Innovative Sales Forecasting: Utilizing Fuzzy Neural Networks for Enhanced Sales Prediction

Marcin Bartkowiak
Poznań University of Economics and Business
Poznań, Poland *m.bartkowiak@ue.poznan.pl*

Piotr Cyplik
Poznań University of Technology
Poznań, Poland *piotr.cyplik@put.poznan.pl*

Tomasz Górecki
Adam Mickiewicz University
Poznań, Poland *tomasz.gorecki@amu.edu.pl*

Adam Karolewski
SmartStock Sp. z o.o.
Poznań, Poland *akarolewski@smartstock.cloud*

Abstract

This study aims to improve retail sales forecasting using fuzzy neural networks (FNNs). Traditional methods often miss complex sales patterns. We use accuracy and loss metrics to apply FNNs to the Walmart sales dataset, comparing them to conventional time series models and advanced techniques like LightGBM and LSTM. Comprehensive data preprocessing ensures data quality. FNNs handle uncertainties and complex relationships better, outperforming traditional methods. The findings suggest that FNNs enhance forecasting accuracy, supporting informed decision-making in retail.

Keywords: fuzzy neural networks, sales forecasting, machine learning

1. Introduction

Sales forecasting has traditionally relied on time-series models like ARIMA (AutoRegressive Integrated Moving Average) and ETS (Error, Trend, Seasonal), which predict future demand based on historical data. These forecasting approaches predict future demand based on historical time series. However, these methods have limitations: (1) They require long historical data to capture seasonality, which is often unavailable for new products. (2) Sales data often contains outliers and missing values, requiring preprocessing. (3) Many additional features impacting sales must be considered.

While FNNs have been applied to sales forecasting, early applications were limited to single and short time series and predate advancements in DNNs [16, 18]. Genetic FNNs use genetic algorithms to generate initial weights and learn fuzzy rules from marketing experts [15]. Hybrid FNN models combine FNNs with other methods like time series analysis or genetic algorithms to improve accuracy [14]. Adaptive FNNs adjust parameters in real-time based on data changes [10]. Multi-step forecasting with FNNs projects several steps ahead, aiding long-term forecasts [23]. Some studies incorporate external factors, such as economic indicators, social media sentiment, or weather, to enhance accuracy [20, 1]. Online and incremental learning with FNNs allows models to update incrementally, useful in dynamic data scenarios [6].

Our paper used fuzzy neural networks (FNNs) to predict sales for Walmart data (M5 competition) [11], incorporating a deep learning fuzzy network to handle complex sales prediction scenarios with multiple variables, non-linear relationships, and uncertainty. We examined the impact of external factors such as economic and weather conditions and compared our model with other methods: classic time series models [12]: NAIVE and Seasonal NAIVE (SNAIVE), ETS, Holt-Winters (HW), Theta, ARIMA, the classic machine learning model LightGBM [13], Temporal Fusion Transformer (TFT) model [22] and LSTM model [9]. LSTM networks, noted for their sequence processing capabilities [5], retain information from previous data points, aiding in time series analysis. We used the TFT model for its ability to handle complex temporal dynamics and multi-horizon forecasting [17, 19].

2. Methods

2.1. Dataset preprocessing

If necessary, we imputed missing values using the MICE (Multivariate Imputation by Chained Equations) method [4]. We used well-known and straightforward (fast) Tukey's method [21] to detect outliers. We standardized features using tanh estimators [7].

2.2. Fuzzy neural networks

An FNN is an artificial neural network incorporating fuzzy logic principles. It represents input and output variables with fuzzy sets, allowing for the representation of uncertainty and vagueness in data. FNNs define membership functions to determine the degree of membership of values to specific fuzzy sets. They are beneficial for uncertain or imprecise data, addressing the non-transparency of traditional neural networks, which are often criticized as black-box models [3]. This lack of transparency can create a communication gap between analysts and deep learning networks (DNNs). [8, 2] proposed a DNN combined with fuzzy systems, producing a novel deep neuro-fuzzy system (DNFS) that reduces uncertainty using fuzzy rules.

