ANNOTATION

PhD thesis on specialty 6D070400 - «Computer Systems and Software Engineering» on the topic «Development software and hardware tools with machine learning for monitoring technical systems»

Relevance of the research topic. The idea of the work is to create an artificial intelligence-based system to reduce the risk of accidents or predict a before failure state in technical devices, in particular, in a biogas plant. The paper considers the creation of an integrated approach based on machine learning algorithms for technical devices.

Predicting the accident rate and stable operation of a technical complex is an important task in the problem of maintaining the high-quality operation of various types of equipment. Solving this problem is logically cost-effective compared to the problem of complete replacement, deep repair of the technical complex or equipment downtime. Before accident detection of the technical complex is based on the prediction of the accident rate of the main components of the device, which have a huge impact on the performance of the device. Modern research has shown that monitoring the characteristics of the components of a technical device is still the most important task for making informed decisions in the field of operation and the technical complex.

The problem of security for different types of technical devices is considered in several modern works. Safety and accident studies are studied in the work of Moreno et al., the object of the study is accidents at biogas plants. The authors highlight the following points: the protective measures of biogas projects are insufficient, monitoring methods are relatively backward, real-time data recording and statistical analysis cannot be performed, biogas leaks and explosions occur from time to time. The study shows a higher increase in accidents at biogas plants than the increase in energy production at conventional plants. The assessment of the risk of accidents at biogas plants is described in the different scientific papers, where the risks have seriously affected the progress of biogas projects. Safety issues in the use of biogas are one of the most important topics in the biogas industry. To ensure the safety of workers and the proper management of biogas projects, a monitoring system is needed in real time and for a long time under various conditions. An important component for this purpose is a large amount of data on the parameters.

The method for constructing a model with hidden variables is proposed that describes the generation of a time series based on differential equations. The authors propose a new neural network architecture based on ordinary differential equations. The main idea in the dissertation work is the application of a new architecture of neural networks for practical problems of predicting accidents (the problem of extrapolating time series) and classification (classifying accidents based on historical data). The use of the new architecture has a number of advantages over existing architectures (recurrent neural network) RNN, such as:

- Efficiency. In training neural differential equations, it is not necessary to calculate the gradient through all the operations of the numerical method, and there is also no need to store intermediate results.
- Flexibility in working hours. Compared to solving differential equations, where the grid spacing determines the accuracy of the model and influences the computation time, in neural differential equations we can explicitly control the balance between numerical accuracy and computational cost.
- Number of parameters. Compared to ResNet, the advantage of the proposed architecture is the smaller number of parameters, since no new parameters are required for each layer.

The purpose of the work is to develop an intelligent system for analyzing and evaluating the accident rate of technical systems, built on the basis of modern artificial intelligence technologies.
To achieve this goal, it was necessary to solve the following tasks:

1. Develop a comprehensive platform to improve the efficiency of a biogas plant based on neural differential equations to improve equipment fault tolerance.

2. Build the architecture of neural differential equations for calculating the risk assessment of the accident rate of a technical device using an ordinary differential equation for extrapolating time series and estimating density.
3. Create an automated complex for collecting and processing data from a biogas plant.
4. Determine the criteria for the "optimal" and before emergency mode of operation and maintenance of facilities, minimizing the role of the "human factor".
5. Develop a system for monitoring and "quick" notification of emergency (emergency) situations in the operation of the facility.

The object of the research is the readings of sensors (time series), accident events or deviations in the operation of a technical device, and methods for analyzing time series.

The subject of the research is a technical device, in particular a biogas plant.

Scientific novelty. Scientific novelty consists in developing and obtaining the following conclusions:

1. A mathematical model based on neural differential equations has been developed to predict future values.
2. The stability of neural differential equations is proved.
3. An integrated system has been built for monitoring the performance of units and assemblies of a biogas plant and identifying a per-emergency state of technical devices.

Theoretical and practical significance of the work. The theoretical significance of this work lies in the mathematical proof of the use of the mathematical apparatus of ODEs to create a new architecture of neuronal networks, as well as in the derivation of the property of the sufficiency of the existence of a solution for neural network ODEs, i.e. absence of intersection of integral curves. The practical significance of the work lies in the approbation of neural network ODEs for the problem of predicting accidents at a biogas plant.

