Ek kaldırmaya dayalı Özbekçe kök bilgisinin geliştirilmesi için dil derlemi, ek veri tabanı, Özbekçe morfolexicon, kök veri tabanı kullanılmıştır [13,14]. Kökün belirlenmesi için kelime form grubu (POS etiketi) dikkate alınarak aşağıdaki iki aşamalı işlem gerçekleştirildi.

I adım. Verilerin hazırlanması (veri kümesi).

1) Köke eklenebilecek önek (prefiks) tanımlanmıştır (Çizelge I).

ÇIZELGE I. ÖZBEK DILINDE ÖNEKLER

Önekler		
ba-	be-	ham-
bo-	no-	bar-
bad-	ser-	be-
bar-	g'ayri-	xush-

2) Belirli bir kelime grubunun köklerine eklenebilecek bir dizi önek ve sonek dizisi, her bir kelime grubunun çeşitli (aktif) kelime formları alınarak ve bunların köklerinin kelime grubundan silinmesiyle belirlenir.

3) Ekler dizisi, bu kelime biçimini oluşturmak üzere köke eklenen birçok Özbek ekinin birleşimidir (Çizelge II).

ÇIZELGE II . DIL KÜLLIYATINDAN TANIMLANAN ÖZBEKÇE ÖNEK VE SONEK KOMBINASYONLARININ BIR KOLEKSIYONU

	Isim (N)	Sayı (Num)	Fiil (Vb)
1.	-chiligimizning	-inchi	-dilarki
2.	-dagilarning	-larcha	-nganlaridan
3.	doshlarimiznikidan	-larda	-yapsizlarmi
4.	-korlarimizning	-ovi	ilmayotganidan
5.	-laganlaridagina	ovimiz	-ishimizdan
…	...	...	...
Toplam:	**1773**	**229**	**15870**
	Sıfat (JJ)	**Zamir (P)**	**Zarf (R)**
1.	-liligi	-nga	-larcha
2.	-bardoshligini	-ngacha	-ona
3.	-imizning	-ngagina	chiligimizning
4.	-lamaganidan	nikilarning	-dangina
5.	lashayotganligini	-dek	-gilarining
…	...	...	...
Toplam:	**725**	**509**	**1376**

4) Köke eklenebilecek tüm ekleri dil derleminden almak için özel bir yazılım uygulaması1 kullanılır.

5) Bu yazılım aracının yardımıyla Özbek dili külliyatındaki mevcut kelime şekillerinden verilen düzenli ifadelere dayanarak gerekli kelime kalıplarının bir tabanını oluşturmak mümkündür.

6) Arama sonucunda belirlenen bazı kelime biçimleri önek, kök ve sonek dizisi gibi bölümlere ayrılmıştır. Çizelge III'te bir örnek veriyoruz:

ÇIZELGE III. KELIME BIÇIMLERINI ÖNEKLERE, KÖKLERE VE SON EKLERE BÖLME

Kelime bilgisi	önek	stem	Sonekler satırı
kitobning		kitob	ning
kitoblarning		kitob	larning
shahringdan		shahr	ingdan
shaharda		shahar	da
badavlat	ba	davlat	
berahm	be	rahm	

*7) **p+s+q=w** koşulu karşılayan tüm verilen kelime biçimleri için **g={p,s,q}** değerler kümesi tanımlanır. Burada,*

***w**: kelime formu;*

***s**: stem;*

p ∈P

***P**: önek dize kümesi*

s ∈S

k: kök kelimelerden oluşan bir toplam;

q ∈Q

***Q**: sonek dize kümesi*

g ∈G

***G**: her öğe*

***{p,s,q}** üç parametreden oluşan bir toplam.*

8) Belirlenen p{,s,q} değerlerin bir örneği Çizelge IV'te verilmiştir:

ÇIZELGE IV. PREFIKS, STEM VE SUFFIKSLER

Prefiks (p)	Stem kelime grubu (s)	Son ek zarf dizesi (q)
	isim	-ning
	isim	-larning
	isim	-da
	isim	-lariga
	fiil	-di
	fiil	-ysan
	isim	-ga
	isim	-lari
	isim	-iga
	isim	-ingdan
	isim	-da
ba-	isim	
be-	isim	

Çizelge IV'teki değerler belirlendikten sonra aşağıdaki **UzbStemming algoritması** kullanılarak kelime formunun kökü belirlenebilir.

