

№	Статья и ссылка	Аннотация
1	Aizhulov, D. ., Kurmanseiit, M. ., Shayakhmetov, N., Tungatarova, M., & Suleimenova, A. (2024). SYNTHETIC DATA GENERATION FOR ANN MODELING OF THE HYDRODYNAMIC PROCESSES OF IN-SITU LEACHING. <i>Scientific Journal of Astana IT University</i>, 17(17), 5–15. https://doi.org/10.37943/17STXF5228	The work presents an approach to enhance the forecasting capabilities of In-Situ Leaching processes during both the production stage and early prognosis. ISL, a crucial method for resource extraction, demands rapid on-site forecasting to guide the deployment of new technological blocks. Traditional modeling techniques, though effective, are hindered by their computational demands and network throughput requirements, particularly when dealing with substantial datasets or remote computing needs. The integration of AI technologies, specifically neural networks, offers a promising opportunity for expedited calculations by leveraging the power of forward propagation through pretrained neural models. However, a critical challenge lies in transforming conventional numerical datasets into a format suitable for neural modeling. Furthermore, the scarcity of training data during the production phase, where vital parameters are concealed underground, poses an additional challenge in training AI models for In-Situ Leaching processes. This research addresses these challenges by proposing a methodology for generating training data tailored to the most resource-intensive Computational Fluid Dynamics problems encountered during modeling. Traditional numerical modeling techniques are harnessed to construct training datasets comprising input and corresponding expected output data, with a particular focus on varying well network patterns. Subsequent efforts are directed at the conversion of the acquired data into a format compatible with neural networks. The data is normalized to align with the data ranges stipulated by the activation functions employed within the neural network architecture. This preprocessing step ensures that the neural model can effectively learn from the generated data, facilitating accurate forecasting of In-Situ Leaching processes. An advantage of proposed technique lies in provision of large, reliable datasets to train neural network to predict hydrodynamic properties based on technological regimes currently active or

		expected on ISL site. A major implication of this approach lies in applicability of pre-trained AI technologies to forecast future or determine current hydrodynamic regime in the stratum circumventing cost deterministic simulations currently deployed at mining sites. Hence, innovative approach outlined in this paper holds promise for optimizing forecasting, allowing for quicker and more efficient decision-making in resource extraction operations while getting around the computational barriers associated with traditional methods. Keywords: In-Situ Leaching modelling, neural networks, data preparation, data normalization, computational fluid dynamics, synthetic data generation
2	Amirkumar, M., Orynbekova , K. ., Talasbek, A. ., Ayazbayev , D. ., & Cankurt, S. . (2024). COMPARATIVE EFFECTIVENESS OF RULE-BASED AND MACHINE LEARNING METHODS IN SENTIMENT ANALYSIS OF KAZAKH LANGUAGE TEXTS. <i>Scientific Journal of Astana IT University</i>, 17(17), 16–27. https://doi.org/10.37943/17RHPH9724	Sentiment analysis is increasingly pivotal in natural language processing (NLP), crucial for deciphering public opinions across diverse sectors. This research conducts a comparative examination of rule-based and machine learning (ML) methods in sentiment analysis, specifically targeting the Kazakh language. Given the Kazakh language's limited exposure in computational linguistics, the study meticulously evaluates datasets from news articles, literature, and Amazon product reviews, aiming to compare the efficiency, adaptability, and overall performance of these distinct approaches. Employing a detailed set of evaluation metrics such as accuracy, precision, recall, and computational efficiency, the study provides a comprehensive analysis of the strengths and limitations of rule-based techniques versus ML models like Logistic Regression, Multinomial Naive Bayes, Decision Trees, Random Forest, and XGBoost. The findings suggest rule-based methods excel in identifying nuanced emotional expressions within literary texts, while ML models demonstrate superior adaptability and robustness, particularly effective in handling the linguistic variations found in news and reviews. Despite the strengths identified, the study also reveals significant limitations of the rule-based approach, especially in broader contexts beyond literary analysis. This highlights an

