МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН
НЕКОММЕРЧЕСКОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО
«АЛМАТИНСКИЙ УНИВЕРСИТЕТ ЭНЕРГЕТИКИ И СВЯЗИ ИМЕНИ ГУМАРБЕКА ДАУКЕЕВА»

ISSN 2790-0886

В Е С Т Н И К

АЛМАТИНСКОГО УНИВЕРСИТЕТА ЭНЕРГЕТИКИ И СВЯЗИ

Учрежден в июне 2008 года

Тематическая направленность: энергетика и энергетическое машиностроение, информационные, телекоммуникационные и космические технологии

3 (74) **2026**

Импакт-фактор - 0.076

Научно-технический журнал
Выходит 4 раза в год

Алматы

ВЕСТНИК АЛМАТИНСКОГО УНИВЕРСИТЕТА ЭНЕРГЕТИКИ И СВЯЗИ

СВИДЕТЕЛЬСТВО
о постановке на переучет периодического печатного издания,
информационного агентства и сетевого издания
№ KZ14VPY00024997
выдано
Министерством информации и общественного развития
Республики Казахстан

Подписной индекс – 74108

Бас редакторы – главный редактор
Стояк В.В.
к.т.н., профессор

Заместитель главного редактора — Жауыт Алгазы, доктор PhD
Ответственный секретарь — Алиева К.В.

Редакция алқасы – Редакционная коллегия

Главный редактор – Стояк В.В., кандидат технических наук, профессор Алматинского Университета Энергетики и Связи имени Гумарбека Даукеева, Казахстан;

Заместитель главного редактора – Жауыт А., доктор PhD, ассоциированный профессор Алматинского Университета Энергетики и Связи имени Гумарбека Даукеева, Казахстан;

Сагинтаева С.С., доктор экономических наук, кандидат физико-математических наук, профессор математики, академик МАИН;

Ревалде Г., доктор PhD, член-корреспондент Академии наук, директор Национального Совета науки, Рига, Латвия;

Илиев И.К., доктор технических наук, Русенский университет, Болгария;

Белоев К., доктор технических наук, профессор Русенского университета, Болгария;

Обозов А.Д., доктор технических наук, НАН Кыргызской Республики, заведующий Лабораторией «Возобновляемые источники энергии», Кыргызская Республика;

Кузнецов А.А., доктор технических наук, профессор Омского государственного технического университета, ОмГУПС, Российская Федерация, г. Омск;

Алипбаев К.А., PhD, доцент Алматинского Университета Энергетики и Связи имени Гумарбека Даукеева, Казахстан;

Зверева Э.Р., доктор технических наук, профессор Казанского государственного энергетического университета, Российская Федерация, г. Казань;

Лахно В.А., доктор технических наук, профессор Национального университета биоресурсов и природопользования Украины, кафедра компьютерных систем, сетей и кибербезопасности, Украина, Киев;

Омаров Ч.Т., кандидат физико-математических наук, директор Астрофизического института имени В.Г. Фесенкова, Казахстан;

Коньшин С.В., кандидат технических наук, профессор Алматинского Университета Энергетики и Связи имени Гумарбека Даукеева, Казахстан;

Тынымбаев С.Т., кандидат технических наук, профессор Алматинского Университета Энергетики и Связи имени Гумарбека Даукеева, Казахстан.

За достоверность материалов ответственность несут авторы.
При использовании материалов журнала ссылка на «Вестник АУЭС» обязательна.

ЭНЕРГЕТИКА И ЭНЕРГЕТИЧЕСКОЕ МАШИНОСТРОЕНИЕ

UDC 004.85 DOI 10.51775/2790-0886_2026_74_3_35

DEVELOPMENT OF A REAL-TIME ELECTRICITY DEMAND FORECASTING PLATFORM BASED ON MACHINE LEARNING AND APACHE KAFKA STREAM PROCESSING

[1]I. Khussainov, [1]D. Astapenko, [1]*G. Kim, [2]A. Kabdirova

1 NJSC "Manash Kozybayev North Kazakhstan University", Petropavl, Kazakhstan.
2 The branch of Joint Stock Company “National center of professional development for pedagogical staff «Orleu», Petropavl, Kazakhstan.

E-mail: kabdirova_a@orleu.edu.kz, gakim@ku.edu.kz, wilentineqsq@gmail.com, husainovilnur451@gmail.com

Поступила в редакцию 07.07.2026
Получена исправленная версия 19.08.2026
Принята 18.09.2026
Доступна на сайте 15.09.2026

Abstract: *The goal of this research is to create GridSense, a real-time digital platform that uses stream data processing and machine learning to forecast electricity consumption in Petropavlovsk city. The study tackles the issue of manual load planning in the North Kazakhstan regional power system, where dispatch services face considerable unpredictability due to significantly continental climate conditions, wind influence, heating demand, and noticeable daily consumption peaks. The primary goal was to develop an economically accessible, low-latency, and interpretable decision-support platform that could generate short-term forecasts up to 48 hours in advance. The work's methodological foundation consists of the creation of an event-driven software architecture based on Apache Kafka and system-structural modeling using Data Flow Diagram (DFD) and Integration Definition for Function Modeling (IDEF0). A Linear Regression model trained on climatic and temporal parameters, such as temperature, wind speed, humidity, present load, forecast horizon, and goal hour, is used by the forecasting module. A browser-based single page application, a Node.js and Express.js backend, a Kafka message broker, a Python machine learning consumer, a PostgreSQL database, and WebSocket real-time delivery are all integrated into the implemented platform. The model attained an operational forecasting accuracy of 88.8% with $R^2 = 0.77$ and Mean Absolute Error of 23.7 MW, according to experimental evaluation. With a mean response time of 8.9 ms, p95 latency of 12.1 ms, p99 latency of 18 ms, and no detected errors, load testing with Artillery showed robust system operation. The suggested approach serves as a foundation for intelligent energy planning systems in Kazakhstan and offers useful assistance for regional dispatch management.*

Keywords: *electricity demand forecasting, machine learning, Linear Regression, Apache Kafka, stream data processing, PostgreSQL, Node.js, WebSocket, GridSense, dispatch management.*

INTRODUCTION

The electric power industry represents a critical component of socio-economic infrastructure. Public safety, municipal services, industrial productivity, and population quality of life are all directly impacted by the dependability of the electrical supply. The issue of short-term electricity demand forecasting is particularly significant for regional power distribution systems since dispatch choices must be made prior to overloads. Operators can activate reserve capacity, delay non-critical maintenance, get substations ready for peak load periods, and cut down on needless operating expenses if the anticipated demand is known several hours in advance [1, 2].

