

ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ  
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН  
MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN



**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ  
КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР  
ЖУРНАЛЫ**

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ  
ИНФОРМАЦИОННЫХ И  
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

**INTERNATIONAL JOURNAL OF INFORMATION  
AND COMMUNICATION TECHNOLOGIES**

**2025 (24) 4**

*қазан- желтоқсан*

ISSN 2708–2032 (print)  
ISSN 2708–2040 (online)

## БАС РЕДАКТОР:

**Исахов Асылбек Абдишимович** — есептеу теориясы саласында математика бойынша PhD доктор, "Компьютерлік ғылымдар және информатика" бағыты бойынша қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университетінің Басқарма Төрағасы – Ректор (Қазақстан)

## БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:

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

## ҒАЛЫМ ХАТШЫ:

**Ипалакова Мадина Түлегеновна** — 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)  
**Лучино 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 Мұхтарбай Өтелбайұлы** — физика-математика ғылымдарының докторы, профессор, КР ҰҒА академигі, Халықаралық ақпараттық технологиялар университеті математика және компьютерлік модельдеу кафедрасының профессоры (Қазақстан)

**Рысбайұлы Болатбек** — физика-математика ғылымдарының докторы, профессор, Есептеу және деректер ғылымдары департаментінің профессоры, Astana IT University (Қазақстан)

**Дайнеко Евгения Александровна** — PhD, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының профессор-зерттеушісі (Қазақстан)

**Дузаев Нуржан Токсужаевич** — 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 Абдуллаевна** — техника ғылымдарының докторы, Халықаралық ақпараттық технологиялар университеті компьютерлік технологиялар және киберқауіпсіздік факультетінің деканы (Қазақстан)

**Мұхамедиева Ардак Габитовна** — экономика ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті бизнес медиа және басқару факультетінің деканы (Қазақстан)

**Абдикаликова Замира Тұрсынбаевна** — 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)

**Дамелия Максумовна Ескендірова** — 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)

**Айтмағамбетов Алтай Зуфарович** — техника ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті радиотехника, электроника және телекоммуникация кафедрасының профессоры (Қазақстан)

**Бахтиярова Елена Азизбековна** — 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 Сахиябай** — техника ғылымдарының кандидаты, профессор, Халықаралық ақпараттық технологиялар университеті компьютерлік инженерия кафедрасының профессор-зерттеушісі (Қазақстан)

**Алимереб 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)

**Талеуш Валдас** — PhD, Адам Мицкевич атындағы (Польша) университеттің проректоры

**Мамырбаев Оркен Жұмажанович** — PhD, КР 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)

**Белошицкая Светлана Васильевна** — техника ғылымдарының докторы, доцент, Astana IT University есептеу және деректер ғылымы кафедрасының профессоры (Қазақстан)

## 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)

Меншік иесі: АҚ «Халықаралық ақпараттық технологиялар университеті» (Алматы қ.).

Қазақстан Республикасы Ақпарат және қоғамдық даму министрлігіне мерзімді баспасөз басылымын есепке қою туралы куәлік № KZ82VPY00020475, 20.02.2020 ж. берілген

Тақырып бағыты: ақпараттық технологиялар, ақпараттық қауіпсіздік және коммуникациялық технологиялар, әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технология.

Мерзімділігі: жылына 4 рет.

Тираж: 100 дана.

Редакция мекенжайы: 050040 Алматы қ., Манас к., 34/1, каб. 709, тел: +7 (727) 244-51-09.

E-mail: [ijict@iitu.edu.kz](mailto:ijict@iitu.edu.kz)

Журнал сайты: <https://journal.iitu.edu.kz>

© Халықаралық ақпараттық технологиялар университеті АҚ, 2025

Журнал сайты: <https://journal.iitu.edu.kz> © Авторлар ұжымы, 2025

---

## ГЛАВНЫЙ РЕДАКТОР

**Исахов Асылбек Абдиашимович** — доктор PhD по математике в области теории вычислимости, ассоциированный профессор по направлению "Компьютерные науки и информатика", Председатель Правления – Ректор Международного университета информационных технологий (Казахстан)

## ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

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

## УЧЕНЫЙ СЕКРЕТАРЬ:

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

**Рысбайұлы 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 Максумовна Ескендирова** — кандидат технических наук, ассоциированный профессор, заведующая кафедрой кибербезопасности Международного университета информационных технологий (Казахстан)

**Ниязгулова Айгуль Аскарбековна** — кандидат филологических наук, доцент, профессор, заведующая кафедрой медиакоммуникации 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 Сахиябай** — 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)

**Талеуш 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)

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

## 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)

Собственник: АО «Международный университет информационных технологий» (г. Алматы).

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

Тематическая направленность: информационные технологии, информационная безопасность и коммуникационные технологии, цифровые технологии в развитии социально-экономических систем.

Периодичность: 4 раза в год.

Тираж: 100 экземпляров.

Адрес редакции: 050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09.

E-mail: [ijict@iitu.edu.kz](mailto:ijict@iitu.edu.kz)

Сайт журнала: <https://journal.iitu.edu.kz>

© АО Международный университет информационных технологий, 2025

© Коллектив авторов, 2025

## EDITOR-IN-CHIEF

**Assylbek Issakhov** — PhD in Mathematics in Computability Theory, associate professor in “Computer Science and Informatics,” Chairman of the Board – Rector of the International Information Technology University (Kazakhstan)

## DEPUTY EDITOR-IN-CHIEF

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

## ACADEMIC SECRETARY

**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 Rysbauly** — 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)

**Bakhtgerai 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)

## EDITOR

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

«International Journal of Information and Communication Technologies»

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Owner: International Information Technology University JSC (Almaty).

The certificate of registration of a periodical printed publication in the Ministry of Information and Social Development of the Republic of Kazakhstan, Information Committee No. KZ82VPY00020475, issued on 20.02.2020.

Thematic focus: information technology, digital technologies in the development of socio-economic systems, information security and communication technologies

Periodicity: 4 times a year.

Circulation: 100 copies.

Editorial address: 050040. Manas st. 34/1, Almaty. +7 (727) 244-51-09. E-mail: [ijict@iitu.edu.kz](mailto:ijict@iitu.edu.kz)

Journal website: <https://journal.iitu.edu.kz>

© International Information Technology University JSC, 2025

© Group of authors, 2025

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Vol. 6. Is.4. Number 24 (2025). Pp. 20–42

Journal homepage: <https://journal.iitu.edu.kz><https://doi.org/10.54309/IJICT.2025.24.4.002>

УДК 004.931

## DEVELOPMENT OF A MODEL FOR SOIL SALINITY SEGMENTATION BASED ON REMOTE SENSING DATA AND CLIMATE PARAMETERS

*G.B. Abdikerimova<sup>1</sup>, M.B. Yessenova<sup>\*1</sup>, G. Yusupova<sup>2</sup>, A.S. Tynykulova<sup>3</sup>,  
A.M. Nedzved<sup>4</sup>*

Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan;

<sup>2</sup>ALT UNIVERSITY, Almaty, Republic of Kazakhstan;

<sup>3</sup>Astana International University, Astana, Kazakhstan;

<sup>4</sup>Belarusian State University, Minsk, Republic of Belarus.

E-mail: [moldir\\_11.92@mail.ru](mailto:moldir_11.92@mail.ru)

**Abdikerimova Gulzira** — Associate Professor, Department of Information Systems, Eurasian National University named after L.N. Gumilyov, PhD, Astana, Kazakhstan <https://orcid.org/0000-0002-4953-0737>;

**Yessenova Moldir** — Senior Lecturer, Department of Information Systems, L.N. Gumilyov Eurasian National University, PhD, Astana, Kazakhstan

E-mail: [moldir\\_11.92@mail.ru](mailto:moldir_11.92@mail.ru), <https://orcid.org/0000-0002-2644-0966>;

**Yusupova Gulbahar** — ALT UNIVERSITY. Departments «Radio Engineering and Telecommunications», Almaty, Kazakhstan

<https://orcid.org/0000-0001-9765-2221>;

**Tynykulova Assemgul** — Astana International University, PhD, Senior Lecturer at the Higher School of Information Technology and Engineering, Astana, Kazakhstan <https://orcid.org/0000-0002-4557-6869>;

**Nedzved Alexander Mikhailovich** — Doctor of Technical Sciences, Associate Professor, Professor of the Department of Computer Technologies and Systems, Faculty of Applied Mathematics and Informatics, Minsk, Belarus

<https://orcid.org/0000-0001-6367-5900>.

© G.B. Abdikerimova, M.B. Yessenova, G. Yusupova, A.S. Tynykulova, A.M. Nedzved

**Abstract.** The paper presents a hybrid machine learning model for the spatial segmentation of soils by salinity using multispectral satellite data from Sentinel-2 and climate parameters of the ERA5-Land model. The proposed method aims to solve the problem of accurate soil cover segmentation under climate change and high spatial heterogeneity of data. The approach includes the sequential application of unsupervised learning algorithms (KMeans, hierarchical clustering, DBSCAN),



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

the XGBoost model, and a multi-tasking neural network that performs simultaneous classification and regression. At the first stage, pseudo-labels are formed using KMeans, then a probabilistic assessment of object membership in classes and ensemble voting of clustering algorithms are carried out. The final model is trained on an extended feature space and demonstrates improved results compared to traditional approaches. Experiments on a sample of 33,624 observations (23,536 — training sample, 10,088 — test sample) showed an increase in the Silhouette Score value from 0.7840 to 0.8156 and a decrease in the Davies-Bouldin Score from 0.3567 to 0.3022. The classification accuracy was 99.99 %, with only one error in more than 10,000 test objects. The results confirmed the proposed method's high efficiency and applicability for remote monitoring, environmental analysis, and sustainable land management.

**Keywords:** soil salinity segmentation; remote sensing; ensemble clustering; Sentinel-2; ERA5-Land; XGBoost; multi-task neural network

**For citation:** G.B. Abdikerimova, M.B. Yessenova, G. Yusupova, A.S. Tynykulova, A.M. Nedzved. Development of a model for soil salinity segmentation based on remote sensing data and climate parameters. 2025. Vol. 6. No. 21. Pp. 20–42 (In Kaz.). <https://doi.org/10.54309/IJICT.2025.21.1.002>.

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

## ЖЕРДІҢ ТҮЗДЫҒЫН СЕГМЕНТАЦИЯЛАУ МОДЕЛІН ДАЛАЛЫҚ ЗЕРТТЕУЛЕР МЕН КЛИМАТТЫҚ ПАРАМЕТРЛЕР НЕГІЗІНДЕ ӘЗІРЛЕУ

*Г.Б. Абдикеримова<sup>1</sup>, М.Б. Есенова<sup>1\*</sup>, Г.М. Юсупов<sup>2</sup>, А.С. Тыныкулова<sup>3</sup>, А.М. Недзьведь<sup>4</sup>*

<sup>1</sup>Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;

<sup>2</sup>ALT UNIVERSITY, Алматы, Қазақстан Республикасы;

<sup>3</sup>Астана Халықаралық университет, Астана, Қазақстан;

<sup>4</sup>Беларусь мемлекеттік университеті, Минск, Беларусь Республикасы.

