MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN

ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН

KAZAKHSTAN

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

Published since 2020.
Volume 7. 2 (26). 2026
April–June

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

2020 жылдан бері шығарылады
Том 7. 2 (26). 2026
Сәуір-Маусым

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

Издается с 2020 г.
Том 7. 2 (26). 2026
Апрель-Июнь

Свидетельство о постановке на учет периодического печатного издания в Министерство информации и общественного развития Республики Казахстан № KZ82VPY00020475, выданное от 20.02.2020 г.

Зарегистрировано в Международном центре регистрации серийных изданий ISSN (ЮНЕСКО, Париж, Франция). ISSN 2708–2032 (print), ISSN 2708–2040 (online)

Журнал входит в Перечень научных изданий, рекомендуемых КОКНВО МНВО РК для публикации основных результатов научной деятельности.

EDITOR-IN-CHIEF:

Kateryna Kolesnikova — Doctor of Technical Sciences, professor, Vice-Rector for Research, International Information Technology University (Kazakhstan)

DEPUTY EDITOR-IN-CHIEF:

Madina Ipalakova — Candidate of Technical Sciences, associate professor, Director of the Research Department, International Information Technology University (Kazakhstan)

EDITORIAL BOARD:

Abdul Razak — PhD, professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)
Lucio Tommaso De Paolis — Director of the R&D Department of the AVR Laboratory, Department of Engineering for Innovation, University of Salento (Italy)
Liz Bacon — Professor, Deputy Vice-Chancellor, Abertay University (United Kingdom)
Michele Pagano — PhD, Professor, University of Pisa (Italy)
Mukhtarbay Otelbayev — Doctor of Physical and Mathematical Sciences, professor, academician of the National Academy of Sciences of the Republic of Kazakhstan, professor of the Department of Mathematical and Computer Modeling, International Information Technology University (Kazakhstan)
Bolatbek Rysbaiuly — Doctor of Physical and Mathematical Sciences, professor, professor of the Department of Computing and Data Science, Astana IT University (Kazakhstan)
Yevgeniya Daineko — PhD, research professor, Department of Information Systems, International Information Technology University (Kazakhstan)
Nurzhan Duzbayev — PhD, associate professor, Vice-Rector for Digitalization and Innovation, International Information Technology University (Kazakhstan)
Bakhtgerei Sinchev — Doctor of Technical Sciences, professor, Department of Information Systems, International Information Technology University (Kazakhstan)
Nurgul Seilova — Candidate of Technical Sciences, Dean of the Faculty of Computer Technologies and Cybersecurity, International Information Technology University (Kazakhstan)
Ardak Mukhamediyeva — Candidate of Economic Sciences, Dean of the Faculty of Business, Media and Management, International Information Technology University (Kazakhstan)
Zamira Abdikalikova — PhD, associate professor, Head of the Department of Mathematical and Computer Modeling, International Information Technology University (Kazakhstan)
Yerlan Shildibekov — PhD, associate professor, Head of the Department of Economics and Business, International Information Technology University (Kazakhstan)
Damilya Yeskendirova — Candidate of Technical Sciences, associate professor, Head of the Department of Cybersecurity, International Information Technology University (Kazakhstan)
Aigul Niyazgulova — Candidate of Philological Sciences, Professor, Head of the Department of Media Communications and History of Kazakhstan, International Information Technology University (Kazakhstan)
Altai Aitmagambetov — Candidate of Technical Sciences, Professor, Department of Radio Engineering, Electronics and Telecommunications, International Information Technology University (Kazakhstan)
Yelena Bakhtiyarova — Candidate of Technical Sciences, associate professor, Head of the Department of Radio Engineering, Electronics and Telecommunications, International Information Technology University (Kazakhstan)
Kanibek Sansyzbay — PhD, research professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)
Sakhybay Tynymbayev — Candidate of Technical Sciences, Professor, Research Professor, Department of Computer Engineering, International Information Technology University (Kazakhstan)
Ali Abd Almisreb — PhD, associate professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)
Mohamed Ahmed Hamada — PhD, associate professor, Department of Information Systems, International Information Technology University (Kazakhstan)
Yang Im Chu — PhD, Professor, Gachon University (South Korea)
Tadeusz Wallas — PhD, Vice-Rector, Adam Mickiewicz University (Poland)
Orken Mamyrbayev — PhD, Deputy Director for Science, RSE Institute of Information and Computational Technologies, Committee for Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Kazakhstan)
Sergey Bushuyev — Doctor of Technical Sciences, professor, Director of the Ukrainian Project Management Association "UKRNET," Head of the Department of Project Management, Kyiv National University of Construction and Architecture (Ukraine)
Svetlana Beloshitskaya — Doctor of Technical Sciences, professor, Department of Computing and Data Science, Astana IT University (Kazakhstan)

MANAGING EDITOR

Raushan Mrzabayeva — Master of Science, editor, International Information Technology University (Kazakhstan)

International Journal of Information and Communication Technologies
Periodicity: 4 times a year.
Languages: Kazakh, Russian, English
DOI prefix: 10.54309
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Thematic focus: "Information technology"; "Digital technologies in the development of socio-economic systems"; "Information security and communication technologies".
Distribution: Materials are distributed under the Creative Commons Attribution 4.0
Journal website: https://journal.iitu.edu.kz
Owner: International Information Technology University JSC (Almaty).
Copyright: © International Journal of Information and Communication Technologies, 2026

РЕДАКЦИЯ

БАС РЕДАКТОР:

Колесникова Катерина Викторовна — техника ғылымдарының докторы, профессор, Халықаралық ақпараттық технологиялар университетінің ғылыми-зерттеу қызметі жөніндегі проректор (Қазақстан)

БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:

Ипалакова Мадина Тулегеновна — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университетінің ғылыми-зерттеу қызметі жөніндегі департамент директоры (Қазақстан)

РЕДАКЦИЯЛЫҚ АЛҚА:

Разак Абдул — PhD, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының профессоры (Қазақстан)
Лучио Томмазо де Паолис — Саленто Университеті (Италия) инновация және технологиялық инжиниринг департаменті AVR зертханасының зерттеу және әзірлеу бөлімінің директоры
Лиз Бэкон — профессор, Абертей Университеті (Ұлыбритания) вице-канцлерінің орынбасары
Микеле Пагано — PhD, Пиза Университетінің (Италия) профессоры
Өтелбаев Мухтарбай Өтелбайұлы — физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Халықаралық ақпараттық технологиялар университеті математика және компьютерлік модельдеу кафедрасының профессоры (Қазақстан)
Рысбайулы Болатбек — физика-математика ғылымдарының докторы, профессор, Есептеу және деректер ғылымдары департаментінің профессоры, Astana IT University (Казахстан)
Дайнеко Евгения Александровна — PhD, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының профессор-зерттеушісі (Қазақстан)
Дузбаев Нуржан Токкужаевич — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті цифрландыру және инновациялар жөніндегі проректор (Қазақстан)
Синчев Бахтгерей Куспанович — техника ғылымдарының докторы, профессор, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының профессоры (Қазақстан)
Сейлова Нургуль Абадуллаевна — техника ғылымдарының докторы, Халықаралық ақпараттық технологиялар университеті компьютерлік технологиялар және киберқауіпсіздік факультетінің деканы (Қазақстан)
Мухамедиева Ардак Габитовна — экономика ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті бизнеса медиа және басқару факультетінің деканы (Қазақстан)
Абдикаликова Замира Турсынбаевна — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті математика және компьютерлік модельдеу кафедрасының меңгерушісі (Қазақстан)
Шильдибеков Ерлан Жаржанович — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті экономика және бизнес кафедрасының меңгерушісі (Қазақстан)
Дамеля Максутовна Ескендирова — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының меңгерушісі (Қазақстан)
Ниязгулова Айгуль Аскарбековна — филология ғылымдарының кандидаты, доцент, профессор, Халықаралық ақпараттық технологиялар университеті медиакоммуникация және Қазақстан тарихы кафедрасының меңгерушісі (Қазақстан)
Айтмагамбетов Алтай Зуфарович — техника ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті радиотехника, электроника және телекоммуникация кафедрасының профессоры (Қазақстан)
Бахтиярова Елена Ажибековна — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті радиотехника, электроника және телекоммуникация кафедрасының меңгерушісі (Қазақстан)
Канибек Сансызбай — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының профессор-зерттеушісі (Қазақстан)
Тынымбаев Сахибай — техника ғылымдарының кандидаты, профессор, Халықаралық ақпараттық технологиялар университеті компьютерлік инженерия кафедрасының профессор-зерттеушісі (Қазақстан)
Алмисреб Али Абд — PhD, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының қауымдастырылған профессоры (Қазақстан)
Мохамед Ахмед Хамада — PhD, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының қауымдастырылған профессоры (Қазақстан)
Янг Им Чу — PhD, Гачон университетінің профессоры (Оңтүстік Корея)
Тадеуш Валлас – PhD, Адам Мицкевич атындағы (Польша) университеттің проректоры
Мамырбаев Оркен Жумажанович — PhD, ҚР ҒЖБМ Ғылым комитеті ақпараттық және есептеу технологиялары институты ӨМК директорының ғылым жөніндегі орынбасары (Қазақстан)
Бушуев Сергей Дмитриевич — техника ғылымдарының докторы, профессор, Украинаның "УКРНЕТ" жобаларды басқару қауымдастығының директоры, Киев ұлттық құрылыс және сәулет университеті жобаларды басқару кафедрасының меңгерушісі (Украина)
Белощицкая Светлана Васильевна — техника ғылымдарының докторы, доцент, Astana IT University есептеу және деректер ғылымы кафедрасының профессоры (Қазақстан)

ЖАУАПТЫ РЕДАКТОР:

Мрзабаева Раушан Жалиевна — магистр, Халықаралық ақпараттық технологиялар университетінің редакторы (Қазақстан)

Халықаралық ақпараттық және коммуникациялық технологиялар журналы
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Префикс DOI: 10.54309
Мерзімділігі: жылына 4 рет.
Басылым тілі: қазақ, орыс, ағылшын.
Тақырып бағыты: "Ақпараттық технологиялар"; "Ақпараттық қауіпсіздік және коммуникациялық технологиялар"; "Әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технология".
Журнал сайты: https://journal.iitu.edu.kz
Тарату: материалдар Creative Commons Attribution 4.0 лицензиясы бойынша таратылады
Меншік иесі: АҚ «Халықаралық ақпараттық технологиялар университеті» (Алматы қ.).
Авторлық құқық: © Халықаралық ақпараттық және коммуникациялық технологиялар журналы, 2026

РЕДАКЦИЯ

ГЛАВНЫЙ РЕДАКТОР:

Колесникова Катерина Викторовна — доктор технических наук, профессор, проректор по научно-исследовательской деятельности Международного университета информационных технологий (Казахстан)

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

Ипалакова Мадина Тулегеновна — кандидат технических наук, ассоциированный профессор, директор департамента по научно-исследовательской деятельности Международного университета информационных технологий (Казахстан)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Разак Абдул — PhD, профессор кафедры кибербезопасности Международного университета информационных технологий (Казахстан)

Лучио Томмазо де Паолис — директор отдела исследований и разработок лаборатории AVR департамента инноваций и технологического инжиниринга Университета Саленто (Италия)

Лиз Бэкон — профессор, заместитель вице-канцлера Университета Абертей (Великобритания)

Микеле Пагано — PhD, профессор Университета Пизы (Италия)

