ISSN 2518-170X (Online)
ISSN 2224-5278 (Print)

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ
АКАДЕМИЯСЫ» РҚБ

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ

# ХАБАРЛАРЫ

# ИЗВЕСТИЯ

РОО «НАЦИОНАЛЬНОЙ
АКАДЕМИИ НАУК РЕСПУБЛИКИ
КАЗАХСТАН»

# NEWS

OF THE NATIONAL ACADEMY
OF SCIENCES OF THE REPUBLIC
OF KAZAKHSTAN

**SERIES**

**OF GEOLOGY AND TECHNICAL SCIENCES**

**1 (469)**

**JANUARY – FEBRUARY 2025**

THE JOURNAL WAS FOUNDED IN 1940

PUBLISHED 6 TIMES A YEAR

ALMATY, NAS RK

*NAS RK is pleased to announce that News of NAS RK. Series of geology and technical sciences scientific journal has been accepted for indexing in the Emerging Sources Citation Index, a new edition of Web of Science. Content in this index is under consideration by Clarivate Analytics to be accepted in the Science Citation Index Expanded, the Social Sciences Citation Index, and the Arts & Humanities Citation Index. The quality and depth of content Web of Science offers to researchers, authors, publishers, and institutions sets it apart from other research databases. The inclusion of News of NAS RK. Series of geology and technical sciences in the Emerging Sources Citation Index demonstrates our dedication to providing the most relevant and influential content of geology and engineering sciences to our community.*

*Қазақстан Республикасы Ұлттық ғылым академиясы «ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы» ғылыми журналының Web of Science-тің жаңаланған нұсқасы Emerging Sources Citation Index-те индекстелуге қабылданғанын хабарлайды. Бұл индекстелу барысында Clarivate Analytics компаниясы журналды одан әрі the Science Citation Index Expanded, the Social Sciences Citation Index және the Arts & Humanities Citation Index-ке қабылдау мәселесін қарастыруда. Webof Science зерттеушілер, авторлар, баспашылар мен мекемелерге контент тереңдігі мен сапасын ұсынады. ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы Emerging Sources Citation Index-ке енуі біздің қоғамдастық үшін ең өзекті және беделді геология және техникалық ғылымдар бойынша контентке адалдығымызды білдіреді.*

*НАН РК сообщает, что научный журнал «Известия НАН РК. Серия геологии и технических наук» был принят для индексирования в Emerging Sources Citation Index, обновленной версии Web of Science. Содержание в этом индексировании находится в стадии рассмотрения компанией Clarivate Analytics для дальнейшего принятия журнала в the Science Citation Index Expanded, the Social Sciences Citation Index и the Arts & Humanities Citation Index. Web of Science предлагает качество и глубину контента для исследователей, авторов, издателей и учреждений. Включение Известия НАН РК. Серия геологии и технических наук в Emerging Sources Citation Index демонстрирует нашу приверженность к наиболее актуальному и влиятельному контенту по геологии и техническим наукам для нашего сообщества.*

---

**«ҚР ҰҒА» РҚБ Хабарлары. Геология және техникалық ғылымдар сериясы».**
**ISSN 2518-170X (Online),**
**ISSN 2224-5278 (Print)**

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.).

Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 29.07.2020 ж. берілген **№ KZ39VPY00025420** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *Геология, гидрогеология, география, тау-кен ісі, мұнай, газ және металдардың химиялық технологиялары*

Мерзімділігі: жылына 6 рет.

Тиражы: 300 дана.

Редакцияның мекен-жайы: 050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19

http://www.geolog-technical.kz/index.php/en/

---

**«Известия РОО «НАН РК». Серия геологии и технических наук».**
**ISSN 2518-170X (Online),**
**ISSN 2224-5278 (Print)**

Собственник: Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).

Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ KZ39VPY00025420,** выданное 29.07.2020 г.

Тематическая направленность: *геология, гидрогеология, география, горное дело и химические технологии нефти, газа и металлов*

Периодичность: 6 раз в год.

Тираж: 300 экземпляров.

Адрес редакции: 050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19

http://www.geolog-technical.kz/index.php/en/

---

**News of the National Academy of Sciences of the Republic of Kazakhstan. Series of geology and technology sciences.**
**ISSN 2518-170X (Online),**
**ISSN 2224-5278 (Print)**
Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty).
The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. KZ39VPY00025420,** issued 29.07.2020.
Thematic scope: *geology, hydrogeology, geography, mining and chemical technologies of oil, gas and metals*
Periodicity: 6 times a year.
Circulation: 300 copies.
Editorial address: 28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19
http://www.geolog-technical.kz/index.php/en/

NEWS of the National Academy of Sciences of the Republic of Kazakhstan
SERIES OF GEOLOGY AND TECHNICAL SCIENCES
ISSN 2224–5278
Volume 1. Number 469 (2025), 107–127

***https://doi.org/10.32014/2025.2518-170X.479***

**© O.A. Kolenchukov*, V.A. Fayfer, V.V. Bukhtoyarov, 2025.**
Siberian Federal University, Krasnoyarsk, Russia.
E-mail: okolenchukov@sfu-kras.ru

# PREDICTION OF THE REMAINING SERVICE LIFE OF PUMPING UNIT ELEMENTS BASED ON REGULARIZATION OF RECURRENT NEURAL NETWORKS

**Oleg A. Kolenchukov** – Cand. Tech. Sc., Siberian Federal University, Krasnoyarsk, Russia, e-mail: okolenchukov@sfu-kras.ru, ORCID: https://orcid.org/0000-0001-8644-7302;
**Vyacheslav A. Fayfer** – lab. asst, Siberian Federal University Krasnoyarsk, Russia, e-mail: mega.fayfertt@gmail.com, ORCID: https://orcid.org/0009-0000-8901-6241;
**Vladimir V. Bukhtoyarov** – Cand. Tech. Sc., Siberian Federal University, Krasnoyarsk, Russia, e-mail: vbukhtoyarov@sfu-kras.ru, ORCID: https://orcid.org/0000-0003-4505-2851.

**Abstract:** *Relevance.* The oil industry is one of the key areas of the mining industry. It has a significant impact on the economy of the state, providing various branches of industrial production with oil refining products. To ensure the normal functioning of enterprises it is important to organise continuous transportation of hydrocarbons using pumping units. This approach allows to determine the residual life of the pump unit and to carry out its timely maintenance before failure. *Objective.* The aim of the work is to ensure uninterrupted transportation by means of predictive maintenance based on big data processing technology and machine learning methods. *Methods.* Recurrent neural networks LSTM and GRU were used as mathematical models to determine the residual life of the pump unit in this paper. *Results and Conclusions.* The effectiveness of the modernised recurrent neural networks was demonstrated by comparing them with traditional machine learning methods (PCR and Random Forest) on different data variations. The comparative analysis demonstrated the significant performance of recurrent neural network based models, especially the LSTM model. An improvement in prediction accuracy was shown in comparison with PCR and Random Forest. The average percentage improvement in the four metrics was 36.35 % and 25.21 % using a smaller sample (n = 10000), and 31.86 % and 25.64 % using a larger sample (n = 20000).

**Keywords:** oil and gas industry, pumping unit, forecasting, remaining equipment life, machine learning, recurrent neural network, regularization.

**© О.А. Коленчуков, В.А. Файфер, В.В. Бухтояров, 2025.**
Сібір федералды университеті, Красноярск, Ресей.
E-mail: okolenchukov@sfu-kras.ru

# ҚАЙТАЛАНАТЫН НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ РЕТТЕУ НЕГІЗІНДЕ СОРҒЫ ҚОНДЫРҒЫЛАРЫНЫҢ ЭЛЕМЕНТТЕРІНІҢ ҚАЛҒАН ҚЫЗМЕТ ЕТУ МЕРЗІМІН БОЛЖАУ

**Олег Александрович Коленчуков** – т.ғ.к., Сібір федералды университеті, Красноярск, Ресей, e-mail: okolenchukov@sfu-kras.ru, ORCID: https://orcid.org/0000-0001-8644-7302;
**Вячеслав Алексеевич Файфер** – лаборант, Сібір федералды университеті, Красноярск, Ресей, e-mail: mega.fayfertt@gmail.com, ORCID: https://orcid.org/0009-0000-8901-6241;
**Владимир Викторович Бухтояров** – т.ғ.к., Сібір федералды университеті, Красноярск, Ресей, e-mail: vbukhtoyarov@sfu-kras.ru, ORCID: https://orcid.org/0000-0003-4505-2851.

