Sales forecasting of goods in shoe retail

2 authors:

Ivan Kaipov
Belarusian State University

Alexander Nedzved
Belarusian State University

Sales forecasting of goods in shoe retail

Ivan Kaipov, Alexander Nedzved

***Abstract*—**Offline sales in the fashion industry are currently facing intense competition from online sales. In order to successfully exist, offline retailers need to optimize their processes. Sales forecasting is a critical factor for retailers in the shoe industry. Sales forecast close to real values allows the company to prevent sales losses due to a shortage of goods or overcrowded stores with goods. This article discusses methods and approaches for sales forecasting based on machine learning and time series forecasting. The forecasting results obtained using the Prophet turned out to be the most accurate to reality compared to other methods considered. The results obtained allow us to move on to the next step in the transformation of production from a push to a pull strategy, namely forecasting demand as the sum of projected sales and projected lost sales.

***Keywords*—**Forecasting, machine learning, sales, shoe retail, time series.

I. INTRODUCTION

Online commerce with the development of the Internet has received a tremendous impetus to development. This is well seen in companies like Amazon, Alibaba, Ebay, etc. Consequently, there is an outflow of buyers from offline retail to online, which leads to a drop in revenues of companies that only work online. The decline of offline retail is especially noticeable last time. Because many retailers did not see the benefits of digital opportunities. Over the past few years, consumers have realized that going to the store is a much less productive process compared to online shopping. And it does not depend on which online stores are discussed in comparison. Below in Fig. 1 are the indicators that the analytical group PYMNTS received [1]. They compared data from the Census Bureau and other sources and built their own model based on them. Analysts say that stores selling clothing, sports goods and electronics cannot confirm that 90% of retail sales still occur in their premises, if they occur at all.

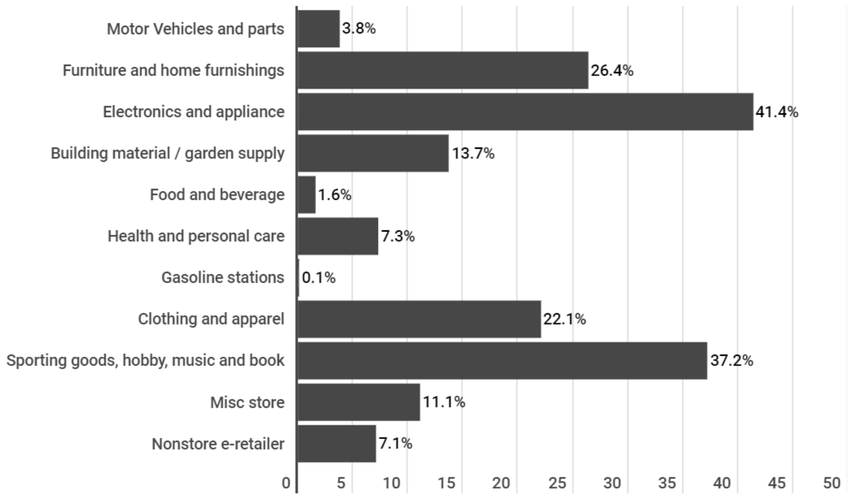

Fig. 1 E-commerce share in retail

Mobile applications and logistics innovations have significantly improved online shopping practices. But the in-store experience has become less reliable. Consumers, who call time the

I. K. Kaipov, Belarusian State University, Minsk, Belarus (e-mail: kaipovIK@bsu.by).
A. M. Nedzved, Belarusian State University, United Institute for Informatics Problems, Minsk, Belarus (e-mail: nedzveda@tut.by).

most valuable resource, want the buying process to be convenient and free from any uncertainties. Trips to the physical store do not always meet these two requirements. Therefore, people prefer to buy online those things that were once more convenient to buy only in physical retail - clothes, jewelry, sports goods, electronics. Increasingly, this also includes home goods and car parts.

