УДК 330.4:004.8
C32, C52, C53
ORCID ID 0000-0002-0197-4179
ORCID ID 0000-0003-2574-3815
DOI https://doi.org/10.17721/tppe.2023.47.11

Liashenko O., Doctor of Economics, Professor,
Taras Shevchenko National University of Kyiv, Ukraine
Yakymchuk B., PhD Student,
Taras Shevchenko National University of Kyiv, Ukraine

THE APPLICATION OF TIME-SERIES FORECASTING MODELS IN GROCERY RETAIL INDUSTRY

The grocery retail landscape in Ukraine has witnessed profound transformations, driven by disruptions like the COVID-19 pandemic and full-scale invasion, leading to unstable consumer behavior and market dynamics. In response, forecasting models must evolve to consider stochastic exogenous factors, such as blackout periods and air alarms. This study explores advanced time series forecasting models and proposes a comprehensive framework for optimal model selection.

The study introduces the Neural Prophet, a model that combines interpretability and predictive power by incorporating components like non-periodic trends, periodic seasonality, holiday effects, and regressors. The research methodology involves a comparative analysis of classical time series forecasting methods, machine learning regression approaches, and neural networks. Noteworthy models include LightGBM, RNN, TCN, and Neural Hierarchical Interpolation for Time Series (N-HiTS). Optuna hyperparameter optimization and k-fold cross-validation enhance model accuracy.

The study applies the proposed framework to forecast order quantities in the e-commerce segment of the Ukrainian grocery retail company. The system accommodates diverse factors like weather, holidays, and promotions, providing robust decision support. Anomalies are detected using the IQR method, and missing values are filled using Exponentially Weighted Moving Average. Results show the Neural Prophet consistently outperforming other models in 65% of cases, emphasizing its superiority. However, a complete transition to neural models results in reduced accuracy, highlighting the need for a nuanced approach based on data characteristics. The study presents a sophisticated framework for forecasting accuracy, supporting effective operational decision-making. Future research should explore ensemble methods while maintaining computational efficiency, aligning with the ongoing pursuit of optimized forecasting accuracy for informed decision-making in grocery retail.

Keywords: *time-series forecasting, machine learning, grocery retail, artificial neural networks.*

Problem statement. In recent years, the Ukrainian grocery retail landscape has undergone profound transformations, initially spurred by the unprecedented disruptions caused by the COVID-19 pandemic and subsequently exacerbated by a full-scale invasion. These events have induced substantial shifts in consumer behavior, giving rise to novel trends within the grocery retail market. Consequently, the market dynamics have become markedly less predictable, imposing heightened complexity levels on operational and strategic planning endeavors.

In light of these transformative influences, the traditional focus on forecasting internal initiatives, projects, and local strategies must now be expanded to encompass the nuanced

impacts of stochastic exogenous factors. This broader scope requires a careful analysis of external factors, encompassing various elements such as blackout periods and air alarms, among others. Such stochastic elements introduce an additional layer of uncertainty, demanding a more sophisticated and comprehensive approach to analytical methodologies within the grocery retail sector. The integration of these diverse considerations into forecasting models becomes imperative for fostering resilience and adaptability in the face of an increasingly volatile and unpredictable market environment.

Literature review. Forecasting stands as an extensively researched and applied discipline in both academic and industrial spheres. The precision of forecasting plays a pivotal role in enhancing the accuracy of predicting future events, a particularly crucial facet in industries like retail. This significance arises from the inherent unpredictability characterizing future product sales in the retail sector. The need for accurate forecasting is paramount in navigating the dynamic nature of consumer demand, where uncertainties can exhibit sudden and pronounced variations. As such, the pursuit of precision in forecasting methodologies becomes not just a pursuit of academic interest but an operational imperative for industries, especially in the retail domain, to effectively anticipate and respond to the dynamic shifts in market conditions.

The study of N. Arunraj, D. Ahrens, M. Fernandes, and M. Müller [1] emphasizes the importance of implementing an efficient sales forecasting system that addresses both accuracy and uncertainty as a fundamental measure to mitigate food waste. The authors propose the development of a Seasonal Autoregressive Integrated Moving Average with external variables (SARIMAX) model. This model aims to account for various demand-influencing factors, such as promotions, discounts, holidays, and weather, in addition to the traditional time series components like level, trend, and seasonality.