II adım. UzbStemming algoritmı

Algoritmada aşağıdaki tanımlar kullanılmaktadır:

***R**: kök kümesi;*

***G{p,s,q}**: değerler kümesi;*

***T_G**: toplam elementlerin P ve q değerleri azalan uzunluk sırasına göre sıralanan bir küme;*

***V**: bir dizi sesli harf;*

***C**: bir dizi ünsüz değer;*

***W**:kelime formu;*

***replacelastletter(x,y)_x**: satırın son harfini y harfiyle değiştirme işlevi;*

***removelastvowel(x)_x**: satırın sonundaki 2. sesli harfi silme işlevi;*

***right(x,n)_x**: dizenin sonundaki n harfini bulma işlevi;*

***left(x,n)_x**: dizenin başlangıcındaki n harfini bulma işlevi;*

***RPOS(r)_r:** kökteki kelime gruplarının listesini belirleyen bir işlev. Köklerin tabanına ve kelime gruplarına göre belirlenir;*

***QPOS(p,q)_p ve Q;** son ek dizesine eklenebilecek tümceciklerin listesini belirleyen bir işlev. Bu fonksiyon Çizelge 5'e göre belirlenir.*

1. **w** bir satır girin.
2. **t** kümenin her elementi için
   - 2.1. Eğer **w** satırı **p** ile başlayarak **q** ile biterse, 2.2.adıma geçilsin, aksi takdirde 2.6.adıma geçilsin.
   - 2.2. **w** satırın **p** ve **q** parçaları keselim, gerisini **r'** ile ifade ediyoruz. Burada, ***p+r'+q=w*** eşitliği uygundur.
   - 2.3. **r'** dize en az bir sesli harf içeriyorsa, 2.4.adıma geçilsin, aksi takdirde 2.6.adıma geçilsin.
   - 2.4. Eğer ***r'∈R*** olursa, 4.adıma geçilsin, aksi takdirde 2.5.adıma geçilsin.
   - 2.5. **R** toplamın her elementi **r** için
     - 2.5.1. Eğer ***RPOS(r)∩QPOS(p,q)=∅*** olursa, 2.5.19.adıma geçilsin, aksi takdirde ***q1=left(q,1)***, 2.5.2.adıma geçilsin.
     - 2.5.2. Eğer ***q1=right(r,1)="k"*** ve ***q1="i"*** olursa, 2.5.3.adıma geçilsin, aksi takdirde 2.5.5.adıma geçilsin.
     - *2.5.3.* ***z=replacelastletter(r",g")***
     - 2.5.4. Eğer ***r'=z*** olursa, 4.adıma geçilsin.
     - 2.5.5. Eğer right(r,1)="q" ve q1="i" olursa, 2.5.6.adıma geçilsin, aksi takdirde 2.5.8.adıma geçilsin.
     - *2.5.6.* ***z=replacelastletter(r",g")***
     - 2.5.7. Eğer ***r'=z*** olursa, 4.adıma geçilsin.
     - *2.5.8.* ***l=length(r)***
     - 2.5.9. Eğer ***q1="i"*** ve ***l>4*** ve ***right(r,1) ∈ C*** ve ***right(r,2) ∈ V*** olursa, 2.5.10.adıma geçilsin, aksi takdirde 2.5.12.adıma geçilsin.
     - *2.5.10.* ***z=replacelastvowel(r)***
     - 2.5.11. Eğer ***r'=z*** olursa, 4.adıma geçilsin.
     - 2.5.12. Eğer ***"Verb" ∈PROS(r)*** olursa, 2.5.13.adıma geçilsin, aksi takdirde 2.5.19.adıma geçilsin.
     - 2.5.13. Eğer ***right(r,1)"a"*** olursa, 2.5.14.adıma geçilsin, aksi takdirde 2.5.16.adıma geçilsin.