		imperative for future research to integrate sentiment dictionaries or domain-specific lexicons that cater to a wider array of linguistic styles, potentially enhancing sentiment analysis tools' applicability in Kazakh and similar less-studied languages. This investigation contributes significantly to the sentiment analysis discourse, offering invaluable insights for both researchers and practitioners by elucidating the complexities of applying NLP technologies across diverse linguistic landscapes, thus advancing the understanding and methodologies of sentiment analysis in the Kazakh language context. Keywords: sentiment analysis, machine learning , rule-based approach , Logistic Regression , Multinomial Naive Bayes
3	Mimenbayeva, A., Issakova, G., Tanykpayeva , B., Tursumbayeva, A. ., Suleimenova, R. . ., & Tulkibaev, A. . (2024). APPLYING MACHINE LEARNING FOR ANALYSIS AND FORECASTING OF AGRICULTURAL CROP YIELDS. <i>Scientific Journal of Astana IT University</i>, 17(17), 28–42. https://doi.org/10.37943/17LKYP9288	Analysis and improvement of crop productivity is one of the most important areas in precision agriculture in the world, including Kazakhstan. In the context of Kazakhstan, agriculture plays a pivotal role in the economy and sustenance of its population. Accurate forecasting of agricultural yields, therefore, becomes paramount in ensuring food security, optimizing resource utilization, and planning for adverse climatic conditions. In-depth analysis and high-quality forecasts can be achieved using machine learning tools. This paper embarks on a critical journey to unravel the intricate relationship between weather conditions and agricultural outputs. Utilizing extensive datasets covering a period from 1990 to 2023, the project aims to deploy advanced data analytics and machine learning techniques to enhance the accuracy and predictability of agricultural yield forecasts. At the heart of this endeavor lies the challenge of integrating and analyzing two distinct types of datasets: historical agricultural yield data and detailed daily weather records of North Kazakhstan for 1990-2023. The intricate task involves not only understanding the patterns within each dataset but also deciphering the complex interactions between them. Our primary objective is to develop models that can accurately predict crop yields based on various weather parameters, a crucial aspect for effective agricultural planning and

		resource allocation. Using the capabilities of statistical and mathematical analysis in machine learning, a Time series analysis of the main weather factors supposedly affecting crop yields was carried out and a correlation matrix between the factors and crops was demonstrated and analyzed. The study evaluated regression metrics such as Root Mean Squared Error (RMSE) and R^2 for Random Forest, Decision Tree, Support Vector Machine (SVM) algorithms. The results indicated that Random Forest generally outperformed the Decision Tree and SVM in terms of predictive accuracy for potato yield forecasting in North Kazakhstan Region. Random Forest Regressor showed the best performance with an $R^2=0.97865$. The RMSE values ranged from 0.25 to 0.46, indicating relatively low error rates, and the R^2 values were generally positive, indicating a good fit of the model to the data. This paper seeks to address these needs by providing insights and predictive models that can guide farmers, policymakers, and stakeholders in making informed decisions. Keywords: machine learning, time series analysis, linear regression, correlation matrix, crop yield
4	Uvaliyeva I., Ismukhamedova, A., Belginova, S., & Shaikhanova , A. (2024). IMPLEMENTATION OF A BASE OF RULES FOR DIFFERENTIAL DIAGNOSIS OF CLINICAL AND HEMATOLOGICAL SYNDROMES BASED ON MORPHOLOGICAL CLASSIFICATION ALGORITHM . <i>Scientific Journal of Astana IT University</i>, 17(17), 43–56. https://doi.org/10.37943/17BRIJ4866	The evolving landscape of modern medicine underscores the growing importance of automating diagnostic processes. This advancement is not merely a convenience but a necessity to harness the full potential of technological progress, aiming to elevate research and clinical outcomes to new heights. Among the innovative strides in this field, the development of diagnostic systems based on morphological classification algorithms stands out. Such systems, rooted in comprehensive rule bases for differential diagnosis, promise to revolutionize the way we approach complex medical conditions. This paper introduces a cutting-edge system that epitomizes this evolution. Designed to harness the power of data analysis, it paves the way for groundbreaking research opportunities. At the heart of this system is a sophisticated set of rules derived from a morphological classification algorithm. This foundation enables the system to perform automated diagnoses of a wide array of

