A Hybrid Model for Demand Forecasting Based on the Combination of Statistical and Machine Learning Methods

Fadoua Ouamani[1(✉)], Asma Ben Fredj[1,2], Mohamed Rayen Fekih[2], Anwar Msahli[2], and Narjès Bellamine Ben Saoud[1]

[1] RIADI Research Laboratory, National School of Computer Science, University of Manouba, Manouba, Tunisia
fadoua.ouamani@ensi-uma.tn
[2] COGNIRA, Imm. PREMIUM, Les Berges du Lac, Tunis, Tunisia
https://cognira.com/

Abstract. Demand Forecasting (DF) is nowadays a key component of successful businesses in retailing field. In fact, accurate customer's demand forecasts and insights into the reasons driving the forecasts may increase confidence, assist decision-making and therefore boost's the retailer's profit. It is then crucial for an accurate DF model to not only understand the retail time-series repeated patterns but also the impacts of factors such as the promotions on the data behavior. The literature review of existing research works has shown that statistical models gave good results in accurately detecting time-series components such as seasonality or trend but they fail when it comes to detecting external factors or causal effects compared to machine learning models. Moreover, the combination of both models either focused only on the trend component and neglected the seasonality or considered both of them but used sophisticated neural networks, which are computationally expensive. To this end, in this paper, we propose an approach that combines statistical and machine learning models to take advantages of their aforementioned properties. We used first Multiple Linear Regression (served as the baseline model as well) and linear interpolation to remove the promotions effect from the data and compute promotional multipliers. Then, each resulting data was fed to two statistical models (Prophet and Exponential Triple Smoothing). Finally, the combination step consisted in reintegrating the promotions effect into the forecasting results of each statistical model. Quantitative and qualitative evaluations of the hybrid models' performance showed that the hybrid models outperformed the baseline model.

Keywords: Retailing · Data mining · Demand forecasting · Time-series · Statistical models · Machine learning · Hybrid model · Prophet

1 Introduction

In today's fast-paced world, data collection, understanding and mining has become the key to staying ahead of the competition, achieving success and boosting profits in many

business fields. In retailing for example, it helps retailers gathering the most useful information about their products and customers, forecasting the customer's demand, understanding the reasons behind sales actions and behaviors and making of the right business decision. In retailing, the data is daily recorded and is organized into three-dimensional hierarchies: time hierarchy (year, half of year, month, week, and day), product hierarchy (: group, department, class, subclass and Stock Keeping Unit-SKU) and location hierarchy (chain, area, region, district and store). Moreover, Retail data is usually time-series data which is a collection of data sorted chronologically by a time index. It has three components:1)trend component (T(t)) which shows the overall tendency of the data to increase, decrease or remain stable through long periods, 2) seasonal component (S(t)) which represents regularly spaced fluctuations having almost the same pattern and magnitude during the same period every year and 3) irregular component (I(t)) which refers to the unpredictable sudden changes in data. Time-series (Y(t)) may be additive (Y(t) = T(t) + S(t) + I(t)) or multiplicative (Y(t) = T(t)*S(t)*I(t)). When the magnitude of the seasonal pattern evolves with the trend then the time series is multiplicative.

Demand forecasting make then use of this data history to forecast the amount of goods or services that will be sold in the future or in response to particular causal-effects such as the promotions. In real world scenarios, the demand forecasting is based on the Eq. (1) where the Y is the demand to forecast, Level is the average selling unit in the absence of seasonality and Promotional_mutliplier is the ratio representing the increase in sales within the presence of a promotion.

$$Y = \text{Level} \times \text{Seasonality} \times \text{Promotional\_multiplier} \tag{1}$$

Due to the particularity of retail data, it is crucial for a demand forecasting model to not only understand the time series repeated patterns (trend and seasonality) but also the impact of factors such as the promotions on the behavior of the data. However, the existing research works on retail time-series forecasting propose statistical models, machine learning models or a combination of both models that do not satisfy the two aforementioned requirements simultaneously. In fact, statistical models focus on detecting the time series components such as seasonality and trend, whereas machine learning models have proven high effectiveness in handling causal effects. To this end, this paper proposes a new approach to retail demand forecasting that combines both models to handle causal effects and detect time series components with low complexity and easily interpretable models.