Отелбаев Мухтарбай Отелбайулы — доктор физико-математических наук, профессор, академик НАН РК, профессор кафедры математического и компьютерного моделирования Международного университета информационных технологий (Казахстан)

Рысбайулы Болатбек — доктор физико-математических наук, профессор, профессор Astana IT University (Казахстан)

Дайнеко Евгения Александровна — PhD, профессор-исследователь кафедры информационных систем Международного университета информационных технологий (Казахстан)

Дузбаев Нуржан Токкужаевич — PhD, ассоциированный профессор, проректор по цифровизации и инновациям Международного университета информационных технологий (Казахстан)

Синчев Бахтгерей Куспанович — доктор технических наук, профессор, профессор кафедры информационных систем Международного университета информационных технологий (Казахстан)

Сейлова Нургуль Абадуллаевна — кандидат технических наук, декан факультета компьютерных технологий и кибербезопасности Международного университета информационных технологий (Казахстан)

Мухамедиева Ардак Габитовна — кандидат экономических наук, декан факультета бизнеса медиа и управления Международного университета информационных технологий (Казахстан)

Абдикаликова Замира Турсынбаевна — PhD, ассоциированный профессор, заведующая кафедрой математического и компьютерного моделирования Международного университета информационных технологий (Казахстан)

Шильдибеков Ерлан Жаржанович — PhD, ассоциированный профессор, заведующий кафедрой экономики и бизнеса Международного университета информационных технологий (Казахстан)

Дамеля Максутовна Ескендирова — кандидат технических наук, ассоциированный профессор, заведующая кафедрой кибербезопасности Международного университета информационных технологий (Казахстан)

Ниязгулова Айгуль Аскарбековна — кандидат филологических наук, доцент, профессор, заведующая кафедрой медиакоммуникаций и истории Казахстана Международного университета информационных технологий (Казахстан)

Айтмагамбетов Алтай Зуфарович — кандидат технических наук, профессор кафедры радиотехники, электроники и телекоммуникаций Международного университета информационных технологий (Казахстан)

Бахтиярова Елена Ажибековна — кандидат технических наук, ассоциированный профессор, заведующая кафедрой радиотехники, электроники и телекоммуникаций Международного университета информационных технологий (Казахстан)

Канибек Сансызбай – PhD, ассоциированный профессор, профессор-исследователь кафедры кибербезопасности, Международного университета информационных технологий (Казахстан)

Тынымбаев Сахибай – кандидат технических наук, профессор, профессор-исследователь кафедры компьютерной инженерии, Международного университета информационных технологий (Казахстан)

Алмисреб Али Абд — PhD, ассоциированный профессор кафедры кибербезопасности Международного университета информационных технологий (Казахстан)

Мохамед Ахмед Хамада — PhD, ассоциированный профессор кафедры информационных систем Международного университета информационных технологий (Казахстан)

Янг Им Чу — PhD, профессор университета Гачон (Южная Корея)

Тадеуш Валлас – PhD, проректор университета имен Адама Мицкевича (Польша)

Мамырбаев Оркен Жумажанович — PhD, заместитель директора по науке РГП Института информационных и вычислительных технологий Комитета науки МНВО РК (Казахстан)

Бушуев Сергей Дмитриевич — доктор технических наук, профессор, директор Украинской ассоциации управления проектами «УКРНЕТ», заведующий кафедрой управления проектами Киевского национального университета строительства и архитектуры (Украина)

Белощицкая Светлана Васильевна — доктор технических наук, доцент, профессор кафедры вычислений и науки о данных Astana IT University (Казахстан)

ОТВЕТСТВЕННЫЙ РЕДАКТОР:

Мрзабаева Раушан Жалиевна — магистр, редактор Международного университета информационных технологий (Казахстан)

Международный журнал информационных и коммуникационных технологий
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Префикс DOI: 10.54309
Периодичность: 4 выпусков в год.
Язык издания: казахский, русский, английский.
Тематическая направленность: "Информационные технологии"; "Информационная безопасность и коммуникационные технологии"; "Цифровые технологии в развитии социо-экономических систем".
Сайт журнала: https://journal.iitu.edu.kz
Распространение: материалы распространяются по лицензии Creative Commons Attribution 4.0
Собственник: АО «Международный университет информационных технологий» (г. Алматы).
Авторские права: © Международный журнал информационных и коммуникационных технологий, 2026

CONTENTS

DIGITAL TECHNOLOGIES IN THE DEVELOPMENT OF SOCIO-ECONOMIC SYSTEMS

D. Abzhanova, A. Biloshchytskyi
A MODEL AND METHOD FOR MANAGING DATA ON EMISSIONS FROM STATIONARY SOURCES OF POLLUTION IN AN INTELLIGENT ENVIRONMENTAL MONITORING SYSTEM 11
A. Slanbekova, M. Rakhimzhanova, A. Zhanibekova, A. Alimagambetova, M. Xudoyberganov
EARLY DETECTION OF HYDROLOGICAL HAZARDS BASED ON SPATIOTEMPORAL ANALYSIS 27
A. Sakhipov, R. Seiitbek, D. Yedilkhan
REAL-TIME PASSENGER FLOW PREDICTION USING AUTOMATED DATA PREPARATION 38
K. Tokayev, A. Moldagulova
A REVIEW OF PREDICTIVE MODELS APPROACHES TO STOCK MARKET FORECASTING IN KAZAKHSTAN 60

INFORMATION TECHNOLOGY

F.N. Abdraimova, A.A. Kereibayeva, D.S. Dyussenova, D.A. Aliyeva, T.Zh. Toktarova
AI TECHNOLOGIES IN LANGUAGE EDUCATION: PRACTICAL ASPECTS AND CHALLENGES OF STUDENT USAGE 83
G. Azieva, M. Yessenova, A. Abzhapparova, G. Abdikerimova, P. Schmidt
HYBRID STACKING FRAMEWORK FOR CROP CLASSIFICATION USING UAV DATA 97
A.K. Aitim
JOINT MORPHOLOGICAL DISAMBIGUATION AND POS TAGGING FOR AGGLUTINATIVE LANGUAGES 109
A. Auyezova, Aigerim Aitim, B. Sinchev, W. Wójcik
NP-COMPLETE ALGORITHMS IN LEGALTECH: HYBRID ARCHITECTURE FOR SEMANTIC RETRIEVAL IN KAZAKHSTAN'S LEGAL ACTS 125
Y. Hozak, S. Paliy, T. Honcharenko
EVALUATING OPTIMAL CONFIGURATIONS OF SHALLOW CNNS ON LOW-COMPLEXITY DATASETS 138
S.A. Yesniyazova, S.T. Kaimov
PREDICTIVE MAINTENANCE OF HEAVY-DUTY TRUCKS USING EXPLAINABLE MACHINE LEARNING 153
T. Imanbekova, Zh. Ibrayeva, G. Jakanova, G. Askanbay
DATA COMPRESSION ALGORITHM BASED ON WAVELET TRANSFORMER; ANALYSIS AND IMPLEMENTATION IN MATLAB 167
B.Z. Kenzhegulov, Zh.T. Bilyalova, K.N. Uteuliyeva, L. Nurgaliyeva, Sh.S. Nurzhanova
A MATHEMATICAL AND ALGORITHMIC APPROACH TO THE DEVELOPMENT OF AN INTELLIGENT TEXT-TO-SQL SYSTEM BASED ON LARGE LANGUAGE MODELS 185
M. Kabdullin, A. Kairbekov, A. Kabdullin, L. Naizabayeva, A. Mukhit
DESIGN AND IMPLEMENTATION OF A WEB-BASED PROTOTYPE FOR CARDIOLOGICAL IMAGE MANAGEMENT 206
L.Kurmangaziyeva, Sh. Yelezhanova, S. Iskakova, Sh. Kodanova, H.İ. Bülbül
APPROACHES TO BUILDING A HYBRID INTELLIGENT DECISION SUPPORT SYSTEM BASED ON FUZZY LOGIC, XAI, AND LARGE LANGUAGE MODELS 219
Z. Momynkulov, B. Omarov
DDPG FOR TRAJECTORY GENERATION 236
F. Myrzabekov, N. Mustafin, M. Mukushev
WHY IS THIS NEWS NEGATIVE?: AN ATTENTION-BASED VISUALIZATION TECHNIQUE FOR EXPLAINABLE SENTIMENT ANALYSIS. AN ANALYTICAL SURVEY AND CONCEPTUAL FRAMEWORK 248
N.Sh. Maxutova, J.A. Tussupov, A.A. Shekerbek, Zh.E. Kenzhebayeva, Q.O. Rakhimov
MACHINE LEARNING FOR COMPREHENSIVE EVALUATION OF CARDIOVASCULAR DISEASE RISK AND BIOCHEMICAL ALTERATIONS: FOCUS ON ASPARTATE AMINOTRANSFERASE 260
L. Naizabayeva, G. Turken, R.A. Druba, A.Sh. Yawar
LEVERAGING ENTERPRISE DATA WAREHOUSING TO IMPROVE WEB-BASED SALES

OPERATIONS 275
S. Nurgaliyeva, N. Naiman, K. Makhmetova, M. Kali, N. Zhakiyev
COMPARATIVE ANALYSIS OF MACHINE LEARNING METHODS FOR RAPID DETECTION OF VEGETABLE DISEASE USING COMPUTER VISION 291
O.S. Salykova, V.A. Madin, B.R. Salykov, D.N. Komarov, N.V. Manuilov
INTEGRATION OF MEMS ACCELEROMETER SENSOR MODULES IN INDUSTRIAL MONITORING SYSTEMS 309
R. Taberkhan, M.A. Sambetbayeva, G. Kalman
KAZCAUSAL: THE FIRST CORPUS-BASED ANNOTATION OF CAUSAL RELATIONSHIPS IN THE KAZAKH LANGUAGE 323
S.Tynymbayev, S.E. Mamanova, R. Berdybayev, Zh.E. Temirbekova, T. Chinibayeva
DIVIDING DEVICES WITH PRELIMINARY PREPARATION OF MULTIPLES OF THE DIVISOR 335
K.N. Uteuliyeva, B.Z. Kenzhegulov, T.A. Karazhigitova, H.İ. Bülbül, Z.Zh. Zhanuzakova
MATHEMATICAL AND ALGORITHMIC APPROACHES TO THE DEVELOPMENT OF A COLLABORATIVE FILTERING-BASED RECOMMENDER SYSTEM 351
S. Sharmukhanbet, G. Turmukhanova, O. Findik, V. Makhatova, L. Kurmangaziyeva
HIGH-PRECISION ROBOTIC ASSEMBLY UNDER VARIABLE ILLUMINATION: A ROBUST MECHATRONIC ARCHITECTURE FOR VISUAL SERVOING 372

