Usage of time series forecasting model in Supply chain sales prediction

A Raiyani¹, A Lathigara², and H Mehta³

¹School of Engineering, RK University, Gujarat, India
²School of Engineering, RK University, Gujarat, India
³School of Engineering, RK University, Gujarat, India

¹ ashwin.rkce@gmail.com

Abstract. This paper representing a study of supply chain operation data that was used on 100 different store items from 10 stores using 5 years history of sales through open sources contest to compare the performance of time-series forecasting model mainly, decomposition, Auto-Regressive Integrated Moving Average (ARIMA), Prophet, Box-Cox transformation. Here data is collected from 2013 to 2018 were used in real-time transaction at different store, initially model was applied on 2013 to 2017 data and based on the that predicted for 2018 then again cross checked with actual 2018 with proceed predicted data of 2018. To improve the performance and evaluation of the supply chain management system, scrutiny 3 metrics that will help to make decision on the model selection. The accuracy of the Machine learning model in forecasting future sales of supply chain store. Although the result on comparison indicates that there is no single method gives better and superior result. But present study indicates that prophet and ARIMA hybrid model gives better result compare to individual model.

Keywords: supply chain, time-series, prediction, sales

1. Introduction

A more open and efficient supply chain has been required in terms of evolving and competitive markets, which supply chain uses, suppliers who are working for multiple customers and other suppliers in the same sector and consumer expectations. The bulk of existing metric packages with manufacturers stagnate and contribute to fire struggles. By having a part of the chain of supply transparency, current metric packages add value, and they should be used; however, the whole image is not drawn. The changing supply chain environments are not completely considered in present metric packages. With an increase in the amount of outsourcing, there is a need for a new metric kit to handle larger or more dynamic supply chains and to enhance the supply chain management operations by using predictive analysis techniques. Suppliers operate with more customers than ever and their relationships also shift over time with each customer. Many suppliers have in the past been based on one client and will do anything to make that client happy, predictive and reactive analysis techniques were appropriate if a supplier relationship was built on the surviving supplier because of the customer alone.

The research has been given. The supplier's loyalty to their only customer enables consumers to exploit their suppliers so that they can trust their suppliers more with little screening, other than tracking quality goods on-time delivery. This research comparative study on the use of predictive
2. Related Work.

Sales system data are designed to facilitate the detection of abnormal buying behavior of customer. Time series models have for quite some time been of interest. The time series model allows to do prediction on sell store items based in transaction data available. There were many analysis mechanisms set on historical transactional data to have more accuracy on prediction of sales. A method like exponential smoothing and generalized regression strategies [1,2] were utilized to forecast sales historical data expectation. Addition to that, decomposition methods and multilevel time series models [3,4] were implied on better way on various sales components for prediction. ARIMA models have been rigorously utilized for future forecasting including the pattern sessional, occasion patterns and holiday trends. Machine learning sales prediction models are dependent on artificial neural networks were used to train and test the dataset.

Support vector machines (SVMs) are a new type of approach based on statistical learning theory for machine learning [5, 6]. In practical applications, they may contribute to greater potential and better performance. Typically, since of the concept of basic hazard minimization utilized in SVMs, which has more prominent speculation potential and is prevalent to the rule of experimental chance minimization grasped by customary neural systems. SVMs have been effectively connected to different time arrangement prediction issues, such as yield esteem determining within the apparatus industry[7], motor unwavering quality prediction[8], and financial time arrangement prediction[9],[10]. By utilizing SVM for time series estimating, the effective utilize of SVM in time arrangement forecast propels our research work.

The main motive of this research present paper is to have comparison on current time series methods like SVM, regression, ARIMA and exponential smoothing with reference to their real forecasting efficacy in scourge time series. This comparison may be supportive for the transmission expert to select the foremost appropriate technique in each circumstance.

3. Methodologies

3.1. Materials

Gathering of data from various sales store department for 100 product item level. Supply chain operation data that was used on 100 different store items from 10 stores using 5 years history of sales through open sources contest to compare the performance of time-series forecasting model. Again, these data were collected from 2013 to 2018 duration to measure the different parameter for decision improvement.

To improve data processing capability, the Hadoop and Spark framework is used to achieve the distribution storage and analysis work of the collected big data. Big Data analytics uses a variety of techniques such as statistics, modelling and data mining to analyse current and historical facts and make predictions for the future. [15] Big Data can fulfil correlation analysis. Correlation analysis points us whether or how two things are connected. To get better decision and improve performance, predictive analytics will be suitable. PySpark was used as python library as it supports standard data processing framework. Benefits of these library are speed, ease of use and runs everywhere. Additionally, Spark SQL and SQL used for structured data processing, MLlib for machine learning, GraphX for graph processing.
3.2. Methods

3.2.1. Decomposition methods

The decomposition [11, 12] methods attempt to separate long trend pattern, a seasonal trend and residuals. To make it separate, different indices are used to indicate pattern, trend, and residual things. Decomposition time series is simple to disclose to the end user as it does not contain lots of mathematics. This is a significant bit of advantage since, in such a case that the end user has acknowledgment of how the forecast was developed, the person in question may have more confidence in its use for decision making.

