UDC 004.855.5(045)
DOI:10.18372/1990-5548.77.18002

[1]**V. M. Sineglazov,**
[2]**A. O. Samoshyn**

LONG-TERM DEMAND FORECASTING: USING AN ENSEMBLE OF NEURAL NETWORKS TO IMPROVE ACCURACY

[1]Aviation Computer-Integrated Complexes Department, Faculty of Air Navigation Electronics and Telecommunications, National Aviation University, Kyiv, Ukraine
[2]Educational and Scientific Institute of Applied System Analysis, National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
E-mails: [1]svm@nau.edu.ua ORCID 0000-0002-3297-9060, [2]samoshyn.andriy@lll.kpi.ua

Abstract—*This research paper proposes a method of long-term demand forecasting based on an ensemble of neural networks that considers the novelty of the data. A tool for creating the ensemble was developed that uses a bagging technique as well as a modification that allows for the relevance and novelty of the data to be considered when creating training samples for each model in the ensemble. The study examines and compares the developed method with known approaches to long-term demand forecasting. Experimental results have indicated that the proposed approach allows for obtaining more accurate and reliable demand forecasts compared to existing methods. The results emphasize the importance of data in the demand forecasting process and indicate the potential of the proposed method to eventually improve inventory management strategies and product planning.*

Index Terms—Deep learning; ensemble method; long-term forecasting; demand forecasting; neural networks; multilayer perceptron.

I. INTRODUCTION

In today's economic environment, characterized by rapid change and instability, demand forecasting plays a pivotal role in strategic business management and effective risk management. Long-term demand forecasting is particularly crucial, influencing a company's ability to plan production, manage inventory, develop new products, and adapt to changes in market conditions.

However, existing methods for long-term demand forecasting fall short of ensuring the necessary accuracy and reliability. Unstable market conditions, dynamic changes in consumer preferences, and competitive strategies pose formidable challenges. Traditional methods often struggle to account for the novelty of data, a crucial factor for accuracy and relevance in today's environment [1].

The primary focus of this paper is to address the challenge of developing an effective long-term demand forecasting method that considers the intricate and dynamic changes in the market, as well as the novelty of data. The objective is to create a model that is not only accurate and reliable but also forms the basis for strategic decisions in demand management and business development.

The chosen methodology, based on the use of an ensemble of neural networks, exhibits high potential flexibility and the ability to adapt to complex nonlinear dependencies in demand data. Neural networks facilitate the modeling of interactions between various variables, capturing implicit relationships. Leveraging an ensemble, where multiple neural networks collaborate, helps avoid overtraining and enhances the overall generalizability of the model [2].

A critical challenge in long-term demand forecasting is managing the dynamism and novelty of data. To address this, we modified our chosen methodology by incorporating a mechanism that considers the novelty of the data when creating subsamples for each model in the ensemble. This adaptation enables our model to flexibly adjust to changes in input data, providing stable and accurate forecasts even in dynamic market conditions.

Furthermore, employing neural networks in an ensemble allows us to utilize different architectures and training parameters for each model, expanding the range of potential solutions and increasing the likelihood of identifying optimal forecasting approaches [3].

This innovative approach to long-term demand forecasting holds the potential to emerge as an effective tool for demand management, contributing to more accurate and reliable forecasts, even in unpredictable market conditions.

II. Overview of Existing Forecasting Methods

A. Autoregressive models

The time series method is used to analyze and predict time-dependent data. Autoregressive Integrated Moving Average (ARIMA) models are effective for time series forecasting because they consider trends, seasonality, and random deviations, which can be especially useful in financial forecasting and economic research [4].

B. Exponential smoothing

Exponential smoothing is a method that allows you to consider the rate of change of data and smooth out random fluctuations. By using different types of exponential smoothing, such as single, double, or triple exponential smoothing, you can get accurate predictions for different types of data [5].

Traditional forecasting methods are an important tool for many industries and business segments, but they have their limitations in a fast-paced world.

Machine learning methods have become an important tool in the field of forecasting. Not only do they allow you to analyze large amounts of data, but they also reveal complex relationships and patterns.

C. Linear regression

Linear regression is a method that establishes a linear relationship between a dependent variable (the quantity we are trying to predict) and one or more independent variables (factors that influence the dependent variable). In the case of a single independent variable, linear regression looks like a straight-line equation. Linear regression assumes a linear relationship between the variables. It is fast and simple, but unfortunately too simple for some complex data models. In such cases, we can use polynomial regression [6].