The research of T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu [4] delves into the application of predictive analytics in Supply Chain Management, with a particular focus on demand forecasting within the retail sector. The paper aims to provide an overview of different approaches in retail SCM. It compares the quality of two selected methods, namely SARIMA (Seasonal Autoregressive Integrated Moving Average) and LSTM (Long Short-Term Memory) models. To enhance forecasting accuracy at the store level, the authors propose hybrid approaches, suggesting the training of SARIMA(X) and LSTM on similar, pre-clustered store groups.

The study of B. Kumar Jha and S. Pande [10] focuses on supermarket sales prediction to enhance business performance, aiding decision-making processes within this problem domain. It delves into various forecasting models, including the additive, Autoregressive Integrated Moving Average (ARIMA), and Facebook (FB) Prophet models. The key findings of the research suggest that the FB Prophet model outperforms other models in terms of accuracy, with lower errors, improved prediction capabilities, and better fitting. The authors highlight the potential of FB Prophet as a superior prediction tool for supermarket sales based on the evaluation metrics employed in the study.

The article “A Comparative Study on Forecasting of Retail Sales” by M. Hasan, M.A. Kabir, R.A. Shuvro, and P. Das [6] addresses the challenging task of predicting product sales for large retail companies, acknowledging the complex nature of factors such as

volatile trends, seasonalities, and unforeseen events like the COVID-19 outbreak. The paper offers a comprehensive theoretical overview and analysis of state-of-the-art time-series forecasting models. The models evaluated include the traditional Autoregressive Integrated Moving Average (ARIMA) and advanced models such as the Prophet model developed by Facebook and the Light Gradient Boosting Machine (LightGBM) model. The comparative analysis provides insights into the strengths and weaknesses of different forecasting models, contributing valuable information for improving the accuracy and efficiency of retail sales predictions.

The research of Y. Ensafi, S.H. Amin, G. Zhang, and B. Shah [3] explores the application of various forecasting models on the sales history of a retail store. The research employs both classical time-series forecasting techniques, such as Seasonal Autoregressive Integrated Moving Average (SARIMA) and Triple Exponential Smoothing, and more advanced methods, including Prophet, Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN). The results highlight the superiority of the Stacked LSTM method over other techniques, indicating its effectiveness in accurately predicting sales. Additionally, the study acknowledges the strong performances of the Prophet and CNN models in the context of sales forecasting.

In recent advancements in time series forecasting research, with a specific focus on retail sales, various approaches and models have emerged to enhance forecasting accuracy. A noteworthy observation gleaned from the outcomes of this research lies in the varied classes of models yielding optimal results for distinct datasets. The authors, in their analyses of selected datasets, have discerned superior performances from models representing disparate categories, thereby underscoring the nuanced nature of retail sales forecasting. The identification of optimal forecasting outcomes across divergent models underscores the intricate interplay of factors influencing retail sales, further emphasizing the dynamic and complex nature of the retail landscape. Such variability in model effectiveness not only highlights the necessity for a tailored approach but also underscores the ongoing need for methodological refinement and innovation in the pursuit of heightened forecasting accuracy. The research, by virtue of unveiling these diverse efficacies, contributes significantly to the scientific discourse surrounding time series forecasting in the realm of retail sales.

Unresolved part of the problem. The imperative to advance time series forecasting methodologies in the context of retail sales underscores an exigent requirement for a more robust and stable approach. This mandate arises from the inherent complexity of retail dynamics, characterized by fluctuating trends, dynamic seasonalities, and unforeseen disruptions. Addressing this challenge necessitates not only a refined understanding of the underlying variability within time series but also an exploration of mechanisms enabling the automatic selection of optimal models tailored to the idiosyncrasies of specific datasets.

The quest for stability in forecasting models involves an intricate examination of historical patterns and their persistence over time. This pursuit is underscored by the recognition that temporal variations in retail sales demand a nuanced and adaptable modeling framework capable of accommodating diverse patterns, thereby mitigating the inherent volatility within the retail environment.

Moreover, the notion of automatic model selection introduces a dimension of sophistication in forecasting methodologies. This entails the development of intelligent algorithms or frameworks that can discern the most fitting model for a particular time series based on an amalgamation of factors such as historical context, variability, and stability. Such an automated selection process not only streamlines the forecasting workflow but also ensures a more accurate alignment between model and data characteristics. In a scientific context, the call for a stable approach and automatic model selection necessitates a convergence of statistical rigor and computational intelligence. The synthesis of these elements becomes paramount in developing advanced forecasting frameworks that transcend the limitations of traditional models, paving the way for a more adaptive, resilient, and scientifically grounded approach to time series forecasting in the intricate domain of retail sales.