E-mail: moldir\_11.92@mail.ru

**Абдикеримова Гульзира Бахытбековна** — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық жүйелер кафедрасының доценті, PhD, Астана, Қазақстан, Сатпаев к., 2, 010000  
<https://orcid.org/0000-0002-4953-0737>;

**Есенова Молдир Балкаировна** — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық жүйелер кафедрасының аға оқытушысы, PhD, Астана, Қазақстан

E-mail: moldir\_11.92@mail.ru, <https://orcid.org/0000-0002-2644-0966>;

**Юсупова Гульбахар Мадреймовна** — LT UNIVERSITY. «Радиотехника және телекоммуникациялар» кафедрасының қауымдасқан профессоры, PhD,



Алматы, Қазақстан

<https://orcid.org/0000-0001-9765-2221>;

**Тыныкулова Әсемгүл Серғожақызы** — Астана Халықаралық университеті, PhD, аға оқытушы Ақпараттық технология және инженерия жоғарғы мектебі, Астана қаласы, Қазақстан

<https://orcid.org/0000-0002-4557-6869>;

**Недзьвядь Александр Михайлович** — т.ғ.д., доцент, Қолданбалы математика және информатика факультетінің Компьютерлік технологиялар және жүйелер кафедрасының профессоры, Минск қаласы, Беларусь

<https://orcid.org/0000-0001-6367-5900>.

© Г.Б. Абдикеримова, М.Б. Есенова, Г.М. Юсупова, А.С.Тыныкулова, А.М. Недзьведь

**Аннотация.** Мақалада Sentinel-2 көпспектралды серіктік деректері мен ERA5-Land модельінің климаттық параметрлерін қолдана отырып, топырақтың минералды тұздығына негізделген кеңістік сегментациясына арналған машиналық оқытудың гибридік моделі ұсынылған. Ұсынылған әдіс климаттың өзгеруіне және деректердің жоғары кеңістіктік гетерогенділігіне байланысты топырақ жамылғысын дәл сегментациялау мәселесін шешуге бағытталған. Бұл әдіс бақылаусыз оқыту алгоритмдерін (KMeans, иерархиялық кластерлеу, DBSCAN), XGBoost моделін және бір уақытта классификация мен регрессияны жүзеге асыратын көп тапсырмалы нейрон желісін қолдануды қамтиды. Алғашқы кезеңде KMeans арқылы жалған атаулар құрылады, содан кейін объектілердің сыныптарға жату ықтималдығын бағалау жүргізіледі және кластерлеу алгоритмдерінің біріктірілген дауысы қолданылады. Соңғы модель кеңейтілген белгі кеңістігінде оқытылып, дәстүрлі әдістермен салыстырғанда жақсартылған нәтижелер көрсетеді. 33 624 бақылаудан тұратын таңдамалар бойынша (23 536 — оқу таңдамасы, 10 088 — тест таңдамасы) Silhouette Score мәнінің 0,7840-тан 0,8156-ға дейін өсіп, Davies-Bouldin Score мәнінің 0,3567-ден 0,3022-ге дейін төмендегені көрсетілді. Сыныптау дәлдігі 99,99 %-ды құрады, 10 000 тест объектісінде тек бір қате жасалды. Нәтижелер ұсынылған әдістің қашықтан мониторинг жүргізу, қоршаған ортаны талдау және жер ресурстарын тұрақты басқару үшін жоғары тиімділігі мен қолданылу мүмкіндігін растады.

**Түйін сөздер:** топырақ тұздылығын сегментациялау; алыстан зондтау; ансамбль кластерлеу; Sentinel-2; ERA5-Land; XGBoost; көп міндетті нейрондық желі

**Дәйексөздер үшін:** Г.Б. Абдикеримова, М.Б. Есенова, Г.М. Юсупова, А.С.Тыныкулова, А.М. Недзьведь. Жердің тұздығын сегментациялау моделін далалық зерттеулер мен климаттық параметрлер негізінде әзірлеу// Халықаралық ақпараттық және коммуникалық технологиялар журналы. 2025. Т. 6. No. 21. 20–42 бет. (қазақ тілінде). <https://doi.org/10.54309/IJICT.2025.21.1.002>.

**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

## РАЗРАБОТКА МОДЕЛИ СЕГМЕНТАЦИИ ЗАСОЛЕННОСТИ ПОЧВ НА ОСНОВЕ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ И КЛИМАТИЧЕСКИХ ПАРАМЕТРОВ

**Г.Б. Абдикеримова<sup>1</sup>, М.Б. Есенова<sup>\*1</sup>, Г.М. Юсупов<sup>2</sup>, А.С. Тыныкулова<sup>3</sup>,  
А.М. Недзьведь<sup>4</sup>**

Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан;

<sup>2</sup>ALT university, Алматы, Республика Казахстан;

<sup>3</sup>Международный университет Астана, Астана, Казахстан;

<sup>4</sup>Белорусский государственный университет, Минск, Республика Беларусь.

E-mail: moldir\_11.92@mail.ru

**Абдикеримова Гульзира Бахытбековна** — доцент кафедры информационных систем Евразийского национального университета имени Л.Н. Гумилева, PhD, Астана, Казахстан

E-mail: gulzira1981@mail.ru, <https://orcid.org/0000-0002-4953-0737>;

**Есенова Молдир Балкаировна** — старший преподаватель кафедры информационных систем Евразийского национального университета имени Л.Н. Гумилева, PhD, Астана, Казахстан

<https://orcid.org/0000-0002-2644-0966>;

**Юсупова Гульбахар Мадреймовна** — Ассоциированный профессор кафедры «Радиотехники и телекоммуникации», PhD, Алматы, Казахстан

<https://orcid.org/-0000-0001-9765-2221>;

**Тыныкулова Асемгуль Сергужаевна** — Международный университет Астана, PhD, старший преподаватель Высшей школы информационных технологий и инженерии, Астана, Казахстан

<https://orcid.org/0000-0002-4557-6869>;

**Недзьведь Александр Михайлович** — д.т.н, доцент, профессор кафедры Компьютерных технологии и систем, факультета Прикладной математики и информатики, Республика Беларусь, Минск

<https://orcid.org/0000-0001-6367-5900>.

© Г.Б. Абдикеримова, М.Б. Есенова, Г.М. Юсупова, А.С.Тыныкулова, А.М. Недзьведь

**Аннотация.** В статье представлен гибридная модель машинного обучения для пространственной сегментации почв по солености с использованием многоспектральных спутниковых данных Sentinel-2 и климатических параметров модели ERA5-Land. Предложенный метод направлен на решение проблемы точной сегментации почвенного покрова в условиях изменения климата и высокой пространственной гетерогенности данных. Подход включает последовательное применение алгоритмов неконтролируемого обучения (KMeans, иерархическая кластеризация, DBSCAN), модели XGBoost и многозадачной



нейронной сети, которая выполняет одновременную классификацию и регрессию. На первом этапе формируются псевдоназвания с использованием KMeans, затем осуществляется вероятностная оценка принадлежности объектов к классам и объединяющее голосование алгоритмов кластеризации. Финальная модель обучается на расширенном пространстве признаков и демонстрирует улучшенные результаты по сравнению с традиционными подходами. Эксперименты на выборке из 33 624 наблюдений (23 536 — обучающая выборка, 10 088 — тестовая выборка) показали увеличение значения Silhouette Score с 0,7840 до 0,8156 и снижение значения Davies-Bouldin Score с 0,3567 до 0,3022. Точность классификации составила 99,99 %, было допущено только одна ошибка более чем в 10 000 тестовых объектов. Результаты подтвердили высокую эффективность и применимость предложенного метода для дистанционного мониторинга, анализа окружающей среды и устойчивого управления земельными ресурсами.

**Ключевые слова:** сегментация солености почвы; дистанционное зондирование; ансамблевое кластеризование; Sentinel-2; ERA5-Land; XG-Boost; многоцелевые нейронные сети

**Для цитирования:** Г.Б. Абдикеримова, М.Б. Есенова, Г.М. Юсупова, А.С. Тыныкулова, А.М. Недзведь. Разработка модели сегментации засоленности почв на основе данных дистанционного зондирования и климатических параметров//Международный журнал информационных и коммуникационных технологий. 2025. Т. 6. No. 21. Стр. 20–42. (На каз.). <https://doi.org/10.54309/IJICT.2025.21.1.002>.

**Конфликт интересов:** авторы заявляют об отсутствии конфликта интересов.

## Кіріспе

Құқықтық процестерді автоматтандыру және цифрландыру (Karpenko et. al., 2023; Parysek et. al., 2023) көптеген елдерде, соның ішінде Қазақстанда да, сот жүйесін модернизациялаудың маңызды бөлігіне айнауда. Машиналық оқыту сияқты заманауи технологиялар құқық қолдану тәжірибесінің тиімділігін айтарлықтай арттыруға мүмкіндік береді, әсіресе нақты мәліметтерге шектеулі қолжетімділік жағдайында (Ruslan et. al., 2023; Said et. al., 2023). Сот шешімдерінің нәтижелерін деректер негізінде болжау сот процестерінде бар заңдылықтарды жақсырақ түсінуге, сондай-ақ уақыт пен ресурстарды барынша үнемдей отырып, негізделген шешімдер қабылдауға көмектеседі. Алайда, Қазақстанда сот істеріне қатысты нақты деректерге қолжетімділік құпиялылыққа байланысты шектеулі, бұл ғылыми талдау жүргізуге және олардың негізінде болжам жасау модельдерін әзірлеуге айтарлықтай кедергілер туғызады (Bianchini et. al., 2024; Setiawan et. al., 2024). Осы мәселені шешу Топырақтың тұздануы – ауыл шаруашылығы дақылдарына, экожүйелердің жұмысына және аймақтардың су режиміне тікелей әсер ететін негізгі экологиялық проблемалардың бірі (Aksoy