A few years ago, analysts expressed the opinion that the future of physical retail as a category and physical stores as points of interaction with consumers will be similar to the evolution in the media sphere: only the most large-scale or highly specialized projects will survive. The largest retailers will be able to take advantage of their branching to provide a range and efficient logistics. This will satisfy the needs of consumers across all channels, including digital.
Physical retail of the future will pass the road of modern media: there will be only players who have adopted new technologies, and business models offering a new look at familiar processes [1]. Thus, sales forecasting also includes many factors that are formed from heterogeneous information. Effective analysis of such data sets requires the use of machine learning methods.

II. Modern Sale Forecasting Models

Shoe sales are some of the hallmarks. First of all, shoe sales have a strong seasonality. Seasonality is noticeable as during the year, there are surges in sales when the season changes. Fig. 2 shows the average change in shoe sales of fifty stores during 2018 relative to the average value of sales. There is a general trend, expressed in a surge in sales in mid-spring, as well as slightly less significant surges during the fall. Sales fall in summer and winter. This is due to the fact that new shoe collections, spring-summer and autumn-winter appear in stores in mid-spring and early autumn, at this time the store shelves are the most complete and the collection sale cycle is at its peak.There is also seasonality during the week, peak sales occur on Friday and Saturday (Fig. 3).

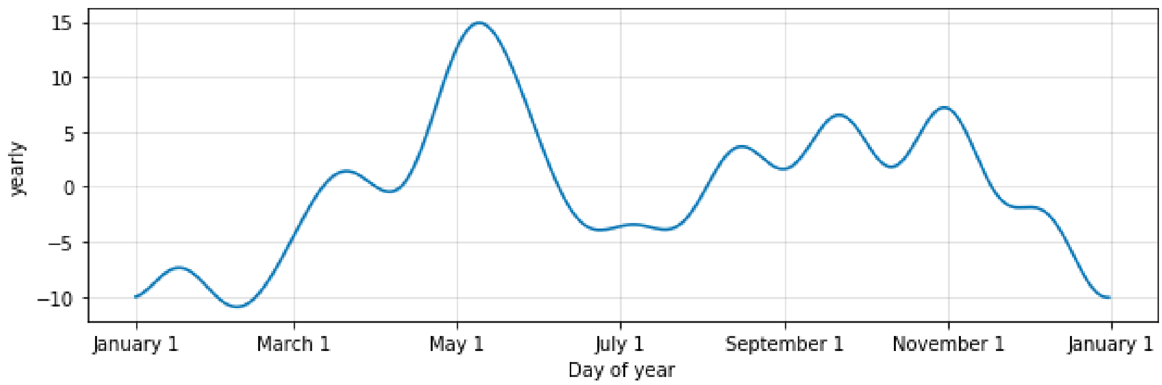

Fig. 2 Annual seasonality of shoe sales

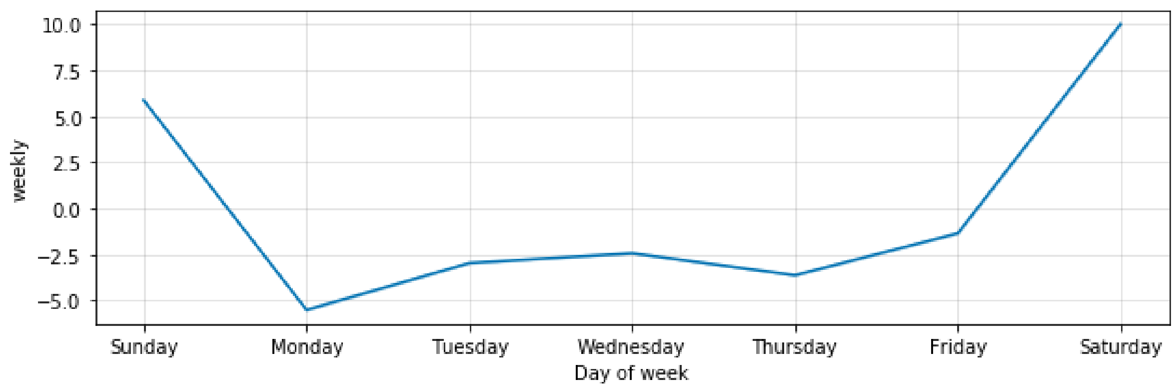

Fig. 3 Weekly seasonality of shoe sales

There are also significant changes in sales during the holidays. On holidays, there can be a sharp increase in sales, or a significant drop. Fig. 4 clearly shows fluctuations in sales during the holidays relative to the average daily value of sales.