The purpose of the research is to proffer a methodological framework for the comparative evaluation of time series models, facilitating the judicious selection of the most optimal model to enhance forecasting accuracy.

Research methodology. This study undertakes a comparative analysis encompassing conventional time series forecasting methodologies and their advanced counterparts, classified as follows:

1. Classical Time Series Forecasting Methods;
2. Machine Learning Regression Approaches;
3. Integration of Artificial Neural Networks;
4. Fusion of Time Series Decomposition and Neural Networks;

Within the ambit of classical methodologies employed in this investigation, notable techniques include Exponential Smoothing, ARIMA (Autoregressive Integrated Moving Average), and the Theta model. Exponential smoothing is a time series forecasting method that assigns exponentially decreasing weights to past observations, emphasizing recent data more heavily. This adaptive approach enables the model to capture evolving trends and patterns, focusing on more recent occurrences. The Simple Exponential Smoothing is apt for predicting data that lacks evident trends or discernible seasonal patterns [9]:

$$\hat{y}_{T+1|T} = \alpha y_T + (1-\alpha)\hat{y}_{T|T-1} = \sum_{t=0}^{T-1} \alpha (1-\alpha)^t y_{T-t} + (1-\alpha)^T l_0 \qquad (1)$$

where y_T - forecasted value at period T, the parameters: $\alpha \in [0;1]$ - the smoothing parameter and l_0 - initial fitted value are obtained using maximum likelihood estimation (MLE) method.

Holt and Winters expanded the Simple Exponential Smoothing method to accommodate trend and seasonality. The Holt-Winters seasonal method involves a forecast equation and three smoothing equations for the level, trend, and seasonal components, with two variations based on the nature of the seasonal component. The additive method is suitable when seasonal variations remain relatively constant, whereas the multiplicative method is preferable when seasonal variations change proportionally to the series level. The structure of Holt-Winters' additive method can be described in the following way [9]:

$$\hat{y}_{T+h|t} = l_t + hb_t + s_{t+h-m(k+1)}$$

$$\underbrace{l_t = \alpha\left(y_t - s_{t-m}\right) + \left(1-\alpha\right)\left(l_{t-1} + b_{t-1}\right)}_{Level\ component}$$

$$\underbrace{b_t = \beta\left(l_t - l_{t-1}\right) + \left(1-\beta\right)b_{t-1}}_{Trend\ component} \quad (2)$$

$$\underbrace{s_t = \gamma\left(y_t - l_{t-1} - b_{t-1}\right) + \left(1-\gamma\right)s_{t-m}}_{Seasonal\ component}$$

where h - number of steps to be forecasted, b_t - trend parameter, $\alpha, \beta, \gamma \in [0;1]$ - smoothing parameters for level, trend, and seasonality component accordingly, m - the frequency of data, $k = \left[\left(h-1\right)/m\right]$.

ARIMA (Autoregressive Integrated Moving Average) represents a sophisticated and widely utilized statistical approach for time series forecasting. Comprising autoregressive and moving average components, this method involves differencing the time series to achieve stationarity, identifying autoregressive and moving average parameters, and subsequently forecasting based on these determined parameters:

$$y_t = c + \underbrace{\sum_{p=1}^{P} \varphi_t y_{t-p}}_{AR(AutoRegression)} + \underbrace{\sum_{q=1}^{Q} \theta_q \varepsilon_{t-q}}_{MA(Moving\ Average)} + \varepsilon_t \quad (3)$$

where y_t - forecasted value at period t, ε_t - white noise, P - order of the autoregressive part, Q - order of the moving average part.

Variations of ARIMA models extend the traditional framework to accommodate additional complexities and factors influencing time series data. SARIMA extends ARIMA to address seasonality in time series data. SARIMAX incorporates external or exogenous variables that may impact the time series. These variations offer enhanced flexibility for modeling time series data by considering both seasonality and external factors.