et all., 2024; Bo et all 2024). Климаттың өзгеруі және ауылшаруашылық жерлерін пайдаланудың интенсификациясы жағдайында топырақтың деградация дәрежесі (Bougiouklis et all., 2025) керемет жылдам қарқынмен артады. Сондықтан тұздылықтың тиімді мониторингі мен болжау құралдарын әзірлеу қажет. Спутниктік деректер, атап айтқанда, Sentinel-2 платформаларынан алынған мульти-диапазонды суреттер (Cui et all., 2025), ERA5-Land сияқты климаттық модельдермен біріктіріліп, топырақ жамылғысының күйі туралы кең ақпарат береді. Дегенмен, оларды өңдеу деректерді талдаудың заманауи әдістерін қажет етеді. Қазіргі уақытта ғылыми әдебиеттерде NDVI, NDWI және NDSI (Gao et all., 2025; Gao et all., 2025), сияқты спектрлік индекстерді қоса алғанда, топырақтың тұздылығын бағалаудың әртүрлі тәсілдері, сонымен қатарты спектрлік и (Geoderma et all., 2024) және регрессия (Jia et all., 2024; Jiang et all., 2025) әдістерін қолданатын машиналық оқыту ұсынылған. Дегенмен, бақыланатын оқытуға негізделген классикалық әдістер сенімді деректер белгілерінің болмауына және мүмкіндіктерді таратудың гет-эрогенділігіне байланысты бірнеше шектеулерге тап болады. Осыған байланысты, соңғы жылдары болжау дәлдігін жақсарту үшін бақыланбайтын алгоритмдерді, ансамбльдік әдістерді және терең нейрондық желілерді біріктіретін гибридік әдістерге қызығушылық артты. Бұл саладағы өзекті мәселелердің бірі әр түрлі топырақ түрлерінің, соның ішінде сортаң, қалыпты және құмды топырақтардың спектрлік ерекшеліктерін ескеру қажеттілігі болып табылады, өйткені құмды топырақтардың спектрлік сипаттамалары сортаңданған аймақтардан келетін сигналдармен қабаттасуы мүмкін. Сонымен қатар, маңызды міндет - жоғары кеңістіктік және спектрлік өзгергіштік жағдайында деректерді бейімдеуге болатын алгоритмдерді әзірлеу. Бұл жұмыста біз бақыланбайтын алгоритмдерді (KMeans (Kaliyeva et all., 2025; Land et all., 2024), агломеративті кластерлеу, DBSCAN (Liu et all., 2024; Luo et all., 2025)) басқарылатын әдістермен (XGBoost (Lusiana et all., 2025., Lusiana et all., 2025), көп тапсырмалы нейрондық желі) біріктіретін гибридік модельді ұсынамыз. Бастапқы кезең жалған белгілерді қалыптастыру үшін KMeans әдісін пайдаланып деректерді кластерлеуді қамтиды, содан кейін олар нақты сыныптарға жататын нысандардың ықтималдығын есептейтін XGBoost үлгісін үйрету үшін пайдаланылады. Әрі қарай, бірнеше кластерлік алгоритмдердің дауыс беруімен ансамбльдік тәсіл қолданылады, содан кейін түпкілікті белгілер топырақтың тұздылығын жіктеу және регрессиялау мәселелерін бір уақытта шешетін көп тапсырмалы нейрондық желіні оқыту үшін пайдаланылады.

Зерттеу ұсынылып отырған әдіс дәстүрлі кластерлеу және жіктеу әдістерімен салыстырғанда топырақты тұздылық бойынша дәлірек сегменттеуді қамтамасыз етеді деп болжайды. Эксперименттерде модель сапасының негізгі көрсеткіштері бағаланады, соның ішінде Silhouette Score, Davies-Bouldin Score, Calinski-Harabasz Score және гибридік ансамбльдің соңғы болжамды үлгінің сапасына әсері. Зерттеудің негізгі нәтижелері гибридік тәсілді қолдану тұзды әсер ететін топырақ сегментациясының дәлдігін жақсартады, әртүрлі топырақ

жамылғысының түрлері арасындағы шекаралық әсерлердің әсерін азайтады және модельді деректердің кеңістіктегі біркелкі еместігіне бейімдейді. Бұл зерттеу топырақты қашықтықтан зондтау әдістерін жақсартуға көмектеседі және тұзданған аймақтағы жердің деградациясын бақылау үшін пайдаланылуы мүмкін. Топырақтың тұздануы – егістік алқаптарының өнімділігіне, экожүйе функцияларына және суға үлкен әсер ететін экологиялық маңызды мәселе. Климаттың өзгеруі және ауыл шаруашылығының қарқынды дамуы туралы, топырақтың тұздануын бақылау және болжау құралдарын құрудың өзектілігі біртіндеп артып келеді. Ғалымдар соңғы жылдары топырақтың тұздылығын болжау мен картаның дәлдігін жақсарту үшін спутниктік бейнелеуді, машиналық оқыту алгоритмдерін және географиялық ақпараттық жүйелерді біріктіруге тырысты. Зерттеудің жаңа бағыттарының бірі топырақтың тұздылығын анықтау үшін қашықтықтан зондтау және спектрлік көрсеткіштерді пайдалану болып табылады. Зерттеуде (Paramonova et al., 2025) Sentinel 1-2 деректері мен терең оқыту алгоритмдері арқылы тұзды жерлерді автоматты түрде анықтау әдістері талданды. Нәтижелер спектрлік индекстерді (NDVI, NDSI, BI) көпқабатты нейрондық желілермен біріктіру тұзды топырақтарды жоғары классификациялық дәлдікке (91,2 % дейін) мүмкіндік беретінін көрсетті. Басқа зерттеу (Sadyrova et al., 2025) топырақ жағдайын талдау үшін гиперспектрлік деректерді пайдаланудың жетілдірілген әдістемесін ұсынды. Тұздылық өзгерістерінің кеңістік-уақыттық заңдылықтарын дәлірек модельдеуге мүмкіндік беретін деректер трансформаторлық әдістерді қолдану арқылы өңделді. Сонымен қатар, топырақтың тұздану динамикасын болжау үшін климаттық деректер мен спутниктік суреттерді біріктірудің әртүрлі тәсілдері зерттелуде. Мысалы, жұмыс (Tashpolat et al., 2025) топырақ температурасы мен ылғалдылығының өзгеруін ескере отырып, топырақтың деградациясының өзгеруін болжауға мүмкіндік беретін машиналық оқытуға және ERA5-Land моделінің деректеріне негізделген әдісті ұсынады. Градиентті күшейту және кездейсоқ орман сияқты жиынтық әдістер дәстүрлі сызықтық регрессия үлгілерімен салыстырғанда болжау дәлдігін 15%-ға арттырды. Зерттеудің тағы бір маңызды бағыты топырақтың сортаңдануына байланысты ауылшаруашылық жерлерінің деградациясын бағалау болып табылады. Зерттеу (Tussupov et al., 2024) аридті аймақтардағы топырақтың тұздану дәрежесін сандық түрде анықтауға мүмкіндік беретін мультиспектрлік зондтау және машиналық оқытуға негізделген тәсілді ұсынады. Авторлар бақыланбайтын әдістерді (KMeans кластерлеу, DBSCAN) бақыланатын әдістермен (XGBoost, нейрондық желілер) біріктіру классификация қателерін азайтуға және топырақ сипаттамаларындағы күрделі сызықтық емес тәуелділіктерді қарастыруға мүмкіндік беретінін атап көрсетеді.

Қашықтықтан зондтаумен қатар, жергілікті өлшемдерге және тұздану процестерін модельдеуге негізделген тәсілдер белсенді түрде әзірленуде. Зерттеуде (Tussupov et al., 2024) Топырақтың тұздылығын картаға түсіру үшін экипажсыз ұшу аппаратының (ҰҰА) деректерін геостатистикалық модельдермен



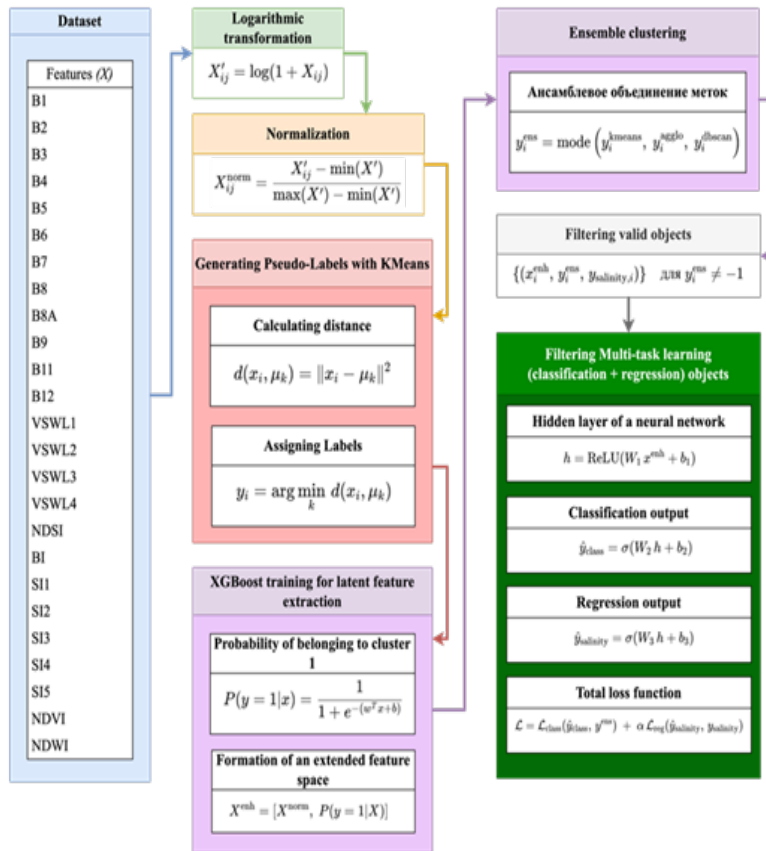
біріктіру мүмкіндіктері талданды. Жұмыс ұшқышсыз ұшу аппараттары терең оқыту әдістерімен үйлесе отырып, жоғары кеңістіктік дәлдікті қамтамасыз ете отырып, бақылау шығындарын азайта алатынын көрсетеді. Соңында, зерттеу (Velilla et al., 2025) әртүрлі аймақтардағы топырақ тұздылығын болжау үшін спутниктік бақылау, жердегі өлшеулер және машиналық оқытуды біріктіретін кешенді тәсілді қарастырады. Авторлар ансамбльдік үлгілерді пайдалану, әсіресе далалық деректердің шектеулі қолжетімділігі жағдайында болжамдардың сенімділігін арттыратынын атап көрсетеді. Осылайша, қазіргі заманғы зерттеулер топырақтың тұздануын бақылау және болжауда айтарлықтай прогресті көрсетеді. Қашықтықтан зондау деректерін, климаттық модельдерді және машиналық оқыту алгоритмдерін біріктіру тұзданған жерлерді жіктеудің дәлдігін арттыруға және топырақ қасиеттерінің кеңістік-уақыттық динамикасына негізделген болжамды модельдерді жасауға мүмкіндік береді. Бұл жерді тұрақты басқару, топырақтың деградациясының әсерін азайту және ауыл шаруашылығы өндірісінің тиімділігін арттыру үшін қажет.