The remainder of the paper is then organized as follows: Sect. 2 will be dedicated to the review of existing research works on retail time-series forecasting and discussion of the findings. This will lead us to the paper's contributions described in Sect. 3 in terms of a revisited KDD (Knowledge Discovery from Databases) that goes through different steps: from data collection and preprocessing to model evaluation. Section 4 will be then devoted to the presentation of the efforts made to carry out the first step. Once data is ready, it's fed to two machine learning techniques (Multiple Linear regression and Linear Interpolation) to remove the promotions effect and compute promotional multipliers; which will be described in Sect. 5. Section 6 will present the statistical models used, namely Prophet and Exponential Triple Smoothing (ETS), their configurations and

results of their trainings. Section 7 will finally present the model combination process and comparison results of the performance of the different hybrid models based on both quantitative and qualitative evaluations.

2 Related Work

Several retail time-series forecasting models have been proposed in the literature by following two different approaches: single model-based approach and model combination based approach. In this section, we will briefly present the research works using each approach, which help us to draw important conclusions and come up with our contribution to the field.

2.1 Statistical Models

Different types of statistical models were utilized for retail time-series forecasting. MA(Moving Average), ARIMA (Auto Regression Moving Average), SARIMA(Seasonal ARIMA), ETS(Exponential Triple Smoothing), Prophet, to cite but a few. In [1], the authors have used ARIMA which is an optimized version of MA, as MA cannot handle all time-series components and thus the high dimensionality variable's space and business related information about promotion and price fluctuation. ARIMA was deployed for demand forecasting as an important component of supply chain management process but it was incapable of capturing the seasonality pattern within the time-series data. In an attempt to solve this problem, researchers in [2] have used SARIMA for red lentils market price forecasting. Reference [3] proposes a ETS model to forecast the monthly retail sales of consumer goods. The model has shown good results with seasonal data. In [4], the researchers focused on forecasting monthly sales of different items using Prophet and creating a product portfolio based on the forecast reliability. Reference [5] has compared the performance of Prophet to ARIMA based on the MAPE (Mean Absolute Percentage Error) evaluation metric calculation. The performance of Prophet was slightly better than ARIMA in forecasting stock prices of a unique individual. In [6], the authors compared classical Prophet model to its optimized version by modifying the default Fourier order used to detect the seasonality to meet the particularity of the training data. Both models were used for daily sales forecasting.

2.2 Machine Learning Models

An alternative to statistical methods for retail time-series forecasting, are machine learning method, including deep learning methods. In [7], the authors compared the forecasting results of prophet and SARIMA to machine learning model LightGBM (Light Gradient Boosting Machine) on daily sales of Walmart stores. It has been shown that the latter model outperformed the former models based on the calculation of RMSE (Root Mean Squared Error) evaluation metric. The authors relate this difference to the fact that the data used is highly affected by external factors, which statistical models failed to capture. Whereas in [8], SARIMA outperformed ARIMA, Holt-Winters and ANN when forecasting Amazon's daily sales. The authors brought these results to the fact

that the date is characterized by a strong seasonality pattern and was less affected by external factors. More sophisticated deep learning techniques were also proposed in [9], in which the authors compared LSTM (Long-Short Term Memory), RNN (Recurrent Neural Networks) and CNN (Convolutional Neural Network) models' performance on stock price prediction. As CNN use the information of the particular instant when forecasting unlike the others models and since stock markets are subjects to sudden changes, it captured the effect of external factors and outperformed the other models.

2.3   Combined Models

Reference [10] provides a literature review on retail sales forecasting, in which single models and combined models were compared to show better results among combined models. In [11], the authors propose an ensemble model for demand forecasting that combines Regression, Exponential smoothing, Holt-Winters and ARIMA models. The method examines the performance of each model over time and combines the weighted forecasts of the best performing models. In [12], the researchers used both MA and ANN models. MA was used to remove the trend effect from the sale time-series data to show only the differences in values from the trend. This has enhanced the performance of ANN model as it will be able to understand and capture accurately causal effects. Moreover, in [13], the authors have also proposed an hybrid model that uses LSTM and RF (Random Forest) for demand forecasting in multi-channel retail. The authors started by applying LSTM but they found that a residual caused by non-temporal explanatory causes remained. Thus, they applied RF on these unexplained components which allow them to outperform the single models applied on the same data.