INFORMATION SECURITY AND COMMUNICATIONTECHNOLOGIES

A. Amirbay, Z. Amanbaikyzy, K. Maxutova, A. Mukhanova, M. Kassim
MACHINE LEARNING ALGORITHM FOR EARLY DETECTION OF AUTISM SPECTRUM DISORDERS IN CHILDREN BASED ON MULTIMODAL ANALYSIS OF EYE MOVEMENTS AND FACIAL EXPRESSIONS 390
K. Baisylbayeva, Sh. Mussiraliyeva, Zh. Yeltay
DETECTION OF EXTREMIST IDEOLOGY IN THE KAZAKH LANGUAGE: ANNOTATION CHALLENGES AND DEEP LEARNING APPROACHES 405
M.A. Bolatbek, A.M.Usmanova, K.B. Bagitova, G.B. Baispay
DEVELOPMENT AND RESEARCH OF A METHOD FOR ANALYZING NETWORK TRAFFIC TO IDENTIFY A CYBER THREAT 424
D.I. Prokopovych-Tkachenko, N.K. Zhumagalieva, D.N. Shchytov, N.F. Mormul, D.A. Cherkaskyi
FUZZY MODEL FOR EVALUATING INFORMATION SECURITY PARAMETERS OF INFORMATION SYSTEMS UNDER INCOMPLETE AND QUALITATIVE DATA: CONSTRUCTION METHODOLOGY, RULE BASE TUNING, AND DEMONSTRATION CASE FOR ORGANIZATIONS 442
E.A. Pustovoy, O.A. Pustovaya, A.N. Raushanova, I.S. Zaurbekov
EVALUATION OF THE EFFECTIVENESS OF SYNTHESIS OF STOCHASTIC MODELS WITH CONTROLLED PROPERTIES 468
Y. Serzhan, T. Umarov, A. Abilbayeva
FRAUD DETECTION IN CREDIT CARD TRANSACTIONS USING MACHINE LEARNING: A COMPARATIVE ANALYSIS 484
T. Babenko, Ye. Bakhtiyarova, D. Yeskendirova, A. Sysoev, Sh. Makilenov
FLEXIBLE IDPS BASED ON MACHINE LEARNING ALGORITHMS FOR DIPTYCHS AND IOT EDGE DEVICE PROTECTION 496

МАЗМҰНЫ
ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖҮЙЕЛЕРДІ ДАМЫТУДАҒЫ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

Д.Е. Абжанов, А.А. Белощицкий
ЭКОЛОГИЯЛЫҚ МОНИТОРИНГТІҢ ЗИЯТКЕРЛІК ЖҮЙЕСІНДЕГІ СТАЦИОНАРЛЫҚ ЛАСТАНУ КӨЗДЕРІНІҢ ШЫҒАРЫНДЫЛАРЫ ТУРАЛЫ ДЕРЕКТЕРДІ БАСҚАРУДЫҢ МОДЕЛІ МЕН ӘДІСІ 11
А.Е. Сланбекова, М.Б. Рахимжанова, А.И. Жанибекова, А.З. Алимагамбетова, М. Худойберганов
КЕҢІСТІКТІК-УАҚЫТТЫҚ (SPATIOTEMPORAL) ТАЛДАУ НЕГІЗІНДЕ ГИДРОЛОГИЯЛЫҚ ҚАУІПҚАТЕРДІ ЕРТЕ АНЫҚТАУ 27
А. Сахипов, Р. Сеитбек, Д. Едильхан
АВТОМАТТАНДЫРЫЛҒАН ДЕРЕКТЕРДІ ДАЙЫНДАУДЫ ПАЙДАЛАНА ОТЫРЫП, ЖОЛАУШЫЛАР

АҒЫНЫН НАҚТЫ УАҚЫТ РЕЖИМІНДЕ БОЛЖАУ 38
К. Токаев, А. Молдагулова
ҚАЗАҚСТАНДАҒЫ ҚОР НАРЫҒЫН БОЛЖАУДЫҢ БОЛЖАМДЫҚ МОДЕЛЬДЕРІНЕ ШОЛУ 60

АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

Ф.Н. Абдраимова, А.А. Керейбаева, Д.С. Дюсенова, Д.А. Алиева, Т.Ж. Токтарова
ТІЛ БІЛІМІНДЕ ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫ: СТУДЕНТТЕР ҚОЛДАНУЫНЫҢ ПРАКТИКАЛЫҚ АСПЕКТІЛЕРІ МЕН МӘСЕЛЕЛЕРІ 83
Г.Т. Азиева, М.Б. Есенова, А.К. Абжаппарова, Г.Б. Абдикеримова, P. Schmidt
UAV ДЕРЕКТЕРІ НЕГІЗІНДЕ АУЫЛ ШАРУАШЫЛЫҒЫ ДАҚЫЛДАРЫН ЖІКТЕУГЕ АРНАЛҒАН ГИБРИДТІ СТЕКИНГ МОДЕЛІ 97
Ә.Қ. Әйтім
АГГЛЮТИНАТИВТІ ТІЛДЕРГЕ АРНАЛҒАН МОРФОЛОГИЯЛЫҚ ДИЗАМБИГУАЦИЯ МЕН POS-ТАҢБАЛАУДЫ БІРЛЕСІП МОДЕЛЬДЕУ 109
А. Әуезова, Ә. Әйтім, Б. Синчев, В. Вуйчик
ҚАЗАҚСТАННЫҢ ҚҰҚЫҚТЫҚ АКТІЛЕРІНДЕГІ NP-ТОЛЫҚ АЛГОРИТМДЕР: SEMANTIC ІЗДЕУГЕ АРНАЛҒАН ГИБРИДТІ АРХИТЕКТУРА 125
Я. Гозак, С. Палий, Т. Гончаренко
КҮРДЕЛІЛІГІ ТӨМЕН ДЕРЕКТЕР ЖИЫНТЫҚТАРЫНДА ТЕРЕҢ ЕМЕС ОРАМДЫ НЕЙРОНДЫҚ ЖЕЛІЛЕРДІҢ (CNN) ОҢТАЙЛЫ КОНФИГУРАЦИЯЛАРЫН БАҒАЛАУ 138
С.А. Есниязова, С.Т. Каимов
ТҮСІНДІРІЛЕТІН МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНА ОТЫРЫП АУЫР ЖҮК КӨЛІКТЕРІНЕ БОЛЖАМДЫ ТЕХНИКАЛЫҚ ҚЫЗМЕТ КӨРСЕТУ 153
Т.Д. Иманбекова, Ж.Б. Ибраева, Г.Т. Джаканова, Г.Т. Асқанбай
МӘЛІМЕТТЕРДІ ВЕЙВЛЕТ-ТҮРЛЕНДІРГІШТІҢ НЕГІЗІНДЕ ҚЫСУ АЛГОРИТМІ; MATLAB ОРТАСЫНДА ТАЛДАУ ЖӘНЕ ІСКЕ АСЫРУ 167
Б.З. Кенжегулов, Ж.Т. Билялова, К.Н. Утеулиева, Л. Нургалиева, Ш.С. Нуржанова
ҮЛКЕН ТІЛДІК МОДЕЛЬДЕР НЕГІЗІНДЕ ИНТЕЛЛЕКТУАЛДЫ TEXT-TO-SQL ЖҮЙЕСІН ӘЗІРЛЕУДІҢ МАТЕМАТИКАЛЫҚ-АЛГОРИТМДІК ТӘСІЛІ 185
М. Кабдуллин, А. Каирбеков, А. Кабдуллин, Л. Найзабаева, А. Мухит
КАРДИОЛОГИЯЛЫҚ КЕСКІНДЕРДІ БАСҚАРУҒА АРНАЛҒАН ВЕБ-НЕГІЗДЕЛГЕН ПРОТОТИПТІ ӘЗІРЛЕУ ЖӘНЕ ІСКЕ АСЫРУ 206
Л. Курмангазиева, Ш. Ележанова, С. Искакова, Ш. Коданова, Х. И. Бұлбұл
АЙҚЫН ЕМЕС ЛОГИКА, ХАІ ЖӘНЕ ҮЛКЕН ТІЛДІК МОДЕЛЬДЕР НЕГІЗІНДЕ ГИБРИДТІ ЗИЯТКЕРЛІК ШЕШІМ ҚАБЫЛДАУДЫ ҚОЛДАУ ЖҮЙЕСІН ҚҰРУ ТӘСІЛДЕРІ 219
З. Момынкулов, Б. Омаров
ТРАЕКТОРИЯЛАРДЫ ГЕНЕРАЦИЯЛАУ ҮШІН DDPG 236
Ф. Мырзабеков, Н. Мустафин, М. Мукушев
БҰЛ ЖАҢАЛЫҚ НЕЛІКТЕН ТЕРІС?: ТҮСІНДІРІЛЕТІН СЕНТИМЕНТТІК ТАЛДАУҒА АРНАЛҒАН НАЗАР АУДАРУ МЕХАНИЗМІНЕ НЕГІЗДЕЛГЕН ВИЗУАЛИЗАЦИЯ ӘДІСІ. ТАЛДАУЛЫҚ ШОЛУ ЖӘНЕ ТҰЖЫРЫМДАМАЛЫҚ НЕГІЗ 248
Н.Ш. Максутова, Ж.А. Тусупов, А.Ә. Шекербек, Ж.Е. Кенжебаева, К.О. Рахимов
ЖҮРЕК-ҚАН ТАМЫРЛАРЫ АУРУЛАРЫНЫҢ ҚАУІП-ҚАТЕРІН ЖӘНЕ БИОХИМИЯЛЫҚ ӨЗГЕРІСТЕРДІ КЕШЕНДІ БАҒАЛАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ: АСПАРТАТАМИНОТРАНСФЕРАЗАҒА ЕРЕКШЕ НАЗАР 260
Л. Найзабаева, Г. Түркен, Р.А. Друба, А.Ш. Явар
ВЕБ НЕГІЗІНДЕГІ САТУ ОПЕРАЦИЯЛАРЫН ЖЕТІЛДІРУ ҮШІН КӘСІПОРЫНДЫҚ ДЕРЕКТЕР ҚОЙМАСЫН ПАЙДАЛАНУ 275
С. Нургалиева, Н. Найман, К. Махметова, М. Кали, Н. Жакиев
КОМПЬЮТЕРЛІК КӨРУ ТЕХНОЛОГИЯСЫН ПАЙДАЛАНА ОТЫРЫП, КӨКӨНІС ДАҚЫЛДАРЫНЫҢ АУРУЛАРЫН ЖЕДЕЛ АНЫҚТАУҒА АРНАЛҒАН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН САЛЫСТЫРМАЛЫ ТАЛДАУ 291
О.С. Салыкова, В.А. Мадин, Б.Р. Салыков, Д.Н. Комаров, Н.В. Мануйлов
ӨНЕРКӘСІПТІК МОНИТОРИНГ ЖҮЙЕЛЕРІНДЕГІ MEMS-АКСЕЛЕРОМЕТРЛЕРДІҢ СЕНСОРЛЫҚ МОДУЛЬДЕРІН ИНТЕГРАЦИЯЛАУ 309
Р. Таберхан, М.А. Самбетбаева, Г. Қалман

KAZCAUSAL: ҚАЗАҚ ТІЛІНДЕГІ СЕБЕП-САЛДАРЛЫҚ ҚАТЫНАСТАРДЫҢ АЛҒАШҚЫ КОРПУСТЫҚ АННОТАЦИЯСЫ 323
С. Тынымбаев, С.Е. Маманова, Р. Бердібаев, Ж.Е.Темірбекова, Т. Чинибаева
БӨЛГІШТІҢ ЕСЕЛІ МӘНДЕРІН АЛДЫН АЛА ДАЙЫНДАУМЕН ЖҮЗЕГЕ АСЫРЫЛАТЫН БӨЛУ ҚҰРЫЛҒЫЛАРЫ 335
К.Н. Утеулиева, Б.З. Кенжегулов, Т.А. Каражигитова, Х. Булбул, З.Ж. Жанузакова
КОЛЛАБОРАТИВТІК СҮЗГІЛЕУ НЕГІЗІНДЕГІ ҰСЫНЫМДЫҚ ЖҮЙЕНІ ӘЗІРЛЕУДІҢ МАТЕМАТИКАЛЫҚ-АЛГОРИТМДІК ТӘСІЛДЕРІ 351
С. Шармұханбет, Г. Тұрмұханова, О. Финдик, В. Махатова, Л. Курмангазиева
АЙНЫМАЛЫ ЖАРЫҚ ЖАҒДАЙЫНДАҒЫ ЖОҒАРЫ ДӘЛДІКТІ РОБОТТЫҚ ҚҰРАСТЫРУ: ВИЗУАЛДЫ СЕРВОТЕЖЕУДІҢ ТӨЗІМДІ МЕХАТРОНИКАЛЫҚ АРХИТЕКТУРАСЫ 372