### **Әдістер мен материалдар**

Зерттеуде көпжолалы Sentinel-2 спутниктік суреттері және ERA5-Land климаттық моделі қолданылды. Sentinel-2 спектрлік деректеріне B2, B3, B4, B8, B8A, B11, B12 арналары, сондай-ақ алынған спектрлік индекстер (NDVI, NDWI, NDSI, SI1–SI5, BI) кірді, бұл өсімдік жамылғысын, топырақ ылғалдылығын және жер бетінің минералдану дәрежесін талдауға мүмкіндік берді. Топырақ суының көлемдік көрсеткіштері (VSWL1–VSWL4) ERA5-Land моделінен төрт қабатқа (0–7 см, 7–28 см, 28–100 см, 100–289 см) алынды, бұл топырақтың сортаңдануына ылғалдың әсерін есепке алуға мүмкіндік береді. Талдау үшін әртүрлі топырақ сипаттамалары бар үш аймақ таңдалды: тұздандың жоғары деңгейімен сипатталатын Арал теңізінің бұрынғы түбі; Тұздылығы қалыпты және өсімдіктері типтік Бозайғыр ауылы (Ақмола облысы, Қазақстан); және құмды топырақтардың спектрлік сипаттамаларын бағалау үшін бақылау аймағы ретінде пайдаланылған ең аз өсімдік жамылғысы бар құмды жерлер. Деректер 2019–2024 жылдар аралығын қамтыды, топырақтың тұздану заңдылықтарын және олардың уақыт бойынша динамикасын анықтау үшін жеткілікті ақпарат береді.

1-суретте машиналық оқыту әдістерін біріктіретін топырақ тұздылығын кластерлеуге және болжауға осы тәсілдің әдістемесі суреттелген. Ағындық диаграмма алдын ала өңдеуден және жалған белгіні жасаудан бастап ансамбльді оқытуға және көп тапсырмалы нейрондық модельдеуге дейінгі деректерді өңдеудің дәйекті кезеңдерін көрсетеді.

1. Зерттеудің бірінші кезеңі деректерді алдын ала өңдеуден тұрады. Бұл кезеңде бастапқы деректер одан әрі талдау үшін дайындалады, соның ішінде үлестірімнің асимметриясын жою және мүмкіндіктерді бір шкалаға келтіру. Біріншіден, логарифмдік түрлендіру қолданылады, ол шеткі мәндердің әсерін және үлестірімнің қисаюын азайтады.



Сур. 1. Мәліметтерді өңдеу кезеңдері және топырақтың тұздылығын болжау моделін құру.

Әрі қарай деректер қалыпқа келтіріледі, сондықтан барлық мүмкіндіктер  $[0,1]$  ауқымында мәндерге ие болады. Бұл келесі алгоритмдердің конвергенциясын жақсартады және әрбір мүмкіндіктің модельге біркелкі әсер етуін қамтамасыз етеді.

Логарифмдік түрлендіру (2):

$$X'_{ij} = \log(1 + X_{ij}) \quad (1)$$

where  $X_{ij}$  – is the original value of the  $j$  - th feature of object  $i$ , where  $X \in Rn \times d$ ,  $X'_{ij}$  – prime is the transformed value after applying the logarithm,  $X' \in R_n \times d$ .

2. Нормализация (2):

$$X^{norm}_{ij} = \frac{X'_{ij} - \min(X')}{\max(X') - \min(X')} \quad (2)$$

мұндағы  $X'_{ij}$  – жай – логарифмдік түрлендіруден кейінгі мүмкіндік матрицасы,

$X_{ij}^{norm}$  norm– нормаланған мүмкіндік матрицасы, мұнда мәндер 0-ден 1-ге дейінгі диапазонға жеткізіледі.

Деректерді алдын ала өңдеуден кейін псевдобелгілер KMeans алгоритмі арқылы жасалады. Бұл кезеңде KMeans алгоритмі  $X_{\{norm\}}$  нормаланған деректерді екі кластерге бастапқыда бөлу үшін қолданылады. Алгоритм кластер орталықтарын есептеп, ең жақын орталыққа дейінгі ең аз қашықтық негізінде әрбір нүктеге белгіні тағайындайды. Алынған жалған белгілер бастапқы деректерді бөлуді (3) және белгіні тағайындауды (4) қамтамасыз ететін XGBoost моделін одан әрі оқыту үшін нұсқаулық ретінде пайдаланылады.

$$d(x_i, \mu_k) = |x_i - \mu_k| \quad (3)$$

мұндағы  $x_i - X_{norm}^j$  - объектінің ерекшелік векторы,  $\mu_k$  – кластерінің центрі, мұндағы  $k=1,2$ ,  $d(x_i, \mu_k)$  –  $x_i$  и  $\mu_k$  арасындағы евклидтік қашықтықтың квадраты.

$$y_i = \arg \min_k d(x_i, \mu_k) \quad (4)$$

мұндағы  $d(x_i, \mu_k)$  – әрбір  $x_i$  объектісінің кластер орталықтарына дейінгі есептелген қашықтығы,  $k$  – кластер индексі,  $k \in \{1,2\}$ ,  $y_i - i$  - нысанға тағайындалған жалған белгі,  $Y \in \{0,1\}$ .

Бұл қадамда XGBoost белгілі бір кластерге мүшеліктің ықтималдық бағалауларын алу үшін  $Y$  жалған белгілермен  $X^{norm}$  нормаланған деректері бойынша оқытылады. Бұл ықтималдықтар қарапайым кластерлеу әдістерімен түсірілмеген деректердегі жасырын үлгілерді көрсетеді. Нәтиже (ықтималдық  $(y = 1 | x)$ ) (5)  $X^{enh}$  (6) кеңейтілген мүмкіндік кеңістігін құра отырып, бастапқы мүмкіндіктерге қосылады).

$$P(y = 1 | x) = \frac{1}{1 + e^{-(w^T x + b)}} \quad (5)$$

Мұндағы  $x - X^{norm}$  ерекшелік векторы,  $w$  – модель салмақтарының векторы,  $b$  – (қиғаштық),  $P(y = 1 | x)$  класқа жататын  $x$  объектінің ықтималдығы, мән  $[0,1]$  диапазонында.

$$X^{enh} = X^{norm}, (y = 1 | X) \quad (6)$$

Мұндағы  $X^{norm}$  – нормаланған деректер,  $P(y = 1 | X)$  – формула (6) бойынша алынған әрбір нысан үшін ықтималдық векторы,  $X^{enh}$  – бастапқы мүмкіндіктерді және қосымша ықтималдық мүмкіндігін қамтитын кеңейтілген мүмкіндік матрицасы,  $X^{enh} \in R_n \times (d + 1)$ .

Келесі қадам - ансамбльдік кластерлеуді қолдану. Бұл қадамда  $X^{enh}$  кеңейтілген мүмкіндік кеңістігі үш кластерлеу әдісін қатар қолдану үшін пайдаланылады: KMeans, агломеративті кластерлеу және DBSCAN. Әрбір әдіс нысандарға белгілерді дербес тағайындайды, содан кейін нәтижелер дауыс беру процедурасы арқылы біріктіріледі, мұнда нысан үшін соңғы белгі ең жиі кез-



десетін мән ретінде анықталады. Бұл ансамбльдік тәсіл әрбір әдістің жеке қателерінің әсерін азайтады және деректерді неғұрлым тұрақты және дәлірек бөлуді қамтамасыз етеді (7).

$$y_i^{ens} = \text{mode}(y_i^{kmeans}, y_i^{agglo}, y_i^{dbscan}) \quad (7)$$

Мұндағы  $y_i^{kmeans}$  – нысанға тағайындалған белгі  $i$  алгоритмі Kmeans,  $y_i^{agglo}$  – нысанға тағайындалған белгі  $i$  агломеративті кластерлеу,  $y_i^{dbscan}$  – нысанға тағайындалған белгі  $i$  DBSCAN алгоритмі,  $y_i^{ens} - Y_{ens} \in R_n$  нысанының соңғы ансамблі.

Ансамбльді кластерлеуден кейін кейбір нысандар «шу» (-1) белгісін алуы мүмкін, сондықтан әрі қарай оқыту үшін жарамды белгілері бар нысандар ғана таңдалады. Көп тапсырмалы үлгінің кірісі  $x^{enh}$  кеңейтілген мүмкіндіктері ((таңдалған нысандар үшін), сондай-ақ екі мақсатты айнымалы болады: (1) жіктеу үшін  $y^{ens}$  класы және (2) тұздылықтың жалған белгісі  $y^{salinity}$  біздің жағдайда 1-сынып үшін XGBoost қайтарған ықтималдық немесе басқа қолжетімді тұздылық (8).

$$\{(x_i^{enh}, y_i^{ens}, y_{salinity,i})\} \quad \text{для } y_i^{ens} \neq -1 \quad (8)$$

ұнда  $x^{enh}$  – кеңейтілген мүмкіндіктер,  $y^{ens}$  – ансамбль белгілері,  $y^{salinity}$  – тұздылықтың псевдобелгілері (немесе дәл белгілер).

Қорытынды кезең – жіктеу мен регрессияны біріктіретін көп тапсырмалы оқыту. Бір уақытта екі мәселені шешетін нейрондық желі пайдаланылады: ансамбль белгілері бойынша жіктеу  $y^{ens}$   $y^{salinity}$  псевдобелгілері бойынша регрессия. Модельде стандартты кіріс қабаты, бірнеше жасырын қабаттар (9), содан кейін екі «бас» - классификацияға арналған шығыс (10) және регрессияға арналған шығыс (11). Жаттығу кезінде екі мәселе бойынша жоғалту функциялары (12) бірге ескеріледі.

$$h = \text{ReLU}(W_1 x^{enh} + b_1) \quad (9)$$

мұндағы  $x^{enh}$  – бір нысан үшін кеңейтілген мүмкіндік векторы,  $W_1, b_1$  – бірінші жасырын қабаттың салмақтық матрицасы және қиғаштық векторы,  $h$  – жасырын қабаттың белсендірілуі (жаңа мүмкіндікті көрсету).

$$\hat{y}_{class} = \sigma(W_2 h + b_2) \quad (10)$$

мұндағы  $h$  – жасырын қабат шығысы,  $W_2, b_2$  – классификациялық деңгейдің параметрлері,  $\sigma(z) = \frac{1}{1+e^{-z}}$ ,  $\hat{y}_{class} \in [0,1]$  – 1 классқа жататын болу ықтималдығы.

$$\hat{y}_{salinity} = \sigma(W_3 h + b_3) \quad (11)$$

мұндағы  $h$  – жасырын қабаттың шығысы,  $W_3, b_3$  – регрессия қабатының параметрлері,  $\sigma(z)$  – мұнда болжамды  $[0,1]$  диапазонына әкелетін шектеуші ретінде пайдаланылады,  $\hat{y}_{\text{salinity}} \in [0,1]$  – болжамды тұздылық мәні.

$$\mathcal{L} = \mathcal{L}_{\text{class}}(\hat{y}_{\text{class}}, y^{\text{ans}}) + \alpha \mathcal{L}_{\text{reg}}(\hat{y}_{\text{salinity}}, y_{\text{salinity}}) \quad (12)$$

мұндағы  $\hat{y}_{\text{class}}, \hat{y}_{\text{class}}$  – жіктеу үшін болжамды және нақты белгі  $\hat{y}_{\text{salinity}}, y_{\text{salinity}}$  – болжамды және нақты (немесе жалған) тұздылық белгісі,  $\alpha$  – регрессия мәселесінің жалпы қатеге қосқан үлесін анықтайтын салмақ тұрақтысы,  $\mathcal{L}$  – жаттығуда қолданылатын соңғы жоғалту функциясы.