To sum up, the research works studied have shown that statistical models gave good results in accurately detecting time-series data components, namely trend and seasonality but they fail when it comes to detecting external factors or causal effects compared to machine learning models. The reason why some research works ([12] and [13]) have combined them. However in [12], the authors focused only on the trend component and neglected the seasonality component. Whereas in [13], the authors have considered all the time-series data components but they have used sophisticated Neural Networks, which are computationally expensive.

In fact, for optimization and interpretability sake, and as stated by the authors of [14, 15], and due to the huge scale of item-level sales in the retail business, it is crucial to apply simple forecast methods and time-applicable approaches based on a decomposable models that can be updated gradually with only a limited amount of new sales data. Based on these findings, we propose a new approach for retail demand forecasting that will ensure a trade-off between the following three objectives: Handling causal effects, detecting all time-series data components and proposing a decomposable approach with low complexity and explainable results. To the best of our knowledge, this is the first investigation in such an approach that satisfies simultaneously these three criteria.

3   The Proposed Approach

The proposed approach (see Fig. 1) is based on a hybrid model that will allow us to capture both time-series data components and causal effects, and mainly the promotion

effects on demand forecasting. The hybrid model will combine machine learning models and statistical models. We will then use Multiple Linear Regression (MLR) Model, not only for its capability to detect causal effects, but also for its easily interpretable outputs. This model will serve also as a baseline model. Linear interpolation will be the second model that we will use to eliminate the promotion's effect. The idea is to remove the actual value of the data in presence of the promotions; then linearly link the rest of the data points in order to have an approximation of the data with absence of promotion. Statistical models like ETS, Prophet and SARIMA will be then used on the resulting data (without promotion's effect). The choices were based on the findings from the literature review and motivated by the fact that these models were commonly used and have given accurate results when handling seasonality and trend components of time-series data where external factors are removed.

After preparing the data, we will first apply MLR to get the coefficients that will be used to compute the promotions multipliers. The multipliers will be then used to remove the promotion effect from the data. We will also apply Interpolation to get another version of the data without promotional effect. The resulting data will be fed to statistical models after being preprocessed accordingly to meet the requirements of each model. The combination step will reintegrate the promotional effect into the time-series data by multiplying it with the promotion multipliers whenever there is a promotion to reconstitute the predicted time-series with the correct magnitude. Finally, the results of each combination will be compared to the baseline model.

4   Data Generation, Preprocessing and Exploration

The real data was inaccessible due to confidentiality issues. That being said, the data on which the process was applied is a generated data that represents a simulation of a grocer.
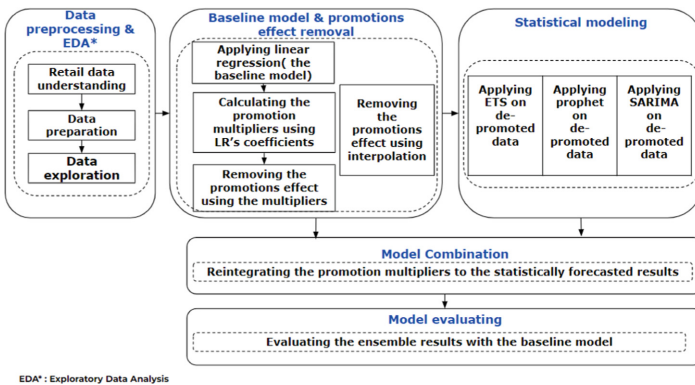

Fig. 1. The proposed approach for demand forecasting

4.1   The Dataset

The generated data is described by 5 parquet files as follows:

1. The product file: data about 100 products, 12 classes, 24 subclasses and 5 different departments;
2. The stores file: data about 30 stores, 4 regions in 1 country;
3. The customers file: data about 15000 customers and their purchasing frequency;
4. The calendar file: data about the time hierarchy of 1093 days (order of the week, the month in a year, the name of the year, the days of weeks and the weeks of years (WOY))
5. The Tlog file: data about the transactions (4702056 transactions) containing the action of buying a product by a customer from a store in a certain week within a promotion category: ′promo_cat′ and its discount ′promo_discount′.