D. K-Nearest Neighbors algorithm

The K-Nearest Neighbors (KNN) algorithm is a versatile and intuitive machine learning technique used for both classification and regression tasks. Unlike parametric models, KNN is a non-parametric, instance-based learning algorithm, meaning it makes predictions based on the similarity between a new data point and its neighboring data points [7].

E. Decision trees and their ensembles

A decision tree is a hierarchical structure where each node branches into two or more branches depending on the value of a particular attribute. The nodes of the tree contain data partitioning conditions that determine which branch will be taken for a particular object. The leaves of the tree contain predicted values that are the result of the recursive branching of attributes and conditions.

The main advantage of decision trees is their ability to model complex dependencies in data, including nonlinearity and interactions between attributes. They are also easy to interpret because you can track each step in the decision-making process [8].

Decision tree ensembles include several trees that are combined to improve the accuracy and stability of the model. The two main types of ensembles are random forests and gradient boosting. In a random forest, several trees are trained independently of each other, and then their predictions are averaged or selected by voting. Gradient boosting, on the other hand, builds successive trees, each based on the errors of the previous ones.

F. Artificial neural networks

Artificial Neural Networks (ANNs) are a powerful class of machine learning algorithms that model the behavior of the human brain to solve various tasks, such as classification, regression, pattern recognition, and prediction. Neural networks are trained using algorithms that determine the optimal weights between neurons. During training, the networks adapt to the input data, improving their weights to best suit the task at hand [9].

G. Recurrent neural networks

Recurrent neural networks (RNNs) are a class of neural networks that are designed to work with sequential data and other time-dependent data. They differ from conventional feed-forward neural networks in that they have internal memory, which allows them to store and use information about previous inputs as context for more complex computations [10].

III. Ensemble Methods

Ensemble methods in machine learning are an approach that combines multiple models to improve the overall accuracy of predictions. This approach is based on the idea that combined solutions from multiple models can be more reliable and accurate than individual models.

Ensemble methods are commonly used for various types of tasks, such as classification, regression, and clustering. The main idea is to create a combination of models that complement each other and compensate for the shortcomings of each model.

Ensemble methods can combine different machine learning models, such as decision trees, neural networks, or regression-based methods, into a single ensemble. This can be done by averaging the results (as in bootstrap aggregation) or by voting, where each model is entitled to one vote in determining the final result [11].

Many ensemble methods use randomization, such as bootstrap sampling, to create multiple training sets for each model. This reduces overfitting and ensures more robust results. In some ensemble methods, models can have weights that determine their importance in the final solution. During voting, models with higher weights have a greater influence on the final result.

Ensemble methods have significant advantages that make them popular in machine learning, but they also have disadvantages that require careful consideration when using them.

Advantages

- One of the key advantages of ensemble methods is their ability to significantly improve the accuracy of predictions. Combining the opinions of multiple models can allow the ensemble to identify complex dependencies in the data that may be missed by individual models.
- Ensemble methods are less prone to overfitting because they consider the opinions of multiple models, which in general can contribute to better generalizability.
- Ensemble methods are usually more robust to noise and random deviations in the training data due to the use of randomization and bootstrap sampling.
- Using different learning algorithms as part of an ensemble can allow you to adapt the approach to different aspects of the data, which improves overall performance.

Disadvantages

- Ensemble methods have many parameters that need to be tuned to achieve the best performance. This can be time-consuming and requires a lot of experimentation.
- Since ensemble methods use multiple models, they can require significantly more computing resources than single models, especially if the data volume is large.

A. Bootstrap aggregation

Bootstrap aggregation, also known as bagging, is one of the most common ensemble methods in machine learning. This method improves model robustness and accuracy by generating multiple training sets using bootstrap sampling [12].

How it works

1) First, a different training set is created for each model in the ensemble by randomly selecting elements from the stored data set and returning them. This process is known as bootstrap sampling, and it allows you to create several training sets for each model.

2) A separate model is trained on each bootstrap set. Since training takes place on different data sets, the models become independent of each other.

3) After all models are trained, their results are aggregated. In the case of classification tasks, this can be done by voting (each model votes for its prediction), and in regression tasks, by averaging the predictions of each model.

B. Gradient boosting

Gradient boosting is a powerful ensemble machine learning method that allows you to create extremely accurate predictions by combining the opinions of several weak models. This method is especially effective for regression and classification tasks [13].

How it works

1) Gradient boosting starts with the creation of a first base model, which can be a simple constant or a small model that predicts the target well.