Іске асыру үшін төмендегі бағдарламалық құрал пайдаланылды: нейрондық желілер үшін Python 3.9.12, TensorFlow 2.12.0, градиентті күшейту үшін XGBoost 1.6.2, кластерлеу және қалыпқа келтіру үшін Scikit-learn 1.1.3, геодефектерді өңдеу үшін GDAL 3.4.0, Matplot2 және визуализация үшін Sea.5.1.2. Спутниктік кескін деректері Google Earth Engine API арқылы алынды және өңделді. Модельдер жоғары есептеу жылдамдығы мен үлкен деректер сыйымдылығын қамтамасыз ету үшін NVIDIA RTX 3090 графикалық картасы (24 ГБ VRAM) және 128 ГБ жады бар есептеу серверінде орындалды. Sentinel-2 деректері пайдаланылған спутниктік деректерге Коперник ашық қолжетімділік хабы (<https://scihub.copernicus.eu>) арқылы қол жеткізіледі. ERA5-Жер климаты деректеріне ECMWF Climate Data Store (<https://cds.climate.copernicus.eu>) арқылы қол жеткізуге болады. Осылайша, ұсынылған әдісті басқа аймақтар үшін қайта шығаруға және бейімдеуге болады, бұл оны топырақтың тұздылығын талдаудың әмбебап құралына айналдырады.

### Нәтижелер және талқылау.

Ұсынылған әдіс топырақтың тұздануын болжау мен кеңістіктік сегменттеуде жоғары тиімділікті көрсетті. Тек KMeans алгоритміне негізделген негізгі тәсілден айырмашылығы, ансамбльді кластерлеуді, ықтималды мүмкіндіктерді байытуды және көп тапсырмалы нейрондық желінің архитектурасын біріктіретін әзірленген жүйе барлық негізгі көрсеткіштерде айтарлықтай жақсартуларды көрсетті. Бұл қашықтан бақылау тапсырмаларына тән күрделі гетерогенді деректерді өңдеу кезінде машиналық оқыту әдістерін біріктірудің орындылығын растайды. Модель екі қосымша деректер көзіне негізделген: Sentinel-2 мультиспектрлі кескіні және ERA5-Land моделінен алынған климаттық деректер. Біз тұздануға әсер ететін беттік шағылысу қасиеттерін де, гидрологиясын да түсіру үшін екеуін де қолдандық. Коперник бастамасы бойынша Еуропалық ғарыш агенттігінің Sentinel-2 спутниктік жүйесі спектрдің үлкен бөлігінде (10–60 м) жоғары кеңістіктік ажыратымдылығы бар бейнелеуді ұсынатын MSI мультиспектрлі сенсорын алып жүреді. Талдау 10 метрлік ажыратымдылық диапазондарын пайдаланды: B2 (көк), B3 (жасыл), B4 (қызыл), B8 және B8A (қысқа толқынды инфрақызыл) және B11 және B12 (орта инфрақызыл). Бұл арналар өсімдік жамылғысы, ылғалдылық және жер бетінің шағылыстыру сипаттамалары туралы ақпаратты, соның ішінде минералды құрамы мен ықтимал тұздылық туралы кеңестерді береді. Модель екі қосымша

ақпарат көзіне негізделген: Sentinel-2 мультиспектрлі бейнелеу және климаттық ERA5-Жер үлгісі деректері. Біз бетінің шағылысу қасиеттерін және тұздануға әсер ететін гидрологияны ескеру үшін қолдандық. Еуропалық ғарыш агенттігінің «Коперник» бағдарламасының Sentinel-2 спутниктік жүйесінде спектрдің кең бөлігінде (10–60 м) жоғары ажыратымдылықты бейнелеуі бар мультиспектрлік сенсор MSI бар. Талдау 10 метрлік ажыратымдылық жолақтарын пайдаланды: B2 (көк), B3 (жасыл), B4 (қызыл), B8 және B8A (қысқа толқынды инфрақызыл) және B11 және B12 (ортаңғы инфрақызыл). Бұл деректер тұзды тік тасымалдау процестерін бағалау үшін қажет, яғни тереңдік пен маусымның функциясы ретінде шаймалау немесе жиналу. Негізгі спектрлік және гидрологиялық қасиеттерден басқа, модельдің өсімдіктер мен топырақтың ылғалдылық қасиеттеріне сезімталдығын арттыратын туынды спектрлік индекстер (SI1–SI5, NDVI, NDWI, NDSI, BI) болды. Осылайша, мүмкіндіктер жиынтығы тұздылықты модельдеуге кешенді тәсілді қамтамасыз ететін беткі және терең параметрлерді қамтиды. Бұл талдау мүмкіндіктерін кеңейтеді және әртүрлі табиғи-климаттық жағдайларда қолданғанда әдістің әмбебаптығын арттырады. Деректер Google Earth Engine арқылы алынды. Бұл жұмыста келесі көрсеткіштер пайдаланылды: 1. Нормалданған дифференциалды тұздылық индексі (NDSI) (13):

$$NDSI = \frac{R - NIR}{R + NIR} \quad (13)$$

мұндағы  $R$  – қызыл арнаның мәні (B4),  $NIR$  – жақын инфрақызыл арна (B8). Бұл көрсеткіш топырақтың тұздану деңгейін бағалау үшін қолданылады. 2. Жарықтық индексі (BI) (14):

$$BI = \sqrt{R^2 + NIR^2} \quad (14)$$

Бұл көрсеткіш беттің жарықтығын көрсетеді және топырақ құрылымын бағалауға көмектеседі. 3. Тұздылық индексі 1 (SI1) (15):

$$SI1 = \sqrt{B + R} \quad (15)$$

мұндағы  $B$  – көк арнаның мәні (B2).

4. Тұздылық индексі 2 (SI2) (16):

$$SI2 = \sqrt{G^2 + R^2 + NIR^2} \quad (16)$$

мұндағы  $G$  – жасыл арнаның мәні (B3).

5. Тұздылық индексі 3 (SI3) (17):

$$SI3 = \frac{SWIR1 - (SWIR1 \cdot SWIR2)}{SWIR1} \quad (17)$$

мұндағы **SWIR1** және **SWIR2** – орташа инфрақызыл арналардың мәндері (B11 және B12).

6. Тұздылық индексі 4 (SI4) (18):

$$SI4 = \sqrt{R \cdot NIR} \quad (18)$$

7. Тұздылық индексі 5 (SI5) (19):

$$SI5 = \frac{R}{\frac{R}{G} + \frac{R}{B}} \quad (19)$$

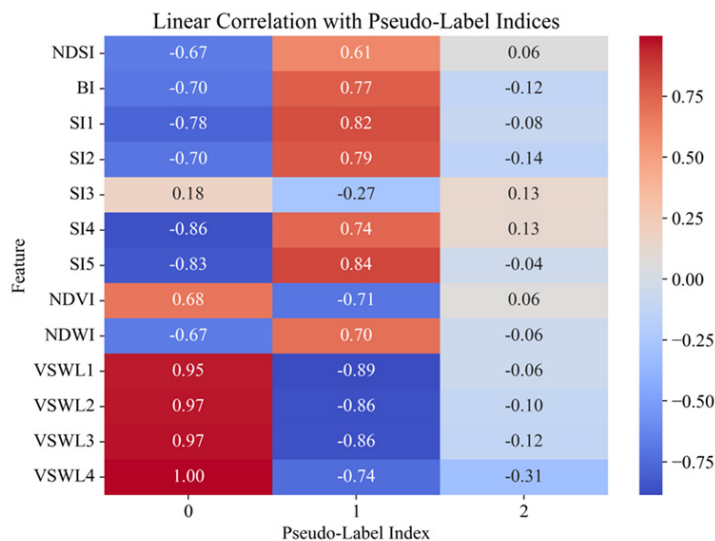
8. Өсімдік жамылғысының нормаланған айырмашылық индексі (NDVI) (20):

$$NDVI = \frac{NIR - R}{NIR + R} \quad (20)$$

9. Судың нормаланған айырма индексі (NDWI) (21):

$$NDWI = \frac{G - NIR}{G + NIR} \quad (21)$$

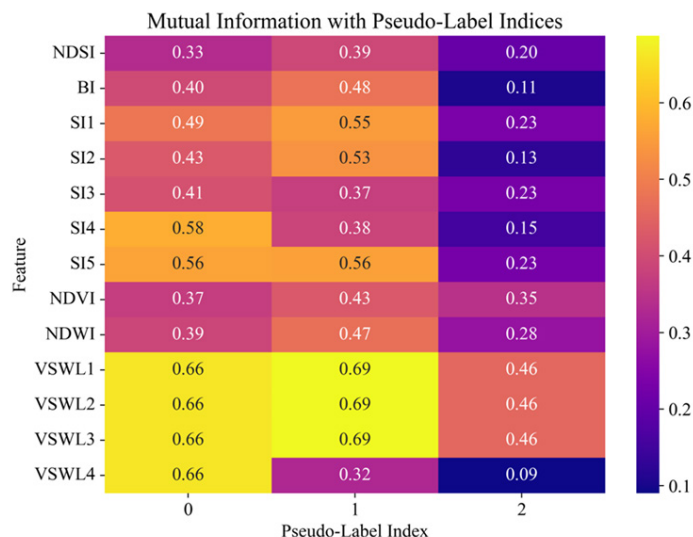
Олар топырақты бағалауда басты орынды алады, өйткені олар оның күйінің әртүрлі аспектілерін, соның ішінде тұздылық пен су балансын анықтайды. Зерттеу жалпы 33 624 деректер жолын өңдеді. Оның ішінде 23 536 жол модельді үйрету үшін пайдаланылды, ал қалған 10 088 жол кластерлеу сапасын сынау және бағалау үшін пайдаланылды. 2-суретте кластерлеуден алынған кластерлердің белгілері мен псевдобелгілері арасындағы сызықтық корреляцияның жылу картасы (Псевдо-белгі индексі) көрсетілген. Y осі спектрлік индекстерді (NDSI, NDVI, NDWI, SI1–SI5, BI) және гидрологиялық ERA5-Жер үлгісінің айнымалы мәндерін (VSWL1–VSWL4) қоса алғанда, пайдаланылған мүмкіндіктерді көрсетеді, ал X осі кластерлік индекстерді (0, 1, 2) көрсетеді. Түс шкаласы корреляция деңгейін білдіреді: Оң мәндер қызыл реңктермен көрсетіледі, бұл тікелей қатынасты көрсетеді, ал теріс мәндер көк түспен көрсетіледі, бұл кері қатынасты білдіреді. 0 кластерімен ең күшті оң корреляция барлық гидрологиялық айнымалылар үшін (VSWL1–VSWL4) байқалады, VSWL4 (1,00) бұл кластер түрімен топырақ ылғалдылығының жоғары деңгейінің байланысын көрсетеді. Сонымен қатар, 1 кластер үшін көптеген белгілер (атап айтқанда, SI1, SI2, SI5, NDVI) оң корреляцияны көрсетеді, бұл тұздылықтың спектрлік ерекшеліктерін көрсетуі мүмкін. 2-кластер әлсіз немесе шамалы теріс корреляциямен сипатталады, бұл оның аралық немесе аралас сипатын көрсетуі мүмкін. Мұндай визуализация кластер құруға жеке белгілердің үлесін жақсырақ түсінуге және кластерлеу кезінде анықталған әртүрлі топырақ түрлерінің физикалық қасиеттерін түсіндіруге мүмкіндік береді.