4.2   Data Preprocessing

A data pipeline of three steps (joining, aggregation and cleaning) was applied on the raw data to transform it into exploitable one. Specific additional techniques were used depending on the model to meet its requirements (see Subsects. 6.1 and 6.2).

In the first step, we joined the transactional data in the Tlog file with their corresponding time information from the calendar file, the region and country information, and the products data. Then, we aggregated the transactions having the same combination of SKU, store, and week to the same row and filled it with the corresponding features in addition to a sales feature containing the sum of these transactions.

The common resulting input for the next steps is composed of categorical features and numerical features. We have as categorical features, the promotion category, at the product level, we have subclasses and 100 SKU, at the location level, we have 4 Regions an 30 stores, at the time level, we have the year (2018 to 2020), the WOY (52) and the Week (from 01-01-2018 to 12-28-2020). The numerical features are the target variable which is the Units and the feature variable which is the promotion discount.

4.3   Exploratory Data Analysis (EDA)

In an attempt to gain insight into time-series data components and their behaviors a plot is proposed (see Fig. 2.). The seasonality peaks in the plot occurred in the same period, a pattern that was observed in most SKUs. Same plots for the other subclasses have shown almost the same results. The seasonality phenomenon was observable at the subclass level and it is additive as its magnitude doesn't change significantly over time. Whereas, the trend component was absent.

The promotion effect (compared to a baseline plot that represents the median of non-promoted sales; green line in Fig. 2) was observable through the variation of behavior of the time-series data, with a peak in the fourth week of February. In fact, promotion effect varies according to the promotion categories (BXGY, BXGX, 50%_off, 25%_off, 10%_off). The average of estimated promotion multipliers has shown that 50%_off had the biggest effect on sales. This allows us to conclude that there is a positive correlation between the estimated promotion multiplier and the promotion category.
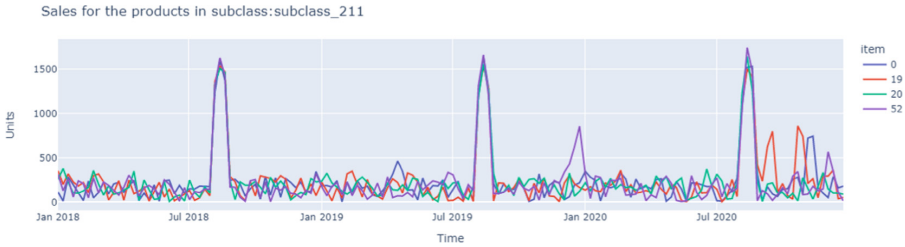
Fig. 2. Plot showing the sum of sales for the products in subclass 211

5 Removing the Promotions Effect Using MLR and Linear Interpolation

Since statistical models are univariate and unable to model causal effects, the promotions effect was first removed from data using two different techniques: MLR and linear interpolation. Customized data preparation was done before feeding the data to the MLR model. First, data was sorted by time to assure a non-erroneous forecast. Then, the discount was transformed into a discount elasticity using the Eq. (2)

$$\text{discount\_elasticity} = 1 - \text{discount} \tag{2}$$

After that, log based data scaling was applied to reduce the vast range of the data and to fit it to the additive model using the Eq. (3). Finally, one-hot encoding was applied to the categorical features to make them understandable by the model.

$$log((\widehat{y_t})) = \beta_1 log(x_{1t}) + \beta_2 log(x_{2t}) + \beta_3 log(x_{3t}) + \cdots + \beta_n log(x_{nt}) + log(e_t) \tag{3}$$

$$(\widehat{y_t}) = e^{(\beta_1 log(x_{1t}) + \beta_2 logx_{2t} + \beta_3 log(x_{3t}) + \cdots + \beta_n log(x_{nt}) + log(e_t))}$$

5.1 Multiple Linear Regression Coefficients Generation and Promotions Effect Removal

MLR has more than one predictor variable as shown by the Eq. (4) unlike simple Linear regression.