Theta Model, a more straightforward yet effective forecasting technique, involves capturing the trend and seasonality patterns within a time series. Particularly suited for data exhibiting consistent seasonality, the theta model incorporates a level parameter and a seasonality parameter, offering a straightforward yet valuable tool for forecasting in certain contexts [8]:

$$\hat{y}_{T+H|T} = l_t + hb = l_0 + b\left(h - 1 + \frac{1}{\alpha} - \frac{\left(1-\alpha\right)^n}{\alpha}\right) \quad (4)$$

where h - number of steps to be forecasted

Machine Learning Regression Approaches are represented in the study with the LightGBM (Light Gradient Boosting Machine), which stands as a cutting-edge gradient boosting framework within the realm of machine learning. Developed by Microsoft, it has garnered substantial attention and popularity for its exceptional performance, scalability, and efficiency, especially when handling extensive datasets. This framework is built on the principles of gradient boosting, employing tree-based learning algorithms to enhance

predictive modeling capabilities. The method is based on a search for the optimal model, which can be initialized [5]:

$$F_0 = \arg\min_{\gamma} \sum_{i=1}^{n} L\left(y_i, F\left(x\right)\right) \tag{5}$$

where $L\left(y, F\left(x\right)\right)$ - the differential loss function.

Over the subsequent M steps, pseudo-residues are computed, and a one-dimensional optimization problem is resolved to ascertain the multiplier for executing the requisite descent and find the optimal model $\hat{F}_M\left(x\right)$:

$$F_m\left(x\right) = F_{m-1}\left(x\right) + \gamma_m h_m\left(x_i\right) \tag{6}$$

One of the distinctive features of LightGBM lies in its ability to distribute and efficiently train on large datasets, making it well-suited for applications where scalability is crucial. The underlying algorithm optimizes training efficiency by utilizing a histogram-based approach for binning data, resulting in reduced computational costs and accelerated model training times. This efficiency is particularly advantageous in scenarios where expeditious model development and deployment are paramount.

Leveraging neural networks provides an additional capacity for effectively modeling time series data. Recurrent Neural Networks (RNNs) are a class of neural networks designed explicitly for handling sequential or time series data. RNNs are particularly well-suited for processing sequences of data points. In the context of time series modeling, each element in the sequence corresponds to a specific time point, making it a natural fit for temporal data [7].

Temporal Convolutional Networks (TCN) is another algorithm that represents a class of neural networks designed for time series modeling. The TCN algorithm incorporates temporal convolutions to capture and process sequential patterns. TCN often employs dilated causal convolutions. “Causal” implies that the output at any time step depends only on the present and past inputs, not the future. “Dilated” convolutions allow the network to have a broader receptive field without significantly increasing computational cost. TCN typically consists of multiple stacked convolutional layers. Each layer captures different levels of temporal abstractions, with higher layers learning more complex and longer-range dependencies [11].

Another algorithm included in the research is Neural Hierarchical Interpolation for Time Series (N-HiTS), which combines multi-rate input sampling and hierarchical interpolation techniques to enhance long-horizon time-series predictions. The model operates in the univariate regime, accepting only the predicted time-series history, and demonstrates superior performance compared to previous Transformer-based multi-variate models while requiring significantly less computation [2].

Within the mentioned models, a mix of conventional interpolation methods and models employing artificial neural networks exists. The NeuralProphet is a method that adeptly merges interpretability with predictive prowess. The Facebook Prophet algorithm excels in simultaneously capturing seasonality, the impact of external factors, and the effects of initiatives promoting organic growth to generate forecasts. The model relies on a

decomposable time-series structure, incorporating components such as a non-periodic trend $g(t)$, periodic seasonality (weekly, monthly, yearly) $s(t)$, the influence of holidays, and additional regressors $r(t)$ [12]:

$$y(t) = g(t) + s(t) + h(t) + r(t) + e_t \tag{7}$$

1. The trend component represents the overarching trajectory of the data throughout time:

$$g(t) = \frac{c(t)}{1 + e^{-\left(\beta + \delta\alpha(t)^T\right)\left(t - \left(m + \gamma\alpha(t)^T\right)\right)}} \tag{8}$$

where $\alpha(t)^T$ a binary operator for events, indicating the shift in the growth rate with an offset parameter:

$$\gamma_i = \left(s_i - m - \sum_{l<i}\gamma_i\right)\left(1 - {\beta + \sum_{l<i}\delta_l} \Big/ {\beta + \sum_{l\le i}\delta_l}\right) \tag{9}$$

2. The seasonality component adeptly captures recurring patterns inherent in the data, encompassing cycles that may manifest on a daily, weekly, or monthly basis. This capturing process involves decomposition utilizing a Fourier series, allowing for a comprehensive representation of these repetitive patterns within the time series:

$$s(t) = \sum_{n=1}^{N}\left(a_n \cos(2\pi n t) + b_n \sin(2\pi n t)\right) \tag{10}$$