Сур. 2. Кластерлердің белгілері мен псевдобелгілері арасындағы сызықтық корреляцияның жылу картасы.

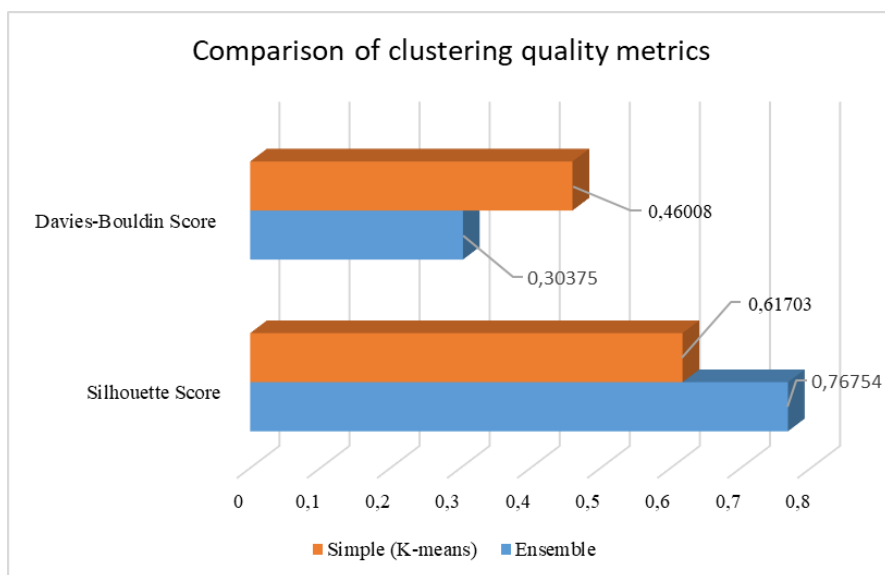
3-суретте кластерлеу процесінде алынған кластерлердің белгілері мен псевдобелгілерінің индекстері арасындағы өзара ақпараттың жылу картасы көрсетілген. Тік ось спектрлік индекстерді (NDSI, SI1–SI5, NDVI, NDWI, BI) және гидрологиялық параметрлерді (VSWL1–VSWL4) қамтитын мүмкіндіктерді, ал көлденең ось кластер индекстерін (0, 1, 2) көрсетеді. Түс шкаласы өзара ақпараттың мәнін көрсетеді, мұнда ашық сары реңктер кластер мүшелігіне қатысты мүмкіндіктің ақпараттылығының жоғары деңгейіне сәйкес келеді. 0 және 1 кластерлері үшін VSWL1, VSWL2 және VSWL3 гидрологиялық параметрлері (0,69) топырақ ылғалдылығы деңгейі бойынша деректерді бөлуге қатысты олардың маңызды мәнін растайтын максималды өзара ақпараттық мәндерге жетеді. Спектрлік сипаттамалардың ішінде SI4, SI5 және SI1 таңбалары бар ең жоғары өзара ақпаратқа ие, әсіресе 0 кластері мен 1 кластеріне қатысты, сондықтан топырақтың тұздылығы мен құрылымын анықтауда ең маңызды болып табылады. 2-кластермен салыстырғанда әрбір мүмкіндік үшін өзара ақпарат аз, бұл 2-кластердің құрылымының азырақ немесе аралас сипатын көрсетуі мүмкін. Осылайша, өзара ақпаратты визуализациялау бізге кластер құрылымын құруға ықпал ететін ең маңызды мүмкіндіктерді көруге мүмкіндік береді және қосымша мүмкіндіктерді таңдау және үлгіні түсіндіру үшін пайдаланылуы мүмкін.

4-суретте екі үлгінің кластерлік сапасы салыстырылады: біреуі KMeans алгоритмімен қарапайымырақ (қызғылт сары жолақтар) және бірнеше әдістерді біріктіретін ансамбльдік үлгі (көк жолақтар). Екі ішкі бағалау көрсеткіші пайдаланылды: Силуэт бағасы және Дэвис-Боулдин бағасы.



Сур. 3. Кластерлердің белгілері мен псевдобелгілері арасындағы өзара ақпараттың жылу картасы

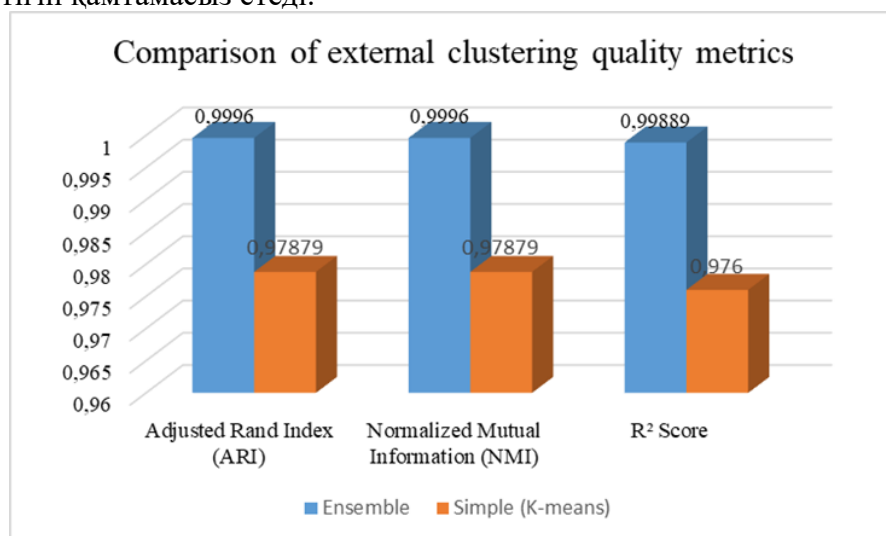
Кластерлердің қаншалықты оқшауланғанын сипаттайтын «Силуэт» көрсеткіші қарапайым үлгіге (0,61703) қарағанда ансамбльдік үлгіде (0,76754) жоғары болды, бұл нысандар кластерлері арасындағы ұлғайтылған бөлуді көрсетеді. Сонымен қатар, Da-vies-Bouldin Score, неғұрлым төмен болса, жақсырақ, ансамбль техникасында да кішірек болды (0,30375 қарсы 0,46008), бұл кластердің ықшамдылығын және аз қабаттасуды бейнелейді. Бұл өлшемдер базалық KMeans алгоритмімен салыстырғанда ансамбль моделінің жақсырақ кластерлеуге қол жеткізетінін растайды.



Сур. 4. Ішкі бағалау метрикасы.

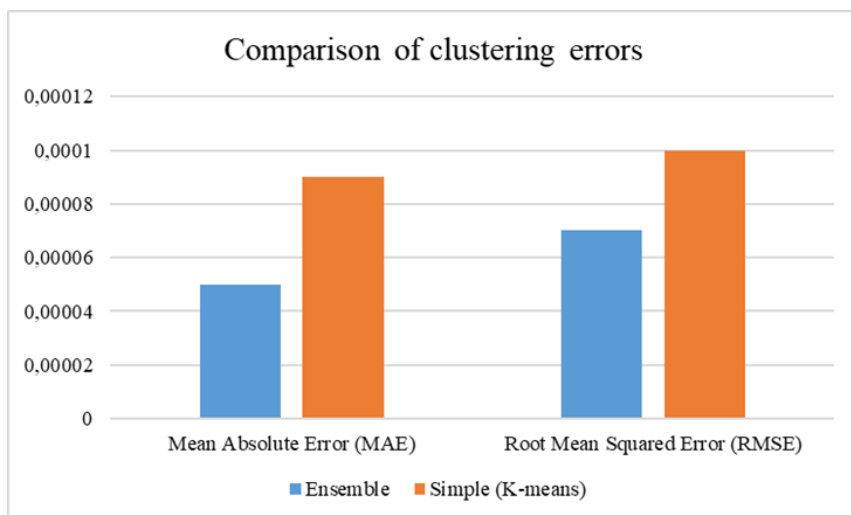


5-суретте екі тәсіл үшін сыртқы кластерлеу сапасының көрсеткіштері салыстырылады: KMeans алгоритміне негізделген қарапайым үлгі (қызғылт сары жолақтар) және біріктірілген кластерлеу тәсілін пайдаланатын ансамбльдік үлгі (көк жолақтар). Бағалау көрсеткіштері ретінде Adjusted Rand Index (ARI), Normized Mutual Information (NMI) және детерминация коэффициенті ( $R^2$  Score) пайдаланылды. Барлық көрсеткіштер ансамбль үлгісінің маңызды артықшылығын көрсетеді. Осылайша, ARI және NMI мәндері ансамбль үшін 0,9996 болса, KMeans үшін 0,97879 болды, бұл тиісті белгілермен жоғары сәйкестікті және кластерлеудің үлкен ақпараттандырғыштығын көрсетеді. Сол сияқты,  $R^2$  Score мәні — ансамбль үшін 0,99889 және KMeans үшін 0,976 — үлгінің жақсырақ түсіндіру күшін көрсетеді. Осылайша, ансамбльдік тәсіл барлық үш көрсеткіш растайтын нақты деректер құрылымына кластерлердің дәлірек сәйкестігін қамтамасыз етеді.



Сур. 5. Сыртқы бағалау метрикасы.

6-суретте екі тәсіл үшін кластерлеу қателері салыстырылады: KMeans алгоритмін пайдаланатын негізгі модель (қызғылт сары жолақтар) және бірнеше әдістерді біріктіретін ансамбльдік модель (көк жолақтар). Өлшем ретінде орташа квадрат қатесі (RMSE) және орташа абсолютті қателік (MAE) қолданылды, бұл бізге болжамды және анықтамалық мәндер арасындағы айырмашылықты салыстыруға мүмкіндік берді. Ансамбльдік модельдің қателік мәндерінің едәуір аз екенін байқауға болады: MAE шамамен 0,00005-ке қарсы 0,00009-ға қарсы қарапайым модель үшін, ал RMSE шамамен 0,00007-ге қарсы 0,0001. Бұл дәлірек болжауды және модельдің нақты деректерді таратуға жақынырақ жақындауын көрсетеді. Осылайша, ансамбльдік тәсіл құрылымдық тұрғыдан анық кластерлеуді қамтамасыз етеді және сандық қателерді азайтады, бұл модельді регрессиялық талдау мәселелеріне немесе топырақтың тұздану дәрежесі сияқты ерекшеліктерді сандық бағалауға қолдану кезінде әсіресе маңызды.