Forecasting electricity demand in the North Kazakhstan Region presents complex multifaceted challenges. With harsh winter temperatures, a significant yearly temperature amplitude, windy steppe conditions, and a high demand for electricity due to heating, the area has a distinctly continental climate. Because household heating, district heating circulation pumps, and industrial infrastructure need more power

during cold seasons, consumption rises quickly. While everyday rhythms are still influenced by job schedules and household behavior, the load drastically lessens in the summer. As a result, the same distribution network may see both large winter peak loads and extremely low usage at night [3].

Manual planning, historical average tables, and accumulated professional knowledge are all major components of traditional dispatch practice. In dynamic situations, these approaches are helpful yet insufficient. Although dispatchers may qualitatively anticipate the risks associated with winter peak loads, manual planning is unable to rapidly determine how a particular combination of temperature, wind speed, humidity, present load, and predicted horizon would affect demand three, twelve, or forty-eight hours in advance. As a result, there is a practical disconnect between operational decision-making and the information that is available.

While quantitative projections can be produced from historical and contextual data using contemporary machine learning techniques, the approach must align with the actual working environment. Despite the widespread adoption of machine learning in global power systems, a significant gap exists in accessible, real-time forecasting solutions for regional grids operating under sharply continental climates. Classical statistical methods, such as ARIMA, struggle to effectively incorporate exogenous climatic variables and rapidly lose accuracy beyond a 12-hour horizon. Conversely, advanced models like Long Short-Term Memory (LSTM) neural networks [3] or Gradient Boosting (XGBoost) [4] may show excellent accuracy on big datasets, but their "black-box" nature limits operator trust in safety-critical dispatch environments, and their high inference latency complicates real-time stream processing. Furthermore, commercial enterprise platforms (e.g., Oracle Utilities) are prohibitively expensive and require advanced smart-metering infrastructure that is currently unavailable in the North Kazakhstan Region [3].

Therefore, this study bridges the gap by developing and validating an interpretable predictive model of electricity demand that can process multivariate climate and temporal data in real time with ultra-low latency. The primary goal of this project is to propose a custom, open-source platform combining an event-driven Apache Kafka architecture with a transparent Linear Regression model. This approach ensures horizontal scalability, ultra-low inference latency, and clear connections between input elements and anticipated load—delivering structured forecasting, risk classification, load type, and operational recommendations directly to regional dispatch services [5].

METHODS

System analysis, structural modeling, supervised machine learning, and software performance testing are all part of the study technique. The topic was examined from the viewpoints of operational dispatch management and regional electricity distribution. The forecasting horizon, current load, calendar variables, and meteorological characteristics were the primary components of the forecasting process [1, 2].

The need for real-time operation served as the foundation for the hardware and software architecture. Because of this, the system was not developed as a monolithic application but rather as a distributed pipeline. For the client interface, the primary technological components are HTML5, CSS3, JavaScript, and jQuery; for the backend, Node.js and Express.js; for message routing, Apache Kafka; for machine learning inference, Python and scikit-learn; for data storage, PostgreSQL; and for real-time result delivery, socket.io WebSocket [6].

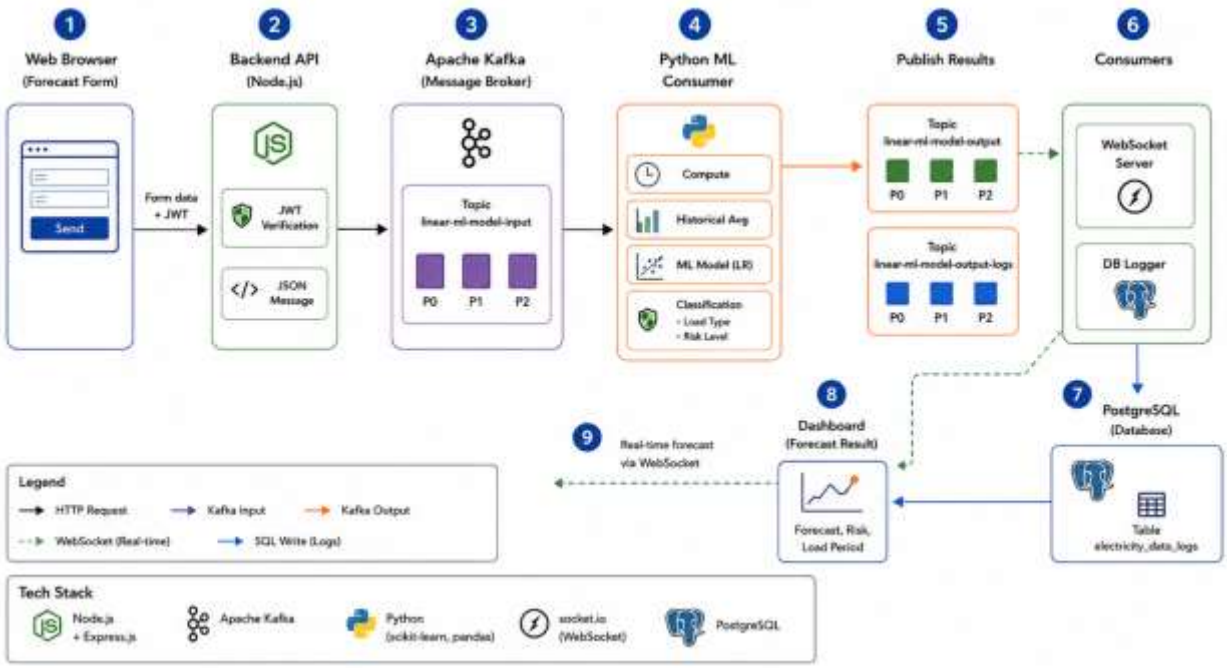


Figure 1 - Overall System Architecture and Processing Chain

Figure 1 shows the processing chain and overall system design. The flow of a forecast request from the dispatcher interface to the final dashboard output is depicted in the diagram. The dispatcher inputs the date, time, temperature, wind speed, humidity, present load, and forecast horizon in the browser-based Single Page Application to start the procedure. The backend API receives the request as an HTTP POST message.