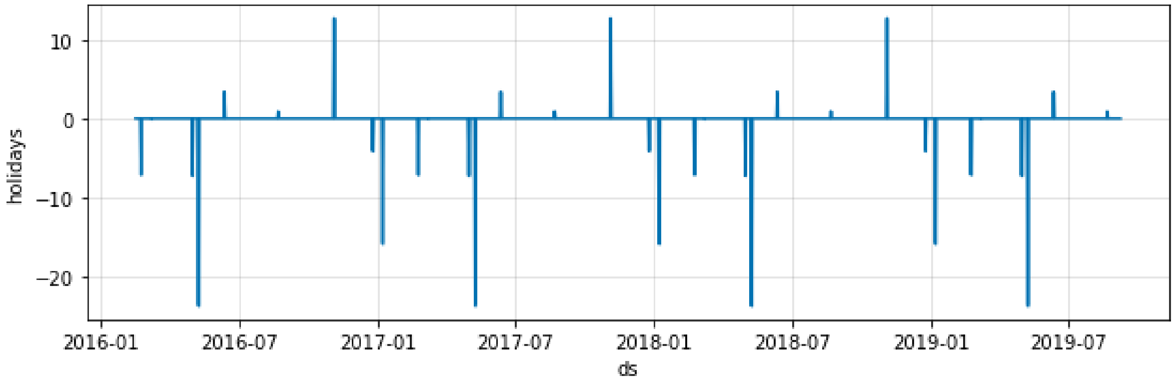

Fig. 4 The impact of holidays on shoe sales

This case examines the sale of shoes in each individual store separately because the stores have a unique level of sales, unique assortment and a unique geographical location. Based on the above, need to choose a model that responds well to seasonal changes, holidays and which can show a good result with a limited dataset.

Today, there are several types of sales models. In paper [2], the authors suggested using greedy heuristics to predict sales in fashion retailing. They work with a real case for one fashion retailer in Singapore. The essence of heuristics is to predict sales of a certain sku (stock keeping unit) without taking into account other products that are sold in the store. Sales forecast is calculated for each sku in each store for each day.

The complexity of the such algorithm is O (n3), which makes it inapplicable to our case because of the large number of stores and sku. The Mean Absolute Percentage Error when using the heuristics proposed by the authors was 20% with a planning horizon of 15 days. With a time, horizon of 275 days, MAPE was 77.2%. This work did not take into account the possibility that the goods may not remain in the store on one of the days. The article stipulated that the entire range of stores is always available for purchase. In the real world, this is an impossible condition.

The last few years it has become a trend to use neural networks and artificial intelligence to forecast demand in retail. One of the landmark works in this direction is the study given in paper [3]. This method demonstrates the results of WNN (Wavelets Neural Network) and TS (Takagi-Sugeno Fuzzy System). The authors of the article use these neural networks to predict sales in food retail. Three categories of products were selected: yogurt, milk and dairy dessert. This was done not only to simplify the presentation of the technical part of the article, but also because the prediction methods proposed in the article are computationally complex. Although they show minimal error with a short forecast horizon.

As a result of the analysis of the article and the results presented in it, we came to the conclusion that these methods are difficult to apply. This is due to the extremely difficult implementation of the neural network architecture and the operating time of the neural networks. In addition, the prediction error increases greatly with the increase in the forecast horizon.