So, performing the decomposition method on neural network and running the Additive Model and Multiplication Model, observed figure 1 (a) and (b) that there is not much difference on applying methods.

![Fig. 1(a). Decomposition – Additive Model](image1)

![Fig. 1(b). Decomposition - Multiplication](image2)

3.2.2. The Prophet Forecasting Model

Additive regression model \( x(t) = np(t) + s(t) + ho(t) + \text{err} \), where \( np(t) \) is non-periodic changes, \( s(t) \) is seasonal changes, \( ho(t) \) customized holiday list with irregular schedules, \( \text{err} \) error term accounts for any unusual changes not accommodated by the model Mean Absolute Percentage Error (MAPE) as a forecast precision measurement. In this research [18], Prophet has appeared critical lower determining blunder than the other models can see from figure 2.

![Figure 2. – prophet has lower error.](image3)
Again, By using Prophet.plot components (pyspark function library) figure out and experimented on different components prophet model and taken each component like trend, yearly and weekly seasonality in separate consideration.

![Graph of Prophet plot components – trend, yearly and weekly seasonality.](image)

It seems like People go to Shopping Mostly in July, they are peaking on Sunday and Saturday. So we should add the holidays effect to make Prophet perform better. Now adding parameter types of Holidays like Less-Important public holidays and weekends. And Festive Holidays with high importance. There may occurs a high increase or decrease in sales due to these holidays e.g. New Year, Christmas etc.

![Graph of Prophet plot components – trend, yearly and weekly seasonality.](image)

Figure 3 – prophet plot components – trend, yearly and weekly seasonality. 

Figure 4 – with weekend and festival holiday.
Above figure 4. result shows that there is spike on the days around less important days holiday appearances, with an especially large spike for the festival holidays. With this configuration running and calculating Forecasting error.

```python
def calculate_forecast_errors(df, prediction_size):
    df = df.copy()
    df['e'] - df['y'] - df['what']
    df['p'] = 188 * df['e'] / df['y']
    predicted_part = df['predicted_size']
    # Define the function that averages absolute error values over the predicted part.
    error_mean = lambda error_name: np.mean(np.abs(predicted_part[error_name]))
    # Now we can calculate MAPE and MAE and return the resulting dictionary of errors.
    return {'MAPE': error_mean('p'), 'MAE': error_mean('e')}
```

As a result, MAPE is about 27.5%, and ME is about 3.54 predicts. It indicates that relative error is low and on an average 3.54% model goes wrong. Again, to make improvement on machine learning model, prophet combine with Box Cox Transformation model.

3.2.3. Box-Cox Transformation

A Box Cox transformation is a way to transform non-normal dependent variables into a normal shape. Normality is an important assumption for many statistical techniques; if data isn’t normal, applying a Box-Cox means that able to run a broader number of tests.

![Figure 5 (a) no transformation.](image1)

![Figure 5 (b) with Box Cox transformation](image2)

In box cox transformation, as shown in above figure 5 (a) and (b) we have upper bound predict value and lower bound predict value which gives closer to our actual sales values. It will help in nearby decision making for sales store items.
As result is not that much improved compare to no-transformation model but yes it indicates slightly positive result at the end.

3.2.4. Proposed Approached

ARIMA models are denoted with the notation ARIMA(p, d, q). These three parameters account for seasonality, trend, and noise in data.

AR: Auto-Regressive (p): AR terms are just lags of dependent variable. For example lets say p is 3, we will use x(t-1), x(t-2) and x(t-3) to predict x(t)

I: Integrated (d): These are the number of nonseasonal differences. For example, in our case we take the first order difference. So we pass that variable and put d=0

MA: Moving Averages (q): MA terms are lagged forecast errors in prediction equation.

The ARIMA [13,14] models are inclusion of 3 models, autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA). The current estimation of the time arrangement is a direct capacity of its past qualities and arbitrary clamor in the AR model; though the current estimation of the time arrangement is a straight capacity of its current and past estimations of residuals in the MA model. The ARMA model is the blend of AR and MA, which takes into consideration of historical values and residuals. Stationary processes [16] are the time series required in the AR, MA, and ARMA models. This can be interpreted and implies that the mean value of result and the covariance value do not change with time. For non-stationary time series, the transformation of the series into a stationary one must be achieved first. In general, the ARIMA algorithm matches the time series data based on the ARMA model and a separation method that converts the non-stationary data into a stationary one efficiently.
By applying proposed approached forecasting model found following results:

| Dep. Variable: | sales | No. Observations: | 60 |
|---------------|-------|-------------------|----|
| Model:        | SARIMAX(2, 0, 1)x(2, 2, 0, 12) | Log Likelihood | -1.683 |
| Date:         | Fri, 07 Aug 2020 | AIC | 15.366 |
| Time:         | 13:51:08 | BIC | 17.181 |
| Sample:       | 01-01-2013 | HQIC | 13.374 |
|               | - 12-01-2017 |     |    |
| Covariance Type: | opg |                      |    |