2.3. Quality measures

During finding and optimizing models, we used the following measures of quality [12]: MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), WAPE (Weighted Absolute Percentage Error), MAD (Median Absolute Deviation) and RMSE (Root Mean Squared Error).

3. Machine learning framework for the forecasting sales

The retail market is highly competitive, with sales influenced by various factors like seasonality, trends, promotions, price, and even temperature. Sales data varies in client behavior and volume, requiring complex solutions for accurate forecasting and KPI targets. However, modeling for each product is time and resource-intensive. Therefore, the modeling phase must include profiling and tuning to balance these conflicting requirements. A detailed presentation of the process flow is presented in Fig. 1.

4. Application to financial market

4.1. Data description and preprocessing

We used M5 forecasting challenge data from Kaggle [11]. The data represent the retail perspective well in various products, sales points, and sales nature (types of time series). There is a history of daily sales from 10 stores Walmart from 2011 to 2016. The data are daily, and the forecast is to be made per day (date) based on the total sales of the item in the store. The data comprises the following features: (1) Product (item) and product category. (2) Stores and state (location). (3) Amount of sales and price. (4) Public holidays, cultural and sporting events, reli-

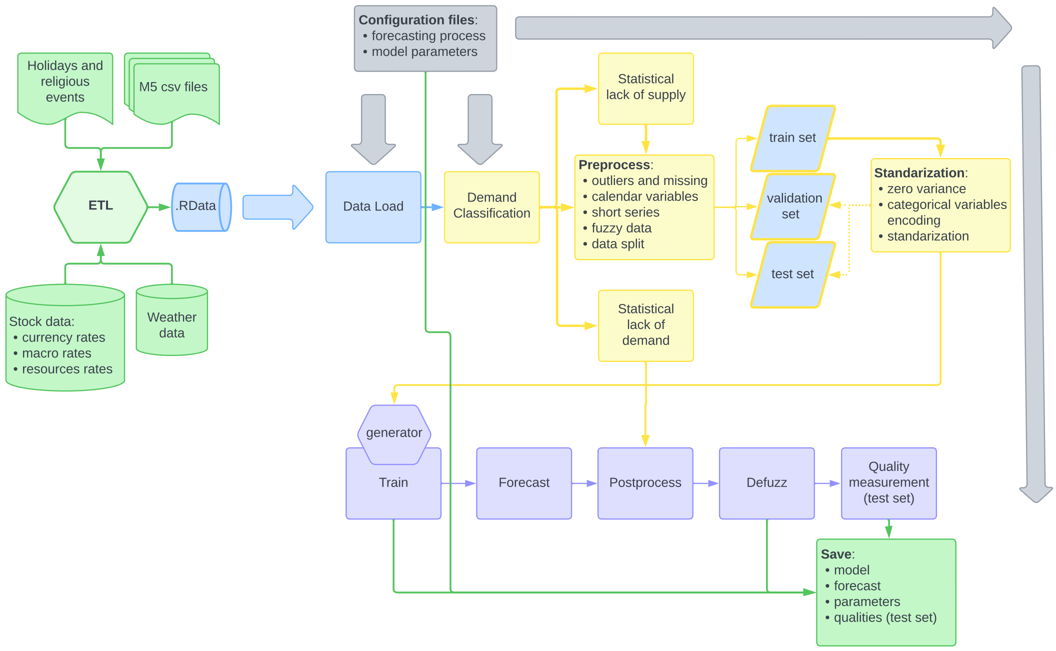

Fig. 1. Process flowchart

gious holidays and observances. (5) SNAP (Supplemental Nutrition Assistance Program) – The program is administered at the state level, and each state has its own monthly deposit schedule.