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_n x_{n,t} + \varepsilon_t \tag{4}$$

β_1, β_2, β_3,.., β_n are the coefficients for the predictor variables $x_{1,t}$, $x_{2,t}$, $x_{3,t}$,.., $x_{n,t}$. For ϵ_t, it represents the unobserved random variable

Before training the model, we selected the following three parameters: 1) fit_intercept: True (β0 must be calculated for each execution of the model); 2) normalize: False (data is already scaled using the log); 3) positive: False (the coefficients can be negative or positive). The model was applied with subclass pooling. Prediction results (see Fig. 3) using MLR showed infinite values prediction (after july 2020). This
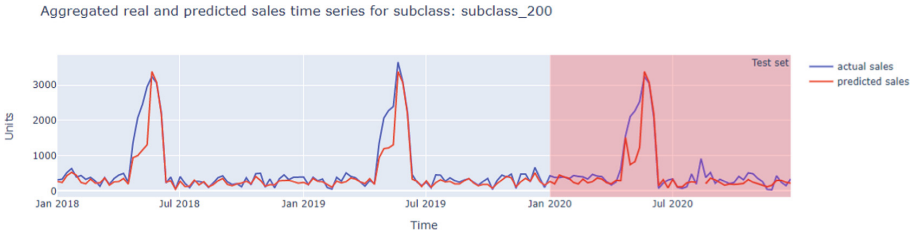
Fig. 3. Excerpt from the actual and predicted aggregated sum of sales within the subclass 200 using MLR with subclass pooling

is due to the correlation between input features as depicted by EDA. The problem was solved by using only three values for promo_cat: percentage, BXGX and BXGY.

The MLR coefficients obtained were then used to compute the promotion multipliers. The promotion multipliers will be used in turn to remove the promotions effect from data by dividing the sales data by their corresponding ones. Figure 4. Shows that the higher the discount value is, the stronger the promotional effect gets as well as its coefficient. The coefficients are capturing accurately the effects and are leading to accurate promotion multipliers; thus an effective promotions effect removal.

Fig. 4. Excerpt from the actual time-series and time-series without promotions effect using MLR coefficients for the product 65 in store 27

5.2   Interpolation Based Promotion Effect Removal

The promotion effect removal was also applied using the linear interpolation. The results of the two techniques look similar. However, on the one hand, using only linear interpolation is risky since it is basically linking linearly the data without understanding the category, discount, or any other information about the promotion effect. On the other hand, this linear linking can be useful as it outperforms the MLR based approach in some cases where the coefficients are relatively not accurate. Moreover, there are no specific metrics to evaluate the two approaches. Thus, we decided to use both of them in the next steps and integrate them into the final benchmark.

6   Statistical Modelling

Statistical models require as input a univariate time series containing only two columns, one for the ordered dates and the other for their corresponding values. To do so, the data is first sorted by week feature. Then, log scaling is applied on the units if the statistical model is additive. Finally, time-series data version is created using the data and units columns.

6.1   Prophet Model

Prophet model is capable of handling trend, seasonality, and holiday effects, which occur on an irregular basis over a day or a period of days. To do so, it decomposes the time-series y(t) into trend (g(t)), seasonality (s(t)), holiday (h(t)) and error using the Eq. (5)

$$y(t) = g(t) + s(t) + h(t) + e_t \tag{5}$$

The model was then configured as follows: Seasonality_mode was set to additive (default value) as we previously detected additive seasonality within the data; Yearly_seasonality was set to True as the presence of repeated pattern of seasonality will be detected in each year; growth was set to its default value (Linear) as the data do not follow a logistic growth; changepoint_range was set also to its default minimum value (0.6) as our data doesn't contain the trend component, so the model don't need to be flexible to it; prior_scale_changepoints was set to 0.05 as the model will focus less on trend component; changepoint_num was not specified to make the model investigate by itself the number of changepoint in the above fixed range. After training the model, the results show the prophet capability in capturing the seasonality pattern (see Fig. 5). However, they also show a case of under-forecasting. In fact, Prophet gives significant importance to the trend component when forecasting. When in the first period of sales, the seasonal peak was higher than that of the second period; the model assumed the presence of a decreasing trend. The altitude of this trend was high enough to cause an under-forecasting in the test set.