АҚПАРАТТЫҚ ҚАУІПСІЗДІК ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАРҒА АРНАЛҒАН

А. Амирбай, З. Аманбайқызы, К. Максутова, А. Муханова, M. Kassim
КӨЗ ҚОЗҒАЛЫСТАРЫ МЕН БЕТ МИМИКА БЕЛГІЛЕРІН МУЛЬТИМОДАЛЬДЫ ТАЛДАУҒА НЕГІЗДЕЛГЕН БАЛАЛАРДАҒЫ АУТИЗМ СПЕКТРІНІҢ БҰЗЫЛЫСТАРЫН ЕРТЕ АНЫҚТАУҒА АРНАЛҒАН МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМІ 390
К.Д. Байсылбаева, Ш.Ж. Мусиралиева, Ж. Елтай
ҚАЗАҚ ТІЛІНДЕГІ ЭКСТРЕМИСТІК ИДЕОЛОГИЯНЫ АНЫҚТАУ: АННОТАЦИЯЛАУ МӘСЕЛЕЛЕРІ ЖӘНЕ ТЕРЕҢ ОҚЫТУ ТӘСІЛДЕРІ 405
М.А. Болатбек, А.М. Усманова, Қ.Б. Багитова, Г.Б. Байспай
КИБЕР ҚАУІПТІ АНЫҚТАУ ҮШІН ЖЕЛІЛІК ТРАФИКТІ ТАЛДАУ ӘДІСІН ӘЗІРЛЕУ ЖӘНЕ ЗЕРТТЕУ 424
Д.И. Прокопович-Ткаченко, Н.К. Жумагалиева, Д.Н. Щитов, Н.Ф. Мормуль, Д.А. Черкасский
ТОЛЫҚ ЕМЕС ЖӘНЕ САПАЛЫҚ ДЕРЕКТЕР ЖАҒДАЙЫНДА АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІҢ АҚПАРАТТЫҚ ҚАУІПСІЗДІК ПАРАМЕТРЛЕРІН БАҒАЛАУДЫҢ БҰЛЫҢҒЫР МОДЕЛІ: ҚҰРУ ӘДІСТЕМЕСІ, ЕРЕЖЕЛЕР БАЗАСЫН БАПТАУ ЖӘНЕ ҰЙЫМДАРҒА АРНАЛҒАН ДЕМОНСТРАЦИЯЛЫҚ КЕЙС 442
Е.А. Пустовой, О.А. Пустовая, А.Н. Раушанова, И.С. Заурбеков
БАСҚАРЫЛАТЫН ҚАСИЕТТЕРІ БАР СТОХАСТИКАЛЫҚ МОДЕЛЬДЕРДІ СИНТЕЗДЕУДІҢ ТИІМДІЛІГІН БАҒАЛАУ 468
Е. Сержан, Т. Умаров, А. Әбілбаева
МАШИНАЛЫҚ ОҚУ ӘДІСІ АРҚЫЛЫ КРЕДИТ КАРТА ОПЕРАЦИЯЛАРЫНДАҒЫ АЛАЯҚТЫҚТЫ АНЫҚТАУ: САЛЫСТЫРМАЛЫ ТАЛДАУ 484
Т.В. Бабенко, Е.А. Бахтиярова, Д.М. Ескендирова, А. Сысоев, Ш.Н. Макиленов
ДИПТИХ ЖӘНЕ IOT ШЕТКІ ҚҰРЫЛҒЫЛАРЫН ҚОРҒАУҒА АРНАЛҒАН МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІНЕ НЕГІЗДЕЛГЕН ИКЕМДІ IDPS 496

СОДЕРЖАНИЕ
ЦИФРОВЫЕ ТЕХНОЛОГИИ В РАЗВИТИИСОЦИО-ЭКОНОМИЧЕСКИХ СИСТЕМ

Д.Е. Абжанова, А.А. Белощицкий
МОДЕЛЬ И МЕТОД УПРАВЛЕНИЯ ДАННЫМИ О ВЫБРОСАХ СТАЦИОНАРНЫХ ИСТОЧНИКОВ ЗАГРЯЗНЕНИЯ В ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЕ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА 11
А.Е. Сланбекова, М.Б. Рахимжанова, А.И. Жанибекова, А.З. Алимагамбетова, М. Худойберганов
РАННЕЕ ВЫЯВЛЕНИЕ ГИДРОЛОГИЧЕСКИХ ОПАСНОСТЕЙ НА ОСНОВЕ ПРОСТРАНСТВЕННО-ВРЕМЕННОГО (SPATIOTEMPORAL) АНАЛИЗА 27
А. Сахипов, Р. Сеитбек, Д. Едильхан
ПРОГНОЗИРОВАНИЕ ПАССАЖИРОПОТОКА В РЕАЛЬНОМ ВРЕМЕНИ С ИСПОЛЬЗОВАНИЕМ АВТОМАТИЗИРОВАННОЙ ПОДГОТОВКИ ДАННЫХ 38
К. Токаев, А. Молдагулова
ОБЗОР ПРОГНОСТИЧЕСКИХ МОДЕЛЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ФОНДОВОГО РЫНКА В КАЗАХСТАНЕ 60

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Ф.Н. Абдраимова, А.А. Керейбаева, Д.С. Дюсенова, Д.А. Алиева, Т.Ж. Токтарова
ТЕХНОЛОГИИ ИИ В ЯЗЫКОВОМ ОБРАЗОВАНИИ: ПРАКТИЧЕСКИЕ АСПЕКТЫ И ПРОБЛЕМЫ ПРИМЕНЕНИЯ СТУДЕНТАМИ 83
Г.Т. Азиева, М.Б. Есенова, А.К. Абжаппарова, Г.Б. Абдикеримова, P. Schmidt
ГИБРИДНАЯ МОДЕЛЬ СТЕКИНГА ДЛЯ КЛАССИФИКАЦИИ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР ПО ДАННЫМ UAV 97
Ә.Қ. Әйтім
СОВМЕСТНАЯ МОРФОЛОГИЧЕСКАЯ ДИЗАМБИГУАЦИЯ И POS-РАЗМЕТКА ДЛЯ АГГЛЮТИНАТИВНЫХ ЯЗЫКОВ 109
А. Ауезова, А. Айтим, Б. Синчев, В. Вуйчик
NP-ПОЛНЫЕ АЛГОРИТМЫ В LEGALTECH: ГИБРИДНАЯ АРХИТЕКТУРА ДЛЯ СЕМАНТИЧЕСКОГО ПОИСКА В ПРАВОВЫХ АКТАХ КАЗАХСТАНА 125
Я. Гозак, С. Палий, Т. Гончаренко
ОЦЕНКА ОПТИМАЛЬНЫХ КОНФИГУРАЦИЙ НЕГЛУБОКИХ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ НА ДАТАСЕТАХ НИЗКОЙ СЛОЖНОСТИ 138
С.А. Есниязова, С.Т. Каимов
ПРЕДИКТИВНОЕ ТЕХНИЧЕСКОЕ ОБСЛУЖИВАНИЕ ТЯЖЁЛЫХ ГРУЗОВИКОВ С ИСПОЛЬЗОВАНИЕМ ОБЪЯСНИМОГО МАШИННОГО ОБУЧЕНИЯ 153
Т.Д. Иманбекова, Ж.Б. Ибраева, Г.Т. Джаканова, Г.Т. Асқанбай
АЛГОРИТМ СЖАТИЯ ДАННЫХ НА ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАТЕЛЯ: АНАЛИЗ И РЕАЛИЗАЦИЯ В MATLAB 167
Б.З. Кенжегулов, Ж.Т. Билялова, К.Н. Утеулиева, Л. Нургалиева, Ш.С. Нуржанова
МАТЕМАТИКО-АЛГОРИТМИЧЕСКИЙ ПОДХОД К РАЗРАБОТКЕ ИНТЕЛЛЕКТУАЛЬНОЙ TEXT-TO-SQL СИСТЕМЫ НА ОСНОВЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ 185
М. Кабдуллин, А. Каирбеков, А. Кабдуллин, Л. Найзабаева, А. Мухит
РАЗРАБОТКА И РЕАЛИЗАЦИЯ ВЕБ-ПРОТОТИПА ДЛЯ УПРАВЛЕНИЯ КАРДИОЛОГИЧЕСКИМИ ИЗОБРАЖЕНИЯМИ 206
Л. Курмангазиева, Ш. Ележанова, С. Искакова, Ш. Коданова, Х. И. Бюльбюль
ПОДХОДЫ К ПОСТРОЕНИЮ ГИБРИДНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ НЕЧЕТКОЙ ЛОГИКИ, XAI И БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ 219
З. Момынкулов, Б. Омаров
DDPG ДЛЯ ГЕНЕРАЦИИ ТРАЕКТОРИЙ 236
Ф. Мырзабеков, Н. Мустафин, М. Мукушев
ПОЧЕМУ ЭТА НОВОСТЬ НЕГАТИВНА?: МЕТОД ВИЗУАЛИЗАЦИИ НА ОСНОВЕ МЕХАНИЗМА ВНИМАНИЯ ДЛЯ ОБЪЯСНИМОГО АНАЛИЗА ТОНАЛЬНОСТИ. АНАЛИТИЧЕСКИЙ ОБЗОР И КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ 248
Н.Ш. Максутова, Д.А. Тусупов, А.А. Шекербек, Ж.Е. Кенжебаева, К.О. Рахимов
МАШИННОЕ ОБУЧЕНИЕ ДЛЯ КОМПЛЕКСНОЙ ОЦЕНКИ РИСКА СЕРДЕЧНО-СОСУДИСТЫХ ЗАБОЛЕВАНИЙ И БИОХИМИЧЕСКИХ ИЗМЕНЕНИЙ: АКЦЕНТ НА АСПАРТАТАМИНОТРАНСФЕРАЗЕ 260
Л. Найзабаева, Г. Туркен, Р.А. Друба, А.Ш. Явар
ИСПОЛЬЗОВАНИЕ КОРПОРАТИВНОГО ХРАНИЛИЩА ДАННЫХ ДЛЯ УЛУЧШЕНИЯ ВЕБ-ОРИЕНТИРОВАННЫХ ОПЕРАЦИЙ ПРОДАЖ 275
С. Нургалиева, Н. Найман, К. Махметова, М. Кали, Н. Жакиев
СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ БЫСТРОГО ОБНАРУЖЕНИЯ ЗАБОЛЕВАНИЙ ОВОЩНЫХ КУЛЬТУР С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНОГО ЗРЕНИЯ 291
О.С. Салыкова, В.А. Мадин, Б.Р. Салыков, Д.Н. Комаров, Н.В. Мануйлов
ИНТЕГРАЦИЯ СЕНСОРНЫХ МОДУЛЕЙ MEMS-АКСЕЛЕРОМЕТРОВ В СИСТЕМАХ ПРОМЫШЛЕННОГО МОНИТОРИНГА 309
Р. Таберхан, М.А. Самбетбаева, Г. Калман
KAZCAUSAL: ПЕРВАЯ КОРПУСНАЯ АННОТАЦИЯ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ НА КАЗАХСКОМ ЯЗЫКЕ 323
С. Тынымбаев, С.Е. Маманова, Р. Бердибаев, Ж.Е.Темирбекова, Т. Чинибаева