2) After creating the first model, the prediction error is calculated – the difference between the actual values and the predictions of the first model.

3) The next model tries to correct the residual error left by the first model. This new model learns to predict the residual error and adds its prediction to the results of the first model.

4) This process is iteratively repeated; each subsequent model tries to correct the residual error left by the previous models. This is done through gradient descent, where new models try to reduce the gradient (the difference between actual values and predictions) by learning from the residual error.

C. Stacking

Stacking is an ensemble machine learning method that combines the predictions of several base models with another model called a metamodel [14]. The metamodel acts as a higher-level learner, effectively integrating the outputs of the base models and refining predictions to achieve improved accuracy and generalizability.

How it works

1) First, several base models are created and trained independently on the training data.

2) Each base model makes its own prediction for the input data. These predictions form a prediction matrix, where each column corresponds to a

different model and each row to a different data example.

3) A metamodel is created that takes the prediction matrix of the base models as input. This metamodel is trained on the forecasts of the base models and tries to make a final forecast by reconciling the results of the different base models.

4) The metamodel learns by considering the interaction and connections between the predictions of the underlying models, helping to find the best way to combine their results.

IV. PROPOSED METHOD

The proposed forecasting method is based on an improved bagging that uses an ensemble of neural networks. We use adaptive data selection based on their age to train the base models. Typically, in conventional bagging method, the data for each base model is randomly selected with equal probability. However, in our method, we decided to use adaptive data selection, where older data have a lower probability of being selected for training.

The proposed forecasting method is based on an improved bagging that uses an ensemble of neural networks. We use adaptive data selection based on their age to train the base models. Typically, in conventional bagging method, the data for each base model is randomly selected with equal probability. However, in our method, we decided to use adaptive data selection, where older data have a lower probability of being selected for training.

This is achieved by using the row selection function, which reduces the likelihood of selecting old data in the training subsamples for the base models. Suppose we have a training set X consisting of vectors x with n features, $x = (x_1, x_2, \ldots, x_n)$. And let there be a time feature among these features x_{date} for the corresponding vector, x since we are considering data with time dependence. We use a descending function that allows us to control the probability of selecting an observation into the subsample based on its age, in the form:

$$\text{date\_min} = \min_{x \in X}(x_{\text{date}}),$$

$$w = e^{\frac{x_{\text{date}} - \text{date\_min}}{100}}, \qquad (1)$$

$$p = \frac{w}{\sum_{i=0}^{n} w_i}.$$

The obtained probability values p for each observation x allow us to balance the number of old observations in the training subsamples and have a predominant majority of newer ones.

We have implemented this method as a class that takes the weights for each observation as input parameters and the configuration of the base model for bagging. This way, it is possible to customize the weights for each observation and use any available algorithm as a base model, providing flexibility in choosing models for the ensemble. Figure 1 shows the general scheme of the proposed method.

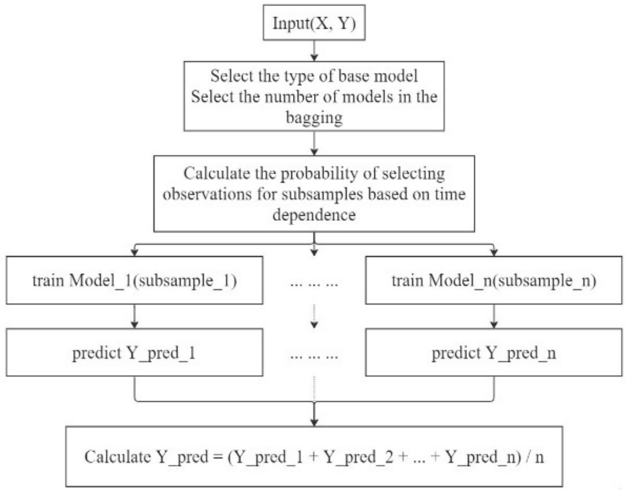

Fig. 1. A general diagram of the time-dependent begging

A multilayer perceptron (MLP) with 25 hidden layers was chosen as the basic model.

V. RESULTS AND COMPARISON OF ALGORITHMS

A. Data processing

For long-term forecasting, the following data granularity was chosen: "administrative region of the country"–"product"–"week". The forecasting horizon is 52 weeks. Forecasting the target variable of total sales. An example of the target variable is shown in Fig. 2.