We prepared our calendar version, adding new events and implementing lag and lead effects based on data exploration and common sense. Calendar variables included day of the week, week number, month, quarter, and year. We incorporated weather data by averaging parameters (temperature, precipitation, wind) for states (CA, TX, WI). Additional variables included currency rates (USDGBP, USDJPY, USDCHF, USDEUR), commodity prices (copper, oil, wheat, milk, sugar, cotton, pork, lumber), and stock/index rates (UMCCUS.M, GDPYUS.M, WHIMUS.M, RSAMUS.M, HONSUS.M, UNRTUS.M, CPIYUS.M, DL.F, XR.F, SW.F, CB.F, LS.F, HG.F, CT.F, ZW.F, 10USY.B). This resulted in 30 490 products (time series) with 82 variables each. The recursive model can predict multiple variables (1 to 7 in our case). In the post-process, we defuzzified variables into one variable, yielding 853 720 forecasts for each product over a 28-day horizon.

4.2. Experimental results

The quality measures for all compared models are shown in Table 1. We observe that we do not have a clear winner. From a business point of view, it seems that MAE and MAPE are the most important measures. The fuzzy model wins in terms of MAE, and LSTM wins MAPE. Notably, classic time series models perform significantly worse than machine learning models. The results from the Fuzzy approach are promising. The Fuzzy model has the lowest MAE, the second-best result in MAPE, and is among the best in the case of RMSE. It is worth noting that several fuzzy models were tested, differing in fuzzification and defuzzification. It was decided on the most auspicious – we use the triangular membership function for fuzzification and the center of area method (COA), also commonly referred to as the centroid method, for defuzzification. The explanation for the good results of the fuzzy model can be seen in the fact

Table 1. Model comparison

Model	MAE	RMSE	MAPE	WAPE	MAD
ARIMA	1.058	1.376	0.630	0.641	0.594
ETS	1.063	1.380	0.640	0.642	**0.588**
HW	1.189	1.560	0.611	0.684	0.723
NAIVE	1.373	1.788	0.522	0.789	**0.588**
SNAIVE	1.396	1.964	0.486	0.836	0.884
THETA	1.085	1.411	0.630	0.649	0.607
LightGBM	1.035	**1.370**	0.584	0.650	0.648
LSTM	1.063	1.468	**0.443**	0.736	0.690
TFT	1.146	1.418	0.557	**0.548**	0.646
FUZZY	**1.027**	1.411	0.479	0.727	0.651

that the fuzzy approach helps neural networks perform better in unexpected situations.

5. Conclusions

For further research, the authors plan to use sales forecasts in inventory management. Determining the safety stock level requires the sales forecast and the RMSE error, which correlates positively with the safety stock level. Due to the difficulty in identifying the best forecasting model from the results (Table 1), a hybrid analysis method will be used. This method involves recalculating forecasts and RMSE errors for each product using all models in Table 1. The model with the smallest RMSE error will be chosen. This approach will identify conditions where Fuzzy Neural Networks outperform other models.

Acknowledgment

The paper is a result of the study conducted under the project “Supply chain management software system for improvement of the accuracy of forecasts and optimization of inventory from the perspective of recipient and supplier with the usage of fuzzy, deep neural networks” funded by The National Centre for Research and Development [POIR.01.01.01-00-1140/19].

References

[1] Asghar, M. Z., Subhan, F., Ahmad, H., Khan, W. Z., Hakak, S., Gadekallu, T. R., and Alazab, M.: Senti-eSystem: A sentiment-based eSystem-using hybridized fuzzy and deep neural network for measuring customer satisfaction. In: *Software: Practice and Experience* 51.3 (2021), pp. 571–594.

[2] Aviles, A. I., Alsaleh, S. M., Montseny, E., Sobrevilla, P., and Casals, A.: A Deep-Neuro-Fuzzy approach for estimating the interaction forces in Robotic surgery. In: *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. 2016, pp. 1113–1119.