2.5.14. ***z=replacelastletter(r,"o")***

2.5.15. Eğer ***q1∈{"v","q"}*** ve ***r'=z*** olursa, 4.adıma geçilsin.

2.5.16. Eğer right(r,1)="i" olursa, 2.5.17.adıma geçilsin, aksi takdirde 2.5.19.adıma geçilsin.

2.5.17. ***z=replacelastletter(r,"u")***

2.5.18. Eğer ***q1∈{"v","q"}*** ve ***r'=z*** olursa, 4.adıma geçilsin.

2.5.19. ***R*** toplamın bir sonraki elementi alınsın, 2.5.1.adıma geçilsin.

2.6. ***T*** toplamın bir sonraki elementi alınsın ve 2.1.adıma geçilsin.

2. "Stem bulunamadı". 5.adıma geçilsin.

3. Stem ***r'***.

4. Bu kadar.

IV. Özbek Dilinde Lemming Kelimelerinin Algoritması

Lemming algoritması bir kelimenin kökünü bulmanın karmaşık bir yoludur. Bu süreç, kelime kümesinin (POS) bağlamının anlaşılmasını ve her POS için farklı normalleştirme kurallarının uygulanmasını gerektirir. Özbek dili eklemeli diller ailesine ait olduğundan cümlelerde belirteç sayısı kadar lemma da bulunabilir.

Özbek dilindeki kelime yapılarına göre temel olarak şu yapıya sahiptirler: Basit kök kelime, basit yapay kelime, birleşik kelime, çift kelime, tekrarlanan kelime ve karmaşık kelime [15].

Basit sözcük. Basit sözlük birimler tarihsel olarak bölünemez. Tek biçimbirimden oluşurlar ve geleneksel dilbilimde "kök", "gövde" olarak adlandırılan birimlere karşılık gelirler. [bog'], [dala], [kel-], [qoch-], [oq], [sariq], [men], [u], [ikki], [o'n] gibi basit sözcüklerdir.

Basit kelimeler de iki türdür: Kök kelimeler (biçimlendirilmemiş kelimeler) ve yapay kelimeler (eklenerek oluşturulan kelimeler):

1) kök (root): universitet, kitob, bola, gul;

2) kök + {der.aff} kitobxon, bolalik, gulchi.

Özbek dilinde lemma, bir kelimenin gramer zarfları olmayan, ancak kelime oluşturan zarfların bulunduğu bağlamla ilgili durumudur. Örneğin, *işçi, işçiler, işçiler çalıştı* sözcüklerinin kökü *iş'*tir, ancak *işçi, işçiler, işçilerin* sözcüklerinde lemma *işçi*dir ve *çalıştı* fiilinde ise lemma *çalışmaktır [16]*. Ancak Özbek dilinin kelime biçimine göre sadece basit olmayıp kök+kök şeklinde karmaşık bir yapıya sahip olması nedeniyle lemma belirleme yöntemlerinin incelenmesi gerekmektedir.

Çift sözcük. Bir kelime çiftinin unsurları hükümdar-ast ilişkisinde değil, eşitlik ilişkisindedir. Kelime çiftleri halinde, bileşenleri genellikle anlamlarını korur. Örneğin, hol-ahvol, baxt-saodat, opa-singil, qozon-tovoq, aql-hush, katta-kichik, baland-past, oyoq-qo'l bu gibi durumlarda iki kelimenin birleşmesinden üçüncü bir kelime oluşmaz.

Bu özelliğe göre bir kelime çiftinde lemmalar sayısı kadar belirteç bulunur.