**Аннотация.** *Өзектілігі.* Мұнай саласы тау-кен өнеркәсібі қызметінің негізгі салаларының бірі. Ол өнеркәсіптік өндірістің әртүрлі салаларын мұнай өңдеу өнімдерімен қамтамасыз ете отырып, мемлекет экономикасына айтарлықтай әсер етеді. Кәсіпорындардың қалыпты жұмыс істеуін қамтамасыз ету үшін сорғы қондырғыларын қолдана отырып, көмірсутектерді үздіксіз тасымалдауды ұйымдастыру маңызды. Бұл тәсіл сорғы қондырғысының қалдық ресурсын анықтауға және істен шыққанға дейін уақтылы қызмет көрсетуге мүмкіндік береді. Жұмыстың *мақсаты* – үлкен көлемдегі деректерді өңдеу технологиясы мен машиналық оқыту әдістеріне негізделген болжамды техникалық қызмет көрсету арқылы үздіксіз тасымалдауды қамтамасыз ету. Бұл жұмыста сорғы қондырғысының қалдық ресурсын анықтау үшін математикалық модельдер ретінде қайталанатын lstm және Gru нейрондық желілері қолданылды. *Нәтижелер мен қорытындылар.* Жаңартылған қайталанатын қолдану тиімділігі нейрондық желілер дәстүрлі Машиналық оқыту әдістерімен (PCR және Random Forest) деректердің әртүрлі вариацияларымен салыстырылып көрсетілді. Салыстырмалы талдау қайталанатын нейрондық желілерге, әсіресе LSTM модельдеріне негізделген модельдердің айтарлықтай тиімділігін көрсетті. PCR және Random Forest-пен салыстырғанда болжау дәлдігінің жақсарғаны төрт метрика бойынша орташа пайыздық жақсару 36,35% және 25,21% болды, кішігірім үлгіні қолданғанда (N = 10000), ал үлкен үлгіні қолданғанда 31,86% және 25,64% (N = 20000). Сонымен қатар, сорғы қондырғысының жағдайына интеллектуалды мониторинг жүргізуге мүмкіндік беретін интеллектуалды жүйе ұсынылды, бұл қызметкерлерге еңбек жүктемесін азайтуға мүмкіндік береді.

**Түйін сөздер:** мұнай-газ өнеркәсібі, сорғы қондырғысы, болжау, қалған жабдықтың қызмет ету мерзімі, машинаны оқыту, қайталанатын нейрондық желі, регуляризация.

**© О.А. Коленчуков, В.А. Файфер, В.В. Бухтояров, 2025.**
Сибирский федеральный университет, Красноярск, Россия.
E-mail: okolenchukov@sfu-kras.ru, vbukhtoyarov@sfu-kras.ru

# ПРОГНОЗИРОВАНИЕ ОСТАТОЧНОГО СРОКА СЛУЖБЫ ЭЛЕМЕНТОВ НАСОСНОГО АГРЕГАТА НА ОСНОВЕ РЕГУЛЯРИЗАЦИИ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

**Олег Александрович Коленчуков** – к.т.н., Сибирский федеральный университет, Красноярск, Россия, e-mail: okolenchukov@sfu-kras.ru, ORCID: https://orcid.org/0000-0001-8644-7302;
**Вячеслав Алексеевич Файфер** – лаборант, Сибирский федеральный университет, Красноярск, Россия, e-mail: mega.fayfertt@gmail.com, ORCID: https://orcid.org/0009-0000-8901-6241;
**Владимир Викторович Бухтояров** – к.т.н., Сибирский федеральный университет, Красноярск, Россия, e-mail: vbukhtoyarov@sfu-kras.ru, ORCID: https://orcid.org/0000-0003-4505-2851.

**Аннотация.** *Актуальность.* Нефтяная отрасль является одной из ключевых сфер деятельности горнодобывающей промышленности. Она оказывает существенное влияние на экономику государства, обеспечивая различные отрасли промышленного производства продуктами переработки нефти. Для обеспечения нормального функционирования предприятий важно организовать непрерывную транспортировку углеводородов с использованием насосных агрегатов. Данный подход позволяет определить остаточный ресурс насосного агрегата и провести его своевременное обслуживание до наступления отказа. *Цель.* Целью работы является обеспечение бесперебойной транспортировки с помощью прогностического технического обслуживания на основе технологии обработки большого объема данных и методов машинного обучения. *Методы.* В качестве математических моделей для определения остаточного ресурса насосного агрегата в данной работе использовались рекуррентные нейронные сети LSTM и GRU. *Результаты и выводы.* Эффективность применения модернизированных рекуррентных нейронный сетей была продемонстрирована при их сравнении с традиционными методами машинного обучения (PCR и Random Forest) на различных вариациях данных. Сравнительный анализ продемонстрировал значительную эффективность моделей на основе рекуррентных нейронных сетей, в особенности модели LSTM. В сравнении с PCR и Random Forest было показано улучшение точности прогнозирования. Среднее процентное улучшение по четырем метрикам составило 36,35 % и 25,21 % при использовании меньшей выборки (n = 10000), а также 31,86 % и 25,64 % при использовании большей выборки (n = 20000). Помимо этого, была предложена интеллектуальная система, позволяющая проводить интеллектуальный мониторинг состояния насосного агрегата, что позволяет снизить трудовую нагрузку на персонал.

**Ключевые слова:** нефтегазовая отрасль, насосный агрегат, прогнозирование, остаточный ресурс оборудования, машинное обучение, рекуррентная нейронная сеть, регуляризация

***Acknowledgments.*** *Some of the presented results were obtained during the research carried out under the state assignment of the Ministry of Science and Education of the Russian Federation «Development of a complex of scientific and technical solutions in the field of intensification and optimization of processes in the production of biofuel compositions, ensuring the possibility of transforming the consumed types of energy carriers in accordance with the trends of energy efficiency, reducing the carbon footprint and using fuels alternative to fossil fuels», scientific topic FSRZ-024-0005.*

**Introduction.** Today, we are experiencing the Fourth Industrial Revolution, also known as Industry 4.0, which includes technologies such as the Internet of Things, cyber-physical systems, and cognitive computing. The development of material culture through the implementation of the ideology of the Fourth Industrial Revolution will depend primarily on the efficiency of production processes (Litvinenko, 2022). In this regard, a trend has currently emerged aimed at systematically replacing hard human labor with automation, including the use of artificial intelligence (AI). This kind of transposition ultimately allows for increased efficiency of production processes, reduced operating costs, and increased accuracy of operational decision-making.

Issues related to automation have recently been increasingly discussed by governments of many countries at the level of forums, congresses and conferences. As a result of such discussions, strategic directions are formed that determine the future development of the state. On July 20, 2017, the government of the People's Republic of China published a plan for a program for the internal development of new-generation AI technologies for their subsequent implementation in the manufacturing sector, as well as in the public administration and defense sectors. On December 14, 2017, the first three-year plan to promote the development of the AI industry was published. From this date, the formation of «smart manufacturing» and «smart equipment» began. Almost two years later, this development direction was supported by two other superpowers, the United States of America (USA) and the Russian Federation (RF). On February 11, 2019, the US President signed Executive Order No. 13859, which consolidates the country's scientific, technological and economic leadership in R&D and the implementation of AI based on a coordinated strategy of the federal government (Selyanin, 2020). On October 10, 2019, the Decree of the President of the Russian Federation No. 490 also announced the National Strategy for the Development of AI for the Period up to 2030. This strategy is aimed at accelerating the development of AI on a national scale, conducting scientific research in the field of AI, and also implies improving the system of training personnel in this area of scientific knowledge.