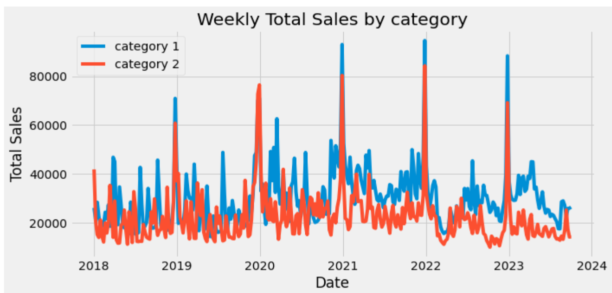

Fig. 2. An example of a target variable

Given the importance of a detailed consideration of the attributes in a long-term demand forecasting system for a chain of stores, here is a description of the basic input attributes:

- the total area of stores in a given area can indicate capacity;

- the total number of cash registers in stores in a given region can indicate their capacity and service capacity;
- the start date of the week is used to aggregate the data and determine the analysis period;
- the number of promotional activities in a given period for the selected product;
- the average base price of the product in the region;
- the average promotional price of a product in the region can be a key factor in determining the benefits for the buyer;
- descriptive features of the product: size, category, brand.

These attributes are important for our study and will allow us to better understand and create new attributes to predict the demand for goods in a chain of stores.

After performing the stationarity test, it was concluded that the series was stationary.

To improve the quality of algorithms, you should try to develop or collect a dataset with the most representative characteristics. Primary data cannot immediately serve as training data [15]. The following features were additionally created:

- the sequential number of the week within the year;
- the sequential number of days within the year;
- the number of the month;
- the number of the calendar quarter;
- cyclic representation of date functions to illustrate the cyclical nature of time (e.g., 12 months is closest to 1 month);
- name of the holiday that was during this week;
- coordinates of the regional center;
- distance from the regional center to the capital;
- population of the region;
- total sales of the product for the previous period;
- average sales of products of the same category and size for the previous period;
- price index based on average category prices.

The studied data contains a set of categorical characteristics that cannot be correctly transferred to machine learning models without certain transformations to a numerical form. For ordinary tabular data, it seems possible to consider most of the popular methods for encoding categorical variables, such as one-hot, label encoding, or methods using a target variable, such as target encoding, etc. That's why one-hot doesn't meet our needs because we have many categories. Using this method will lead to the so-called curse of dimensionality. It is also necessary to consider the nature of the data, namely its time dependence, i.e., only past information should be used in the transformation to avoid data leakage. For this purpose, we will use the CatBoost encoder. It is similar to target encoding but includes the principle of ordering to overcome the problem of looking into the future. It uses a principle similar to time series data validation. The values of the target statistics depend on the observed history, i.e., the target value for the current object is calculated only from the series of observations before it [16].

Thereafter, the data was standardized to allow the use of algorithms that depend on the scale of the features, in particular, various variations of neural networks.

B. Metrics for evaluating the quality of the forecast

To test the accuracy of the forecasting model, this study uses estimation methods such as mean absolute error (MAE), weighted bias (wBias) and weighted mean absolute percentage error (wMAPE) [17]. They are defined by the following formulas:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_{\mathrm{true},i} - y_{\mathrm{pred},i}\right|,$$

$$\mathrm{wBias} = \frac{\sum_{i=1}^{n}\left(y_{\mathrm{true},i} - y_{\mathrm{pred},i}\right)\times \mathrm{weight}_i}{\sum_{i=1}^{n} y_{\mathrm{true},i}\times \mathrm{weight}_i}, \qquad (2)$$

$$\mathrm{wMAPE} = \frac{\sum_{i=1}^{n}\left|y_{\mathrm{true},i} - y_{\mathrm{pred},i}\right|\times \mathrm{weight}_i}{\sum_{i=1}^{n} y_{\mathrm{true},i}\times \mathrm{weight}_i},$$

where $y_{\mathrm{true},i}$ is the true value of the ith observation, $y_{\mathrm{pred},i}$ is the predicted value for the ith observation, and weight_i is the realized price of the ith observation.

C. Results

Tables 1–3 show the quality scores of the implemented bagging and other algorithms on the test sample for different product categories.

Let's take a look at the aggregate forecasts for the developed algorithm, which are shown in Fig. 3 (weeks on the X-axis, total sales on the Y-axis). We can see that the model can generally track changes in sales levels quite well, which is critical when creating a model, and there is no strong bias.