[3] Buhrmester, V., Münch, D., and Arens, M.: Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey. In: *Machine Learning and Knowledge Extraction* 3.4 (2021), pp. 966–989.

[4] Buuren, S. van: Multiple imputation of discrete and continuous data by fully conditional specification. In: *Statistical Methods in Medical Research* 16 (2007), pp. 219–242.

[5] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., and Schmidhuber, J. In: *IEEE Trans. Neural Netw. Learn. Syst.* 28.10 (2017), pp. 2222–2232.

[6] Guan, H.: Neural Network Model of Dynamic Prediction of Cross-Border E-Commerce Sales for Virtual Community Knowledge Sharing. In: *Computational Intelligence and Neuroscience* (2022), p. 4286148.

[7] Hampel, F. R., Ronchetti, E., Rousseeuw, P. J., and Stahel, W. A.: Robust statistics: the approach based on influence functions. Wiley, 1986.

[8] Hayashi, Y.: Black Box Nature of Deep Learning for Digital Pathology: Beyond Quantitative to Qualitative Algorithmic Performances. In: *Lecture Notes in Computer Science*. Springer, 2020, pp. 95–101.

[9] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory. In: *Neural Computation* 9.8 (1997), pp. 1735–1780.

[10] Hou, S., Fei, J., Chen, C., and Chu, Y.: Finite-Time Adaptive Fuzzy-Neural-Network Control of Active Power Filter. In: *IEEE Trans. Power Electron.* 34.10 (2019), pp. 10298–10313.

[11] Howard, A. and Makridakis, S.: *M5 Forecasting – Accuracy*. 2020. URL: `https://kaggle.com/competitions/m5-forecasting-accuracy`.

[12] Hyndman, R. J. and Athanasopoulos, G.: Forecasting: Principles and Practice. OTexts, 2021.

[13] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y.: LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: *Advances in Neural Information Processing Systems*. NIPS'17. 2017, pp. 3149–3157.

[14] Khashei, M., Reza Hejazi, S., and Bijari, M.: A new hybrid artificial neural networks and fuzzy regression model for time series forecasting. In: *Fuzzy Sets and Systems* 159.7 (2008), pp. 769–786.

[15] Kuo, R.: A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm. In: *European Journal of Operational Research* 129.3 (2001), pp. 496–517.

[16] Kuo, R. and Xue, K.: Fuzzy neural networks with application to sales forecasting. In: *Fuzzy Sets and Systems* 108.2 (1999), pp. 123–143.

[17] Lim, B., Arık, S. O., Loeff, N., and Pfister, T.: Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. In: *International Journal of Forecasting* 37.4 (2021), pp. 1748–1764.

[18] Liu, Y. and Liu, L.: Sales Forecasting through Fuzzy Neural Networks. In: *International Conference on Electronic Computer Technology*. 2009, pp. 511–515.

[19] Song, H., Rajan, D., Thiagarajan, J., and Spanias, A.: Attend and Diagnose: Clinical Time Series Analysis Using Attention Models. In: *AAAI Conf. on Artificial Intelligence* 32.1 (2018).

[20] Tu, Q., Zhang, Q., Zhang, Z. j., Gong, D., and Tang, M. c.: A Deep Spatio-Temporal Fuzzy Neural Network for Subway Passenger Flow Prediction with COVID-19 Search Engine Data. In: *IEEE Transactions on Fuzzy Systems* (2022), pp. 1–1.

[21] Tukey, J. W.: Exploratory Data Analysis. Addison-Wesley, 1977.

[22] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is All you Need. In: *Advances in Neural Information Processing Systems*. Vol. 30. 2017.

[23] Yao, J., Lu, B., and Zhang, J.: Multi-Step-Ahead Tool State Monitoring Using Clustering Feature-Based Recurrent Fuzzy Neural Networks. In: *IEEE Access* 9 (2021), pp. 113443–113453.