According to research by Canadian scientists, it was found that a significant amount of money for the development and research of AI comes from two major participants – the USA and the RF. Looking at statistical information on the scale of all countries of the world, one can conclude that funding in the field of automation, including using AI, is increasing. According to the forecasts of the authors of the study, by 2030 there will be a more than twofold increase in funding in this area (Fig. 1).

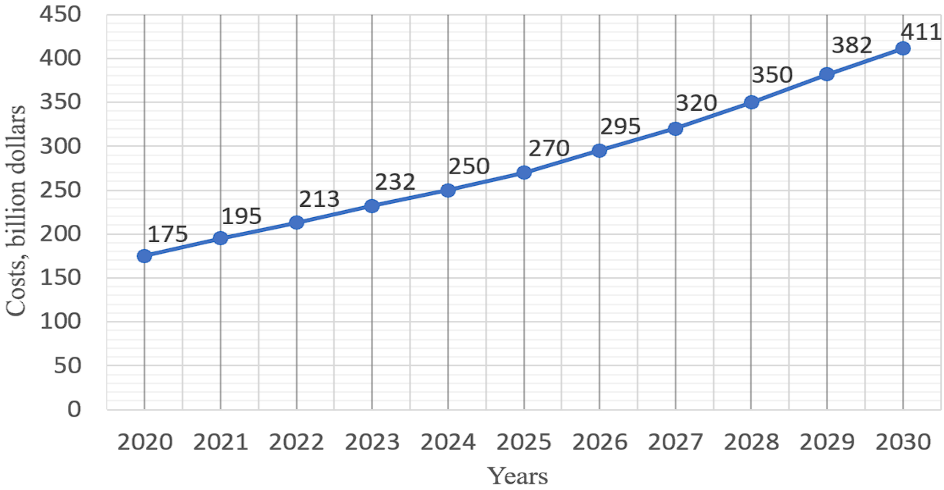


Fig.1. Forecast of changes in costs for the automation industry

The data presented in Fig. 1 show that at present, the development and research of methods and tools for automation are relevant areas that determine the development of the digital economy and economic security of the modern state.

**Materials and equipment.**

The mining industry of the RF is one of the most important industries in terms of its impact on the economy and technical development of the state. Among the groups of mining industries, the oil industry occupies a key position in the financial support of the country. Based on the data of the Ministry of Finance, the share of oil and gas revenues for 2023 is about 30.9 % of all revenues of the RF. In view of this, it is advisable to ensure the uninterrupted operation of oil production and refining equipment (Chernyavskiy, 2014). In this context, the reliability and efficiency of the equipment comes to the fore.

Pumping units are among the most important equipment that ensures the continuity of processes in the oil and gas industry. Their correct operation is directly related to the continuity of production and the safety of technological operations. Optimization of pump operation is inextricably linked to the maintenance of this equipment. The desire to increase reliability, reduce maintenance costs (MT) and prevent accidents has become an important task for the industry. Statistics show that a significant part of the enterprise budget is spent on MT, which emphasizes the relevance of introducing new approaches that allow for the optimization of costs and ensure production efficiency.

Today, there are two main strategies (methods) of MT: «post-factum maintenance after equipment failure» (hereinafter referred to as corrective maintenance) and «preventive equipment maintenance strategy» (hereinafter referred to as preventive maintenance). Touching upon each method of MT, it can be noted that they have significant drawbacks that do not allow them to be used as the «cornerstone» of modern MT. Thus, corrective maintenance has a limitation consisting in untimely response to equipment failures, which can lead to emergency situations and failures in ensuring production processes. It is also necessary to note the presence of unplanned downtime, which seriously affects the efficiency of production and the service life of equipment. Although preventive maintenance is aimed at preventing failures, it can sometimes encounter the fact that there is a premature replacement of equipment components or repair of the unit before the onset of the limit state. This, in turn, can be ineffective from the point of view of resource management. At the same time, there is a need to store a large number of spare parts, which also increases the costs associated with MT.

Over decades of fault diagnosis practice, vast engineering experience has been accumulated in identifying and analyzing machine vibration characteristics. Equipment specialists have already identified the causes of equipment failures and the mechanisms by which vibrations occur. From about 2008 to 2016, specialists have been actively developing the predictive maintenance (PM) methodology based on the signal detection and processing method by continuously monitoring and analyzing the equipment condition. PM integrates management strategies such as equipment life cycle, scheduled maintenance, inventory status, inspection, and lubrication, which saves MT costs on labor and spare parts. With the development of signal detection and processing technologies, these methods have been used to analyze and identify characteristic information from regularly operating equipment to determine the type of failure that could occur next time, as well as the time of its occurrence, and create a management system for MT based on the diagnostic results. The popularity of the method has grown, and it has significantly improved the efficiency of production and equipment.

The PM strategy allows, on the one hand, to minimize the waste of human and material resources due to excessive MT in other methods, and on the other hand, to identify potential safety threats in equipment operation caused by insufficient MT work, with special attention paid to tracking and analyzing the causes of malfunctions and reducing their frequency. The use of machine learning (ML) methods, especially artificial neural networks (ANN), allows for the accurate and prompt identification of potential malfunctions and maintenance only when necessary. In the context of cognitive computing, ANNs are an essential tool. Their ability to model complex phenomena has made them relevant for solving many problems, such as natural language processing and time series modeling. These tasks are important for many industrial processes, which makes the use of ANNs relevant in solving a number of problems. ANN-based models can also be used to analyze the operating parameters of pumping equipment, such as pressure, temperature, and vibration, as well as predict the time ranges of failures. However, PM requires a certain amount of technical knowledge and a significant amount of human and financial resources. There is a high probability of making an incorrect detection of a malfunction, missing a detection, making an erroneous assessment or making an erroneous judgment during the verification of the actual situation. Limitations in technical means, in particular the speed of transmission, analysis and use of information, also leave their mark.

During operation, the machine generates vibration signals with unique signatures that change depending on its operating condition. The technical condition of key components such as bearing assemblies plays a key role in the safety and efficiency of pumps in the oil and gas industry. The reasons why rolling bearings may not operate normally are complex and varied. During equipment operation, rolling bearings are prone to plastic deformation and peeling of the metal surface of the separator, cracks, fatigue pitting, burning, etc., due to such difficult to detect factors in the early stages as fretting corrosion, fatigue wear and material fatigue, abrasive wear, aggravated by possible severe operating conditions of the bearings or their improper installation. If the rolling bearing is damaged, the accuracy, reliability and service life of the whole machine will be directly or indirectly affected to a certain extent. According to statistics, almost 30 % of failures in rotary machines are caused by failures of rolling bearings (Efremenkov, 2024), which leads to increased vibrations, impacts, shaft bending and failure of the machine as a whole. To prevent such failures and increase the service life of pump units, it is important to use effective diagnostic methods. These methods are based on one of the key technologies of PM, which consists in determining the remaining useful life of equipment (RUL) (Han, 2021). Monitoring the operating conditions of rolling bearings in real time allows predicting the RUL of a bearing with a certain accuracy.

Existing technologies used in most cases, such as multivariate regression analysis methods and other basic ML methods, demonstrate certain advantages (Kang, 2020; Yousuf, 2022; Shaheen, 2023). However, they also have certain disadvantages. For example, when using Principal Component Regression (PCR), this method is sensitive to outliers and noise in the sample, which are very abundant in bearing vibration data. Also, if there is a high correlation between the predictors, this can lead to a problem of multicollinearity, which makes the estimate of the variable weights less reliable. These phenomena can lead to model distortion and deterioration in forecasting accuracy (Shitikov, 2017). When using the widespread Random Forest method in ML, overfitting is often observed on the training set. This phenomenon is especially often observed when using a large number of trees or deep neural trees, which also imposes some limitations when using this method in predictive diagnostics.