Scope and structure of work. includes an introduction, 4 sections, a list of references and appendices.

In the introduction, the substantiation of the relevance of the chosen topic of the dissertation work is given. The purpose, object, subject and tasks of the research work are formulated. The results of the conducted studies are described, their scientific novelty and practical significance are shown.

In the first section of the dissertation work, the problem of the per-emergency state of technical devices is considered. The description of the object of study, sensors for collecting information about the biogas plant is presented. The collected data was analyzed and visualized on interactive charts. A description is given of mathematical modeling of the state of objects of a biogas plant, in particular a flat solar collector, which is a device for heating a biogas plant.

The second section contains a description of a new family of deep neural network models based on differential equations (ODEs) or neural network ODEs. At the beginning of the section, it is formally defined what a time series is and what it means to draw conclusions from time series. An overview of classical time series forecasting models such as ARIMA, LSTM and Prophet is provided. For these models, it is shown what weaknesses it has in the case of non-linear dynamics of the origin of temporal data. A brief introduction to differential equations is presented, which is a building block for neural network ODEs. The mathematical apparatus of differential equations is used to parameterize the neural network. This section describes the approximation of the ODE solution using the mathematical apparatus of differential equations, the solution is in the form of a "black box". It also describes scalable back propagation through ODE solution without the use of internal operations, which allows for end-to-end learning. This approach makes it possible to construct models of latent variables with continuous time, in which the latent state progresses as a continuous function in time. The section shows the properties of the sufficient existence of a solution for neural network ODEs, i.e. absence of intersection of integral curves.

The third section contains an introduction to differential equations, which is the building block for neural network ODEs. The mathematical apparatus of differential equations is used to parameterize the neural network. This section describes the approximation of the ODE solution

using the mathematical apparatus of differential equations, the solution is in the form of a "black box". It also describes scalable back propagation through ODE solution without the use of internal operations, which allows for end-to-end learning. This approach makes it possible to construct models of latent variables with continuous time, in which the latent state progresses as a continuous function in time. The section shows the properties of the sufficient existence of a solution for neural network ODEs, i.e. absence of intersection of integral curves.

The fourth section contains a description of the proposed and developed solution architecture, starting with the digitalization of a biogas plant and ending with data visualization on a dashboard. As part of the architecture description, open source solutions for various stages of data collection, storage and processing are presented. Using the collected dataset from the biogas plant, it was shown that ODE-based models are a powerful tool for irregular, sparse and noisy time series and are superior to recurrent neural networks in their extrapolation and interpolation properties.
In conclusion, the main results of the work, the conclusions of the dissertation and future steps in the study of this direction are outlined.

In conclusion, the main results of the work, the conclusions of the dissertation and future steps in the study of this direction are outlined.

Relationship with government programs. The tasks of the dissertation under consideration are of great practical importance and are directly related to the tasks of digitalization of production. Significant attention is paid to digitalization issues in the speech of President Tokayev K.K. and in government documents: the Strategy "Kazakhstan-2050", the State Program "Digital Kazakhstan", the Address of the Head of State "Unity of the people and systemic reforms - a solid foundation for the country's prosperity" dated September 1, 2021, the Concept of Kazakhstan's transition to a "green" economy, the State the program of industrial and innovative development of the Republic of Kazakhstan for 2020-2025, the Law of the Republic of Kazakhstan "On Energy Saving", where a special place is given to measures to involve the energy balance of RES; Kyoto Protocol on Clean Development.

The dissertation work was carried out as part of research work on the commercialization of the Ministry of Education and Science of the Republic of Kazakhstan 0365-18-GK - "Production and sale of biogas, biofertilizers based on the development and construction of a modular automated biogas complex with digital control and operation technologies" of the Institute of Information and Computing Technologies of the Committee of Science of the Ministry Education and Science of the Republic of Kazakhstan (IIVT MS and HE RK), approbation of technical, theoretical methods and hypotheses was carried out on a technical complex - a biogas plant. The results of research on this thesis are included in the project reports for 2019, 2020 and included in the final report for 2021.
Also, within the framework of the dissertation, work was carried out in the direction of creating sensors for collecting data from technical devices (solar collector), a patent was obtained for utility model No. Mathematical and computer models, software and hardware tools and experimental developments to create a network of combined efficient two-circuit solar collectors with thermosyphon circulation and monitoring their operation. The results of research on this thesis are included in the project reports for 2019 and included in the final report for 2020.