An interesting case is considered in article by [4]. This is important because the methods that are considered in it make it possible to forecast demand during the validity of promotions. This method take into account how the demand for products of one category will affect the demand of another category. The advantages of this technique include the simplicity of the data that is needed for the model to work. The data set for this method include such characteristics as sku, mean unit sold per week, median unit sold per week, price reduction, percentages of week for promotional activities. The second advantage was the clear algorithm scheme. To forecast demand, the Multistage LASSO process is used, presented in the diagram in fig. 2.

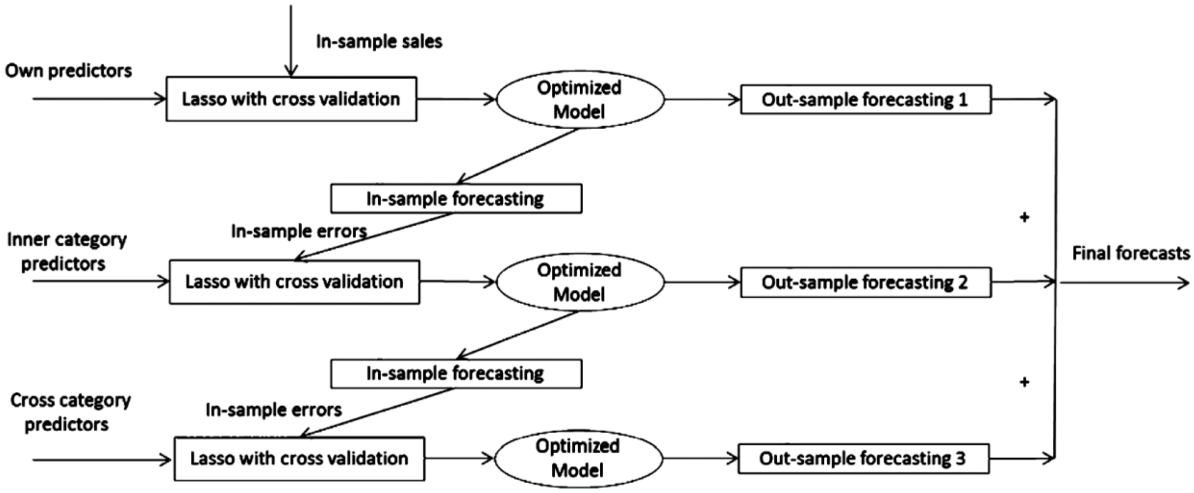

Fig. 5 Multistage LASSO model for sale forecasts

Common analysis of methods and approaches that used for prediction of sales in shoe retail:

– Econometric approaches;
– Methods of machine learning;
– Deep learning, using artificial intelligence.

Machine learning methods include probabilistic and statistical approaches such as regressions (linear, nonlinear, autoregression), the support vector method and its variations, random forest regression, model ensemble combining, and so on. Deep learning includes prediction using recurrent and convolutional neural networks of various architectures.

It is clear that in the modern world with the current level of technology development, econometric approaches are outdated. They were very popular in the 80-90s, but now the mood in business circles and the scientific literature is that machine learning and artificial intelligence have a future of forecasting.

III. Forecasting Model on Base Prophet

Depth learning methods require a huge amount of data and computational resources. This direction for research seems very promising. One of the indicators of the effectiveness of deep learning methods is the success of companies that use them in sales forecasting. Since forecasting is long-term, it is not possible to use methods that based on short data to forecast sales. These methods can work well with a time horizon of about a week, then the quality of the forecast deteriorates dramatically. In world practice for quite a long time there is a trend to use time series to forecast sales with a time horizon of up to a year.

There are many approaches to predict time series, such as ARIMA [5, 6], ARCH [7], regression models [8], neural networks [8]. Implementation from scratch of the above methods (except for regression models that make a prediction at the level of “what happened yesterday, it will be tomorrow”) requires at least six months.