At present, scientific research in the field of intelligent diagnostics of centrifugal pump failures mainly focuses on the methods of ML, information coding and pattern recognition (Konyukhov, 2024). When predicting failures of rolling bearings of pump units, the classical regression problem is often solved. In order to predict RUL, it is necessary to go through two main stages sequentially:

1. Extraction of features. At this stage, the analysis of the vibration signals that arise is carried out. Depending on the area of signal origin, methods used for their processing are distinguished. Analysis in the frequency domain involves the use of the fast Fourier transform, spectral analysis methods and various methods based on the Fourier transform. During analysis in the time domain, the parameters of the root-mean-square vibration acceleration, shock pulses, excess, etc. are calculated. Time-frequency analysis is an extended version of frequency analysis. This method can use data processing methods in the form of the short-term Fourier transform, the Hilbert-Huang transform, the wavelet transform, etc..

2. RUL prediction. The features extracted in the first stage are used as input data for various ANN models. The most basic methods include the Support Vector Machines (SVM) and the k-Nearest Neighbors Algorithm (k-NN). Although these methods provide good results, they are inferior to the most accurate and productive deep ML methods. Currently, the most widely used ANNs for rolling bearing diagnostics are convolutional (CNN) and recurrent (RNN) ANNs (Nussipali, 2024).

The selection of the required algorithm and its improvement at the second stage of RUL forecasting is the most important procedure determining its accuracy and efficiency. To determine the most preferable method for forecasting RUL of centrifugal pump rolling bearings, it is necessary to determine the advantages and disadvantages of the most common models. Further actions should be aimed at improving the algorithm of the selected method, aimed at increasing the accuracy of forecasting.

Convolutional Neural Network (CNN). CNN is a typical feedforward neural network with a multi-layer network architecture. This ANN is similar to a biological neural network (nervous network) and is made to resemble the connections of neurons in the human brain (Shrestha, 2019). CNN has high adaptability and can combine multi-layer perception through structural reorganization and reduction of the weights of neuron connections (Shrestha, 2019; Alzubaidi, 2021).

The peculiarity of CNN is that neurons use a local method of connecting the network structure, namely a layer with incomplete connection and weight sharing. The CNN architecture allows to reduce the number of assigned weights, reduce the complexity of the neural network model, ensure the convenience of network optimization and reduce the risk of overfitting. CNN is a neural network specifically used to process data that can be represented as a grid structure, for example, a graph used to encode information associated with information from several sources. CNNs have achieved excellent results in various fields, such as image pattern recognition and classification, object detection and recognition, etc. (Lindsay, 2021).

Recurrent Neural Network (RNN). RNN is a neural network model in which hidden layers are connected to each other and also connected to the previous output, as shown in Fig. 2. If RNN is deployed in time (creating several copies of the recurrent network), it can be represented as a network with a feedforward signal distribution and shared weights (Sherstinsky, 2020). Using the Backpropagation Through Time method, information about a neuron at a previous point in time will be connected to information about a neuron at the next point in time, thus forming a feedback network structure (Lillicrap, 2019). The backpropagation algorithm must constantly use the chain rule (the rules for differentiating a complex function). The activation value of the target function, depending on the hidden layer, can not only be transmitted to the output layer of the network, but also affect the next hidden layer.

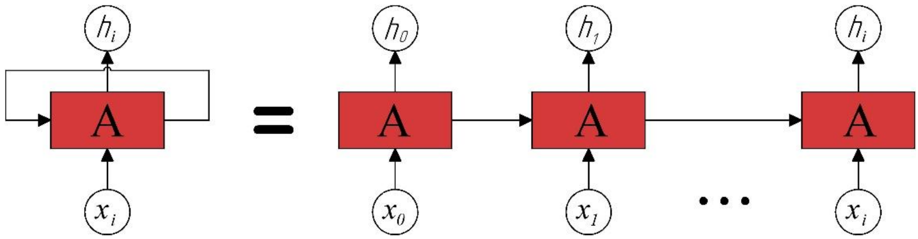


Fig.2. RNN Model Structure

Mathematically, the classical RNN network (Vanilla RNN) is expressed as follows in terms of the hidden layer update and training update expressions (Mehrish, 2023):

$$\mathbf{h}_t = f(\mathbf{W}_{hx}\mathbf{x}_t + \mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{b}_h) \quad (1)$$

$$\mathbf{y}_t = f(\mathbf{W}_{yh}\mathbf{h}_t + \mathbf{b}_y) \quad (2)$$

where f(·) is the activation function, which is usually a nonlinear hyperbolic tangent or sigmoid function, or a continuous piecewise linear function; $\mathbf{W}_{hx}$, $\mathbf{W}_y$ are the learnable weight matrices from the input layer and output layer to the candidate state, respectively; $\mathbf{W}_{hh}$ is the cyclic connection weight matrix; **b** are the shift parameters.

The most significant advantages and disadvantages of the above-described ANN models are given in Table 1.

Table 1. Advantages and Disadvantages of Convolutional and Recurrent Neural Networks

| A type of neural network | Advantages | Disadvantages |
|---|---|---|
| CNN | 1. Suitable for processing input data collected from various sources.<br>2. High efficiency and noise reduction when processing large amounts of data.<br>3. Adaptive feature extraction and rational weight distribution.<br>4. Using spatial and temporal downsampling to reduce network parameters and the likelihood of overfitting. | 1. To ensure high accuracy of the model, it is necessary to use a multi-level architecture, which increases the complexity of the network and affects its performance.<br>2. The need to use a large volume of data marked by features.<br>3. There is a possibility of inadequate behavior (distortion of results) of the neural network in the case of successful training, which can negatively affect the prediction of malfunctions of expensive equipment. |

| | | |
|---|---|---|
| RNN | 1. High efficiency in processing data sequences of different lengths.<br>2. Possibility of dynamic monitoring and forecasting of industrial processes.<br>3. Good analytical ability of the model in analyzing data obtained from dynamic systems.<br>4. Wide applicability and higher accuracy in analyzing and forecasting time series. | 1. The probability of encountering a vanishing or exploding gradient problem during training (small or large values), which can distort predictions or hinder training.<br>2. Low efficiency when processing large amounts of data simultaneously (parallel data).<br>3. The model has limited memory and when it is full, it begins to forget information about previous inputs. |

The pump is a dynamic system, the information about the bearing units of which is most often presented in the form of several time series of received and successively accumulated data of process parameters, such as vibration speed, temperature of the bearing unit, etc. An increase in these parameters over time, as a rule, leads to a gradual failure of the rolling bearing and the pump itself. Analyzing these features, we can conclude that it is advisable to use the RNN model as a method for determining RUL. Although it is inferior to the CNN model in the size of the initial sample, it should nevertheless show better results on a sample of average size, which is justified by the specifics of the model itself.

Considering the ambiguity of existing approaches to the ML in assessing the technical condition of process equipment, including pump units, it is important to evaluate the possibility of using new methods of predictive diagnostics. In view of this, the purpose of this work is to expand the arsenal of ML methods used for predictive diagnostics using neural networks, which make it possible to improve the accuracy of the analysis of unstructured data on bearing unit malfunctions of centrifugal pumps.

**Methods.** Compared with other methods, particularly with the CNN method, the gradient of RNN is distributed not only over the spatial structure but also over the time channel. When the volume of calculations increases, the gradient easily disappears, which usually leads to a failure in model training. Each neuron in RNN is associated with information about the state of the previous moment, but as the time interval increases, the RNN network may experience Vanishing Gradient or Exploding Gradients problems (Muñoz-Zavala, 2024). These phenomena lead to the inability to update the ANN parameters, which affects the training speed and quality of the model. One of the solutions to this problem is to modernize the RNN architecture. As a result, two variants of ANN were created: LSTM (Long Short-Term Memory) and RNN–GRU (Gated Recurrent Unit). The LSTM network was proposed first, but the LSTM model is complex in shape and requires more time to train. To overcome these shortcomings, the GRU network was developed, the architecture of which is simplified, which allows to significantly increase the learning speed.