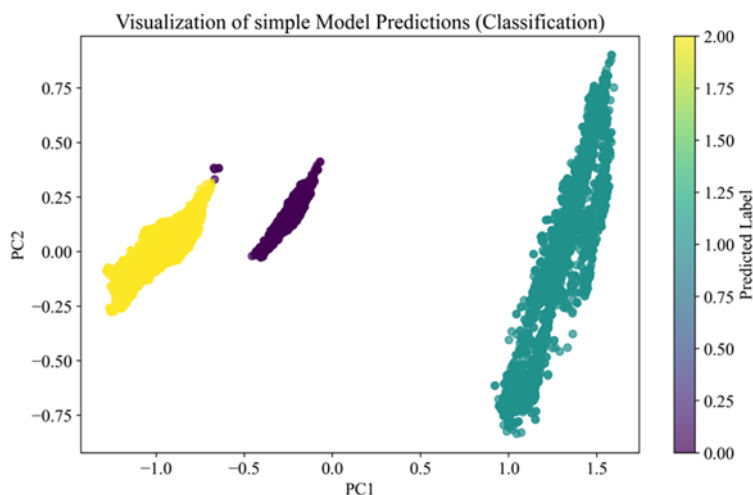


Сур. 6. Өртүрлі кластерлеу әдістері үшін MAE және RMSE талдауы.

Ансамбль үлгісінің жоғары Silhouette Score мәні (қарапайым KMeans үлгісі үшін 0,76754 және 0,61703) жақсы кластерлеу сапасын көрсетеді — кластердегі нысандар ықшамырақ орналасады және кластерлер арасындағы қашықтық ұлғаяды. Бұл ұсынылған алгоритм деректерді бөлуді жақсырақ жобалайтынын білдіреді: бірдей кластері бар нысандар айтарлықтай ұқсастыққа ие, ал әртүрлі кластерлер арасындағы айырмашылықтар жақсарады. Сонымен қатар, KMeans бойынша Дэвис-Боулдин ұпайының 0,46008 ұпайынан ансамбльдік шешімнің 0,30375-ке дейін төмендеуі кластерлер арасындағы тығыз қабаттасуды және көп өлшемді және шулы деректер жиынын қарастырған кезде әсіресе қажет болатын кластер ішілік біртектіліктің жоғарылауын көрсетеді. Сыртқы көрсеткіштер де кластерлеу сапасын жақсарттады. Ансамбль үлгісінің Adjusted Rand Index (ARI) және Norized Mutual Information (NMI) мәндері 0,9996 болды, ал қарапайым үлгінікі 0,97879 болды, бұл болжанған кластерлер мен нақты белгілер арасындағы жоғары сәйкестікті көрсетті. Сол сияқты,  $R^2$  ұпайы KMeans үшін 0,976-дан ансамбль үшін 0,99889-ға дейін өсті, бұл модельдің жақсырақ түсіндіру күшін көрсетеді. Ансамбльдік модель қателер бойынша да жақсырақ мәндерді көрсетті: MAE қарапайым модель үшін 0,00009-дан 0,00005-ке дейін және RMSE 0,0001-ден 0,00007-ге дейін төмендеді. Бұл дәлірек сандық болжамдарды және анықтамалық мәндерден аз ауытқуды көрсетеді. Ғылыми тұрғыдан алғанда, барлық көрсеткіштердің жетілдірілуі ұсынылып отырған тәсілдің бір ғана кластерлеу алгоритмімен шектелмейтіндігінде. KMeans псевдобелгілерінде оқытылған XGBoost пайдалану деректердегі жасырын үлгілерді анықтауға және бастапқы мүмкіндік кеңістігін ықтималдық сипаттамалармен байытуға, оның ақпараттық мазмұнын арттыруға мүмкіндік береді. KMeans, агломеративті кластерлеу және DBSCAN біріктіретін ансамбльдік механизм деректер құрылымының әртүрлі көріністерін қарастыруға мүмкіндік береді, ал нәтижелер бойынша

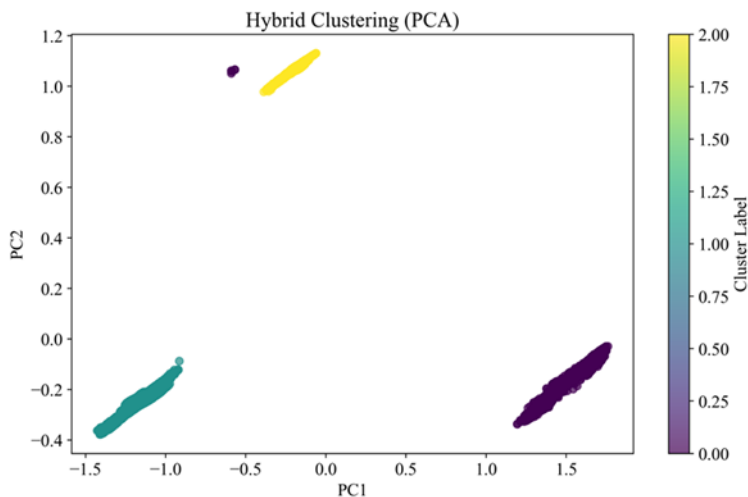
дауыс беру жеке алгоритмдердің кемшіліктерімен байланысты қателердің ықтималдығын азайтады. Кластерлеу кезеңінде алынған белгілердің жоғары сенімділігіне байланысты кейбір объектілерді кейінгі бақылаудағы оқыту үшін пайдалануға болады. Бұл классификациялық есептерді (топырақ түрін анықтау) және регрессияны (тұздану дәрежесін сандық бағалау) бір мезгілде шешетін модель құруға мүмкіндік береді. Осылайша, ұсынылған әдіс деректер құрылымын ерекшелейтін бақыланбайтын әдістердің де, дәлірек, сенімді және түсіндірілетін болжамдарды қамтамасыз ететін сыныптардың шекараларын нақтылайтын бақыланатын әдістердің де күшті жақтарын біріктіреді.

7-суретте PCA арқылы алынған екі негізгі құрамдасқа (PC1 және PC2) қатысты деректер көрсетілген. Нүктелер қарапайым үлгінің болжамды белгілеріне сәйкес боялады, мұнда түс айырмашылықтары әртүрлі сыныптарды көрсетеді. Объектілердің бірнеше сегменттерге топтастырылғанын көруге болады, бірақ кейбір нүктелер арасында кейбір қабаттасулар бар және барлық топтардың нақты белгіленген шекаралары жоқ. Бұл модельдің жалпы бөлуді жасайтынын көрсетеді, бірақ ең алдымен спектрлік кеңістіктегі деректер қабаттасатын немесе жеткілікті түрде дифференциацияланбаған болса, шекті қателер болуы мүмкін.



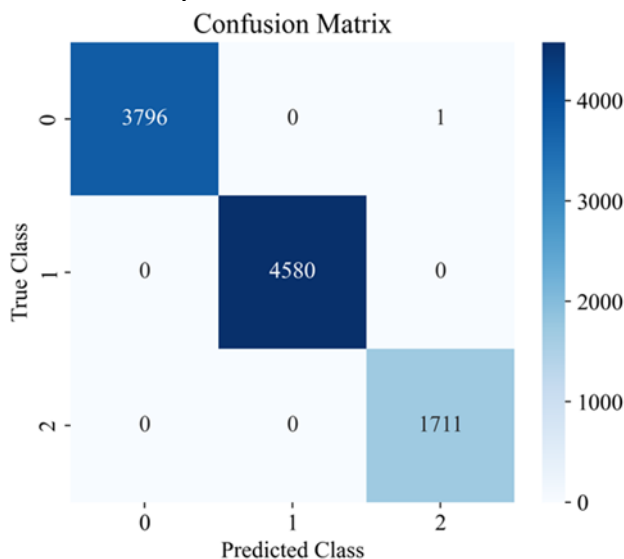
Сур. 7. Қарапайым Үлгі Болжамдарының визуализациясы (Классификация).

8-суретте кластерлеу нәтижелері көрсетілген. Ол сонымен қатар екі өлшемді деректерді проекциялау үшін негізгі құрамдастарды пайдаланады. Дегенмен, соңғы белгілер бірнеше кластерлеу әдістерін (KMeans, агломеративті кластерлеу және DBSCAN) біріктіру және XGBoost арқылы мүмкіндіктерді байыту арқылы алынды. График объектілердің топтарға неғұрлым нақты бөлінуін көрсетеді: кластерлер олардың арасында байқалатын бос орындармен бөлек орналасқан. Бұл ұсынылған тәсіл деректердегі көп өлшемді тәуелділіктерді жақсырақ есептей алатынын көрсетеді, бұл дәлірек және сенімді сегментацияға әкелді.



Сур. 8. Гибридті кластерлеу (PCA).

Ансамбль үлгісі кластерлеу сапасы мен бөлу айқындылығында сенімдірек нәтиже көрсетеді. Ол жақсы ажыратылған кластерлерді қалыптастырып қана қоймайды, сонымен қатар жіктелуі қиын шекаралық нысандардың азырақ санына ие. Бұл ұсынылған тәсіл үшін жоғары Silhouette мәндерін және төменірек Дэвис-Боулдин мәнін көрсететін сандық көрсеткіштермен де расталады. 9-суретте болжамды класс көлденеңінен, ал шынайы класс тігінен салынған. Ұяшықтардың ішіндегі сандар нақты болжанған сыныпқа үлгі тағайындалған әрбір нақты сыныптың (0,1 немесе 2) қанша нысанын көрсетеді. Барлық дерлік нысандар диагональда шоғырланғандығы анық: 0-сыныптан 3796-сы 0,1-сыныптан 4580-і 1, ал 2-сыныптан 1711-і 2 деп дұрыс болжалды. Жалғыз қателік 2-сыныпқа қате тағайындалған (бірінші жолдағы ұяшық, үшінші бағандағы ұяшық) нақты 0-сыныптың бір объектісі болып табылады.



Сур. 9. Классификация моделінің шатастыру матрицасы өнімділік

Осылайша, ұсынылған нәтижелер ұсынылған модель барлық негізгі көрсеткіштер бойынша KMeans алгоритміне негізделген базалық модельден асып, кеңістіктік сегменттеуді және топырақтың тұздылығын сандық бағалауды тиімді жеңетінін көрсетеді. Sil-houette Score және ARI жоғары мәндері, Дэвис-Боулдин ұпайының төмендеуі және MAE және RMSE қателерінің минималды мәндері кластерлердің құрылымдық анықтығын ғана емес, сонымен қатар сандық болжамдардың жоғары дәлдігін растайды. PCA көмегімен деректерді визуализациялау ансамбль үлгісінің сандық көрсеткіштерге сәйкес келетін оқшауланған және түсіндірілетін топтарды құрайтынын көрсетті. Жіктеу қателерінің ең аз саны (10 мыңнан астам жазбаның біреуі ғана) анықталған сыныптардың жоғары сенімділігін көрсетеді. Мұның бәрі бақыланбайтын әдістердің, ықтималды байытудың және нейрондық желіні көп тапсырмалы оқытудың кешенді комбинациясы дәстүрлі тәсілдерге қарағанда дәлірек, сенімді және практикалық нәтижелер беретінін растайды.