In our work, we were based on the Prophet Forecasting Model [9]. Prophet is a procedure for forecasting time series data based on an additive model. It is used non-linear trends that correspond to fit with yearly, weekly, and daily seasonality, plus holiday effects. This model works best with time series with seasonal effects. Prophet model is robust to shifts in the trend and missing data. It typically handles outliers well. The simplest way to use Prophet is to install the package from PyPI (Python) or CRAN (R).

The analysis of forecasting methods for accuracy is presented on fig. 3. As result the Prophet Forecasting Model is more accurate in predicting almost the entire time interval than other most popular forecasting methods.

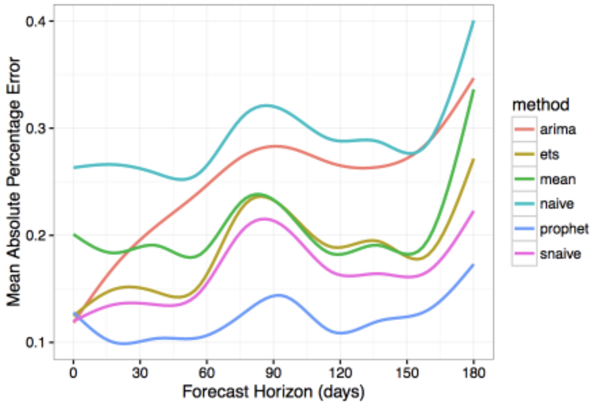

Fig. 6 Comparing Prophet with several other time series forecasting methods

Used here the time series model with three main components of the model: trend, seasonality and holidays. They are combined in the following equation [10]:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon \tag{1}$$

where, *g(t)* is a trend function that models non-periodic changes in the value of a time series, *s(t)* represents periodic changes (for example, weekly and annual seasonality), and *h(t)* represents the effect of holidays that occur on potentially irregular graphs one or more days. The term error represents any specific changes that are not related to the model.

This specification is similar to the generalized additive model (GAM) class of regression models with potentially non-linear smoothers applied to regressors. Here we only use time as a regressor, but perhaps several linear and non-linear functions of time as components[11]. Simulation of seasonality is used as an additive component. It is the same approach used for exponential smoothing [12]. The multiplicative seasonality is a factor of g(t), where the seasonal effect is presented.

The problem of forecasting is an exercise on a curve, which differs from time series models. It explicitly take into account the structure of time dependence in the data. It allow to discard some output advantages of using a generative model, such as ARIMA, this formulation provides a number of practical advantages [9]:

– Flexibility: we can easily adapt the seasonality to several periods and allow us to analyze various assumptions about trends;

– Unlike ARIMA models, measurements should not be distributed regularly, and we do not need to interpolate missing values, for example, from removing outliers;

– The training of the model is carried out very quickly, which allows to immediately explore the result of training with different parameters of the model;

The prediction model has easily interpretable parameters that the analyst can modify to impose assumptions on the forecast. Moreover, the model can easily be expanded to include new components[9].

To use the above model in the research department, we use Python with correspond library PyPI. This library has greatly simplified and accelerated work.

The first step is removing anomalous days from the very beginning form source data. The day is abnormal if the demand on that day exceeded twice the value of the arithmetic average for all days of observations. Next step is creation an empty sales forecasting model by activating the parameter of weekly seasonality. After that holidays and vacations in Russia and Belarus was added to the model data. In our forecast model monthly seasonality is used as lasting 30.5 days. After that, need to choose how often the forecast need.

It is possible spend analysis for any period of time, but in this paper frame the daily forecast is basic. The planning horizon can be configured from one day to one year.

The forecast model is trained on historical data. The output result is represented like as Table 1.

TABLE I
FORECAST RESULTS REPRESENTATION

	ds	yhat	yhat_lower	yhat_upper
1146	2019-04-07	42.945359	31.045040	54.298432
1147	2019-04-08	32.012209	20.560192	43.272301
1148	2019-04-09	34.246160	21.932629	45.579514
1149	2019-04-10	34.541470	23.502869	46.532006
1150	2019-04-11	33.852194	21.417881	45.633005

The **yhat** column predicts sales for the date indicated in the **ds** column. The remaining two columns are the lower and upper intervals, with which, with a 95% probability, the real value of the sales will fall.