Approbation of work. The main results on the topic of the dissertation are presented in 10 papers, 3 articles in journals recommended by the Committee for Quality Assurance in the Sphere of Education of the Ministry of Science and Higher Education of the Republic of Kazakhstan, 4 articles in an international scientific publication included in the Scopus database and 3 materials of international foreign conferences: 14th international conference Electronics Computer and Computation (ICECCO), IEEE. Kazakhstan, SDU, Kaskelen, 2018; International Conference "Computing and Information Technologies in Science, Technology and Education", 2020; 13th International Symposium “Intelligent Systems 2018” (INTELS’18), October 22-24, 2018, St. Petersburg.

Scientific publications: The results obtained on the topic of the dissertation are presented in the following publications:
Journals SCOPUS/WoS:

1. Yedilkhan, A., Murat, K., Beibut, A., Aliya, K., Ainur, K., **Timur, M.**, & Azhibek, D. (2020). Mathematical justification of thermosyphon effect main parameters for solar heating system. Cogent Engineering, 7(1), 1851629. Scopus: 67%, WoS: Q2
2. **Merembayev, T.**, Kurmangaliyev, D., Bekbauov, B., & Amanbek, Y. (2021). A Comparison of machine learning algorithms in predicting lithofacies: Case studies from Norway and Kazakhstan. Energies, 14(7), 1896. Scopus: 85%, WoS: Q3, IF:3.0
3. **Merembayev, T.**, Amirgaliyev, Y., Kunelbayev, M., Yedilkhan, D. (2022). Thermal loss analysis of a solar flat collector using numerical simulation. Computers, Materials and Continua. Accepted. Scopus: 80%, WoS: Q2, IF:3.7
4. Amirgaliyev, Y., Wójcik, W., Kunelbayev, M., & **Merembayev, T.** (2019). Theoretical preheoretical prerequisites of electric water heating. NEWS of National Academy of Sciences of the Republic of Kazakhstan. 6 (438), pp. 54-63. Scopus: 40%, WoS: Q4

The list of journals recommended by committee for quality assurance of education and science of the MES RK:

1. Амиргалиев Е.Н., Юнусов Р., **Мембаев Т.**, Едилхан Д. Проектирование архитектуры хранения данных в сети гелиоколлеторов. Вестник КазНИТУ, Казахский национальный исследовательский технический университет имени К.И. Сатпаева, ISSN 1680-9211, -№4,-2020 г., - С.212-216
2. Amirgaliyev, Y., **Merembayev, T.** (2021). Anomaly Detection in Solar Hot Water System Using Machine Learning. Вестник национальной инженерной академии РК, №3 часть 2, - С.34-44
3. Amirgaliev Y., Kunelbayev M., **Merembayev, T.** Kozbakova A., Sundetov T., Irzhanova A. Control system of controllers of a flat solar collector with a thermosifon circulation. Вестник Казахстанско-британского технического университета, ISSN 1198-6688, -№1,-2019 г., - С.55-61

Patent

1. Патент на полезную модель №5591. Кунелбаев М., Калимолдаев М., Сундетов Т., Даулбаев С., **Мембаев Т.**

Conferences indexing in SCOPUS:

1. Amirgaliyev, Y., **Merembayev, T.**, Kunelbayev, M., Yedilkhan, D. Dynamic Simulation of a Solar Hot Water Heating System for Kazakhstan Climate Conditions. 2018 14th International Conference on Electronics Computer and Computation (ICECCO). IEEE. Electronic ISBN: 978-1-7281-0132-3
2. **Merembayev T.**, Yunusson R., Amirgaliyev Y. Machine Learning Algorithms for Classification Geology Data from Well Logging. 2018 14th International Conference on Electronics Computer and Computation (ICECCO). IEEE. Electronic ISBN: 978-1-7281-0132-3
3. **Merembayev T.**, Yunusson R., Amirgaliyev Y. Machine Learning Algorithms for Stratigraphy Classification on Uranium Deposits. Proceedings of the 13th International Symposium “Intelligent Systems 2018” (INTELS’18), 22-24 October, 2018, St. Petersburg, Russia. Procedia Computer Science. Volume 150, 2019, Pages 46-52. Scopus (CiteScore 2018: 1.48)