| coef | std err | t | P>|z| | [0.025 (0.975) |
|------|---------|---|-----|-----------------|
| ar.L1 | -0.8498 | 0.494 | -1.722 | 0.085 | -1.817 | 0.118 |
| ar.L2 | -0.4951 | 0.383 | -1.293 | 0.105 | -1.246 | 0.255 |
| ma.L1 | -1.0001 | 11161.623 | -0.001 | 0.999 | -2277.739 | 2275.739 |
| ar.S.L12 | -1.0864  | 0.018 | -57.817 | 0.000 | -1.103 | -1.030 |
| ar.S.L24 | -0.0255 | 0.007 | -3.437 | 0.001 | -0.048 | -0.011 |
| sigma2 | 0.0686 | 79.636 | 0.001 | 0.999 | -156.016 | 156.153 |

Including Holidays MAPE 20.563128311735583
Including Holidays MAE 2.495684718273

Figure 6. Proposed model result with MAPE and MAE.

As shown in figure 6, result, the improvement over transformation

![Figure 7](image)

Figure 7 accuracy of proposed model for next 100 month predictions.

In figure 7, we can see overlapping lines for 2018 indicating the accuracy of model and it also indicated that how upper and lower bound predicted values of sales store items for upcoming 100 months.

4. Conclusion

Compared to current models, the new smart business architecture for market prediction is more accurate for future demand. It helps inventory managers better control their output in the supply chain by simultaneously raising days of scope and operation. Prediction using proposed require much prior knowledge or experience of forecasting time series data since it automatically finds seasonal trends beneath the data and offers a set of "easy to understand" parameters. Proposed approach for forecasting
time series dataset dependent on an additive model where non-linear patterns are fit with yearly, week by week, and everyday seasonality, in addition to holiday effects. It works best with time series that have solid seasonal impacts and a few seasons of chronicled data. No single model time-series is perfect as individual so combining technique/model it gives better outcome. In this context, the proposed model is suitable for products that are strongly linked to a part of profile or profile to a significant share of sales. In such cases, the proposed model can be further expanded to consider the migration and immigration dynamics of individuals with specific interest profiles in a region.

5. References

[1] Farrington C, Andrews N 2003 Outbreak detection: application to infectious disease surveillance. Monitoring the Health of Populations: Statistical Principles and Methods for Public Health Surveillance 203–231.

[2] Chadwick D, Arch B, Wilder-Smith A, Paton N 2006 Distinguishing dengue fever from other infections on the basis of simple clinical and laboratory features: application of logistic regression analysis. Journal of Clinical Virology 35: 147–153.

[3] Gonzalez-Parra G, Arenas AJ, Jodar L 2009 Piecewise finite series solutions of seasonal diseases models using multistage Adomian method. Communications in Nonlinear Science and Numerical Simulation 14: 3967–3977.

[4] Spaeder MC, Fackler JC 2012 A multi-tiered time-series modelling approach to forecasting respiratory syncytial virus incidence at the local level. Epidemiology and Infection 140: 602–607.

[5] Chang C-C, Lin C-J 2011 LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2: 27

[6] Thissen U, V B R, De Weijer A, Melssen W, Buydens L 2003 Using support vector machines for time series prediction. Chemometrics and intelligent laboratory systems 69: 35–49.

[7] Pai P-F, Lin C-S 2005 Using support vector machines to forecast the production values of the machinery industry in Taiwan. The International Journal of Advanced Manufacturing Technology 27: 205–210.

[8] Hong W-C, Pai P-F 2006 Predicting engine reliability by support vector machines. The International Journal of Advanced Manufacturing Technology 28: 154–161.

[9] Müller K-R, Smola AJ, Rätsch G, Schölkopf B, Kohlmorgen J, et al. 1997 Predicting time series with support vector machines. Artificial Neural Networks–ICANN’97: Springer. pp. 999–1004.

[10] Tay FE, Cao L 2002 Modified support vector machines in financial time series forecasting. Neurocomputing 48: 847–861.

[11] Bowerman BL, O’Connell RT, Richard T 1993 Forecasting and time series: An applied approach: Belmont CA Wadsworth.

[12] Hamilton JD 1994 Time series analysis: Cambridge Univ Press.

[13] Zhang GP 2003 Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50: 159–175.

[14] Pai P-F, Lin C-S 2005 A hybrid ARIMA and support vector machines model in stock price forecasting. Omega 33: 497–505.

[15] Ashwin R, Amit L, Hardik M, 2018 Supply Chain Management and Big Data Analytics (SCMBDA): Perception to SCM Business International Journal on Future Revolution in Computer Science & Communication Engineering ISSN: 2454-4248 Volume 4 :Issue:5

[16] Waller and Fawcett, 2013 Data Science, Predictive Analytics, and Big Data A Revolution That Will Transform Supply Chain Design and Management,

[17] Ma C, Zhang HH, Wang X. Machine learning for big data analytics in plants. Trends Plant Sci. 2014;19(12):798–808

[18] Taylor SJ, Letham B. 2017. Forecasting at scale. PeerJ Preprints 5:e3190v2