IV. Analysis of Forecasting Results

The prediction error was considered in two ways: cross-validation and demand results in real life. The cross-validation method is to break down historical demand data into 10 parts. Training models to produce on 9 parts of the data, and the model must predict a tenth of the data. The real data and the results of the forecast are compared.

The basic estimation of error is MAPE (mean absolute percentage error). It is the average absolute error of our forecast. Let y_i be real demand, $\bar{y}_i$ be forecasted demand. The $e_i = | y_i - \bar{y}_i |$ is forecast error. Then:

$$p_i = \frac{e_i}{y_i} \quad (2)$$

It is the absolute error of the forecast, expressed as a percentage.

Data was collected using the POS (Point-Of-Sale) system database. This case used Oracle database. Since a shoe sales forecast is needed for each store for each day of the forecasting period, datasets were received for each store separately, where for each date in the history the number of shoe sales was indicated. To obtain datasets, SQL was used, data storage was organized in the csv format.

Different stores have different error values. For example, for store X, the error at the forecast horizon of 25 days was 23.1%. When forecasting for 180 days ahead, the error was 33.9% like as Fig. 4.

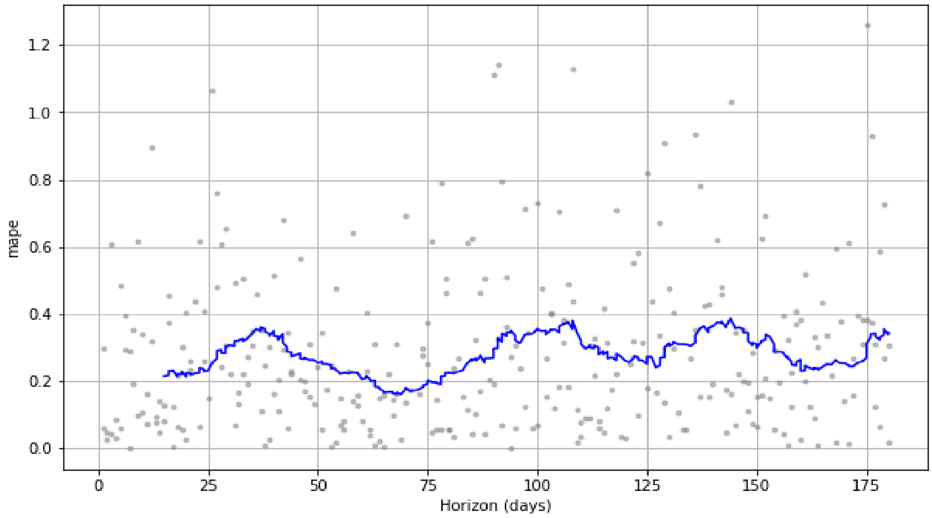

Fig. 7 Forecast error for store X

For Store Y, the prediction error ranged from 28.3% to 34%. The growth of the error is due to the fact that for the Y data store for training the model was 30% less than for the X store like as Fig. 5.