### Қорытынды

Бұл зерттеу Sentinel-2 спутниктік деректері мен ERA5-Land климаттық параметрлерін пайдалана отырып, топырақтың тұздылығын болжау және сегменттеу мәселесіне кешенді тәсілдің тиімділігін көрсетті. Өзірленген модель бақыланбайтын әдістерді (KMeans, Агломеративті кластерлеу, DBSCAN), бақыланатын оқытуды (XGBoost) және гетерогенді кеңістіктік деректермен жұмыс істеу кезінде жоғары дәлдік пен сенімділікке қол жеткізген көп тапсырмалы нейрондық желіні біріктіреді. Негізгі кластерлік көрсеткіштер сапаның жақсарғанын көрсетті: Силуэт ұпайы 0,8156-ға жетті, ал Дэвис-Боулдин ұпайы 0,3022-ге дейін төмендеді. Классификациялық тапсырмада модельдің дәлдігі 99,99 % құрады, 10 000-нан астам сынақ бақылауларында бір рет қате жіберілді.

Ұсынылған әдіс топырақтың әртүрлі типтерінің спектрлік ерекшеліктерін бейімдеуге, шекаралық қателіктерді азайтуға және сортандану процестеріне климаттық факторлардың әсерін ескеруге мүмкіндік береді. Ансамбльдік әдістер мен мүмкіндіктерді ықтималдықпен байыту арқасында дәстүрлі алгоритмдерді пайдаланғаннан гөрі топырақ кластарын дәлірек бөлуге қол жеткізу мүмкін болды.

Алынған нәтижелер ұсынылған тәсілдің ғылыми және қолданбалы маңыздылығын растайды. Оны жердің тозуын бақылау, құнарлылықтың төмендеуі қаупін бағалау және агроэкологиялық жүйелерде шешім қабылдауды қолдау үшін пайдалануға болады. Әдістемені конволюционды нейрондық желілерді қолдану арқылы терең сегменттеу бағытында дамыту, экипажсыз әуе кемелерінің мәліметтерін біріктіру және талдауды басқа аймақтарға және маусымдық интервалдарға кеңейту жоспарлануда. Бұл климаттың өзгеруі жағдайында топырақтың тұздануын бақылаудың әмбебап және ауқымды жүйесін қамтамасыз етеді.



## REFERENCES

- Aksoy S., Sertel E., Roscher R., Tanik A., Hamzehpour N. (2024). Assessment of soil salinity using explainable machine learning methods and Landsat 8 images // *International Journal of Applied Earth Observation and Geoinformation*. — 2024. — Vol. 130. Article 103879. DOI: 10.1016/j.jag.2023.103879.
- Becker S. J., Maloney M. C., Griffin A. W., Lasko K., Sussman H. S. (2025). Bare ground classification using a spectral index ensemble and machine learning models optimized across 12 international study sites // *Geocarto International*. — 2025. — Vol. 40. — No. 1. Article 2465452. DOI: 10.1080/10106049.2023.2465452.
- Bo Q., Lv P., Wang Z., Wang Q., Li Z. (2024). Prediction of the post-mining land use based on random forest and DBSCAN // *PLoS ONE*. — 2024. — Vol. 19. — No. 1. Article e0287079. DOI: 10.1371/journal.pone.0287079.
- Bougiouklis J. N., Barouchas P. E., Petropoulos P., Tsesmelis D. E., Moustakas N. (2025). Precision soil sampling strategy for the delineation of management zones in olive cultivation using unsupervised machine learning methods // *Scientific Reports*. — 2025. — Vol. 15. — No. 1. Article 8253. DOI: 10.1038/s41598-024-41188-9.
- H. Tavakoli Computers and Electronics in Agriculture (2024). Integrating satellite and ground-based measurements for predicting soil salinity // *Computers and Electronics in Agriculture*. — 2024. (Article details not provided).
- Computers and Electronics in Agriculture (2025). Automated soil salinity detection using Sentinel-2 data and deep learning models // *Computers and Electronics in Agriculture*. — 2025. (Article details not provided).
- Cui G., Liu Y., Li X., Wang S., Qu X., Wang L., Zhang W. (2025). Impacts of groundwater storage variability on soil salinization in a semi-arid agricultural plain // *Geoderma*. — 2025. — Vol. 454. Article 117162.
- Environmental Challenges (2024). UAV-based mapping of soil salinity using geostatistical and deep learning approaches // *Environmental Challenges*. — 2024. (Article details not provided).
- Gao L., Xu E. (2025). Mapping cropland soil salinity using multi-cycle classification to mitigate retrieval errors // *Computers and Electronics in Agriculture*. — 2025. — Vol. 231. Article 110055. DOI: 10.1016/j.compag.2024.110055.
- Gao Z., Peña-Arancibia J. L., Siyal A. A. (2025). Three-dimensional soil salinity mapping with uncertainty using Bayesian hierarchical modelling, random forest regression and remote sensing data // *Agricultural Water Management*. — 2025. — Vol. 309. Article 109318. DOI: 10.1016/j.agwat.2024.109318.
- Geoderma (2024). Machine learning-based predictions of soil degradation using ERA5-Land data // *Geoderma*. — 2024. (Article details not provided).
- Jia P., Zhang J., Liang Y., Zhang S., Jia K., Zhao X. (2024). The inversion of arid-coastal cultivated soil salinity using explainable machine learning and Sentinel-2 // *Ecological Indicators*. — 2024. — Vol. 166. Article 112364. DOI: 10.1016/j.ecolind.2023.112364.
- Jiang Z., Ding J., Li Z., Liu J. (2025). Present knowledge and future challenges in remote sensing for soil salinization monitoring: A review of bibliometric analysis // *International Journal of Remote Sensing*. — 2025. — Vol. 46. — No. 1. Pp. 247–272. DOI: 10.1080/01431161.2024.2412804.
- Kaliyeva D., Tokbergenova A., Mirzabaev A., Zulpykharov K., Bissenbayeva S., Taukebayev O., Qadir M. (2025). Assessing soil erosion risk in Kazakhstan: A RUSLE-based approach for land rehabilitation // *Polish Journal of Environmental Studies*. — 2025. — Vol. 34. — No. 3.
- Land (2024). Multispectral remote sensing for soil salinity assessment in arid regions // *Land*. — 2024. (Article details not provided).
- Liu X., Hu Y., Li X., Du R., Xiang Y., Zhang F. (2024). An interpretable model for salinity inversion assessment of the South Bank of the Yellow River based on Optuna hyperparameter optimization and XGBoost // *Agronomy*. — 2024. — Vol. 15. — No. 1. Article 18. DOI: 10.3390/agronomy15010018.
- Luo Z., Deng M., Tang M., Liu R., Feng S., Zhang C., Zheng Z. (2025). Estimating soil profile salinity under vegetation cover based on UAV multi-source remote sensing // *Scientific Reports*. — 2025. — Vol. 15. — No. 1. Article 2713. DOI: 10.1038/s41598-024-40157-9.
- Lusiana E. D., Astutik S., Nurjannah N., Sambah A. B. (2025). Using machine learning approach to cluster marine environmental features of Lesser Sunda Island // *Journal of Applied Data Science*. — 2025. — Vol. 6. — No. 1. Pp. 247–258. DOI: 10.1007/s42488-024-00127-1.
- Lusiana Vediappan D., Godi A., Golla B. (2025). Geospatial mapping of soil properties of forest types using the k-means fuzzy clustering approach to delineate site-specific nutrient management zones in Goa, India // *Journal of the Indian Society of Remote Sensing*. — 2025. (Volume and pages not provided). DOI: 10.1007/s12524-024-01813-5.
- Paramonova T.A., Shynbergenov Y.A., Botavin D.V., Golosov V.N. (2025). Assessment of soil pollution and



erosion processes in the Republic of Kazakhstan according to literature data // Eurasian Soil Science. — 2025. — Vol. 58. — No. 1. P. 11. DOI: 10.1134/S1064229325010083.

Sadyrova G., Tynybekov B., Nazarbekova S., Orazbekova K., Ibragimov T., Shimshikov B., Nurmakhanova A. (2025). Impact of cattle grazing on degradation of mountain pastures in South-East Kazakhstan // ES Energy & Environment. — 2025. (Article details not provided).

Tashpolat N., Rehemat A. (2025). Monitoring of soil salinity in the Weiku Oasis based on feature space models with typical parameters derived from Sentinel-2 MSI images // Land. — 2025. — Vol. 14. — No. 2. Article 251. DOI: 10.3390/land14020251.

Tussupov J., Yessenova M., Abdikerimova G., Aimbetov A., Baktybekov K., Murzabekova G., Aitimova U. (2024). Analysis of formal concepts for verification of pests and diseases of crops using machine learning methods // IEEE Access. — 2024. — Vol. 12. Pp. 19902–19910. DOI: 10.1109/ACCESS.2024.3363886.

Tussupov J., Abdikerimova G., Ismailova A., Kassymova A., Beldeubayeva Z., Aitimov M., Makulov K. (2024). Analyzing disease and pest dynamics in steppe crop using structured data // IEEE Access. — 2024. — Vol. 12.

Velilla E., Snethlage J., Poelman M., van der Meer I.M., van der Werf A., Deolu-Ajayi A.O., van Belzen J. (2025). Too salty to farm: Rethinking coastal land use in response to soil salinization // Restoration Ecology. — 2025. — Vol. 33. Article e70006. DOI: 10.1111/rec.70006.



**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ  
КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И  
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

**INTERNATIONAL JOURNAL OF INFORMATION AND  
COMMUNICATION TECHNOLOGIES**

Правила оформления статьи для публикации в журнале на сайте:

**<https://journal.iitu.edu.kz>**

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

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

ОТВЕТСТВЕННЫЙ РЕДАКТОР  
**Мрзабаева Раушан Жалиқызы**

НАУЧНЫЙ РЕДАКТОР  
**Ермакова Вера Александровна**

ТЕХНИЧЕСКИЙ РЕДАКТОР  
**Рашидинов Дамир Рашидинович**

КОМПЬЮТЕРНАЯ ВЕРСТКА  
**Асанова Жадыра**

Подписано в печать 15.12.2025.

Формат 60х881/8. Бумага офсетная. Печать - ризограф. 9,0 п.л. Тираж 100  
050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).

---

Издание Международного университета информационных технологий  
Издательский центр КБТУ, Алматы, ул. Толе би, 59