Express.js and Node.js are used to implement the backend. The backend creates a JSON message and publishes it to the Kafka topic linear-ml-model-input following JWT authentication. After reading messages from Kafka, the Python ML consumer applies normal scaling, generates a forecasted consumption number, and categorizes the outcome based on load type and risk level. Two output topics are used to publish the prediction result: linear-ml-model-output-logs for PostgreSQL persistence and linear-ml-model-output for WebSocket delivery [7, 8].

The primary forecasting algorithm chosen was linear regression [9]. The model uses a weighted sum of input features to forecast power demand:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n, \quad (1)$$

where $\hat{y}$ is the predicted load in MW, β_0 is the intercept, β_i are model coefficients and x_i are input features [10]. The input feature vector includes temperature (°C), wind speed (m/s), humidity (%), current substation load (MW), forecast horizon (hours ahead) and target hour of the day (0–23). The model was trained on historical load data combined with meteorological observations for Petropavlovsk.

System design was formalised using two complementary modelling methods. The Data Flow Diagram (DFD) approach was used to describe the logical flow of data between platform components. The IDEF0 methodology was applied to model system functions, their inputs, outputs, controls and mechanisms at two levels of abstraction: the context diagram (A0) and the functional decomposition into five blocks.

The Data Flow Diagram of the GridSense platform is shown in Figure 2. It describes the logical flow of information between the user, internal processing modules, Kafka topics, PostgreSQL storage and WebSocket delivery. The external entity is the user operating the SPA browser interface who submits a forecast request containing input data such as date, time, temperature, wind speed, humidity, current load and forecast horizon.

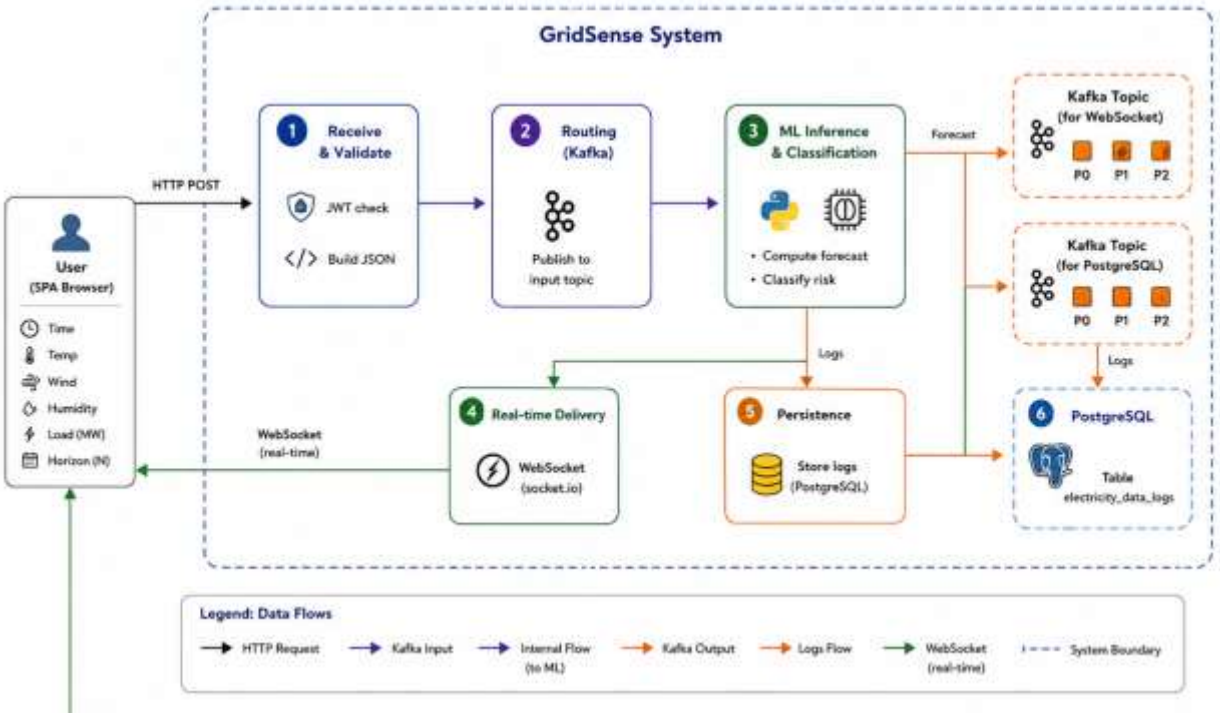


Figure 2 - DFD (Data Flow Diagram)

The upper-level IDEF0 functional model is presented in Figure 3. The central function A0 is defined as GridSense: Electricity Consumption Forecasting. The input block includes climatic data, current time and date, current substation load and forecast horizon from 1 to 48 hours. The control elements include dispatching regulations, risk threshold values and JWT authentication policies. The mechanisms include Node.js, Python, Apache Kafka, PostgreSQL and browser technologies [11, 12].

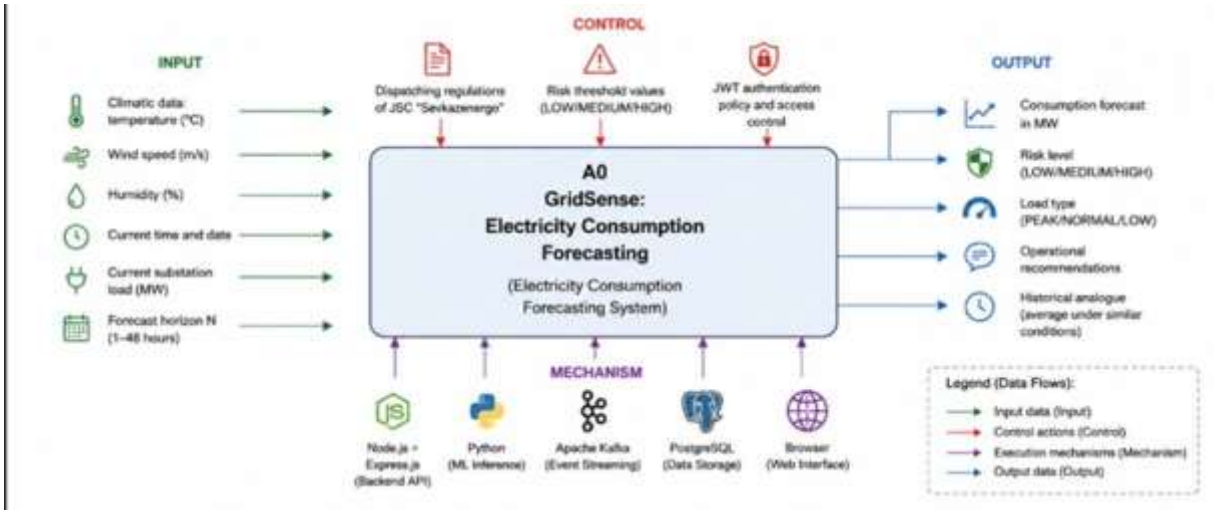


Figure 3 - IDEF0 (Functional Model of the System)

The detailed IDEF0 decomposition is presented in Figure 4. It expands the high-level function A0 into five main functional blocks: request reception and authentication (A1), routing via Apache Kafka (A2), ML inference and classification (A3), persistence and logging (A4), and real-time delivery and visualisation (A5). This decomposition demonstrates that the platform is designed as a chain of controlled functional transformations [13].