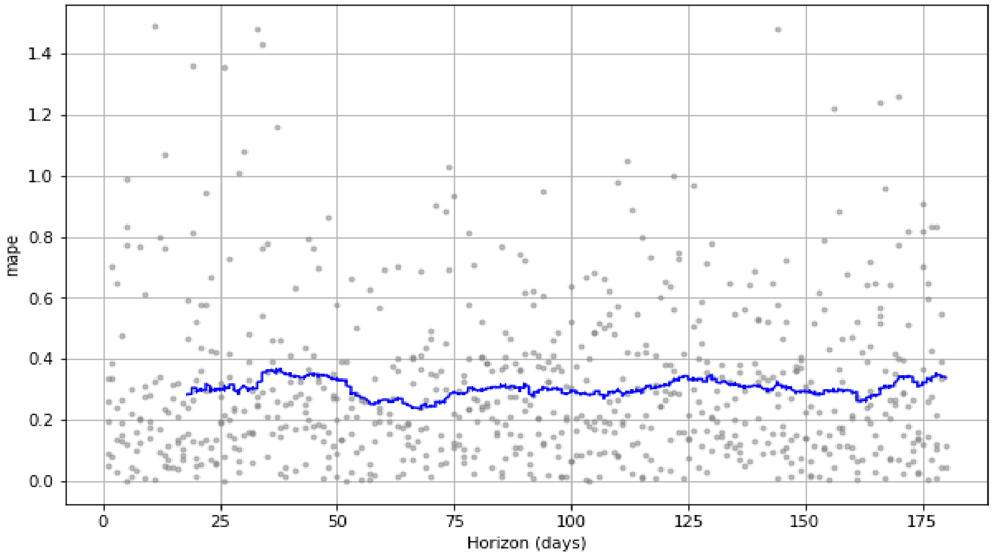

Fig. 8 Forecast error for store Y

V. Conclusion

In this article, we managed to implement sales forecasting method based on time series, which was not previously used in shoe retail. A side effect of solving the forecasting problem was systematized data of the company, which allowed obtaining new analytical data for company employee. In the future, forecasting will be made at the "shop-article" level. This will further reduce the human influence on the distribution of goods.

Acknowledgment

The authors thank BELWEST for access to company data and the possibility of conducting research on real data. This work is supported by bilateral China-Belarus project F20KITG-006 of State Committee on Science and Technology of the Republic of Belarus.

REFERENCES

[1] Why The Physical Store Model Is Dead. Available at https://www.pymnts.com/news/payment-methods/2019/physical-retail-brick-and-mortar-store-model-dead/ (accessed 2019, Nov)

[2] Li, Chongshou, and Andrew Lim. 2018. "A greedy aggregation–decomposition method for intermittent demand forecasting in fashion retailing." European Journal of Operational Research $V 269 (3):860-869

[3] Veiga, Claudimar Pereira da, Cássia Rita Pereira da Veiga, Weslly Puchalski, Leandro dos Santos Coelho, and Ubiratã Tortato. 2016. "Demand forecasting based on natural computing approaches applied to the foodstuff retail segment." Journal of Retailing and Consumer Services $V 31:174-181

[4] Ma, Shaohui, Robert Fildes, and Tao Huang. 2016. "Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra- and inter-category promotional information." European Journal of Operational Research $V 249 (1):245-257

[5] Hyndman, R. J., A. B. Koehler, R. D. Snyder and S. Grose (2002) A state space framework for automatic forecasting using exponential smoothing methods, International Journal of Forecasting, 18(3), 439–454

[6] Asteriou, Dimitros; Hall, Stephen G. (2011). "ARIMA Models and the Box–Jenkins Methodology". Applied Econometrics (Second ed.). Palgrave MacMillan. pp. 265–286. ISBN 978-0-230-27182-1.

[7] Li, D.; Zhang, X.; Zhu, K.; Ling, S. (2018). "The ZD-GARCH model: A new way to study heteroscedasticity" (PDF). Journal of Econometrics. 202 (1): 1–17.

[8] Giang H Nguyen's and al. Sales Forecasting Using Regression and Artificial Neural Networks November 2013 Conference: Midstates Conference for Undergraduate Research in Computer Science and MathematicsAt: Ohio 10p.

[9] Taylor, Sean J. and Benjamin Letham. “Forecasting at Scale.” PeerJ PrePrints 5 (2017): e3190.

[10] Harvey, A. C., and S. Peters. 1990. "Estimation procedures for structural time series models." Journal of Forecasting $V 9 (2):89-108

[11] Hastie, Trevor, and Robert Tibshirani. 1987. Generalized Additive Models: Some Applications. Vol. 82

[12] Gardner, Everette S. 1985. "Exponential smoothing: The state of the art." Journal of Forecasting $V 4 (1):1-28