The LSTM architecture includes the concept of a Memory Cell, also called a

Cell State, and additional components called Gates. The Memory Cell replaces the hidden neurons used in traditional RNN systems and creates a new hidden layer. As shown in Fig. 3, adding a forget gate $\mathbf{f}_t$ (memory clear), input gate $\mathbf{i}_t$ (memory update), and output gate $\mathbf{o}_t$ (output) to the Memory Cell alleviates the gradient decay or exploding problem.

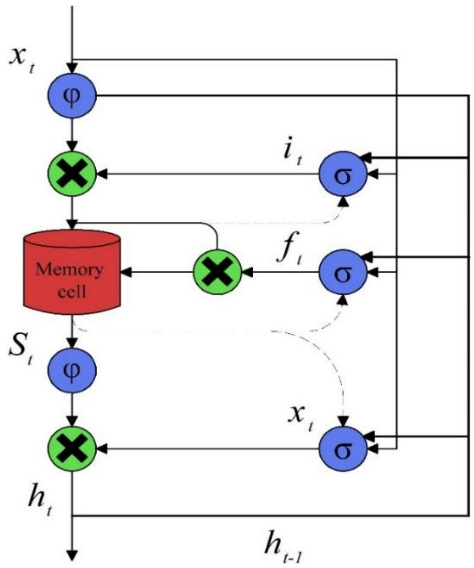


Fig.3. Architecture of recurrent LSTM network

The LSTM neural network controls the memory cell via a forget gate. The input to the forget gate consists of the output of the memory cell $h_{t-1}$ at the previous moment and the input $x_t$ at the current moment. The expression for determining the parameter of the forget gate is as follows (Kong, 2019):

$$\mathbf{f}_t = \sigma(\mathbf{W}_{fx}\mathbf{x}_t + \mathbf{W}_{fh}\mathbf{h}_{t-1} + \mathbf{b}_f) \qquad (3)$$

Next, the information in the memory cell is updated via the input gate, and a possible update of the hidden state vector $\mathbf{g}_t$ occurs. These expressions look like this:

$$\mathbf{i}_t = \sigma(\mathbf{W}_{ix}\mathbf{x}_t + \mathbf{W}_{ih}\mathbf{h}_{t-1} + \mathbf{b}_i) \qquad (4)$$

$$\mathbf{g}_t = \phi(\mathbf{W}_{gx}\mathbf{x}_t + \mathbf{W}_{gh}\mathbf{h}_{t-1} + \mathbf{b}_g) \qquad (5)$$

$$\mathbf{s}_t = \mathbf{g}_t \odot \mathbf{i}_t + \mathbf{s}_{t-1} \odot \mathbf{f}_t \qquad (6)$$

The output of the updated state is controlled through the output gate, and the expressions for its determination can be represented as follows:

$$\mathbf{o}_t = \sigma(\mathbf{W}_{ox}\mathbf{x}_t + \mathbf{W}_{oh}\mathbf{h}_{t-1} + \mathbf{b}_o) \qquad (7)$$

$$\mathbf{h}_t = \phi(\mathbf{s}_t) \odot \mathbf{o}_t \qquad (8)$$

In the given expressions (3)–(8): $\mathbf{W}_{fx}$, $\mathbf{W}_{ix}$, $\mathbf{W}_{gx}$, $\mathbf{W}_{ox}$ are the matrices of weights at the input of the neural network for the forget gate, input gate, hidden state update, and output gate, respectively; $\mathbf{W}_{fh}$, $\mathbf{W}_{ih}$, $\mathbf{W}_{gh}$, $\mathbf{W}_{oh}$ are the matrices representing the weights at the output of the neural network corresponding to the inputs; (•) is the Hadamard product; σ(•) is the sigmoid activation function; ϕ(•) is the hyperbolic tangent activation function.

The GRU neural network is a simplified version of LSTM and combines the input gate and the forget gate into the update gate $\mathbf{z}_t$. In addition, the GRU structure combines the memory cell and the hidden state into a single whole, thereby optimizing the architecture, transforming it from three gates to a two-gate one. The GRU architecture is shown in Fig. 4.

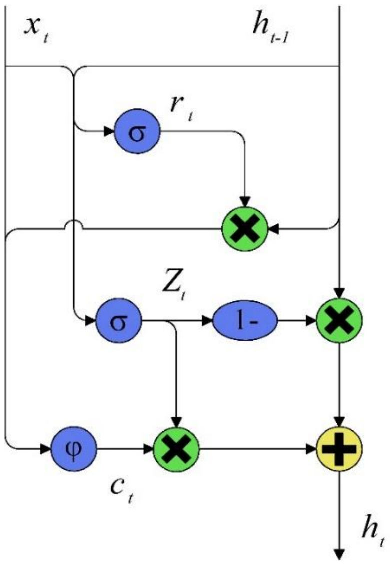


Fig.4. Architecture of recurrent GRU network

The update gate is determined based on the state value of the previous moment and the input state value of the current moment. When the update gate value is large (close to 1), the ANN decides to keep most of the previous hidden state, which is useful in situations where past information is important for predicting or understanding the next step of the sequence, i.e., information about the previous moment is preserved. If the update gate value is close to 0, the ANN decides to update its hidden state with new information, paying less attention to what was learned previously, which is useful when the new input contains enough information for the current prediction and the previous history is less important. Mathematically, this is as follows:

$$\mathbf{z}_t = \sigma(\mathbf{W}_{xz}\mathbf{x}_t + \mathbf{W}_{hz}\mathbf{h}_{t-1} + \mathbf{b}_z) \quad (9)$$

The reset gate $\mathbf{r}_t$ is also determined by the state value of the previous instant and the input value of the current instant. When the value of the reset gate is large enough, the previous hidden state is considered important and more information about the previous instant is stored. Mathematically, this looks like this:

$$\mathbf{r}_t = \sigma(\mathbf{W}_{xr}\mathbf{x}_t + \mathbf{W}_{hr}\mathbf{h}_{t-1} + \mathbf{b}_r) \quad (10)$$

Finally, the update of the hidden state vector $\mathbf{h}_t$ of the memory cell can be expressed using the following expressions:

$$\mathbf{c}_t = \phi[(\mathbf{W}_{xc}\mathbf{x}_t) + \mathbf{W}_{hc}(\mathbf{r}_t \odot \mathbf{h}_{t-1})] \quad (11)$$

$$\mathbf{h}_t = (1 - \mathbf{z}_t) \odot \mathbf{h}_{t-1} + \mathbf{z}_t \odot \mathbf{c}_t \quad (12)$$

In the above expressions (9)–(12): $\mathbf{c}_t$ is a possible update of the hidden state vector; $\mathbf{W}_{xr}$, $\mathbf{W}_{xx}$, $\mathbf{W}_{xc}$ are the matrices of weights at the input of the neural network from the input layer to the reset gate, update gate and hidden states, respectively; $\mathbf{W}_{hr}$, $\mathbf{W}_{hx}$, $\mathbf{W}_{hc}$ are the matrices representing the weights of the network output corresponding to the inputs.

In the experimental studies, we will use these modernized ANNs (LSTM and GRU) as the main methods for determining the RUL of pumps. At the same time, it is also necessary to evaluate the effectiveness of the used deep ML methods with classical methods such as PCR and Random Forest.

The initial data set was statistical information on the operating parameters of the 2NK-E200/120-210 booster pump unit. The data set contained 13 features distributed across the corresponding columns and 526,310 rows with an interval of 1 minute. A fragment of the data set is presented in Table 2.