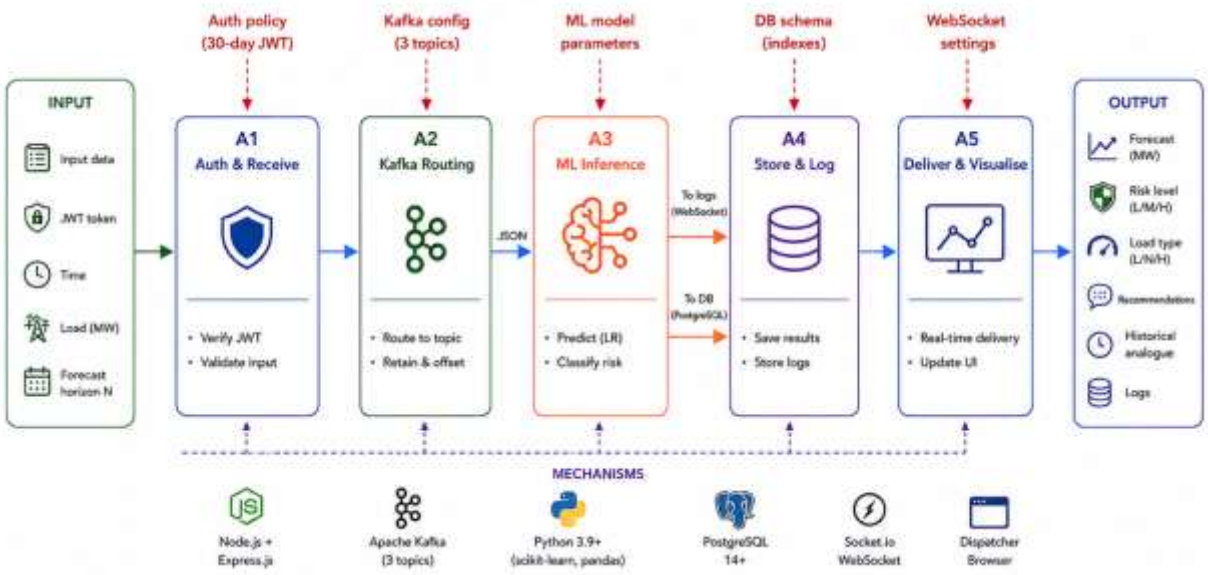


Figure 4 - IDEF0 (Decomposition Diagram of the GridSense System)

The key scientific contribution of the study is the development of a feature space for the N-hour-ahead architecture. The input vector includes 11 features, but two are critical:

a. hours_ahead (target horizon) — allows the unified model to approximate the increase in uncertainty with increasing forecast horizon.

b. target_hour — calculated as $(current_hour + hours_ahead)\ mod\ 24$. This dynamic feature rigidly determines the position of the target moment in the daily cycle, solving the problem of blind extrapolation. The implementation of target_hour significantly reduces modeling error during peak hours compared to models lacking this temporal feature.

All features were standardized (StandardScaler) to eliminate scale imbalances (e.g., temperature and consumption in MW). Multiple linear regression with ordinary least squares (OLS) optimization was chosen as the primary regressor.

To test the model's applicability in real time, an event-driven architecture (EDA) was deployed using the Apache Kafka message broker. Splitting streams into linear-ml-model-input and linear-ml-model-output topics ensured the isolation of the mathematical inference process (implemented in Python) from the web service's I/O operations (Node.js).

RESULTS

The Linear Regression model achieved a coefficient of determination $R^2 = 0.77$, Mean Absolute Error of 23.7 MW and operational forecasting accuracy of 88.8%. These results confirm that the additive structure of climatic and temporal load factors can be effectively represented by an interpretable regression model. The main experimental results are summarized in Table 1.

Table 1. Main experimental results of the GridSense platform

Indicator	Value
Forecasting method	Linear Regression
Forecast horizon	1–48 hours
Coefficient of determination	$R^2 = 0.77$
Mean Absolute Error	23.7 MW
Operational forecasting accuracy	88.8%
Mean response time	8.9 ms
p95 latency	12.1 ms

p99 latency	18 ms
Error rate	0%

To validate the choice of algorithm, a comparative experiment was conducted with the gradient boosting algorithm (XGBoost, n_estimators=300) [4]. XGBoost demonstrated a slight improvement in accuracy, but increased inference latency by 5-50 times (up to 5-50 ms) compared to linear regression (< 1 ms on a standard processor) and reduced interpretability.