3. The holiday component accommodates the influence of specific events or holidays that could have an impact on the data. Additionally, the model incorporates regressor components, which can be either internal factors such as capital projects and marketing campaigns or external factors like COVID-19 and air sirens. Both holiday effects and regressor components are integrated into the model using a binary operator. This comprehensive approach allows the Facebook Prophet model to account for a diverse range of factors influencing the time series:

$$h_i(t) = L_{holiday_i}(t)\kappa^T; r_j(t) = L_{regressor_j}(t)\kappa^T \tag{11}$$

While Prophet demonstrates robust handling of missing data and outliers, coupled with a variety of tunable hyperparameters for optimization, it can be further enhanced by complementing it with FB Neural Prophet. The Neural Prophet algorithm expands its array of components by incorporating auto-regression effects $AR(t)$ and regression effects tailored to lagged observations of exogenous variables $L(t)$ [13]:

$$y(t) = g(t) + s(t) + h(t) + r(t) + \boxed{AR(t) + L(t)} + e_t \tag{12}$$

In contrast to Prophet, Neural Prophet employs a neural network architecture for the modeling of time-series data. Specifically, it utilizes a multi-layer feedforward network comprising both fully connected layers and gated recurrent units (GRUs). The model undergoes training through a combination of backpropagation and gradient descent algorithms, optimizing the weights and biases to enhance its predictive capabilities.

The proposed algorithms offer several advantages applicable to grocery retail

scenarios, as they exhibit the capability to:

- Effectively handle irregularly sampled data, a prevalent characteristic in numerous real-world applications.
- Model seasonality with multiple harmonics, allowing for the capture of intricate and complex seasonal patterns.
- Automatically detect and manage changepoints within the underlying data-generating process, providing adaptability to shifts or transitions in the time series.

Results and discussion. The devised methodology aimed at constructing a sophisticated system for forecasting order quantities within the e-commerce segment of an advanced Ukrainian grocery retail company. In the modeling process, various factors were incorporated, with notable emphasis placed on weather conditions, including precipitation levels such as rain and snow. Additionally, the model accounted for the impact of holidays and extensive promotional periods, ensuring a comprehensive approach to capturing the diverse influences on order predictions. The dataset encompasses a span of three years and data from 118 stores and dark stores. This dataset encapsulates the operational dynamics of a grocery retail company, providing insights into its activities across diverse regions.

This system possesses the capability to establish supportive frameworks for facilitating effective operational decision-making. It serves as a valuable tool in developing decision support systems that enhance the efficiency and precision of day-to-day operational decisions within the context of the advanced Ukrainian grocery retail company's e-commerce function [14].

As an initial step (Fig. 1), the historical orders are redistributed to align with the current polygon levels, aiming to assess a more consistent flow of orders. This process aims to mitigate the impact of variability associated with the delivery network polygon map, ensuring a more stable and reliable estimation of order patterns. This strategic approach enhances the accuracy of order flow analysis by harmonizing historical data with the present polygon configuration in the delivery network.

Following the initial redistribution of historical orders, the subsequent step involves identifying anomalies through the utilization of the IQR method, which involves using the Interquartile Range (IQR) as a robust measure of dispersion for identifying and handling anomalies or outliers in a dataset. The IQR is a measure of statistical dispersion, representing the range between the first quartile (25th percentile) and the third quartile (75th percentile) of a dataset. In the context of anomaly clipping, the IQR is leveraged to calculate a range within which most of the data points lie. Observations falling outside this range are then considered potential anomalies. This approach is robust because it is less sensitive to extreme values than measures like the standard deviation. It provides a more resilient way to identify and handle anomalies, making it particularly useful in situations where the data may exhibit skewed distributions or contain outliers that could unduly influence other statistical measures.