Oluşma şekillerine göre aşağıdaki kelime çiftleri vardır:

1) kök– kök: osh-tuz, baxt-saodat;

2) (anlamlı) kök – (anlamsız) kök: kiyim-kechak, upa-elik;

3) (anlamsız) kök – (anlamlı) kök: dov-daraxt, qo'ni-qo'shni;

4) (anlamsız) kök – (anlamsız) kök: g'ala-g'ovur, adi-badi, ikir-chikir;

5) kök– kök +{der.aff}: baxt-saodatli, tinch-totuvlik, kitob-daftarsiz;

6) (anlamlı) kök – (anlamsız) kök+{der.aff}): g'ala-g'ovurli, irim-sirimli.

Bir kelime çiftine kelime oluşturan bir son ek eklendiğinde kısa çizgi korunur: *baxt-saodatli.* Özbek dilindeki kelime çiftlerinin lemmasını belirlerken sözlüğe güveniyoruz, yani bağımsız anlamlara sahip kelime çiftlerinin iki lemması olduğu kabul ediliyor.

Fonetik yöntemle yapılan bir sıfat veya zarfın yükleme düzeyi (-) ile yazılır ve lemma onun anlamlı kısmından oluşur: *sap-sariq(sariq), qip-qizil (qizil), tuppa-tuzuk (tuzuk).*

Tekrar sözcük. Özbek dili ve diğer Türkiy dilleri ile ilgili literatürün çoğunda tekrarlanan kelimeler geniş anlamda çift kelime olarak yer almaktadır. Bu türlerin ayırt edici özellikleri şunlardır:

1. Bir kelime çifti her zaman başka kelimelerden oluşur: *aql-hush*.
2. Bir kelime çiftinin bir veya her iki unsuru da modern Özbekçede bağımsız anlamını kaybetmiş olabilir: *amal-taqal, ikir-chikir* kabi.
3. Bir kelime çifti esas olarak genelleme, bütünlük ve farklı anlam tonları (anlam renkleri) anlamlarını ifade ederken, tekrarlanan kelimeler ise çoğulluk, süreklilik, tekrar gibi anlamları ifade etmektedir.

Çoğaltılmış kelimeler tüm bağımsız kelime gruplarında ve taklit kelimelerde görülür. Tekrarlanan bir kelime, bir kelimenin basit bir tekrarı anlamına gelmez. Yalnızca belirli bir dilbilgisel anlamı ifade eden ve sözcüksel-dilbilgisel bir bütün oluşturan tekrarlar, tekrarlanan sözcükler olarak kabul edilir [17].

Tekrarlanan sözcükler türemiş sözcükler olmadığından, bunların iki belirteç ve iki lemma içerdiğini varsayıyoruz. Ancak tekrarlanan kelimenin ikinci kısmının hiçbir anlam taşımadığı istisnai bir durum vardır ve bu da lemma sayısının belirlenmesinde sorun yaratır. Bu durumda sözlüğe dayalı olarak anlamlı kısım alınır.

(anlamlı) kök +(anlamlı) kök: tez-tez, qator-qator, qop-qop;

(anlamlı)kök + (anlamsız) kök: tuz-puz, don-dun, osh-posh.

Tekrarlanan kelimelere kelime oluşturan bir ek (der.aff) eklenirken yeni bir kelime oluşturulur ve tire bırakılır: *g'uvg'uvladi (g'uv-g'uv), xoxoladi (xo-xo) …*

Bu nedenle dilde, sözcük biçiminin lemmasının belirlenmesinde bağlamda işlev gören, bağımsız anlam taşıyan bir parça alınır. Kelime türlerindeki tüm biçim oluşturan ekler kaldırılır ve sözcük oluşturan ekler bırakılır.