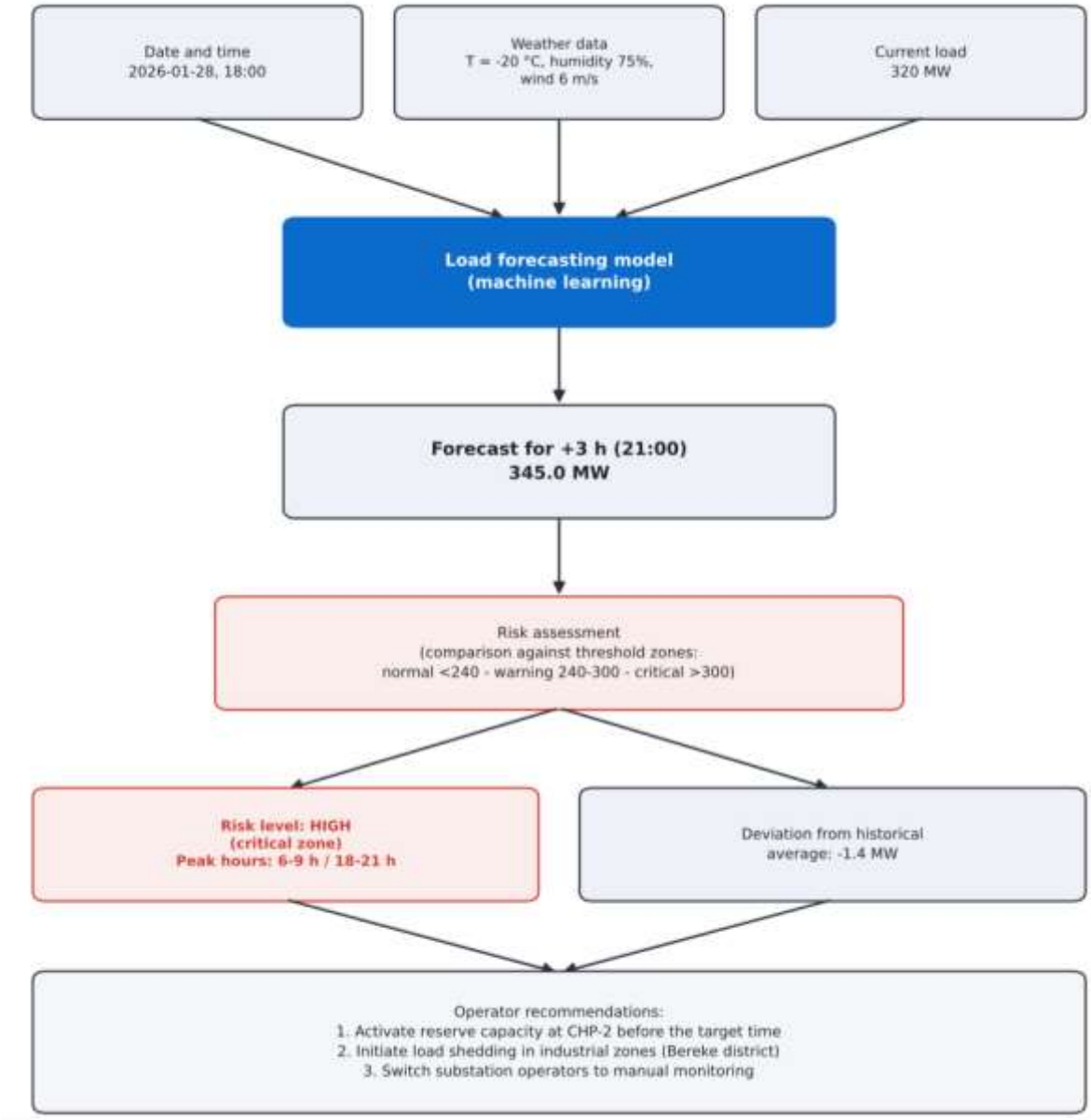


Figure 5 – Structural diagram of the electric load forecasting model (High-Risk Scenario)

The frontend dashboard provides the dispatcher with predicted load, risk level, load period, target time, historical analogue and recommendations. Figure 5 demonstrates the high-risk case with a predicted value of 345.0 MW. In this state, the interface displays critical colour coding and recommends preventive actions such as activating reserve capacity and postponing non-critical maintenance.

Figure 6 demonstrates the same structural diagram under normal conditions. The predicted load is 192.5 MW, and the system classifies the state as safe. Recommendations are focused on standard monitoring, maintenance planning and operational cost optimization.

While achieving reliable predictive accuracy (R^2=0.77) validates the mathematical viability of the approach, real-world dispatch management demands strict adherence to low-latency constraints. To ensure that the event-driven architecture can handle operational message streams without bottlenecks, the system's throughput and response times were rigorously benchmarked.