УСТРОЙСТВА ДЕЛЕНИЯ ЧИСЕЛ С ПРЕДВАРИТЕЛЬНОЙ ПОДГОТОВКОЙ КРАТНЫХ ДЕЛИТЕЛЮ 335
К.Н. Утеулиева, Б.З. Кенжегулов, Т.А. Каражигитова, Х.Бюльбюль, З.Ж. Жанузакова
МАТЕМАТИКО-АЛГОРИТМИЧЕСКИЕ ПОДХОДЫ К РАЗРАБОТКЕ РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ НА ОСНОВЕ КОЛЛАБОРАТИВНОЙ ФИЛЬТРАЦИИ 351
С. Шармуханбет, Г. Турмуханова, О.Финдик, В.Махатова, Л. Курмангазиева
ВЫСОКОТОЧНАЯ РОБОТИЗИРОВАННАЯ СБОРКА ПРИ ПЕРЕМЕННОЙ ОСВЕЩЁННОСТИ: РОБАСТНАЯ МЕХАТРОННАЯ АРХИТЕКТУРА ВИЗУАЛЬНОГО СЕРВОУПРАВЛЕНИЯ 372

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

А. Амирбай, З. Аманбайкызы, К. Максутова, А. Муханова, M. Kassim
АЛГОРИТМ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАННЕГО ВЫЯВЛЕНИЯ РАССТРОЙСТВ АУТИСТИЧЕСКОГО СПЕКТРА У ДЕТЕЙ НА ОСНОВЕ МУЛЬТМОДАЛЬНОГО АНАЛИЗА ДАННЫХ ДВИЖЕНИЯ ГЛАЗ И МИМИЧЕСКИХ СИГНАЛОВ 390
К.Д. Байсылбаева, Ш.Ж. Мусиралиева, Ж.Елтай
ОБНАРУЖЕНИЕ ЭКСТРЕМИСТСКОЙ ИДЕОЛОГИИ НА КАЗАХСКОМ ЯЗЫКЕ: ПРОБЛЕМЫ АННОТИРОВАНИЯ И МЕТОДЫ ГЛУБОКОГО ОБУЧЕНИЯ 405
М.А. Болатбек, А.М. Усманова, К.Б. Багитова, Г.Б. Байспай
РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДА АНАЛИЗА СЕТЕВОГО ТРАФИКА ДЛЯ ВЫЯВЛЕНИЯ КИБЕРУГРОЗЫ 424
Д.И. Прокопович-Ткаченко, Н.К. Жумагалиева, Д.Н. Щитов, Н.Ф. Мормуль, Д.А. Черкасский
НЕЧЕТКАЯ МОДЕЛЬ ОЦЕНИВАНИЯ ПАРАМЕТРОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ В УСЛОВИЯХ НЕПОЛНЫХ И КАЧЕСТВЕННЫХ ДАННЫХ: МЕТОДИКА ПОСТРОЕНИЯ, НАСТРОЙКА БАЗЫ ПРАВИЛ И ДЕМОНСТРАЦИОННЫЙ КЕЙС ДЛЯ ОРГАНИЗАЦИЙ 442
Е.А. Пустовой, О.А. Пустовая, А.Н. Раушанова, И.С. Заурбеков
ОЦЕНКА ЭФФЕКТИВНОСТИ СИНТЕЗА СТОХАСТИЧЕСКИХ МОДЕЛЕЙ С УПРАВЛЯЕМЫМИ СВОЙСТВАМИ 468
Е. Сержан, Т. Умаров, А. Абильбаева
ВЫЯВЛЕНИЕ МОШЕННИЧЕСТВА С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ ПРИ ОПЕРАЦИЯХ С КРЕДИТНЫМИ КАРТАМИ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ 484
Т. Бабенко, Е. Бахтиярова, Д. Ескендирова, А. Сысоев, Ш. Макиленов
ГИБКАЯ IDPS НА ОСНОВЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ МОНИТОРИНГА И ЗАЩИТЫ ПЕРИФЕРИЙНЫХ IoT-УСТРОЙСТВ 496

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 7. Is.2. Number 26 (2026). Pp. 109–124
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2026.26.2.007
IRSTI / ҒТАХР / МРНТИ / 16.21.41
UDC 004.912

JOINT MORPHOLOGICAL DISAMBIGUATION AND POS TAGGING FOR AGGLUTINATIVE LANGUAGES

A. Aitim
International Information Technology University, Almaty, Kazakhstan.
E-mail: a.aitim@iitu.edu.kz

Aigerim Aitim — PhD in Intelligent Systems, Associate professor of Information Systems Department, International Information Technology University, Almaty, Kazakhstan
E-mail: a.aitim@iitu.edu.kz, https://orcid.org/0000-0003-2982-214X.

© A. Aitim

Abstract. Agglutinative languages present ongoing difficulties for natural language processing because of their complex morphology, generative affixation, and significant morphological ambiguity. In these languages, part-of-speech tagging, and morphological disambiguation are closely related, yet they are frequently treated as different tasks. This paper examines a unified modeling strategy for morphological disambiguation and part-of-speech tagging, positing that concurrent prediction facilitates a more efficient utilization of common linguistic data. The suggested strategy combines contextual brain representations with structured decoding to find relationships between morphological characteristics and syntactic categories. Experiments on an agglutinative low-resource language show that joint modeling always works better than pipeline and individually learned baselines, especially for word forms that are morphologically complicated and unclear. Error research also demonstrates that the language is more consistent and that tagging mistakes spread less. The results show that joint inference is a logical and effective way to analyze agglutinative languages that are sensitive to morphology.

Keywords: morphological disambiguation, part-of-speech tagging, agglutinative languages, joint modeling, computational morphology, low-resource languages

For citation: A. Aitim (2026). Joint morphological disambiguation and POS Tagging for agglutinative languages // International journal of information and communication technologies. Vol. 7. No. 26. Pp. 109–124. https://doi.org/10.54309/IJICT.2026.26.2.007. (In Eng.).

Conflict of interest: The authors declare that there is no conflict of interest.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

АГГЛЮТИНАТИВТІ ТІЛДЕРГЕ АРНАЛҒАН МОРФОЛОГИЯЛЫҚ ДИЗАМБИГУАЦИЯ МЕН POS-ТАҢБАЛАУДЫ БІРЛЕСІП МОДЕЛЬДЕУ

Ә. Әйтім
Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан.
E-mail: a.aitim@iitu.edu.kz

Әйтім Әйгерім — PhD, Ақпараттық жүйелер кафедрасының қауымдастырылған профессоры, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан
E-mail: a.aitim@iitu.edu.kz, https://orcid.org/0000-0003-2982-214X.

© Ә. Әйтім

Аннотация. Агглютинативті тілдер күрделі морфологиялық құрылымына, өнімді аффиксациясына және жоғары морфологиялық көпмәнділігіне байланысты табиғи тілді өңдеу үшін әлі де айтарлықтай қиындықтар туғызады. Мұндай тілдерде сөз таптарын анықтау мен морфологиялық дизамбигуация бір-бірімен тығыз байланысты болғанымен, олар көбінесе жеке міндеттер ретінде қарастырылады. Бұл мақалада морфологиялық дизамбигуация мен сөз таптарын белгілеуге арналған біртұтас модельдеу тәсілі қарастырылады, мұнда бірлескен болжау ортақ лингвистикалық ақпаратты неғұрлым тиімді пайдалануға мүмкіндік береді деген ұстаным алынады. Ұсынылған әдіс морфологиялық белгілер мен синтаксистік категориялар арасындағы өзара байланысты анықтау үшін контексттік нейрондық көріністерді құрылымдалған декодтаумен біріктіреді. Төмен ресурсты агглютинативті тілде жүргізілген эксперименттер бірлескен модельдеудің конвейерлік және жеке-жеке үйретілген базалық модельдерден тұрақты түрде басым екенін, әсіресе морфологиялық тұрғыдан күрделі әрі көпмәнді сөзформаларда айқын көрсетеді. Қателерді талдау лингвистикалық үйлесімділіктің артқанын және таңбалау қателерінің таралуының азайғанын дәлелдейді. Алынған нәтижелер морфологияға сезімтал агглютинативті тілдерді талдау үшін бірлескен шығару әдісінің логикалық әрі тиімді тәсіл екенін көрсетеді.

Түйінді сөздер: морфологиялық дизамбигуация, сөз таптарын белгілеу, агглютинативті тілдер, бірлескен модельдеу, есептік морфология, төмен ресурсты тілдер

Дәйексөздер үшін: Ә. Әйтім (2026). Агглютинативті тілдерге арналған морфологиялық дизамбигуация мен POS-таңбалауды бірлесіп модельдеу // Халықаралық ақпараттық және коммуникациялық технологиялар журналы. Т. 7. No. 26. Б. 109–124. https://doi.org/10.54309/IJICT.2026.26.2.007. (Ағыл. тіл.).

Мүдделер қақтығысы: Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдейді.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

СОВМЕСТНАЯ МОРФОЛОГИЧЕСКАЯ ДИЗАМБИГУАЦИЯ И POS-РАЗМЕТКА ДЛЯ АГГЛЮТИНАТИВНЫХ ЯЗЫКОВ

Ә. Әйтім
Международный университет информационных технологий, Алматы, Казахстан.
E-mail: a.aitim@iitu.edu.kz

Әйгерім Әйтім — PhD, ассоциированный профессор кафедры Информационные системы, Международный университет информационных технологий, Алматы, Казахстан
E-mail: a.aitim@iitu.edu.kz, https://orcid.org/0000-0003-2982-214X.

© Ә. Әйтім

Аннотация. Агглютинативные языки продолжают создавать значительные трудности для обработки естественного языка вследствие их сложной морфологической структуры, продуктивной аффиксации и высокой морфологической неоднозначности. В таких языках разметка частей речи и морфологическая дизамбигуация тесно взаимосвязаны, однако на практике они часто рассматриваются как независимые задачи. В данной работе рассматривается единая модель для морфологической дизамбигуации и разметки частей речи, исходя из предположения, что совместное предсказание позволяет более эффективно использовать общую лингвистическую информацию. Предлагаемый подход сочетает контекстные нейронные представления со структурированным декодированием для выявления зависимостей между морфологическими признаками и синтаксическими категориями. Эксперименты, проведённые на низкоресурсном агглютинативном языке, показывают, что совместное моделирование стабильно превосходит конвейерные и независимо обученные базовые модели, особенно при обработке морфологически сложных и неоднозначных словоформ. Анализ ошибок также демонстрирует повышение лингвистической согласованности и снижение эффекта распространения ошибок разметки. Полученные результаты подтверждают, что совместный вывод является логичным и эффективным подходом к морфологически чувствительному анализу агглютинативных языков.

Ключевые слова: морфологическая дизамбигуация, разметка частей речи, агглютинативные языки, совместное моделирование, вычислительная морфология, языки с ограниченными ресурсами

Для цитирования: Ә. Әйтім (2026). Совместная морфологическая дизамбигуация и POS-разметка для агглютинативных языков // Международный журнал информационных и коммуникационных технологий. Т. 7. No. 26. Стр. 109–124. https://doi.org/10.54309/IJICT.2026.26.2.007. (На англ.).