Table 2. Fragment of the dataset for neural network modeling

| № | 41 | 42 | 43 | 44 | 45 | 46 | 47 |
|---|---|---|---|---|---|---|---|
| Time | 00:39 | 00:40 | 00:41 | 00:42 | 00:43 | 00:44 | 00:45 |
| Frequency | 33.7667 | 33.009 | 32.1974 | 31.93 | 31.635 | 31.3386 | 31.5701 |
| Filter difference | 1.0739 | 0.8899 | 0.53 | 0.505 | 0.3875 | 0.4366 | 0.4499 |
| Pin | 57.7399 | 56.9186 | 58.98 | 62.215 | 61.8021 | 64.4225 | 68.2342 |
| Tin | 29.2745 | 29.25 | 29.2255 | 29.201 | 29.1752 | 29.1493 | 29.1235 |
| Pout | 0.7452 | 0.7445 | 0.7438 | 0.7431 | 0.7424 | 0.7416 | 0.7409 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| V1M | 0.5749 | 0.5337 | 0.4887 | 0.4663 | 0.4442 | 0.4362 | 0.5183 |
| V2M | 1.1134 | 1.2951 | 1.0199 | 0.8999 | 0.9054 | 0.8704 | 0.9687 |
| V1P | 0.6499 | 0.5899 | 0.6833 | 0.725 | 0.8949 | 0.7108 | 0.8 |
| V2P | 0.495 | 0.4533 | 0.4975 | 0.49 | 0.5749 | 0.5249 | 0.5649 |
| T1M | 50 | 50 | 50 | 50 | 50 | 49.9 | 49.9 |
| T2M | 38.7999 | 38.7666 | 38.7 | 38.7 | 38.5999 | 38.5 | 38.4 |
| Toil | 32.3441 | 32.3305 | 32.3168 | 32.3031 | 32.2894 | 32.2757 | 32.262 |

Where: Time is the date and time of taking readings from the sensor; Frequency is the speed of the pump rotor; Filter difference is the value of the pressure drop across the filter of the pump unit; Pin is the pressure of the medium at the pump inlet; Tin is the temperature of the medium at the pump inlet; Pout is the pressure of the medium at the pump outlet; V1M is the vibration velocity of the front bearing of the pump motor; V2M is the vibration velocity of the rear bearing of the pump motor; V1P is the vibration velocity of the front bearing of the pump; V2P is the vibration velocity of the rear bearing of the pump; T1M is the temperature of the front bearing of the pump motor; T2M is the temperature of the rear bearing of the pump motor; Toil is the oil temperature in the bearing housing of the pump.

All ML models were tested on two datasets. In the first case, the data sample size n was 10,000 minute-by-minute values, and in the second case, the sample size n was 20,000 minute-by-minute values. The designated datasets were divided into two sets: training and testing. The training set was used to train and fit the models and contained 2,500 and 5,000 checkpoints from the dataset. The testing set was used to evaluate the accuracy and performance of the models and contained 7,500 and 15,000 checkpoints, respectively. RUL prediction was made for 120 checkpoints following the training set. After modeling, the resulting prediction was compared with the testing set to evaluate the accuracy.

For the modeling procedure, we use the Python language and the pandas, numpy, scikit-learn, statsmodels, tensorflow/keras, linearmodels, matplotlib, scipy, sklearn. metrics, MinMaxScaler libraries from sklearn.preprocessin.

In order to evaluate and compare the effectiveness of the considered neural network models and compare them with classical ML methods, the following statistical metrics used for regression problems were determined:

– Mean Square Error (MSE): measures the standard deviation between the predicted values and the actual values, serves to demonstrate the forecasting accuracy of the methods;

– R-squared coefficient ($R^2$): shows what percentage of the variance of the target variable is explained by the model. A value close to 1 indicates an accurate model;

– Mean Absolute Percentage Error (MAPE): measures the average percentage deviation of predictions from actual values;

– Correlation Coefficient: measures the strength and direction of the relationship between predicted and actual values. The value ranges from -1 to 1, where 1 is a perfect positive correlation.

Additionally, we modernize the algorithms of the LSTM and GRU neural networks by introducing Dropout regularization into their composition. The Dropout method is appropriate to use to eliminate the phenomenon of overtraining. The basic scheme for using the Dropout method is shown in Fig. 5.

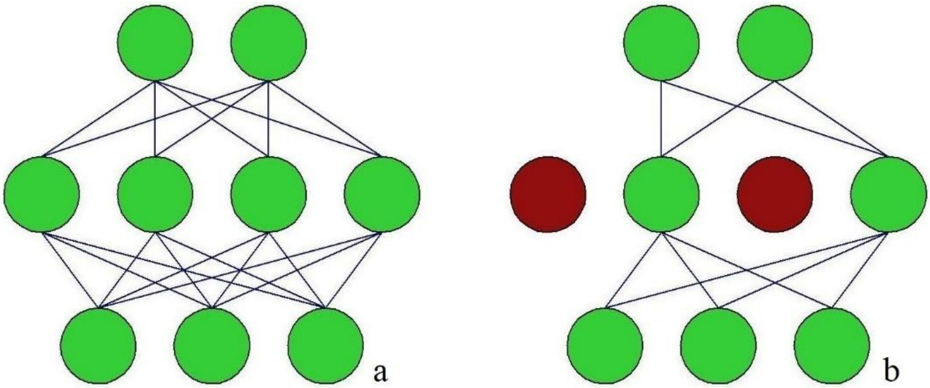


Fig.5. Graphical representation of the structure of neural networks: a – standard neural network; b – neural network after applying Dropout

The method shown in Fig. 5 can be characterized as a method of randomly excluding network nodes with a certain probability. That is, during the training process, some neuron nodes are selected from the neural network layer by probability *p*, obeying the Bernoulli distribution, and then temporarily discarded. During the next training, random selection is again made by this probability and discarded. Discard here means that the weights of the discarded neural nodes do not participate in the neural network calculations, including direct calculation and backpropagation, which also means that these weights are not updated.

Based on the results of experimental data using a sample of statistical information on the parameters of the booster pump (Table 2), it was found that without using Dropout, the ANN is overtrained and the loss function increases (Fig. 6, a). After using Dropout-type regularization (Fig. 6, b), the ANN showed a decrease in the loss function and the achievement of optimality (reaching a plateau of the learning and validation curves on the graph) in a small number of training epochs, which affected the performance of both the LSTM and GRU neural networks.

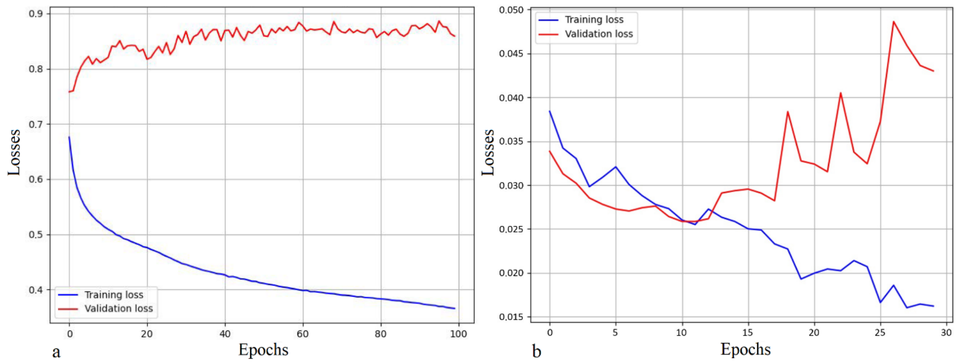


Fig.6. Losses at the stages of training and testing the model: a – without using the Dropout method; b – with using the Dropout method

**Results.** The experimental data results for the first sample (n = 10,000) are presented in Table 3. Analyzing the obtained data, we can conclude that for a smaller sample, the most preferable is to use a neural network of the LSTM type. Although the LSTM and GRU neural networks demonstrate similar good results for all metrics, LSTM has lower values of the mean square error and average absolute percentage deviation (MSE = 0.2906, MAPE = 15.63 %), which indicates more accurate results in forecasting. PCR and Random Forest have higher MSE and MAPE values compared to recurrent networks of the LSTM and GRU types. They also demonstrate significantly lower values of the correlation coefficient. However, it should not be forgotten that these methods are simpler to implement (not a complex algorithm structure) compared to LSTM and GRU neural networks.