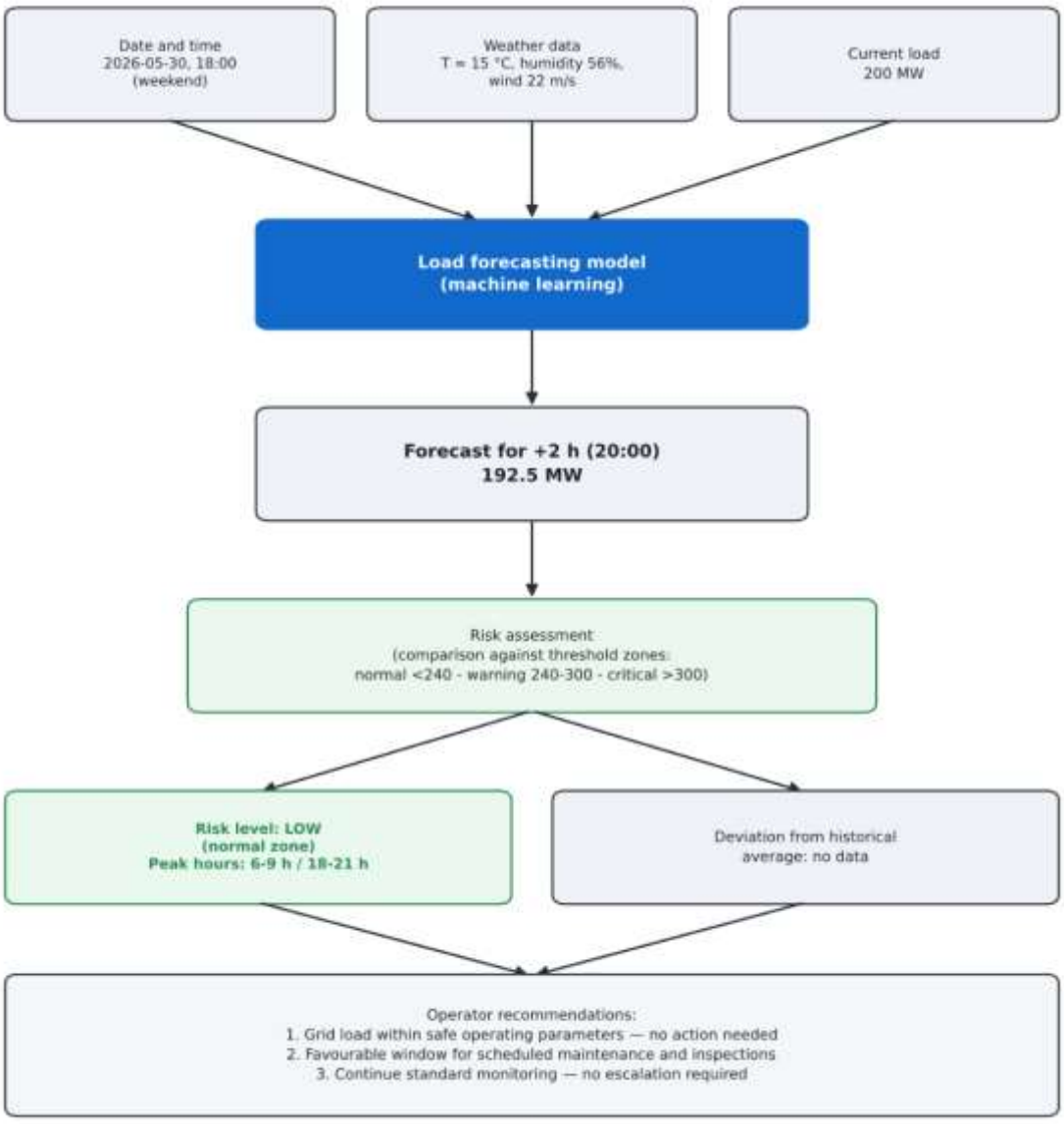


Figure 6 – Structural diagram of the electric load forecasting model (Normal Conditions Scenario)

Figure 6 demonstrates the same structural diagram under normal conditions. The predicted load is 192.5 MW, and the system classifies the state as safe. Recommendations are focused on standard monitoring, maintenance planning and operational cost optimisation.

While achieving reliable predictive accuracy (R^2=0.77) validates the mathematical viability of the approach, real-world dispatch management demands strict adherence to low-latency constraints. To ensure that the event-driven architecture can handle operational message streams without bottlenecks, the system's throughput and response times were rigorously benchmarked.

Performance testing was conducted using Artillery. The minimum HTTP 2xx response time was 5 ms, the maximum was 23 ms, the mean was 8.9 ms and the median was 8.9 ms. The p95 latency was 12.1 ms and the p99 latency was 18 ms. The observed error rate was 0%.

DISCUSSION

The findings of the study show that extremely complicated nonlinear machine learning methods are not necessarily the best option for projecting regional energy usage. A linear model can exhibit great efficiency due to the additive nature of climate parameters (the superposition of temperature delta and wind convection on the base load).

In this situation, linear regression's complete transparency (white-box model) is a major benefit. Similar to engineering heat loss calculations, the weights for the temperature and wind_speed elements have a clear physical meaning that enables operators to verify projections before to making crucial choices (e.g., when surpassing the high-risk threshold of 300 MW). The system's capacity to manage high-load telemetry feeds without blocking was validated by integrating the method with Kafka [14, 15].

CONCLUSION

A real-time electricity demand forecasting platform, GridSense, was created for the Petropavlovsk regional power environment. The system incorporates machine learning, Apache Kafka for stream processing, Node.js for backend services, PostgreSQL for storage, WebSocket for distribution, and a browser-based dashboard.

The Linear Regression forecasting model attained $R^2 = 0.77$, MAE = 23.7 MW, and an operational accuracy of 88.8%. The results indicate that an interpretable model can deliver valuable forecasting accuracy when the feature set encompasses climatic and temporal demand variables. Performance testing utilizing Artillery validated real-time applicability: the average response time was 8.9 ms, the 95th percentile latency was 12.1 ms, the 99th percentile latency was 18 ms, and the error rate was 0%.

The suggested solution can serve as a basis for the advanced development of intelligent energy management systems in Kazakhstan. Future endeavors should concentrate on the integration of actual SCADA measurements, automated meteorological data streams, adaptive model retraining, and comprehensive economic assessment under authentic operational conditions.

LIST OF REFERENCES

1. Zhang, Y., Wang, J. and Wang, X. (2023) 'Recent advances in probabilistic load forecasting for smart grids: A comprehensive review', International Journal of Forecasting, 39(2), pp. 612–635.

2. Hyndman, R. J. and Athanasopoulos, G. (2021) Forecasting: Principles and Practice. 3rd edn. Melbourne: OTexts.

3. Liu, X., Zhang, H. and Li, Y. (2022) 'Short-term electric load forecasting using an improved LSTM neural network with attention mechanism', IEEE Transactions on Industrial Informatics, 18(4), pp. 2410–2419.

4. Ahmadi, M., et al. (2022) 'A comparative evaluation of gradient boosting algorithms for short-term load forecasting in modern distribution systems', Energy Reports, 8, pp. 4521–4535.

5. Sun, M., Zhang, K. and Jiang, T. (2022) 'Data analytics in modern distribution grids: A survey of recent techniques and applications', IEEE Transactions on Smart Grid, 13(5), pp. 3810–3825.

6. Gaber, M. M. and Gama, J. (2022) Stream Data Mining: Algorithms and Architecture. 2nd edn. Cambridge: Cambridge University Press.

7. Vilar, J. A., et al. (2022) 'Day-ahead electricity demand forecasting: A comprehensive comparative study of machine learning and deep learning models', Applied Energy, 310, p. 118501.

8. Goude, Y., et al. (2021) 'Recent advances in generalized additive models for short-term electricity load forecasting', IEEE Transactions on Power Systems, 36(3), pp. 1845–1855.

9. James, G., Witten, D., Hastie, T. and Tibshirani, R. (2021) An Introduction to Statistical Learning: with Applications in R. 2nd edn. New York: Springer.

10. Kuster, C., Rezgui, Y. and Mourshed, M. (2022) 'Electrical load forecasting through artificial intelligence and machine learning: Recent developments and methodologies', Renewable and Sustainable Energy Reviews, 156, p. 111985.

11. Fahad, A. and Al-Hajj, H. (2023) 'Real-time stream processing of energy data using Apache Kafka and machine learning', Journal of Big Data Analytics in Transportation, 5(2), pp. 112–129.

12. Patel, R. and Sharma, V. (2022) 'Performance benchmarking of Apache Kafka and distributed messaging platforms for real-time IoT data streams', Journal of Network and Computer Applications, 201, p. 103320.

13. Méndez, J. I. and Casillas, J. (2022) 'Interpretable machine learning for power systems: A review on the role of linear models and surrogate methods', Expert Systems with Applications, 200, p. 116982.

14. Bhardwaj, A. and Singh, A. (2021) 'Scalable architectures for real-time predictive maintenance in industrial IoT', IEEE Access, 9, pp. 89234–89248.