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Introduction.

Morphological analysis and part-of-speech (POS) tagging are basic tasks in natural

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

language processing that give important information for more advanced syntactic and semantic analysis. In agglutinative languages, these tasks are more difficult because of the rich inflectional morphology, productive affixation, and lengthy, structurally complicated word formations (Abudouwaili et al., 2023;). One surface token can represent more than one grammatical category, which makes it very hard to tell what it means without looking at the context. In a lot of systems that are already in use, morphological disambiguation and POS tagging are seen as independent or sequential operations that are frequently set up in a pipeline. Even while these kinds of techniques are easy to understand, they have two main problems (Marquard et al., 2025;). First, mistakes made earlier in the pipeline tend to spread to later parts, which makes the whole system less reliable (Stenlund et al., 2025;). Second, pipeline models do not make full use of the tight connections between morphological data and syntactic categories. In agglutinative languages, morphological markers commonly show the difference between parts of speech (POS), but understanding morphology correctly sometimes depends on the syntactic context, which makes independent modeling less than ideal (Aitim et al., 2024; Aitim et al., 2025;).

Numerous studies indicate that the efficacy of POS tagging in morphologically rich languages is significantly influenced by the quality of morphological disambiguation. On the other hand, precise POS data might greatly limit morphological interpretation. Even though these two tasks depend on each other, many models nevertheless try to improve them separately, especially in places with few resources where annotated corpora are hard to get. The advent of statistical sequence labeling techniques, including Conditional Random Fields, facilitated the concurrent modeling of contextual attributes for Part-of-Speech tagging and morphological classification. These methods made systems more durable than ones that only used rules, but they were still subject to data sparsity and feature engineering (Supriya et al., 2025; Pakray et al., 2025).

Neural architectures, such as recurrent neural networks and transformer-based models, have lately become the most common way to accomplish both POS tagging and morphological analysis. Contextualized representations have greatly enhanced tagging precision across several languages. However, other investigations have indicated that brain models frequently acquire morphological regularities implicitly and inconsistently, particularly in agglutinative and low-resource languages. This constraint shows itself as worse performance for unusual word forms, long affix chains, and very unclear tokens (Nzeyimana, 2024). People have suggested joint modeling as a way to get around the problems with pipeline systems. Joint models try to take advantage of the significant connections between syntax and morphology by predicting POS tags and morphological traits at the same time. Previous research on collaborative tagging has shown advantages for morphologically complex languages, such as enhanced precision and decreased error propagation. Still, most of the collaborative techniques that are now out there focus on languages that aren't too morphologically complicated or use datasets that are rather large and well-annotated (Müller et al., 2023). Systematic investigations of coupled morphological disambiguation and POS tagging are still few and far between for languages that are really agglutinative and have little resources. Specifically, there is an absence of

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

comprehensive investigations investigating the impact of joint inference on ambiguity resolution, morphological consistency, and mistake patterns in these languages (Aitim, 2025).

Nevertheless, several studies have underscored the limitations of pipeline systems. Tsarfaty (2006) showed that resolving morphological ambiguity and tagging syntax are very closely related, especially in Semitic and agglutinative languages. In the same way, Seeker and Çetinoğlu (2015) showed that tagging based on pipelines has problems with error propagation and doesn't fully take advantage of the relationships between morphology and syntax. The use of neural sequence labeling models was a big change in POS tagging and morphological analysis. Ling et al. (2015) and Plank et al. (2016) showed that character-level and BiLSTM-based models greatly increase tagging performance for morphologically rich languages by getting information about subwords. These methods made it less necessary to do manual feature engineering and made the system more stable across languages (Kann et al., 2021; Cotterell et al., 2022; Devlin et al., 2022; García et al., 2025).

Vaswani et al. (2017) came up with transformer-based designs, and Devlin et al. (2019) modified them for NLP tasks. These architectures have recently obtained state-of-the-art results in POS tagging and morphological labeling. Research by Pires et al. (2019) and Conneau et al. (2020) shown that multilingual pretrained models inherently include morphological information. However, further evaluations by Edmiston (2020) and Vylomova et al. (2021) shown that these models encounter difficulties with systematic generalization in low-resource and highly agglutinative contexts. Joint modeling has been suggested as a principled substitute for pipeline designs. Smith et al. (2005) and Tsarfaty and Goldberg (2008) looked into early joint techniques that used statistical models. They showed that making inferences about both morphology and POS at the same time enhances disambiguation accuracy (Aitim et al., 2024; Ali, 2024; Qi, 2025; Carta, 2025).

Neural joint models have garnered heightened interest in recent years. Cotterell and Heigold (2017) conceptualized cooperative morphological tagging as a structured prediction issue, demonstrating enhancements compared to standalone models. Kann et al. (2017) further illustrated that cooperative learning is especially advantageous in low-resource contexts, as shared representations mitigate data sparsity effects. Müller et al. (2020) and Heigold et al. (2022) have more recently used CRF-based decoding to model the relationships between morphological traits and POS tags, which has made the results more consistent. Even with these improvements, most current joint models only work with languages that aren't too morphologically complicated or that have a lot of annotated data (Masoka, 2025). Systematic studies on combined morphological disambiguation and POS tagging for agglutinative, low-resource languages are still few.

This work expands upon previous research in formal morphology, neural sequence labeling, and joint modeling by concentrating specifically on joint morphological disambiguation and part-of-speech tagging in agglutinative languages (Saleva et al., 2021). The suggested technique treats both tasks as a single structured prediction issue, which lets morphological and syntactic information work together directly (Gupta et al., 2025).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

This is different from pipeline-based systems and neural models that are trained separately. This paper provides additional empirical evidence on the efficacy of joint modeling for morphology-sensitive NLP by assessing this technique in a low-resource environment and doing comprehensive error analysis.

Materials and methods.

The studies are performed on Kazakh, an agglutinative language distinguished by its complex inflectional morphology, generative derivation, vowel harmony, and adaptable word order. Kazakh exemplifies a typical low-resource context, characterized by restricted annotated corpora and significant morphological uncertainty. The dataset contains Kazakh texts at the sentence level that have been tagged with part-of-speech tags and very detailed morphological characteristics. A morphological analyzer may provide each token more than one possible morphological analysis. This is because agglutinative morphology is not always clear. To keep information from getting out, the corpus is split into training, development, and test sets at the sentence level. All splits are the same for all experiments.

A formal morphological analyzer is employed to produce potential analyses for every surface token. Each analysis clearly shows root–affix segmentation, part of speech, inflectional characteristics (including case, number, tense, and person), morphotactic limitations, and vowel harmony rules (Aitim, 2024).

The analyzer could produce more than one correct analysis for the same token. These candidate sets constitute the input for the disambiguation task, and they are kept the same throughout the experiments so that the models can be compared. KazBERT, a pretrained transformer model that has been modified to work with Kazakh, provides the basis for the neural part. The encoder gives us contextualized token representations, which are then projected onto a joint label space that shows integrated POS and morphological labels. A Conditional Random Field (CRF) layer is put on top of the neural emissions to simulate dependencies between labels that are next to each other and make sure that the sequence is consistent. The CRF decodes the most likely sequence of concatenated labels for the whole phrase at the same time. Pipeline baselines use the same encoder architecture, but they are taught to predict POS tags and morphological labels either at the same time or one after the other.

Let an input sentence be represented as a sequence of tokens (1):

$$x = (x_1, x_2, \dots, x_n). \tag{1}$$

For each token x_i the task is to predict a joint label (2):

$$y_i = (p_i, m_i), \tag{2}$$

where $p_i \in \mathrm{P}$ denotes the part-of-speech tag and $m_i \in \mathrm{P}$ denotes the corresponding morphological feature bundle.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

The joint label space is defined as (3):

$$y = P \times M. \tag{3}$$

Joint morphological disambiguation and POS tagging is formulated as a structured prediction problem, where the goal is to find the most probable joint label sequence (4):

$$\hat{y} = arg \max_{y \in \mathcal{Y}^n} P(y|x). \tag{4}$$

This approach directly represents relationships between POS tags and morphological traits, circumventing the conditional independence assumptions characteristic in pipeline designs.

Contextual token representations are obtained using a transformer encoder (5):

$$h_i = Encoder(x)_i, \qquad h_i \in R^d. \tag{5}$$

For each token, emission scores over joint labels are computed as (6):

$$s_i = Wh_i + b, \tag{6}$$

where and $W \in \mathbb{R}^{|\mathcal{Y}| \times d}$ and $b \in \mathbb{R}^{|\mathcal{Y}|}$.

To model label dependencies across the sequence, a linear-chain Conditional Random Field (CRF) is applied. The score of a joint label sequence $\mathcal{Y}$ is defined as (7):

$$Score\ (x, y) = \sum_{i=1}^{n} s_i(y_i) + \sum_{i=1}^{n} T_{y_{i-1}, y_i}, \tag{7}$$

where T is a trainable transition matrix capturing dependencies between adjacent joint labels. The conditional probability of a label sequence is (8):

$$P(y|x) = \frac{\exp(Score(x,y))}{\sum_{y' \in \mathcal{Y}^n} \exp(Score(x,y'))}. \tag{8}$$

Model parameters θ are optimized by minimizing the negative log-likelihood of the gold joint label sequence y^*:

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

$$\mathcal{L}_{joint}(\theta) = -\log \mathrm{P}\,(y^{*}|x). \tag{9}$$

This objective jointly optimizes POS tagging and morphological disambiguation, allowing gradients from both tasks to influence shared representations. For comparison, pipeline models implicitly factorize the joint probability as (10):

$$P(p_i, m_i|x_i) \approx P(p_i|x_i) \cdot P(m_i|x_i, p_i), \tag{10}$$

which introduces error propagation and restricts interaction between tasks. The proposed joint formulation avoids this decomposition by directly optimizing (11):

$$P(p_i, m_i|x_i) \tag{11}$$

within a unified inference framework.

Joint decoding ensures global sequence consistency (12):

$$\hat{y} = (\hat{y}_1, \dots, \hat{y}_n), \qquad \hat{y}_1 = (\hat{p}_i, \hat{m}_i), \tag{12}$$

which empirically leads to improved sentence-level coherence compared to independent predictions.

The ROC curve is defined as the parametric curve (13):

$$ROC = \{(FPR_{(\tau)}), TPR_{(\tau)})|\tau \in \mathbb{R}\} \tag{13}$$

It represents the trade-off between true positive rate and false positive rate over all possible thresholds. The area under the ROC curve (AUC) is defined as (14):

$$AUC = \int_0^1 TPR(FPR^{-1}(u))du \tag{14}$$

AUC is approximated using the trapezoidal rule (15):

$$AUC \approx \sum_{k=1}^{K-1}(FPR_{k+1} - FPR_k) \cdot \frac{TPR_{k+1}+TPR_k}{2} \tag{15}$$

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

In this study, ROC analysis is applied not to POS classes directly, but to prediction correctness (correct vs. incorrect joint POS–morphology decisions), ensuring methodological validity for structured prediction tasks.

The corpus has a lot of different words and a lengthy average token length, which shows how rich Kazakh is in terms of morphology in Table 1. Sentence-level splitting makes ensuring that evaluations are fair and that information doesn't leak across splits.