		clinical and hematological syndromes with unprecedented accuracy. A notable application of this technology is its ability to diagnose anemia by analyzing six distinct blood parameters and further categorize the anemia type based on biochemical criteria. The implications of such diagnostic capabilities are profound. By enabling the systematic collection and analysis of statistical data, the system facilitates in-depth research into the prevalence of diseases across different demographic groups. It aids in identifying disease patterns and supports preventive medicine efforts, potentially shifting the paradigm from treatment to prevention. This study not only highlights the system's capacity for enhancing diagnostic precision but also emphasizes its role as a catalyst for medical research and the improvement of healthcare delivery. The integration of such technologies into the medical field promises to enhance the quality of care, streamline diagnostic processes, and open new avenues for medical research, ultimately contributing to the advancement of global health standards. Keywords: rule base, artificial intelligence, automated systems, data analysis, medical information systems, differential diagnosis, clinical and hematological syndromes, morphological classification
5	Bektemyssova, G., & Bakirova, G. (2024). COMPARATIVE ANALYSIS OF FEDERATED MACHINE LEARNING ALGORITHMS . <i>Scientific Journal of Astana IT University</i>, 17(17), 57–67. https://doi.org/10.37943/17BVCN7579	In this paper, the authors propose a new machine learning paradigm, federated machine learning. This method produces accurate predictions without revealing private data. It requires less network traffic, reduces communication costs and enables private learning from device to device. Federated machine learning helps to build models and further the models are moved to the device. Applications are particularly prevalent in healthcare, finance, retail, etc., as regulations make it difficult to share sensitive information. Note that this method creates an opportunity to build models with huge amounts of data by combining multiple databases and devices. There are many algorithms available in this area of machine learning and new ones are constantly being created. Our paper presents a comparative

		analysis of algorithms: FedAdam, FedYogi and FedSparse. But we need to keep in mind that FedAvg is at the core of many federated machine learning algorithms. Data testing was conducted using the Flower and Kaggle platforms with the above algorithms. Federated machine learning technology is usable in smartphones and other devices where it can create accurate predictions without revealing raw personal data. In organizations, it can reduce network load and enable private learning between devices. Federated machine learning can help develop models for the Internet of Things that adapt to changes in the system while protecting user privacy. And it is also used to develop an AI model to meet the risk requirements of leaking client's personal data. The main aspects to consider are privacy and security of the data, the choice of the client to whom the algorithm itself will be directed to process the data, communication costs as well as its quality, and the platform for model aggregation. Keywords: federated learning, FedAvg , FedAdam , FedYogi , FedSparse , loss , accuracy
6	Myrzakerimova, A., Kolesnikova, K. ., Khlevna, I. ., & Nurmaganbetova, M. . (2024). APPLICATION OF MATHEMATICAL MODELS IN THE DIAGNOSIS OF DISEASES OF INTERNAL ORGANS. <i>Scientific Journal of Astana IT University</i>, 17(17), 68–82. https://doi.org/10.37943/17ODJA2930	The application of diagnostic expert systems in medical technology signifies a notable progression, as they provide a computerized framework for decision-support, assisting healthcare practitioners in the process of disease diagnosis. These systems facilitate the integration of patient data, encompassing symptoms and medical history, with a knowledge base in order to produce a comprehensive compilation of potential diagnoses. Through the utilization of knowledge-based methodologies, they enhance these potentialities in order to ascertain the most probable diagnosis. The present study examines expert systems, investigating their historical development, architectural structure, and the approaches utilized for knowledge representation. There is a significant emphasis placed on the advancement and implementation of these systems within the medical industry of Kazakhstan. This paper provides a comprehensive analysis of the benefits and drawbacks associated with diagnostic expert systems, emphasizing their potential to bring