Table 3. Results of the quality assessment of the methods used, n = 10,000

| A variation of the method | MSE | R-squared ($R^2$) | MAPE, % | Correlation coefficient |
|---|---|---|---|---|
| LSTM | 0.2906 | 0.9993 | 15.63 | 0.9997 |
| GRU | 0.2931 | 0.9994 | 18.53 | 0.9997 |
| PCR | 0.4473 | 0.6879 | 28.34 | 0.8313 |
| Random Forest | 0.3051 | 0.7117 | 28.97 | 0.912 |

The LSTM method, like the GRU method, demonstrates significant results with a high value of the determination coefficient ($R^2$ = 0.9993 for LTSM and $R^2$ = 0.9994 for GRU). High values of the determination coefficient tell us that the used recurrent neural network models effectively establish dependencies when working with time series (sequential data), which makes them a powerful tool for forecasting. At the same time, the method has the lowest $R^2$ indicator, and the Random Forest method is in second place. Lower $R^2$ indicators for these methods indicate a weaker relationship between the variables of the regression model (68.79 % of the variance of the resulting feature is explained for PCR and 71.17 % for GRU).

The results of applying the four presented forecasting methods for the second sample (n = 20,000) are presented in Table 4. In the case of using a larger sample, the LSTM and GRU recurrent neural networks also demonstrate their advantage in the form of higher forecasting accuracy compared to a smaller sample, as well as compared to the classical PCR and Random Forest ML methods. When using a larger sample, there is a decrease in the standard deviation and average absolute percentage deviation for both recurrent neural networks (MSE = 0.2431 for LSTM, MSE = 0.2719 for GRU and MAPE = 5.5 for LSTM, MAPE = 6.30 for GRU). The improvement of the obtained results is associated with a more extensive data set, which provides an opportunity for more accurate training of neural network models. For the convenience of assessing the correlation coefficients, we will use the Chaddock scale. According to the analysis, it can be concluded that the LSTM and GRU methods, as well as the Random Forest method, have a very high strength of connection between variables. In addition, a less pronounced, but nevertheless very high strength of connection is observed in the PCR method.

Table 4. Results of the quality assessment of the methods used, n = 20,000

| A variation of the method | MSE | R-squared ($R^2$) | MAPE, % | Correlation coefficient |
|---|---|---|---|---|
| LSTM | 0.2431 | 0.9995 | 5.5 | 0.9997 |
| GRU | 0.2719 | 0.9995 | 6.30 | 0.9996 |
| PCR | 0.3745 | 0.8483 | 15.58 | 0.9103 |
| Random Forest | 0.3489 | 0.8447 | 11.51 | 0.983 |

In addition to the decrease in MSE and MAPE, there is also an increase in $R^2$ (0.9995 for LSTM and GRU at n = 20,000, 0.9993 and 0.9994 for LSTM and GRU at n = 10,000). The increase in $R^2$ and better MSE and MAPE for the LSTM neural network are achieved due to its ability to better adapt to complex dependencies in the data. GRU also shows good indicators, which allows using both presented methods for forecasting.

The PCR method demonstrates acceptable results with the $R^2$ coefficient value of 0.8483 when using a larger sample (n = 20,000), but still yields to recurrent networks. This is naturally due to the fact that PCR takes into account the time sequence in the data worse. The Random Forest method shows improved results of the $R^2$ coefficient of 0.8447 compared to the previous data set (n = 10,000). This is due to the fact that this method is less prone to overfitting due to its ability to generalize data. The strength of the relationship between the variables remains very high for the LSTM and GRU methods, as indicated by the corresponding correlation coefficients (Table 4). An increase in the correlation coefficient for Random Forest from 0.912 to 0.983 indicates an increase in the strength of the relationship, which is associated with the adaptation of this method to working with high-dimensional data. The PCR method also shows an increase in the correlation coefficient from 0.8313 to 0.9103. Changes in these indicators indicate that using a larger sample size has a positive effect on the quality of forecasting.

Thus, the application of these methods in practice can be realized by introducing neural network algorithms based on recurrent neural networks into the technological process of pumping the working medium by a pumping unit, forming a single intelligent system. At the same time, the dispatcher of the operational-dispatch control does not need to constantly monitor the pump performance readings and track certain patterns preceding its failure. The ability of recurrent neural networks of the LSTM and GRU type to remember the behavior of the system in various situations over a long period of time is capable of predicting various scenarios of events. At the same time, the architecture of such a system can be presented as a graph divided into certain levels of danger (Fig. 7). The advantages of such a system are the ability to conduct intelligent monitoring of the equipment condition without the constant involvement of the dispatcher in this process, which reduces the workload on personnel, as well as the ability to make a more accurate decision to carry out a repair procedure, which allows reducing production costs during equipment downtime.

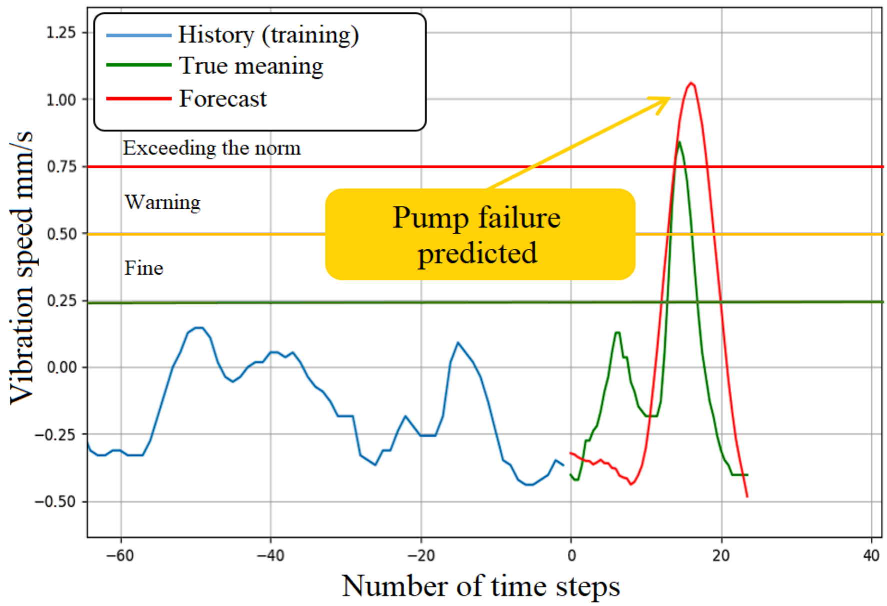


Fig.7. An example of a graphical representation of a predictive system based on recurrent neural networks

**Discussion of results.**

The performance evaluation of PM was conducted using four ML methods. Two of them are recurrent neural networks (LSTM and GRU), and the other two are classical ML methods (PCR and Random Forest). These methods were tested on samples of different sizes (n = 10,000, n = 20,000). The results of the study confirmed the importance of a large sample size, which improves the performance and stability of all methods. As a result of increasing the sample size in the second experiment (n = 20,000), the average forecasting accuracy by MSE of the LSTM, GRU, and PCR methods improved by about 13.72 %. The only exception from the whole variety of methods is Random Forest. An increase in the sample size led to an increase in the MSE indicator from 0.3051 to 0.3489.