15. Schmidt, M. and Weber, T. (2021) 'Comparative evaluation of WebSocket and HTTP/2 for real-time web applications in industrial environments', Journal of Web Engineering, 20(3), pp. 711–736.

Личный вклад авторов рукописи

DEVELOPMENT OF A REAL-TIME ELECTRICITY DEMAND FORECASTING PLATFORM BASED ON MACHINE LEARNING AND APACHE KAFKA STREAM PROCESSING

№	ФИО авторов рукописи	ORCID
1	Хусаинов Ильнур Рашидович	-
2	Астапенко Денис Александрович	-
3	Ким Гульнар Амангельдыевна	0000-0001-8402-8323
4	Кабдирова Айнагуль Амангельдиновна	0009-0001-6710-1311

№	ФИО автора	Авторский вклад	Описание вклада
1	Хусаинов Ильнур Рашидович	Концептуализация,	Разработал системную архитектуру
		Методология,	Провел сравнительный анализ архитектур детекции
		Программное обеспечение,	Реализовал программный модуль на Python
		Курирование данных	Выполнил интеграцию алгоритмов нейросетей
2	Астапенко Денис Александрович	Исследование	Выполнил сбор данных на производстве
		Валидация	Участвовал в эксплуатационных испытаниях системы
		Визуализация	Подготовил визуализацию результатов и сформировал отчетность по производительности.
3	Ким Гульнар Амангельдыевна	Научное руководство	Осуществляла общее научное руководство
		Администрирование проекта	Координировала проектные работы
		Рецензирование и редактирование	Проводила критический анализ методологической структуры исследования и редактирование финальной версии рукописи.
4	Кабдирова Айнагуль Амангельдиновна	Методология,	Разработала методологическую базу для формализации процессов (IDEF0, IDEF3, DFD)
		Формальный анализ	Обеспечила необходимые производственные ресурсы
		Ресурсы	Провела анализ ограничений применимости системы

МАШИНАЛЫҚШИНАЛЫҚ ОҚЫТУ ЖӘНЕ APACHE KAFKA АҒЫНДЫҚ ӨҢДЕУ ТЕХНОЛОГИЯСЫНА НЕГІЗДЕЛГЕН ЭЛЕКТР ЭНЕРГИЯСЫНА СҰРАНЫСТЫ НАҚТЫ УАҚЫТ РЕЖИМІНДЕ БОЛЖАУ ПЛАТФОРМАСЫН ӘЗІРЛЕУ

[1]Хуссаинов И.Р., [1]Астапенко Д.А., [1]Ким Г.А., [2]Кабдирова А.А.

1«Манаш Қозыбаев атындағы Солтүстік Қазақстан университеті» КеАҚ, Петропавл, Қазақстан,
2«Өрлеу» біліктілікті арттыру ұлттық орталығы» АҚ филиалы, Петропавл, Қазақстан,

E-mail: kabdirova_a@orleu.edu.kz, gakim@ku.edu.kz, wilentineqsq@gmail.com, husainovilnur451@gmail.com

Аңдатпа. *Бұл зерттеудің мақсаты – ағындық деректерді өңдеу және машиналық оқыту технологияларын қолдана отырып, Петропавл қаласында электр энергиясын тұтынуды нақты уақыт режимінде болжауға арналған GridSense цифрлық платформасын әзірлеу. Жұмыста Солтүстік Қазақстанның аймақтық энергетикалық жүйесіндегі жүктемені қолмен жоспарлау мәселесі қарастырылады, мұнда күрт континенттік климат жағдайлары, желдің әсері, жылытуға деген сұраныс және тұтынудың айқын тәуліктік шыңдары диспетчерлік қызметтер үшін айтарлықтай белгісіздік тудырады. Негізгі мақсат 48 сағатқа дейінгі қысқа мерзімді болжамдарды жасай алатын экономикалық қолжетімді, кідірісі төмен және түсіндірілетін шешім қабылдауды қолдау платформасын құру болды. Жұмыстың әдіснамалық негізі Apache Kafka базасында оқиғаға бағытталған бағдарламалық архитектураны құрудан және деректер ағыны диаграммасы (DFD) мен қызметтерді модельдеуге арналған интеграциялық анықтаманы (IDEF0) пайдалана отырып, жүйелік-құрылымдық модельдеуден тұрады. Болжау модулі температура, желдің жылдамдығы, ылғалдылық, ағымдағы жүктеме, болжау көкжиегі және мақсатты сағат сияқты климаттық және уақыттық өлшемдерге оқытылған сызықтық регрессия (Linear Regression) моделін пайдаланады. Жүзеге асырылған платформаға браузерлік бір беттік қосымша (SPA), Node.js және Express.js негізіндегі бэкенд, Kafka хабарламалар брокері, Python тіліндегі ML-консьюмер, PostgreSQL дерекқор базасы және WebSocket арқылы нақты уақытта деректерді жеткізу жүйесі біріктірілген. Тәжірибелік бағалауға сәйкес, модель $R^2 = 0.77$ және орташа абсолюттік қателік (MAE) 23.7 MW болғанда 88.8% операциялық дәлдікке қол жеткізді. Artillery көмегімен жүктемелік тестілеу жүйенің тұрақты жұмысын көрсетті: орташа жауап беру уақыты 8.9 ms, p95 кідірісі 12.1 ms, p99 кідірісі 18 ms және ешқандай қате табылмады. Ұсынылып отырған тәсіл Қазақстандағы зияткерлік энергетикалық жоспарлау жүйелерінің негізі болып табылады және өңірлік диспетчерлік басқару үшін пайдалы құрал ұсынады.*

Тірек сөздер*: электр энергиясына сұранысты болжау, машиналық оқыту, Linear Regression, Apache Kafka, ағындық деректерді өңдеу, PostgreSQL, Node.js, WebSocket, GridSense, диспетчерлік басқару.*

РАЗРАБОТКА ПЛАТФОРМЫ ПРОГНОЗИРОВАНИЯ СПРОСА НА ЭЛЕКТРОЭНЕРГИЮ В РЕАЛЬНОМ ВРЕМЕНИ НА БАЗЕ МАШИННОГО ОБУЧЕНИЯ И ПОТОКОВОЙ ОБРАБОТКИ APACHE KAFKA

[1]Хуссаинов И.Р., [1]Астапенко Д.А., [1]Ким Г.А., [2]Кабдирова А.А.