		about significant advancements in medical fields. The study places significant emphasis on the necessity of developing and conducting thorough testing of these systems in order to improve the precision and effectiveness of medical diagnostics. The statement recognizes the importance of continuous research in order to enhance the design and implementation of these systems in various healthcare settings. This research makes a notable addition by examining optimization theory in the field of medical diagnosis. This study presents novel approaches for effectively addressing the intricacies and uncertainties associated with the diagnosis of complicated disorders. The work presents methodology for navigating the complex field of medical diagnostics by utilizing mathematical modeling and optimization approaches, specifically the gradient projection method. The utilization of diverse ways to tackle qualitative ambiguities in this approach signifies a significant progression inside the domain of diagnostic expert systems. Keywords: Diagnosing diseases with the automated system, medical diagnostic expert systems, decision support systems, mathematical modelling
7	Mimenbayeva, A., Bekmagambetova, G., Muratova, G. ., Naizagarayeva, A. ., Ospanova , T. ., & Konyrkhanova , A. (2024). CLASSIFICATION OF KAZAKH MUSIC GENRES USING MACHINE LEARNING TECHNIQUES. <i>Scientific Journal of Astana IT University</i> , 17(17), 83–94. https://doi.org/10.37943/17NZKG3418	This article analysis a Kazakh Music dataset, which consists of 800 audio tracks equally distributed across 5 different genres. The purpose of this research is to classify music genres by using machine learning algorithms Decision Tree Classifier and Logistic regression. Before the classification, the given data was pre-processed, missing or irrelevant data was removed. The given dataset was analyzed using a correlation matrix and data visualization to identify patterns. To reduce the dimension of the original dataset, the PCA method was used while maintaining variance. Several key studies aimed at analyzing and developing machine learning models applied to the classification of musical genres are reviewed. Cumulative explained variance was also plotted, which showed the maximum proportion (90%) of discrete values generated from multiple individual samples taken along the Gaussian curve. A comparison of the

		decision tree model to a logistic regression showed that for f1 Score Logistic regression produced the best result for classical music - 82%, Decision tree classification - 75%. For other genres, the harmonic mean between precision and recall for the logistic regression model is equal to zero, which means that this model completely fails to classify the genres Zazz, Kazakh Rock, Kazakh hip hop, Kazakh pop music. Using the Decision tree classifier algorithm, the Zazz and Kazakh pop music genres were not recognized, but Kazakh Rock with an accuracy and completeness of 33%. Overall, the proposed model achieves an accuracy of 60% for the Decision Tree Classifier and 70% for the Logistic regression model on the training and validation sets. For uniform classification, the data were balanced and assessed using the cross-validation method. The approach used in this study may be useful in classifying different music genres based on audio data without relying on human listening. Keywords: Machine learning algorithms, music genre, Decision Tree Classifier, Logistic regression, cross-validation
8	Muratuly, D., Denissova , N. ., Dyomina, I. ., Tlebaldinova, A. ., Chettykbayev, R. ., & Zuev, V. . (2024). USING STRUCTURAL EQUATION MODELING METHODS TO ASSESS THE UNIVERSITY'S DIGITAL ECOSYSTEM. <i>Scientific Journal of Astana IT University</i>, 17(17), 95–105. https://doi.org/10.37943/17CCXJ5272	This paper explores the construction of a model for evaluating the digital ecosystem within a university, with a focus on identifying key factors influencing satisfaction with the implementation of new digital processes in the educational environment. The study employs mathematical methods, specifically factor analysis, to gauge the impact of these digital processes on the overall educational landscape. A questionnaire was designed to collect relevant data, and structural equation modeling, utilizing the asymptotically distribution-free estimation method with Grammian in STATISTICA software, was employed for survey result processing. The proposed model aims to provide insights into the dynamics of a university's digital ecosystem, offering a systematic approach to assess satisfaction levels and comprehend the implications of integrating novel digital processes within the educational framework. Mathematical methods, including factor analysis, add a quantitative dimension to the evaluation process, enabling a comprehensive