The results of the first experiment with a smaller sample (n = 10,000) showed that LSTM and GRU demonstrate excellent ability to take into account the nature of data in conditions of its scarcity. The average percentage improvement of LSTM in all four metrics relative to PCR was 36.35 %. When using the GRU method, this value is 33.66 %. When comparing LSTM and GRU with the Random Forest method, the average percentage improvement is less significant and is 25.21 % and 22.5 %, respectively. In the case of using a larger sample (n = 20,000), recurrent neural networks still show more accurate forecasting results. As in the previous experiment, the improvement in the quality indicators of the LSTM model stands out the most. The average percentage improvement in all four metrics relative to PCR was 31.86 %, relative to Random Forest – 25.64 %. For GRU, the same values were 28.65 % and 21.84 %, respectively. However, it should be noted that despite the good results of the PCR method, it may not provide the required forecast accuracy, especially when working with large data sets.

It was also shown that the practical implementation of predictive models in production is a powerful tool for improving the efficiency of maintenance. The implementation of predictive models for the implementation of the maintenance approach based on the actual condition can be achieved by introducing their operating algorithms into existing software. This creates a single intelligent system that can not only predict the time of residual operation of equipment, but also show the level of danger, which, in turn, allows the dispatcher to make more accurate decisions on further shutdown and repair of equipment.

**Conclusion.**

The concepts of maintenance of pumping units are considered. It is noted that the corrective type of maintenance is not an optimal solution, since it is carried out after the equipment failure, which entails the shutdown of technological processes, unscheduled financial costs and increases the risk of emergency situations. In the context of preventive maintenance, the resource of pump elements is not taken into account, which prevents an adequate assessment of their current actual condition. At the same time, there is a high probability of equipment failure in the period between scheduled maintenance. Maintenance in accordance with the actual condition with continuous or periodic monitoring of technical condition parameters is a predictive approach to maintenance based on determining the remaining resource of pump elements using ML methods used in predictive analytics. This type of maintenance allows predicting potential failures and equipment failures, allowing for urgent measures to be taken to repair it, which significantly reduces downtime and economic losses. This approach is especially critical in the conditions of highly loaded oil and gas industries, where even short-term downtime can lead to significant losses.

It is proposed to perform predictive diagnostics of the condition of pumping units using ML methods. The effectiveness of using these deep learning methods was studied, as well as their comparison with classical ML methods on real data obtained during the operation of a 2NK-E200/120-210 pump. As a result of the studies, it was found that recurrent networks of the LSTM and GRU types are the most suitable methods for the tasks of predictive diagnostics of pumping units. They showed high accuracy and the ability to take into account the sequential nature of the data. Additional implementation of the Dropout method into neural network algorithms showed an improvement in the quality indicators of the methods used. Random Forest is also of interest, especially if it is important to understand the impact of each feature on the forecast. PCR may be less preferable due to its limited ability to take into account the sequence in the data.

**References**

Chernyavskiy, S.V., Zolotarev, N.A. (2014) Problems of the development of the oil and gas sector of the Russian Federation's economy, Vestnik of the University, 2:91–93.

Efremenkov, E.A., Efremenkova, S.K., Chavrov, E.S. (2024). An Analysis of Power Friction Losses in Gear Engagement with Intermediate Rolling Elements and a Free Cage, Mathematics, 12(6):873. DOI 10.3390/math12060873

Hamasha, M.M., Bani-Irshid, A.H., Al Mashaqbeh, S., Shwaheen, G. (2023) Strategical selection of maintenance type under different conditions, Scientific Reports, 13(1):15560. DOI: 10.1038/s41598-023-42751-5.

Han, X., Wang, Z., Xie, M. (2021) Remaining useful life prediction and predictive maintenance strategies for multi-state manufacturing systems considering functional dependence, Reliability Engineering and System Safety, 210:107560. DOI: 10.1016/j.ress.2021.107560.

He, Y., Zang, C.-Z., Zeng, P. (2023) Few-shot working condition recognition of a sucker-rod pumping system based on a 4-dimensional time-frequency signature and meta-learning convolutional shrinkage neural network, Petroleum Science, 20(2) :1142–1154.

Kang, H., Zhao, H. (2020) Description and application research of multiple regression model optimization algorithm based on data set denoising, Journal of Physics: Conference Series, 1631(1):012063. DOI: 10.1088/1742-6596/1631/1/012063.

Kong, W., Dong, Z.Y., Jia Y. (2019) Short-term residential load forecasting based on LSTM recurrent neural network, IEEE Transactions on Smart Grid, 10(1):841–851. DOI: 10.1109/TSG.2017.2753802.

Konyukhov, V.Yu., Oparina, T.A., Matasova, I.Y., Modina, M.A. (2024) Ecologization of underground coal mining by means of ash use in backfill preparation, Mining Informational and Analytical Bulletin, (10):123–135. DOI: 10.25018/0236_1493_2024_10_0_123

Lillicrap, T.P., Santoro, A. (2019) Backpropagation through time and the brain, Current Opinion in Neurobiology, 55:82–89. DOI: 10.1016/j.conb.2019.01.011.

Lindsay, G.W. (2021) Convolutional neural networks as a model of the visual system: past, present, and future, Journal of Cognitive Neuroscience, 33:2017–2031. DOI: 10.1162/jocn_a_01544.

Litvinenko, V.S., Bowbrick, I., Naumov, I.A., Zaitseva, Z. (2022) Global guidelines and requirements for professional competencies of natural resource extraction engineers: Implications for ESG principles and sustainable development goals. Journal of Cleaner Production, 338, p. 130530. DOI: 10.1016/j.jclepro.2022.130530.

Mehrish, A., Majumder, N., Bharadwaj, R. (2023) A review of deep learning techniques for speech processing, Information Fusion, 99:101869. DOI: 10.1016/j.inffus.2023.101869.

Muñoz-Zavala, A.E., Macías-Díaz, J.E., Alba-Cuéllar, D., Guerrero-Díaz-de-León, J.A. (2024) A literature review on some trends in artificial neural networks for modeling and simulation with time series, Algorithms, 17(2):76. DOI: 10.3390/a17020076.

Nussipali, R., Romanova, V.V., Kononenko, R.V. (2024) Combined Power Generating Complex and Energy Storage System, Electricity, 5(4):31–946. DOI: 10.3390/electricity5040047

Selyanin, Y.V. (2020) The state policy of the USA in the field of artificial intelligence: goals, tasks, prospects for implementation, Problems of National Strategy, 4:140–163.

Shaheen, B., Kocsis, A., Nemeth, I. (2023) Data-driven failure prediction and RUL estimation of mechanical components using accumulative artificial neural networks, Engineering Applications of Artificial Intelligence, 119:105749. DOI: 10.1016/j.engappai.2022.105749.

Sherstinsky, A. (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena, 404:132306. DOI: 10.1016/j.physd.2019.132306.

Shitikov, V.K., Mastitskiy, S.E. Classification, regression and other Data Mining algorithms using R. Tolyatti, London, p. 351.

Shrestha, A., Mahmood, A. (2019) Review of deep learning algorithms and architectures, IEEE Access, 7:53040–53065. DOI: 10.1109/ACCESS.2019.2912200.

Yousuf, S., Khan, S., Khursheed, S.-S., Hutchinson, P., Lieven, N.A.J. (2022) Remaining useful life (RUL) regression using Long–Short Term Memory (LSTM) networks, Microelectronics Reliability, 139:114772. DOI: 10.1016/j.microrel.2022.114772.

## CONTENTS

**Publication Ethics and Publication Malpractice in the journals of the National Academy of Sciences of the Republic of Kazakhstan**

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:
**www:nauka-nanrk.kz**
**http://www.geolog-technical.kz/index.php/en/**
**ISSN 2518-170X (Online),**
**ISSN 2224-5278 (Print)**

Директор отдела издания научных журналов НАН РК *А. Ботанқызы*
Редакторы: *Д.С. Аленов, Ж.Ш.Әден*
Верстка на компьютере *Г.Д.Жадыранова*

Подписано в печать 15.02.2025.
Формат 70x90 $^{1}/_{16}$. Бумага офсетная. Печать – ризограф.
14,5 п.л. Тираж 300. Заказ 1.

---

*РОО «Национальная академия наук РК»*
*050010, Алматы, ул. Шевченко, 28, т. 272-13-19*