Sonuç

Özbek dilinin eğitim külliyatında (http://uzschoolcorpara.uz/uz/Search)deki 100.000'den fazla cümleye eklerin kaldırılmasına dayanan yukarıda belirtilen

UzbStemmer'in uygulanması sonucunda %97,5 doğruluk oranı elde edilmiştir. UzbStemmer algoritmasının etkinliğini arttırmak amacıyla, algoritmanın I aşaması (veri hazırlama) 3.adımında belirlenen ek grubunda bulunmayan eklerin kombinasyonunun bir yöntemle belirlenmesi dil kurallarına dayalı hibrit yaklaşım istenmektedir.

KAYNAKÇA

[1] B.B. Elov, S. M. Khamroeva, S. M., ve Z.Y. Xusainova, Z. Y. (2023). The pipeline processing of NLP. In *E3S Web of Conferences* (Vol. 413, p. 03011). EDP Sciences.

[2] J.B. Lovins, (1968). Development of a stemming algorithm. *Mech. Transl. Comput. Linguistics*, *11*(1-2), 22-31.

[3] W. Kraaij, ve R. Pohlmann, (1996, August). Viewing stemming as recall enhancement. In *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 40-48).

[4] A. Uyar, (2009). Google stemming mechanisms. *Journal of information science*, *35*(5), 499-514.

[5] A. G. Jivani, (2011). A comparative study of stemming algorithms. *Int. J. Comp. Tech. Appl*, *2*(6), 1930-1938.

[6] A. Björkelund, B. Bohnet, L. Hafdell, ve P. Nugues, P. (2010, August). A high-performance syntactic and semantic dependency parser. In *Coling 2010: Demonstrations* (pp. 33-36).

[7] G. Chrupała, (2006). Simple data-driven context-sensitive lemmatization. *Procesamiento del lenguaje natural, nº 37 (sept. 2006), pp. 121-127*.

[8] C. D. Manning, M. Surdeanu, J. Bauer, J.R. Finkel, S. Bethard, ve D. McClosky, (2014, June). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations* (pp. 55-60).

[9] J. Kanerva, F. Ginter ve T. Salakoski, (2021). Universal Lemmatizer: A sequence-to-sequence model for lemmatizing Universal Dependencies treebanks. *Natural Language Engineering*, *27*(5), 545-574.

[10] E. Arslan, E. ve U. Orhan, (2016, August). Graph-based lemmatization of Turkish words by using morphological similarity. In *2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)* (pp. 1-5). IEEE.

[11] B. I. Izatovich, (2021, July). Development of a stemming algorithm based on a linguistic approach for words of the uzbek language. In *E-Conference Globe* (pp. 195-202).

[12] M. Sharipov ve O. Sobirov, (2022). Development of a rule-based lemmatization algorithm through Finite State Machine for Uzbek language. *arXiv preprint arXiv:2210.16006*.

[13] E. B. Boltayevich,, S. S. Samariddinovich, K. S. Mirdjonovna, E. Adalı, ve X. Z. Yuldashevna, (2023, September). Pos Taging of Uzbek Text Using Hidden Markov Model. In *2023 8th International Conference on Computer Science and Engineering (UBMK)* (pp. 63-68). IEEE.

[14] E. B. Boltayevich, E. Adalı, K. S. Mirdjonovna, A. O. Xolmo‘Mınovna, X. Z. Yuldashevna, X. Nizomaddin ve U. O‘g‘li, The Problem of Pos Tagging and Stemming for Agglutinative Languages.

[15] Х. Неъматов, ve Р. Расулов, (1995). Ўзбек тили лексикологияси асослари. *Т.: Укитувчи*.

[16] B. B. Elov, S. M. Khamroeva, R.H. Alayev, Z.Y. Khusainova ve U. S. Yodgorov,. (2023). Methods of processing the uzbek language corpus texts. *International Journal of Open Information Technologies*, *11*(12), 143-151.

[17] А. Ҳожиев, (1963). Ўзбек тилида қўшма, жуфт ва такрорий сўзлар.-Тошкент:" Фан".