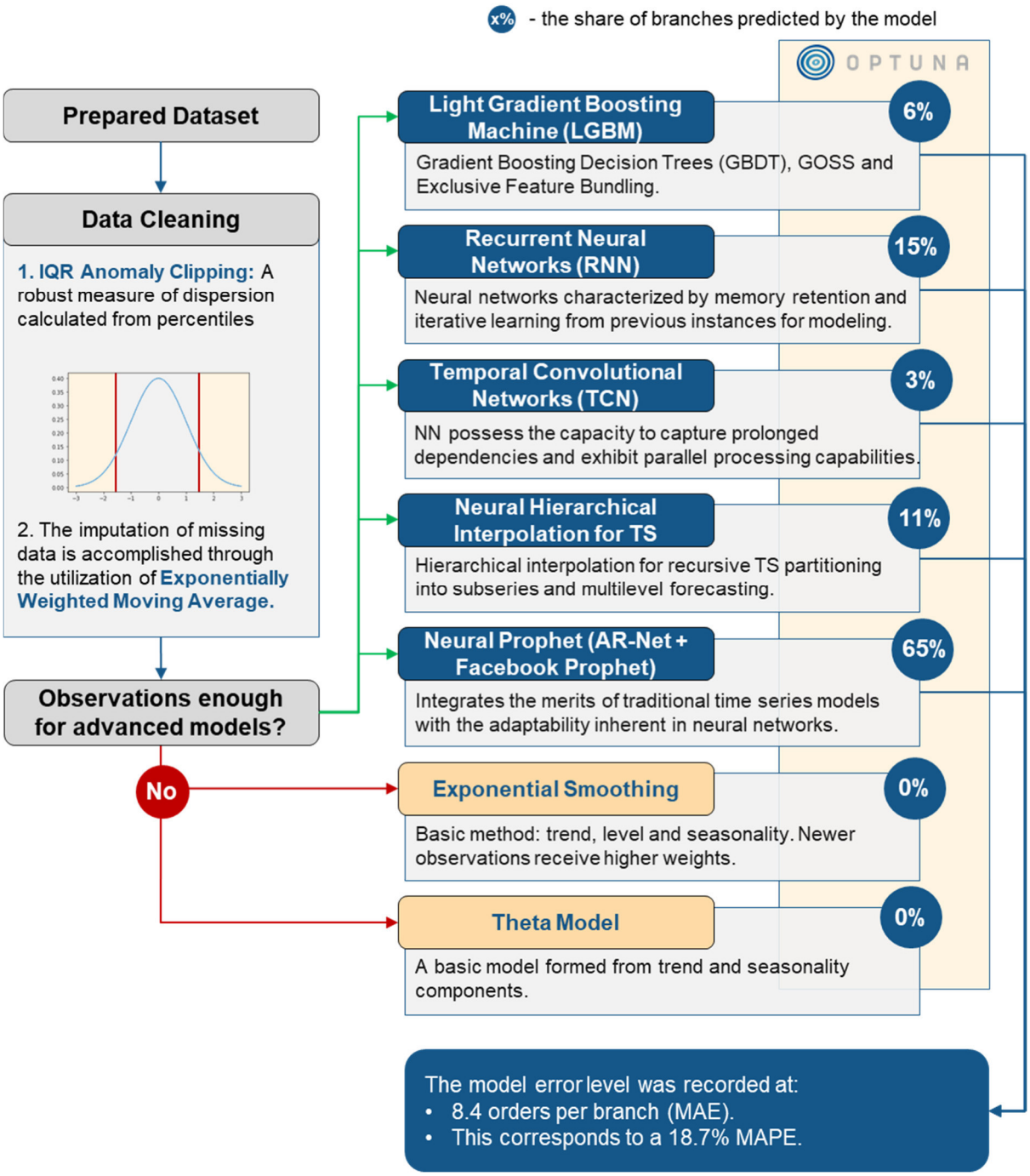


Fig. 1. The framework for selecting the optimal time series forecasting method for evaluating food retailer e-commerce orders

Source: developed by the authors.

Subsequently, the intermediate values are determined and filled in using Exponentially Weighted Moving Average. This two-step process ensures a comprehensive approach to anomaly detection and addresses data gaps through a nuanced filling technique, enhancing the robustness of the overall order flow analysis.

In instances where historical data is limited, the system presents the option to opt for fundamental approximation methods, namely exponential smoothing and the theta model. In the current phase of the study, where all branches possess adequate historical data, advanced time series modeling techniques are employed. However, in scenarios where

the network incorporates a new branch with limited historical data, the system can still provide a forecast, albeit with reduced accuracy, serving as a benchmark for network management.