Fig. 5. Excerpt from the Prophet forecasting results on the resulting data after promotion effect removal using MLR coefficient for the product 66 in store 24

6.2   Triple Exponential Model (ETS)

Known as Holt-Winters Exponential Smoothing, ETS presents an improved version of classical Smoothing models (single and double exponential smoothing) since it explicitly accounts for seasonality in the time series. It allows the configuration of the same parameters: α, β, Trend Type, Dampen Type and Damping coefficient ϕ; and adds a new parameter γ to control the seasonal component's influence.

The ETS model requires specific parameters for both the model definition and fitting steps. For model fitting, the parameter optimized will be set to True so that the model will estimate automatically the best fitting parameters (α, β, ϕ, γ, Trend Type, Dampen Type). The parameters chosen for model definition were based on the EDA findings and are summarized as follows: seasonal was set to additive; trend was set to none as there is no trend in our data, freq was set to W as data is weekly, seasonal_period was set to 52 (52 weeks each year). After training the model, the results have shown that ETS was capable of capturing the seasonal pattern accurately (see Fig. 6). However in another excerpts, an over-forecasting case was spotted out as by default the model gives more importance to the recent values when predicting.

Fig. 6. Excerpt from the results of ETS's forecasting on data without promotions effect using MLR coefficients for the product 43 in store 1.

7   Model Combination and Results

7.1   Model Combination Process

The model combination process will combine the promotion coefficients and the predictions. Based on the Eq. (1) (see Sect. 1), the process will encompass four combination approaches as depicted by Fig. 7.
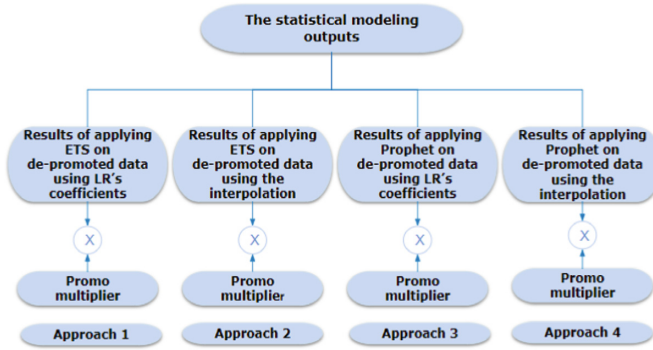
Fig. 7. The model combination process

7.2 Benchmarking Results

The comparison of the four approaches was based on quantitative and qualitative evaluation. A first evaluation effort was based on the use of SMAPE, MAE and RMSE statistical metrics as summarized in Table 1. The four experiments had similar results and all of them outperformed the baseline model. However for approach 3, according to the SMAPE metric, the ensemble model is relatively underperforming the baseline model. Thus, we cannot make a clear assumption based on these results.

Table 1. Results of the quantitative evaluation of models performance

Models	SMAPE	MAE	RMSE
Baseline model	64.394	9.570	5.411
Approach 1	55.616	7.190	4.775
Approach 2	61.450	7.262	4.781
Approach 3	64.480	8.314	4.849
Approach 4	57.200	7.175	4.730

Qualitative evaluation was then carried out since it shows an overview of the model performance on all the SKUs. For the improvement cases, the SKUs and their interpretation are the same for the four approaches as they were capable of detecting both the seasonal component and the promotion effect with almost the same performance. However, for the deterioration cases, they were similar per statistical model. In layman's words, approaches 1 and 2 that used ETS have their common deteriorated SKUs and interpretations. Whereas approaches 3 and 4 which used Prophet have their particular deterioration causes. But, for the majority of deteriorated cases, the main reason is the non-seasonality of the data. This result is understandable since our hybrid models were not designed to accurately forecast this type of data.

8 Conclusion

Through this paper, we have introduced a hybrid modeling approach that combines in different ways MLR and Linear interpolation with Prophet and ETS. The evaluation results of the approach showed that our solution was able to accurately forecast the demand based on both time-series data components and the promotion effect while using a decomposable low complexity models. Even though, the generated data was a good asset to conduct this work, a future work will consider the application of the approach on real world data to obtain more generalized performance evaluation.