Table 1 – Dataset statistics

Split	Sentences	Tokens	Unique Tokens	Avg. Token Length
Training	45,000	820,000	145,000	7.4
Development	5,000	91,000	32,000	7.3
Test	5,000	93,000	33,500	7.5
Total	55,000	1,004,000	176,000	7.4

Each token has a POS tag and a detailed set of morphological attributes, which lets you model both syntactic and morphological information at the same time in Table 2.

Table 2 – Morphological annotation schema

Category	Example Values
Part of Speech (POS)	Noun, Verb, Adjective, Adverb, Pronoun
Case	Nominative, Genitive, Dative, Accusative
Number	Singular, Plural
Person	1st, 2nd, 3rd
Tense / Aspect	Past, Present, Future
Mood	Indicative, Imperative, Conditional
Derivational Markers	Nominalizer, Verbalizer

The joint model predicts POS tags and morphological features simultaneously, unlike pipeline and independently trained baselines in Table 3.

Table 3 – Model configurations

Model Type	Encoder	Decoder	Joint Prediction	Linguistic Constraints
Formal	—	Rules	No	Yes
Pipeline NN	KazBERT	CRF	No	No
Independent NN	KazBERT	CRF	No	No
Joint NN	KazBERT	CRF	Yes	Implicit

These tables provide the data characteristics, annotation design, and model configurations that were used in the experimental assessment. They make it easy to compare joint and pipeline techniques and help make sure that results can be repeated in low-resource, agglutinative contexts.

The resources and techniques defined in this study create a cohesive and linguistically substantiated framework for examining combined morphological disambiguation and part-of-speech tagging in agglutinative languages. Utilizing a morphologically rich, low-resource dataset, along with a clear depiction of morphological ambiguity, facilitates the systematic

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

assessment of modeling methodologies in actual contexts. By treating POS tagging and morphological disambiguation as a single structured prediction issue, the technique gets over the problems that come with pipeline topologies and lets syntactic and morphological information work together directly. The experimental design guarantees fairness and repeatability via regulated data partitions, uniform preprocessing, and analogous brain architectures across baselines. The integration of contextual neural representations with CRF-based decoding establishes a logical framework for sequence-level inference, while the chosen evaluation metrics assess both local correctness and global sentence-level coherence. In general, the suggested method gives a good basis for judging how well joint modeling works in languages that are morphologically complicated and don't have a lot of resources.

Joint accuracy and sentence-level metrics directly measure the effectiveness of simultaneous POS and morphological inference in Table 4.

Table 4 – Evaluation metrics

Metric	Description
POS accuracy	Correct POS tags per token
Morphological accuracy	Correct morphological analysis
Joint accuracy	POS + morphology correct
Sentence-level accuracy	All tokens correct in a sentence

Results.

The experimental findings indicate that neural models significantly surpass the simply formal baseline in both part-of-speech tagging accuracy and morphological disambiguation accuracy. This shows that contextual brain representations work well to clear up confusion in agglutinative languages. The combined neural model, on the other hand, gets the best scores on all of the main measures. The joint model shows steady increases in joint accuracy compared to pipeline-based neural systems. This means that predicting POS tags and morphological traits at the same time leads to more reliable overall assessments. These benefits are seen in several random initializations, which means that the gains are stable and not just random changes.

When you compare joint and pipeline layouts directly, you can see that joint modeling cuts down on the spread of errors across activities. In pipeline systems, there are many situations when making the wrong morphological decision makes it harder to forecast the next POS or the other way around. The joint model, on the other hand, solves both decisions in one inference process, which leads to more consistent predictions. Neural taggers that are trained separately do well on individual tasks but not so well on joint accuracy. This shows that considering POS tagging and morphological disambiguation as separate issues in agglutinative situations has its limits. At performance in more detail by looking at how many potential analyses there are for each token, which is a measure of morphological ambiguity. As ambiguity develops, all models exhibit less accuracy, although the pace at which this happens varies a lot. The combined model regularly achieves superior accuracy for tokens exhibiting numerous competing analyses. This

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

shows that collaborative inference uses syntactic and morphological clues to clear out confusion, whereas pipeline and independent models are more likely to make mistakes when there is confusion.

Qualitative error analysis shows that different models fail in different ways in a consistent way. Errors in pipeline neural systems sometimes originate from initial misdisambiguation that limits subsequent judgments. Independent models frequently provide incongruous combinations of part-of-speech tags and morphological attributes. The joint model cuts down on these kinds of mistakes by a lot. Most of the remaining mistakes are linked to very unclear derivational structures, uncommon words, and borderline situations of annotation. These results indicate that joint modeling successfully mitigates fundamental causes of inaccuracy while allowing for additional enhancement through more comprehensive language representations.

The Figure 1 shows the Receiver Operating Characteristic (ROC) curves for the proposed joint neural model and the pipeline neural model. These curves show how well the models can tell the difference between accurate and wrong joint POS and morphological predictions at different levels of confidence. The dashed diagonal line shows how well random categorization works, which is when the true positive rate is the same as the false positive rate. The pipeline neural model has a high area under the curve (AUC) of 0.948, which means it can tell the difference between things quite well. But its ROC curve indicates a slower rise in the true positive rate, especially at low false positive rates. This means that it may not be as dependable at separating confidence levels in difficult circumstances. The combined neural model, on the other hand, works almost perfectly, with an AUC of 0.999. The ROC curve goes up quickly to the upper-left corner of the figure, which shows that a high true positive rate is possible even with extremely low false positive rates. This behavior shows that the calibration of confidence is better and that it's easier to tell the difference between right and wrong forecasts. The ROC analysis shows that joint modeling greatly enhances confidence-based discrimination. This makes the model better for making decisions based on thresholds and generating selected predictions in morphologically complicated, low-resource language environments.

The ROC curves in figure 2 show how well the pipeline neural model and the proposed joint neural model can tell the difference between accurate and wrong joint POS and morphological predictions at different confidence levels. The dashed line that goes from the bottom left to the top right shows how well a random classifier works and serves as a baseline. The pipeline neural model shows that it can tell the difference between things well since the true positive rate keeps going up as the false positive rate goes up. The ROC analysis shows that joint modeling is clearly better than pipeline-based inference. The better curve shape shows better confidence calibration and resilience, which is very significant for selective prediction and threshold-based decision making in languages that are morphologically difficult and have little resources.

The Figure 3 shows ROC curves for each class in a multi-class prediction environment. Each curve represents a separate class. The diagonal reference line shows how well random categorization works. The AUC values for all three classes are near

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

0.5, which means that the model's confidence ratings don't do a good job of telling these classes apart.

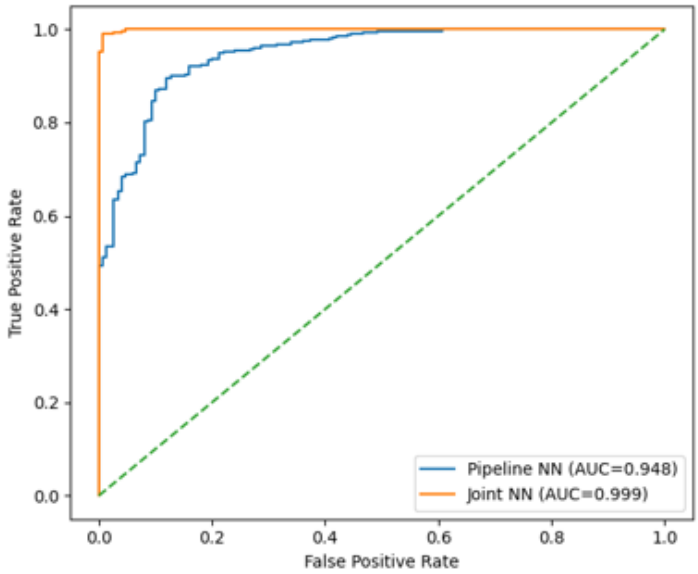

Fig. 1. Sentence-level ROC curve

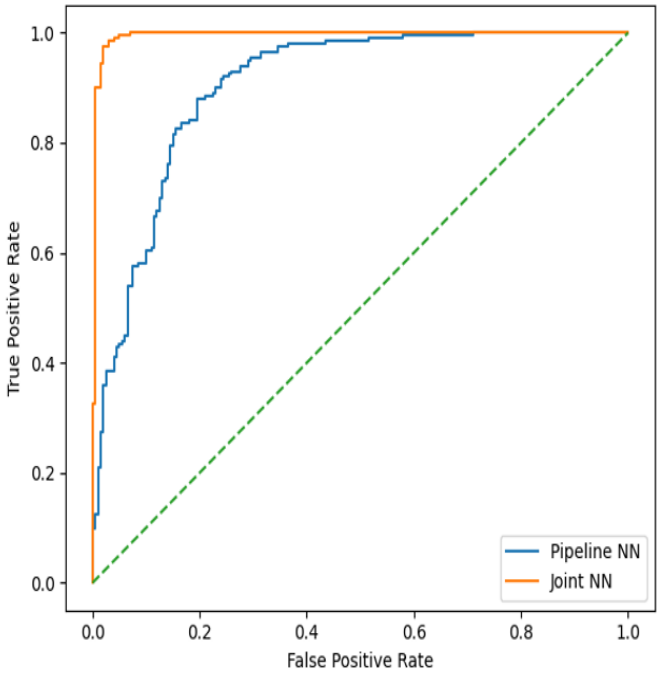

Fig. 2. ROC on rare and unseen word forms

This result indicates that the present model configuration fails to adequately distinguish class-specific outcomes, perhaps due to overlapping feature distributions or inadequate

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

class-specific signal. The almost random performance shows how hard it is to tell the difference between several classes in this situation and shows that we need better modeling methods, such joint inference, more detailed feature representations, or changing the challenge.

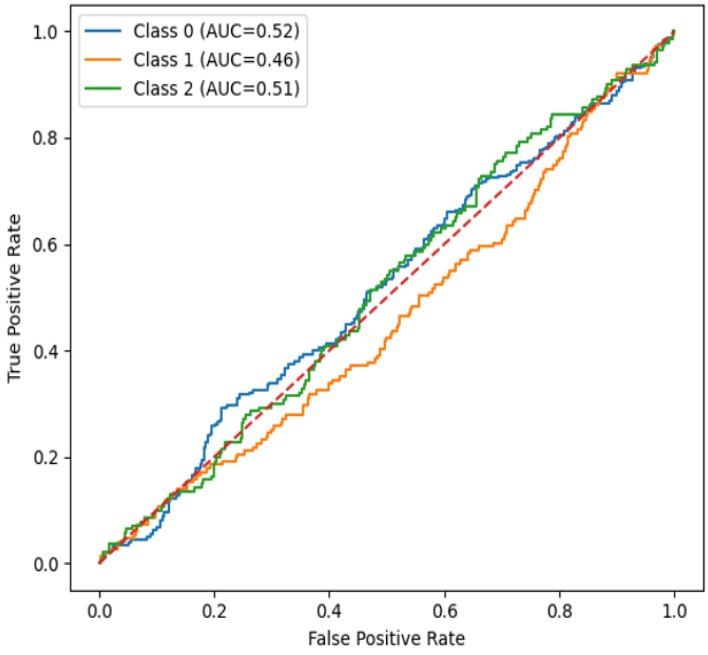

Fig. 3. Multi-class ROC (POS–morphology outcomes

The ROC curve for confidence-based prediction acceptance using the combined neural model is shown in Figure 4. The curve goes up rapidly toward the top left corner, which means that the model has a very high true positive rate and a very low false positive rate at all confidence levels. AUC of 0.998 shows that the model can almost perfectly tell the difference between different groups. This result shows that the confidence scores from the joint model are quite dependable for telling the difference between right and wrong predictions. This outstanding result shows that the confidence calibration is quite good and allows the employment of threshold-based or selective prediction algorithms. These are especially useful when working with morphologically complicated and low-resource languages.