When sufficient history is available, the system conducts an exhaustive analysis by running various types of models to discern the most accurate option for each branch individually. Optuna hyperparameter optimization framework is used for the automated search for optimal hyperparameters, enhancing the efficiency of model training and improving overall performance. Optuna employs an optimization algorithm to explore the hyperparameter space and find configurations that lead to the best model outcomes. The k-fold cross-validation method is used to assess a model's performance and generalizability. It involves partitioning the original dataset into k equally sized subsets, referred to as folds. The model is then trained and evaluated k times, each time using a different fold as the validation set and the remaining k-1 folds for training. Performance evaluation is carried out through metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The resulting ratings are determined proportionally, providing a comprehensive model accuracy assessment.

Table 1

Estimates of the application of the optimal model selection system

Model	Branches Weight	MAE	MSE	MAPE
LGBM	6,1%	6,8	79	16%
NeuralProphet	65,4%	9,4	164,3	18%
N-HiTS	11,1%	6,4	63,4	20%
RNN	14,6%	7,3	94,2	21%
TCN	2,8%	6,4	68,5	18%
Total	100%	8,4	131,9	19%

The analysis reveals that the Neural Prophet model consistently outperforms other models in approximately 65% of cases (Table 1), demonstrating its superiority. However, it's noteworthy that the remaining 35% is covered by alternative models. An interesting observation is that a complete transition to neural models results in a notable decrease in accuracy, as measured by Mean Absolute Percentage Error (MAPE), by more than 5.4 percentage points. This suggests that while excelling in specific classes of time series, deep learning models may not universally outperform other approaches, emphasizing the importance of considering the nature of the data in model selection.

Conclusions. The study presents a comprehensive framework encompassing data processing, evaluation, and the judicious selection of optimal models from the forefront of time series forecasting methods. This meticulously designed approach not only ensures a refined and robust data processing pipeline but also facilitates the identification and deployment of cutting-edge forecasting models. The strategic combination of advanced methodologies in time series forecasting enables the derivation of superior forecast evaluation results. These enhanced results, in turn, empower stakeholders to make more

effective, informed, and nuanced operational and financial decisions. The intricate interplay of data processing, model evaluation, and selection within this framework underscores its significance as a sophisticated tool for elevating forecasting accuracy and, consequently, decision-making processes.

Prospects for further research. Looking forward, a promising avenue for enhancing the system lies in the exploration of ensemble methods. Ensembles involve the combination of multiple forecasting models to exploit the diverse strengths of individual models, thereby potentially amplifying overall predictive performance. However, a paramount consideration is to implement ensemble techniques without disproportionately escalating the computational time required for forecast calculations. Striking a delicate balance between leveraging the benefits of ensemble learning and maintaining computational efficiency will be a key focal point in the continuous refinement of the forecasting system. This strategic evolution aligns with the ongoing pursuit of optimizing forecasting accuracy without compromising the practicality and timeliness essential for effective decision-making processes.

References

1. Arunraj, N., Ahrens, D., Fernandes, M. & Müller, M. (2014). Time series sales forecasting to reduce food waste in retail industry. *The 34th International Symposium on Forecasting, Rotterdam, The Netherlands.* http://dx.doi.org/10.13140/RG.2.1.4829.1607.
2. Challu, C., Olivares, K. G., Oreshkin, B. N., Garza Ramirez, F., Mergenthaler Canseco, M., & Dubrawski, A. (2023). NHITS: Neural Hierarchical Interpolation for Time Series Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(6), 6989-6997. https://doi.org/10.1609/aaai.v37i6.25854
3. Ensafi, Y., Amin, S.H., Zhang, G. & Shah, B. (2022). Time-series forecasting of seasonal items sales using machine learning – A comparative analysis. *International Journal of Information Management Data Insights*, 2 (1). https://doi.org/10.1016/j.jjimei.2022.100058.
4. Falatouri, T., Darbanian, F., Brandtner, P. & Udokwu, C. (2022). Predictive Analytics for Demand Forecasting – A Comparison of SARIMA and LSTM in Retail SCM. *Procedia Computer Science*, 200, 993-1003. https://doi.org/10.1016/j.procs.2022.01.298.
5. Friedman J. H. (2001). Greedy function approximation: a gradient boosting machine, *Ann. Stat*, 29 (5), 1189-1232. http://dx.doi.org/10.1214/aos/1013203451
6. Hasan, M., Kabir, M.A., Shuvro, R.A., & Das, P. (2022). A Comparative Study on Forecasting of Retail Sales. ArXiv. https://doi.org/10.48550/arXiv.2203.06848
7. Hewamalage, H., Bergmeir, C. & Bandara, K. (2021). Recurrent Neural Networks for Time Series Forecasting: Current status and future directions, *International Journal of Forecasting*, Elsevier, 37(1), 388-427. https://doi.org/10.1016/j.ijforecast.2020.06.008
8. Hyndman, R.J. & Billah, B. (2003). Unmasking the Theta method, International Journal of Forecasting, 19 (2), 287-290. https://doi.org/10.1016/S0169-2070(01)00143-1.
9. Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: principles and practice*, 3rd edition, OTexts: Melbourne, Australia. https://otexts.com/fpp3/. Accessed on 30 November 2023.
10. Kumar Jha, B. & Pande S. (2021). Time Series Forecasting Model for Supermarket Sales using FB-Prophet. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India*. 547-554. https://doi.org/10.1109/ICCMC51019.2021.9418033.
11. Lin, Y., Koprinska I. & Rana, M. (2021). Temporal Convolutional Attention Neural Networks for Time Series Forecasting, *2021 International Joint Conference on Neural Networks (IJCNN)*, Shenzhen, China, 1-8. https://doi.org/10.1109/IJCNN52387.2021.9534351.
12. Taylor S.J., Letham B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37-45. https://doi.org/10.1080/00031305.2017.1380080
13. Triebe O., Hewamalage H., Pilyugina P., Laptev N., Bergmeir C. & Rajagopal R. (2021). *NeuralProphet: Explainable Forecasting at Scale*. https://doi.org/10.48550/arXiv.2111.15397.
14. Yakymchuk, B., Liashenko, O. (2023). Modeling the Resource Planning System for Grocery Retail Using