TABLE I. RESULTS ON THE TEST SAMPLE FOR THE OLIVES CATEGORY

Algorithm	MAE	wBias	wMAPE
Linear Regression	21.240999	0.171258	2.089721
Decision Tree	20.523671	0.153452	2.099562
KNN	14.37258	0.002988	1.464464
Light Gradient Boosting Machine	10.265612	0.016345	1.015855
MLP	10.196435	0.023553	0.983567
Our weighted bagging: MLP	10.042824	0.016544	0.964849

TABLE II. RESULTS ON THE TEST SAMPLE FOR THE WATER CATEGORY

Algorithm	MAE	wBias	wMAPE
Linear Regression	84.480005	0.425613	1.390737
Decision Tree	74.053743	0.685274	0.687261
KNN	69.532563	0.22464	0.265432
Light Gradient Boosting Machine	66.621567	0.11069	0.228842
MLP	65.466542	0.14347	0.202356
Our weighted bagging: MLP	65.023453	0.10353	0.188464

TABLE III. RESULTS ON THE TEST SAMPLE FOR THE CHEESE CATEGORY

Algorithm	MAE	wBias	wMAPE
Linear Regression	22.41161	0.583439	1.174249
Decision Tree	26.370057	0.688346	1.373164
KNN	15.777434	0.315182	0.819989
Light Gradient Boosting Machine	9.615076	0.113314	0.494714
MLP	9.605615	0.146724	0.482369
Our weighted bagging: MLP	9.589425	0.127032	0.482391

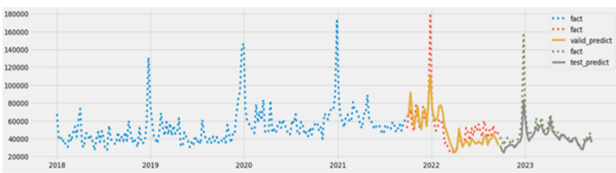

Fig. 3. Prediction on the validation and test sample

VI. CONCLUSIONS

In this research paper, we present a new method for long-term demand forecasting based on an ensemble of neural networks that considers data novelty. Our approach uses a bagging method and modification technique that allows us to consider the relevance and novelty of the data when generating training samples for each model in the ensemble. The results experiments indicated that the proposed method was superior to existing approaches to long-term demand forecasting. The results emphasize the importance of taking the novelty of data into account when developing demand forecasting strategies.

REFERENCES

[1] K. Aaltonen, "Project stakeholder analysis as an environmental interpretation process", International Journal of Project Management, vol. 29, no. 2, pp. 165–183, Feb. 2011 [Online]. Available: https://doi.org/10.1016/j.ijproman.2010.02.001

[2] G. Perin, Ł. Chmielewski, and S. Picek, "Strength in numbers: Improving generalization with ensembles in machine learning-based profiled side-channel analysis", IACR Transactions on Cryptographic Hardware and Embedded Systems, pp. 337–364, Aug. 2020. [Online]. Available: https://doi.org/10.46586/tches.v2020.i4.337-364

[3] P. Flach, "Model ensembles", in Machine Learning. Cambridge: Cambridge Univ. Press, pp. 330–342. [Online]. Available: https://doi.org/10.1017/cbo9780511973000.013

[4] G. M. Jenkins, G. C. Reinsel, and G. E. Box, Time Series Analysis: Forecasting & Control. Pearson Educ. Asia Limited, 2005.

[5] E. S. Gardner, "Exponential smoothing: The state of the art—Part II", International Journal of Forecasting, vol. 22, no. 4, pp. 637–666, Oct. 2006. [Online]. Available: https://doi.org/10.1016/j.ijforecast.2006.03.005

[6] H. F. E. Jr, Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer London, Ltd., 2015.

[7] R. Xu and D. WunschII, "Survey of clustering algorithms", IEEE Transactions on Neural Networks, vol. 16, no. 3, pp. 645–678, May 2005. [Online]. Available: https://doi.org/10.1109/tnn.2005.845141

[8] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward", Plos One, vol. 13, no. 3, Mar. 2018, Art. no. e0194889. [Online]. Available: https://doi.org/10.1371/journal.pone.0194889

[9] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, "Modeling long- and short-term temporal patterns with deep neural networks", in SIGIR '18: The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval, Ann Arbor MI USA. New York, NY, USA: ACM, 2018. [Online]. Available: https://doi.org/10.1145/3209978.3210006

[10] G. P. Zhang and M. Qi, "Neural network forecasting for seasonal and trend time series", European Journal of Operational Research, vol. 160, no. 2, pp. 501–514, Jan. 2005. [Online]. Available: https://doi.org/10.1016/j.ejor.2003.08.037