References

1. Fattah, J., Ezzine, L., Aman, Z., El Moussami, H., Lachhab, A.: Forecasting of demand using ARIMA model. Int. J. Eng. Bus. Manag. **10**, 1847979018808673 (2018). https://doi.org/10.1177/1847979018808673

2. Divisekara, R.W., Jayasinghe, G.J.M.S.R., Kumari, K.W.S.N.: Forecasting the red lentils commodity market price using SARIMA models. SN Bus. Econ. **1**(1), 1–13 (2020). https://doi.org/10.1007/s43546-020-00020-x

3. Aohan, L.: An empirical analysis of total retail sales of consumer goods based on holt-winters. In: Proceedings of the 2020 2nd International Conference on Big Data and Artificial Intelligence, pp. 54–57. Association for Computing Machinery, New York (2020). https://doi.org/10.1145/3436286.3436298

4. Žunić, E., Korjenić, K., Hodžić, K., Đonko, D.: Application of facebook's prophet algorithm for successful sales forecasting based on real-world data. Int. J. Comput. Sci. Inf. Technol. **12**, 23–36 (2020). https://doi.org/10.5121/ijcsit.2020.12203

5. Garlapati, A., Krishna, D.R., Garlapati, K., Yaswanth, N.M.S., Rahul, U., Narayanan, G.: Stock price prediction using facebook prophet and arima models. In: 2021 6th International Conference for Convergence in Technology (I2CT), pp. 1–7 (2021). https://doi.org/10.1109/I2CT51068.2021.9418057

6. Liço, L., Enesi, I., Jaiswal, H.: Predicting customer behavior using prophet algorithm in a real time series dataset. Eur. Sci. J. ESJ. 17, 10 (2021). https://doi.org/10.19044/esj.2021.v17n25p10

7. Jiang, H., Ruan, J., Sun, J.: Application of machine learning model and hybrid model in retail sales forecast. In: 2021 IEEE 6th International Conference on Big Data Analytics (ICBDA), pp. 69–75 (2021). https://doi.org/10.1109/ICBDA51983.2021.9403224

8. Singh, B., Kumar, P., Sharma, N., Sharma, K.P.: Sales forecast for amazon sales with time series modeling. In: 2020 First International Conference on Power, Control and Computing Technologies (ICPC2T), pp. 38–43 (2020). https://doi.org/10.1109/ICPC2T48082.2020.9071463

9. Selvin, S., Vinayakumar, R., Gopalakrishnan, E.A., Menon, V.K., Soman, K.P.: Stock price prediction using LSTM, RNN and CNN-sliding window model. In: 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1643–1647 (2017). https://doi.org/10.1109/ICACCI.2017.8126078

10. Aras, S., Kocakoç, İD., Polat, C.: Comparative study on retail sales forecasting between single and combination methods. J. Bus. Econ. Manag. **18**, 803–832 (2017). https://doi.org/10.3846/16111699.2017.1367324

11. Akyuz, A.O., Uysal, M., Bulbul, B.A., Uysal, M.O.: Ensemble approach for time series analysis in demand forecasting: ensemble learning. In: 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA), pp. 7–12 (2017). https://doi.org/10.1109/INISTA.2017.8001123

12. Nunnari, G., Nunnari, V.: Forecasting monthly sales retail time series: a case study. In: 2017 IEEE 19th Conference on Business Informatics (CBI), pp. 1–6 (2017). https://doi.org/10.1109/CBI.2017.57

13. Punia, S., Nikolopoulos, K., Singh, S.P., Madaan, J.K., Litsiou, K.: Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail. Int. J. Prod. Res. **58**, 4964–4979 (2020). https://doi.org/10.1080/00207543.2020.1735666

14. Beheshti-Kashi, S., Karimi, H.R., Thoben, K.-D., Lütjen, M., Teucke, M.: A survey on retail sales forecasting and prediction in fashion markets. Syst. Sci. Control Eng. **3**, 154–161 (2015). https://doi.org/10.1080/21642583.2014.999389

15. Seaman, B.: Considerations of a retail forecasting practitioner. Int. J. Forecast. **34**, 822–829 (2018). https://doi.org/10.1016/j.ijforecast.2018.03.001