1 НАО «Северо-Казахстанский университет имени Манаша Козыбаева», Петропавловск, Казахстан,
2 «Өрлеу» біліктілікті арттыру ұлттық орталығы» АҚ филиалы, Петропавл, Қазақстан,

E-mail: kabdirova_a@orleu.edu.kz, gakim@ku.edu.kz, wilentineqsq@gmail.com, husainovilnur451@gmail.com

Аннотация. *Настоящее исследование посвящено разработке цифровой платформы GridSense для прогнозирования спроса на электроэнергию в городе Петропавловске в режиме реального времени с использованием потоковой обработки данных и машинного обучения. В работе рассматривается проблема ручного планирования нагрузки в региональной энергосистеме Северного Казахстана, где резко континентальные климатические условия, влияние ветра, отопительная нагрузка и выраженные суточные пики потребления создают значительную неопределенность для диспетчерских служб. Главной целью было создание экономически доступной, низколатентной и интерпретируемой платформы поддержки принятия решений, способной генерировать краткосрочные прогнозы на горизонте до 48 часов. Методологическая база работы включает создание событийно-ориентированной архитектуры программного обеспечения на базе Apache Kafka и системно-структурное моделирование с использованием диаграмм потоков данных (DFD) и методологии функционального моделирования (IDEF0). Модуль прогнозирования использует модель линейной регрессии (Linear Regression), обученную на климатических и временных параметрах, таких как температура, скорость ветра, влажность, текущая нагрузка, горизонт прогнозирования и целевой час. Реализованная платформа*

объединяет браузерное одностраничное приложение (SPA), бэкенд на Node.js и Express.js, брокер сообщений Kafka, ML-консьюмер на Python, базу данных PostgreSQL и доставку данных в реальном времени через WebSocket. Экспериментальная оценка показала, что модель достигла операционной точности прогнозирования 88.8% при $R^2 = 0.77$ и средней абсолютной ошибке (MAE) 23.7 MW. Нагрузочное тестирование с помощью Artillery подтвердило стабильную работу системы: среднее время отклика составило 8.9 ms, p95 — 12.1 ms, p99 — 18 ms, при отсутствии обнаруженных ошибок. Предложенный подход служит основой для систем интеллектуального энергетического планирования в Казахстане и предлагает полезный инструмент для регионального диспетчерского управления.

***Ключевые слова**: прогнозирование спроса на электроэнергию, машинное обучение, Linear Regression, Apache Kafka, потоковая обработка данных, PostgreSQL, Node.js, WebSocket, GridSense, диспетчерское управление.*

Басылымның шығыс деректері

Мерзімді баспасөз басылымының атауы	«Алматы энергетика және байланыс университетінің Хабаршысы» ғылыми-техникалық журналы
Мерзімді баспасөз басылымының меншік иесі	«Ғұмарбек Дәукеев атындағы Алматы энергетика және байланыс университеті» коммерциялық емес акционерлік қоғамы, Алматы, Қазақстан
Бас редактор	Профессор, т.ғ.к., В.В. Стояк
Қайта есепке қою туралы куәліктің нөмірі мен күні және берген органның атауы	№ KZ14VPY00024997, күні 17.07.2020, Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігі
Мерзімділігі	Жылына 4 рет (тоқсан сайын)
Мерзімді баспасөз басылымының реттік нөмірі және жарыққа шыққан күні	Жалпы нөмірі 74, 3-басылым, 2026 жылғы 30 қыркүйек
Басылым индексі	74108
Басылым таралымы	200 дана
Баға	Келісілген
Баспахана атауы, оның мекен-жайы	Satbayev University баспаханасы, Қазақстан Республикасы, Алматы қ., Сәтбаев көшесі, 22, бас ғимарат, 109-кабинет.
Редакцияның мекен-жайы	050013, Алматы қ., «Ғұмарбек Дәукеев атындағы Алматы энергетика және байланыс университеті» КЕАҚ, Байтұрсынұлы к-сі, 126/1 үй, каб. А 224, тел.: 8 (727) 292 58 48, 708 880 77 99, e-mail: vestnik@aues.kz

Выходные данные

Название периодического печатного издания	Научно-технический журнал «Вестник Алматинского университета энергетики и связи»
Собственник периодического печатного издания	Некоммерческое акционерное общество «Алматинский университет энергетики и связи имени Гумарбека Даукеева», Алматы, Казахстан
Главный редактор	Профессор, к.т.н., Стояк В.В.
Номер и дата свидетельства о постановке на переучет и наименование выдавшего органа	№ KZ14VPY00024997 от 17.07.2020 Министерство информации и общественного развития Республики Казахстан
Периодичность	4 раза в год (ежеквартально)
Порядковый номер и дата выхода в свет периодического печатного издания	Валовый номер 71, выпуск 4, 30 декабря 2025
Подписной индекс	74108
Тираж выпуска	200 экз.
Цена	Договорная
Наименование типографии, ее адрес	Типография Satbayev University, Республика Казахстан, г. Алматы, ул. Сатпаева, 22, главный корпус, кабинет 109.
Адрес редакции	050013, г. Алматы, НАО «Алматинский университет энергетики и связи имени Гумарбека Даукеева», ул. Байтурсынулы, дом 126/1, каб. А 224, тел.: 8 (727) 292 58 48, 708 880 77 99, e-mail: vestnik@aues.kz

Issue output

Name of the periodical printed publication	Scientific and technical journal "Bulletin of the Almaty University of Power Engineering and Telecommunications"
Owner of the periodical printed publication	Non-profit joint-stock company "Almaty University of Power Enginnering and Telecommunications named after Gumarbek Daukeyev", Almaty, Kazakhstan
Chief Editor	Professor, candidate of technical sciences Stoyak V.V.
Number and date of the registration certificate and the name of the issuing authority	№ KZ14VPY00024997 from 17.07.2020 Ministry of Information and Social Development of the Republic of Kazakhstan
Periodicity	4 times a year (quarterly)
Serial number and date of publication of a periodical printed publication	Number 74, edition 3, September 30, 2026
Subscription index	74108
Circulation of the issue	200 copies
Price	Negotiable
The name of the printing house, its address	Satbayev University Printing House, 22 Satpayev Street, main building, Room 109, Almaty, Republic of Kazakhstan.
Editorial office address	050013, Non-profit joint-stock company "Almaty University of Power Enginnering and Telecommunications named after Gumarbek Daukeyev", A 224, tel.: 8 (727) 292 58 48, 708 880 77 99, e-mail: vestnik@aues.kz