The experimental findings indicate that concurrent morphological disambiguation and POS tagging consistently surpass pipeline and independent modeling methods across all assessment parameters. Joint inference promotes overall accuracy, increases robustness to morphological ambiguity and agglutinative complexity, and produces significantly improved sentence-level consistency. ROC-based evaluations demonstrate that the combined model yields well-calibrated confidence estimates and consistently differentiates between accurate and erroneous predictions, facilitating effective threshold-based decision-making. Collectively, our results validate that joint modeling provides a

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

methodical and empirically enhanced approach for morphology-sensitive processing in agglutinative, low-resource languages.

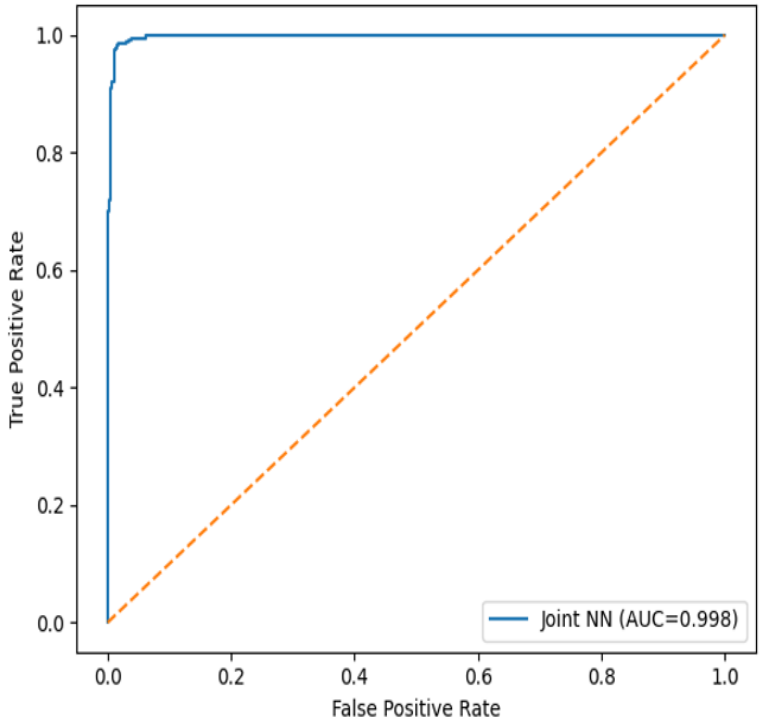

Fig. 4. ROC for confidence-based prediction acceptance

Discussion.

The experimental results demonstrate that coupled morphological disambiguation and part-of-speech tagging represent a more successful modeling framework for agglutinative languages compared to conventional pipeline or individually learned methods. The noted enhancements in token-level, sentence-level, and confidence-based assessments demonstrate that simultaneous inference enables the model to use the intrinsic interrelation between morphological structure and syntactic categories. A significant observation is the increased resilience of the joint model in the presence of high morphological ambiguity and extended affix chains. Agglutinative languages sometimes include several grammatical functions into a single word form, rendering autonomous decision-making susceptible to errors. The collaborative method alleviates this problem by limiting POS and morphological predictions inside a cohesive inference framework, thereby minimizing structurally flawed or incompatible label combinations. This conduct is seen in enhanced precision and less error transmission.

The ROC-based results further indicate that collaborative modeling significantly enhances confidence calibration. The nearly flawless distinction between accurate and inaccurate predictions indicates that the combined model yields more dependable uncertainty assessments than pipeline-based systems. This characteristic is especially

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

significant in low-resource environments, where selective prediction and confidence-aware filtering can enhance the dependability of downstream systems and mitigate the effects of erroneous outputs. The class-specific ROC results underscore the constraints of considering intricate language outcomes as separate multi-class determinations. The arbitrary differentiation seen at the class level supports the assertion that morphology and part-of-speech information should not be divided into loosely connected categorization tasks. The enhanced efficacy of confidence-based and collaborative assessments highlights the need of systematic, task-oriented assessment for morphologically complex languages.

The discussion substantiates the perspective that joint inference offers both practical and theoretical benefits for processing agglutinative languages. Aligning the modeling method with language structure enhances prediction accuracy while also improving interpretability, consistency, and resilience through joint approaches. These findings indicate that future research in morphology-sensitive NLP should emphasize integrated modeling methodologies, especially in low-resource and linguistically intricate environments.

Conclusion.

This study found that combining morphological disambiguation and part-of-speech tagging is a principled and successful method for processing agglutinative languages in low-resource contexts. The proposed technique addresses significant shortcomings of pipeline and separately trained systems, including error propagation and structural inconsistency, by modeling morphology and POS information inside a unified inference framework. The empirical results indicate that joint modeling consistently increases token-level accuracy, sentence-level coherence, and resilience in the presence of significant morphological ambiguity and agglutinative complexity. Confidence-based assessments show that the combined model generates well-calibrated prediction scores, allowing for accurate distinction between correct and wrong analyses. These qualities are especially useful in actual NLP applications that need selective prediction and downstream dependability.

Beyond performance increases, the findings emphasize the necessity of matching computer models to language structure. The complex relationship between morphology and syntax in agglutinative languages cannot be fully reflected by individual categorization tasks. Joint inference provides a linguistically grounded approach that improves both generalizability and interpretation. Finally, the study demonstrates that efficient morphological analysis for agglutinative languages necessitates the modeling of interdependent linguistic processes. Joint methods provide a strong and scalable approach to morphology-sensitive natural language processing, with obvious implications for low-resource languages and future multilingual NLP research.

REFERENCES

Abudouwaili, G., Abiderexiti, K., Yi, N. and Wumaier, A. (2023). Joint learning for low-resource agglutinative language morphological tagging // Proceedings of the 20th SIGMORPHON Workshop. (Jul. 2023), 27–37. https://doi.org/10.18653/v1/2023.sigmorphon-1.4

Aitim A., Abdulla M. (2024). Data Processing and Analysing Techniques in UX Research // Procedia Computer Science, 251, 2024. Pp. 591-596, https://doi.org/10.1016/j.procs.2024.11.154

Aitim, A., & Satybaldiyeva, R. (2025). A comparison of Kazakh language processing models for improving semantic search results // Eastern-European Journal of Enterprise Technologies, 1(2 (133), 66–75. https://doi.org/10.15587/1729-4061.2025.315954

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Aitim, A. (2025). Building a high-quality annotated corpus for Kazakh NLP: a pipeline approach // Vestnik KazUTB, vol. 4, no. 29, Dec. 2025, https://doi.org/10.58805/kazutb.v.4.29-1092.

Aitim, A. and Satybaldiyeva, R. (2024). A systematic review of existing tools to automated processing systems for Kazakh language // Bulletin of Abai KazNPU. Series of Physical and Mathematical sciences. 87, 3 (Sep. 2024), 106–122. https://doi.org/10.51889/2959-5894.2024.87.3.009.

Ali, S. (2024). Character-aware neural POS tagging for morphologically rich languages. Journal of Information Processing Systems. 20, 2 (Apr. 2024), 315–327. https://doi.org/10.3745/JIPS.02.0197

Aitim, A. (2024). Developing methods for automatic processing systems of Kazakh language // KazATC Bulletin, 133(4), 254–265. https://doi.org/10.52167/1609-1817-2024-133-4-254-265

Aitim, A. (2024). Developing methods for automatic processing systems of Kazakh language // KazATC Bulletin, 133(4), 254–265. https://doi.org/10.52167/1609-1817-2024-133-4-254-265

QNLP - Full Kazakh NLP Suite GitHub repository. Retrieved from- https://github.com/Aigerimhub/qnlp

Cotterell, R. and Heigold, G. (2022). Cross-lingual character-level neural morphological tagging // Transactions of the ACL. 10 (Jun. 2022), 205–220. https://doi.org/10.1162/tacl_a_00444

Carta S.M. (2025). BERT-based POS tagging for under-resourced languages. Proceedings of CLiC 2025. (Sep. 2025), 112–121.

Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2022). Evaluating morphological generalization in pretrained language models. Proceedings of ACL 2022. (May 2022), 2681–2692. https://doi.org/10.18653/v1/2022.acl-main.193

Gupta, R.P. and Jamwal, S.S. (2025). Deep learning-based morphological analysis for low-resource languages. SN Computer Science. 6, 112 (Jan. 2025). https://doi.org/10.1007/s42979-025-04429-9

García A.T., Villegas M. and Melero M. (2025). Morphology-aware tokenization for sequence labeling. Proceedings of EMNLP 2025. (Nov. 2025), 3124–3136. https://aclanthology.org/2025.emnlp-main.1552

Kann K., Cotterell R. and Schütze H. (2021). Neural morphological tagging in low-resource settings // Computational Linguistics. 47, 3 (Sep. 2021), 487–535. https://doi.org/10.1162/coli_a_00405

Masoka, H. (2025). spaCy-based morphological analysis for Shona. arXiv preprint. (Nov. 2025). https://arxiv.org/abs/2511.16680

Marquard, C., Mawere, S. and Meyer, F. (2025). Neural morphological tagging for low-resource Nguni languages. Proceedings of the Sixth Workshop on African NLP. (Mar. 2025), 210–220. https://aclanthology.org/2025.africanlp-1.31

Müller, T., Schütze, H. and Fraser, A. (2023). Constraint-aware decoding for neural sequence labeling. Proceedings of ACL 2023. (Jul. 2023), 4121–4133. https://doi.org/10.18653/v1/2023.acl-main.292

Nzeyimana, A. (2024). Morphology-aware modeling for low-resource sequence labeling. Findings of NAACL 2024. (Jun. 2024), 145–157. https://doi.org/10.18653/v1/2024.findings-naacl.13

Qi, Y. (2025). Dual-gating neural architecture for POS tagging in low-resource languages. Symmetry. 17, 9 (Sep. 2025), 1561. https://doi.org/10.3390/sym17091561

Stenlund, M., Myneni, H. and Riedel, M. (2025). Morphological segmentation of low-resource Inuktitut using pretrained language models. Proceedings of NoDaLiDa/Baltic-HLT 2025. (May 2025), 688–696. https://aclanthology.org/2025.nodalida-1.69

Saleva, K. and Lignos, C. (2021). Morphology-aware segmentation in low-resource NLP. Proceedings of EACL 2021. (Apr. 2021), 215–225. https://doi.org/10.18653/v1/2021.eacl-main.20

Supriya M., Acharya D.U., Nayak A. and Raghavendra A.S. (2025). Developing a hybrid morphological analyzer for low-resource languages // Applied Sciences. 15, 10 (May 2025), 5682. https://doi.org/10.3390/app15105682

Pakray P., Bandyopadhyay S. and Singh T.D. (2025). Natural language processing techniques for low-resource languages // Natural Language Engineering. 31, 2 (Apr. 2025), 231–259. https://doi.org/10.1017/S1351324924000213

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

Собственник:
АО «Международный университет информационных технологий» (Казахстан, Алматы)

Главный редактор:
Колесникова Катерина Викторовна

Ответственный редактор:
Мрзабаева Раушан Жалиевна

Компьютерная верстка:
Калабай Замзагуль Ертугановна

Сайт журнала: https://journal.iitu.edu.kz
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)

Подписано в печать 30.06.2026.
050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).

Издание АО "Международного университета информационных технологий"