[11] Z.-H. Zhou, "Ensemble learning", in Machine Learning. Singapore: Springer Singap., 2021, pp. 181–210. [Online]. Available: https://doi.org/10.1007/978-981-15-1967-3_8
[12] F. Petropoulos, R. J. Hyndman, and C. Bergmeir, "Exploring the sources of uncertainty: Why does bagging for time series forecasting work?", European Journal of Operational Research, vol. 268, no. 2, pp. 545–554, Jul. 2018. [Online]. Available: https://doi.org/10.1016/j.ejor.2018.01.045
[13] S. Touzani, J. Granderson, and S. Fernandes, "Gradient boosting machine for modeling the energy consumption of commercial buildings", Energy and Buildings, vol. 158, pp. 1533–1543, Jan. 2018. [Online]. Available: https://doi.org/10.1016/j.enbuild.2017.11.039
[14] "Ensemble learning", in Machine Learning Fundamentals. Cambridge Univ. Press, 2021, pp. 203–218. [Online]. Available: https://doi.org/10.1017/9781108938051.012
[15] X. Wang, Y. Kang, F. Petropoulos, and F. Li, "The uncertainty estimation of feature-based forecast combinations", J. of the Operational Research Society, pp. 1–15, Mar. 2021. [Online]. Available: https://doi.org/10.1080/01605682.2021.1880297
[16] J. Ding, Z. Chen, L. Xiaolong, and B. Lai, "Sales forecasting based on catboost", in 2020 2nd International Conference on Information Technology and Computer Application (ITCA), Guangzhou, China, Dec. 18–20, 2020. IEEE, 2020. [Online]. Available: https://doi.org/10.1109/itca52113.2020.00138
[17] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice. OTexts, 2021.

Received August 31, 2023

Sineglazov Victor. ORCID 0000-0002-3297-9060. Doctor of Engineering Science. Professor. Head of the Department Aviation Computer-Integrated Complexes, Faculty of Air Navigation Electronics and Telecommunications, National Aviation University, Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (1973).
Research area: Air Navigation, Air Traffic Control, Identification of Complex Systems, Wind/Solar power plant, artificial intelligence.
Publications: more than 700 papers.
E-mail: svm@nau.edu.ua

Samoshyn Andrii. Bachelor.
Educational and Scientific Institute of Applied System Analysis, National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," (2022).
Research interests: artificial neural networks, programming.
E-mail: samoshyn.andriy@lll.kpi.ua.

В. М. Синєглазов, А. О. Самошин. Довгострокове прогнозування попиту: використання ансамблю нейромереж для підвищення точності

У статті пропонується метод довгострокового прогнозування попиту на основі ансамблю нейронних мереж, який враховує новизну даних. Було розроблено інструмент для створення ансамблю, який використовує техніку пакетування, а також модифікацію, яка дозволяє враховувати релевантність і новизну даних під час створення навчальних зразків для кожної моделі в ансамблі. У дослідженні розглянуто та порівняно розроблену методику з відомими підходами до довгострокового прогнозування попиту. Експериментальні результати показали, що запропонований підхід дозволяє отримати більш точні та надійні прогнози попиту порівняно з існуючими методами. Результати підкреслюють важливість даних у процесі прогнозування попиту та вказують на потенціал запропонованого методу для остаточного покращення стратегій управління запасами та планування продукції.
Ключові слова: глибоке навчання; ансамблевий метод; довгострокове прогнозування; прогнозування попиту; нейронні мережі; багатошаровий персептрон.

Синєглазов Віктор Михайлович. ORCID 0000-0002-3297-9060.
Доктор технічних наук. Професор. Завідувач кафедрою.
Кафедра авіаційних комп'ютерно-інтегрованих комплексів, Факультет аеронавігації електроніки і телекомунікацій, Національний авіаційний університет, Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (1973).
Напрям наукової діяльності: аеронавігація, управління повітряним рухом, ідентифікація складних систем, вітроенергетичні установки, штучний інтелект.
Кількість публікацій: більше 700 наукових робіт.
E-mail: svm@nau.edu.ua

Самошин Андрій Олександрович. Бакалавр.
Навчально-науковий інститут прикладного системного аналізу, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна.
Освіта: Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна, (2022).
Напрям наукової діяльності: штучні нейронні мережі, програмування.
E-mail: samoshyn.andriy@lll.kpi.ua