		understanding of the relationships between various factors. The study's methodology ensures a rigorous and systematic analysis of survey data, enhancing the reliability of the findings. The developed model and methodology contribute to advancing our understanding of the digitalization of university environments, providing valuable tools for decision-makers in shaping effective strategies for integrating digital processes in education. The study conducted a survey with 350 participants, including university staff and students. A questionnaire with 17 questions, both open and closed-ended, was developed to collect data. The authors employed structural equation modeling, specifically the asymptotically distribution-free estimation method, for data processing. The study's a posteriori model illustrates the structure of interaction factors influencing satisfaction with the university's digital ecosystem. Keywords: digital transformation, digital ecosystem , mathematical statistics , structural equations, structural equation model , latent exogenous and endogenous variables, path diagram, asymptotically distribution-free estimation using Grammian (ADF)
9	Biloshchytskyi, A., Shamgunova, . M. ., & Biloshchytska , S. . (2024). EXPLORATION OF THE THEMATIC CLUSTERING AND COLLABORATION OPPORTUNITIES IN KAZAKHSTANI RESEARCH. <i>Scientific Journal of Astana IT University</i>, 17(17), 106–121. https://doi.org/10.37943/17ALVR8114	In today's academic environment, the rapid growth of research publications calls for advanced methods to organize and understand the extensive collections of academic work. This study aims to systematically categorize a substantial number of research paper abstracts from Kazakhstani institutions, focusing on identifying key themes and potential interdisciplinary collaboration opportunities. The dataset includes 13,356 abstracts from the Scopus database, covering a wide range of academic fields. The methodology of this research goes beyond traditional hand-done analysis by using advanced text analysis tools to organize the text data efficiently. This initial phase is crucial for summarizing each abstract's core content. The next steps of the analysis use this organized data to find and group similar thematic areas, considering the complex and multi-dimensional nature of academic research topics. The results reveal a diverse array of research themes, highlighting

		the dynamic academic contributions from Kazakhstan. Significant areas such as environmental science, technological advancements, linguistics, and cultural studies are among the prominent clusters identified. These insights not only provide an overview of current research directions but also highlight the potential for cross-disciplinary partnerships. Moreover, the findings have important implications for decision-makers, scholars, and educational institutions by illuminating key research areas and collaborative possibilities. This thematic overview acts as a guide for shaping research policies, fostering academic connections, and efficiently distributing resources within the scholarly community. Ultimately, this study adds to the academic conversation by offering a way to navigate and utilize the wealth of information in scientific literature, promoting a more collaborative and integrated research environment. Keywords: data preprocessing, natural language processing , thematic clustering , research abstracts
10	Mary Mojirade AYANTUNJI, Adekunle Emmanuel MAKANJUOLA, & John Olalekan ATANDA. (2024). ASSESSMENT OF VIRTUAL TEAM TEACHING APPLICATION AMONG PRE-SERVICE TEACHERS IN FEDERAL COLLEGE OF EDUCATION ABEOKUTA. <i>Scientific Journal of Astana IT University</i>, 17(17), 122–132. https://doi.org/10.37943/17DMHN7018	A detailed study of pre-service teachers at the Federal College of Education in Abeokuta examines their collaborative skills and virtual team teaching issues. The major goal is to determine how virtual team teaching affects pre-service teachers' ability to collaborate and navigate its complexities. The study aims to show how virtual team teaching affects pre-service teachers' problems and collaboration. The underlying hypothesis posits that participants engaged in virtual team teaching will exhibit heightened levels of collaboration and critical thinking skills compared to their counterparts employing conventional teaching methods. To accumulate robust empirical evidence, a meticulous 20-item Likert scale questionnaire was judiciously administered to a representative sample of pre-service teachers at the Federal College of Education Abeokuta. The questionnaire methodically gauged participants' perceptions regarding the influence of virtual team teaching on collaborative skills and the challenges encountered. In the subsequent analytical phase, the data underwent rigorous

		scrutiny using descriptive statistics, meticulously assessing the levels of agreement with each questionnaire item. This study's discerning discoveries make a substantial scientific contribution, propelling our knowledge of how virtual team teaching molds pre-service teachers' collaboration skills and navigates challenges. Rooted in scientific rigor, these insights bear potential significance for educational institutions and teacher education programs. They furnish a nuanced understanding of the efficacy of virtual team teaching as a transformative pedagogical approach, offering valuable guidance for the optimization of pre-service teachers' skills to meet the evolving demands of the modern educational landscape. Keywords: Virtual team teaching, Pre-service teachers , Collaborative skills , Challenges, Education
--	--	--