Machine Learning. *In: Antoniou, G., et al. Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2023. Communications in Computer and Information Science*, 1980. Springer, Cham. https://doi.org/10.1007/978-3-031-48325-7_22.

О. І. Ляшенко, д.е.н, професор,
Київський національний університет імені Тараса Шевченка
Б. Б. Якимчук, аспірант,
Київський національний університет імені Тараса Шевченка

ЗАСТОСУВАННЯ МЕТОДІВ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ В ІНДУСТРІЇ ПРОДУКТОВОГО РИТЕЙЛУ

Ринок продуктового ритейлу в Україні зазнав глибоких трансформацій, викликаних такими пандемією COVID-19 і повномасштабним вторгненням, що призвело до нестабільної поведінки споживачів і динаміки ринку. У відповідь на волатильність ринку моделі прогнозування повинні розвиватися, щоб враховувати стохастичні екзогенні фактори, такі як періоди блекауту та повітряні тривоги. У цьому дослідженні описані просунуті моделі прогнозування часових рядів і пропонується комплексна основа для вибору оптимальної моделі.

В дослідженні представлено модель NeuralProphet, яка поєднує в собі можливість інтерпретації та прогностичну силу шляхом включення таких компонентів, як неперіодичні тренди, сезонність, ефекти свят і регресори. Методологія дослідження передбачає порівняльний аналіз класичних методів прогнозування часових рядів, регресійних підходів машинного навчання та нейронних мереж. Варті уваги моделі включають LightGBM, RNN, TCN і нейронну ієрархічну інтерполяцію для часових рядів (N-HiTS). Оптимізація гіперпараметрів за допомоги Optuna та k-fold крос-валідаційної перевірки підвищують точність моделі.

Дослідження застосовує запропоновану структуру для прогнозування кількості замовлень у сегменті електронної комерції української компанії на ринку продуктового ритейлу. Система враховує різноманітні фактори, такі як погода, свята та акції, забезпечуючи надійну підтримку прийняття рішень. Аномалії виявляються за допомогою методу IQR, а відсутні значення заповнюються за допомогою експоненційно зваженого ковзного середнього. Результати показують, що NeuralProphet стабільно перевершує інші моделі в 65% випадків, підкреслюючи його прогностичну перевагу. Однак повний перехід від нейронних моделей призводить до зниження точності, підкреслюючи необхідність індивідуального підходу на основі характеристик даних. Дослідження представляє складну структуру для точності прогнозування, що підтримує ефективне прийняття оперативних рішень. У майбутніх дослідженнях слід вивчити ансамблеві методи, зберігаючи при цьому ефективність обчислень, узгоджуючи з поточним прагненням оптимізувати точність прогнозування для прийняття обґрунтованих рішень у роздрібній торгівлі продуктами харчування.

Ключові слова: *прогнозування часових рядів, машинне навчання, продуктовий ритейл, штучні нейронні мережі.*