International Journal of Hybrid Information Technology
Vol.13, No.1 (2020), pp.7-18
http://dx.doi.org/10.21742/ijhit.2020.13.1.02

Comparison of Time Series ARIMA Model and Support Vector Regression

Yekta S. Amirkhalili[1], Amir Aghsami[2], and Fariborz Jolai[3]

[1,3]*School of Industrial & Systems Engineering, College of Engineering, University of Tehran, Iran*

[2]*School of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran*

[1]*yektaakhalili@ut.ac.ir*, [2]*a.aghsami@ut.ac.ir*, [3]*fjolai@ut.ac.ir*

Abstract

As one of the most important and costly functions of any business, sales analytics has been the target of many studies for some time now. Knowing and tracking the sales of a business proves useful in all data-driven decisions made from inventory management to shelf layouts in a supermarket. However, forecasting sales rely heavily on data and algorithms strong enough to handle unseen data. Since sales data are in nature time series datasets one of such predictive methods is time series analytics. In this paper, the ARIMA modeling to the seasonality of the data is compared with a machine learning technique, support vector regression. These comparisons are carried out on three different and unrelated datasets and these algorithms' errors when predicting future sales are compared. The results obtained from our analysis show poor results in general due to datasets having large numbers of oscillation and outliers, but for comparison purposes these datasets and results are fine. We conclude that support vector regression produces better results in comparison with time series analytics on all datasets used in this paper.

Keywords: *Time series analytics, Support vector regression, Sales forecasting*

1. Introduction

Retailers, and especially in recent years, e-tailers rely heavily on forecasting and customer demand analytics to provide the best services to their customers consistently. One example of the use case for forecasting demand for products is inventory management [1] the process which involves the need to know of what products sell larger quantities and on a more frequent basis. This way, the best-sellers are stacked strategically, in larger quantities, so the store minimizes their chances of running out of this product as well as strategic product placement on the shelves. With the help of data analytics and machine learning techniques, customer purchasing habits are studied. One example of this purchasing analytics is mentioned in [2] which looks at the Australian retail industry. This article mentions the use of Big Data Analytics to see what customers are looking for online, keywords such as "Paleo diet" help retailers know what the market requires. These data-driven decisions do not stop at just purchasing behavior but extend to areas such as marketing, production, finance and even accounting [3]. All of this is possible through quantitative analytics on customer purchasing habits, demand and sales.

Article history:
Received (October 30, 2019), Review Result (December 25, 2019), Accepted (January 30, 2020)

Print ISSN: 1738-9968, eISSN: 2652-2233 IJHIT

The focus of this paper is on forecasting sales in the future mainly different techniques. This is useful for inventory management as well as to track how well businesses are doing and what areas need improvement. The use case of such studies is planning and making data-driven decisions with a lower probability of mistakes. The importance of experience and expert knowledge is not lost, but the combination of the two methods would give businesses a tremendous boost in profits. In other words, there are qualitative and quantitative forecasting methods. The qualitative methods are the Delphi method, surveys, and as previously stated experts' knowledge. A major drawback of sales forecasting methods is the inability to account for all the many external factors that influence the purchasing behavior of customers that cannot be measured as well as large volumes of data with irregular fluctuations. Thankfully with the development of new technologies and more powerful computers, demanding computations are possible and excellent machine learning algorithms have been developed to combat such problems optimally. Another reason for the rise in popularity of artificial intelligence-based forecasting methods is because these methods do not require strong model assumptions. Modeling the sales of an e-commerce store is extremely intricate given the many products that are vastly different from use-case to demographics to logistics. This is the reason that in this study a comparison between traditional techniques, namely time series analytics and more recent advances in machine learning, support vector regression, is carried out to try and find the universally best technique.

The sales data is a time series dataset in nature since every sale data is collected through time. Traditional time series analytics methods, namely ARIMA modeling, is used in this paper. The seasonality of the data is not removed, although the most popular traditional approach is to remove seasonality variations. However, adjusting seasonality is an approximation method that would cause larger errors. For this reason, the SARIMA model is employed to do time series analytics. In this paper, the traditional time series forecasting models are compared with machine learning techniques such as regression analysis, namely Support Vector Regression. SV algorithms can generalize excellently to unseen data [4], hence they are well suited for forecasting purposes. Support Vector Regression is the extension of Support Vector Machine to handle regression models since SVM algorithms are suited for classification [5]. Such global comparisons on different and unrelated datasets give important insight into the applications of these methods, and to the best of our knowledge has never been carried out to this extend.

2. Literature review

[6] is a comprehensive look at the history of support vector machines which was first introduced in the paper "on a class of perceptrons" by Vapnik and Chervonenkis in 1964 and further developed in [7] based on VC theory. The Support Vector algorithm is a non-linear generalization of the Generalized Portrait algorithm and can generalize well to unseen data. Support Vector algorithms were first used at AT&T Bell Laboratories by Vapnik, and were largely focused on optical character recognition efforts. [8] Since the central focus of this paper is on comparing two distinct methods of forecasting one via time series analysis and the other through the use of machine learning algorithms, comparative studies were taken into consideration. In [9] a comparative study of both linear and non-linear forecasting models was carried out in the retail industry accounting for seasonality issues. Non-linear forecasting methods utilized in this paper were neural networks as generalized non-linear functional approximators. The conclusion of this paper was in favor non-linear models with even better

results obtained through prior seasonal adjustments, in other words, deseasonalized time series data.

[10] conducted a systematic optimization-based method to utilize support vector regression as the approach for demand forecasting. This is a three-step algorithm using non-linear programming (NLP) and linear programming (LP) to formulate the regression function, and finally solving the mathematical model through the employment of a recursive method to find the underlying customer demand patterns from available train data. This falls into the category of customer behavior analysis. In this study demand attributes are categorized as follows: (1) Past demand attributes which represent the demand for a predetermined time period in the past, and (2) Calendar attributes which depict a characteristic of the specific time period being investigated. The second kind is seldom thought of as binary parameters. [5] investigated the drawbacks of regression and time series methods in forecasting temperatures in short-term load, one of which is their inability to adapt to sudden changes of loads. In this paper, the Mahalanobis kernel was used for parameter declaration, and SMOreg was utilized in solving the optimization problems within the SVR algorithm efficiently. In order to discuss the results, the proposed method was compared to artificial neural networks (ANN) and the results proved in favor of the method proposed. Specifically, the proposed method of this paper shows better results on the absolute average errors; therefore, it allows operators to forecast daily maximum load and more accurate predicted maximum temperatures.

[11] used a hybrid regression-based forecasting method through support vector machine and dynamic feature selection. This system iteratively selects the most relevant features and builds the regression model and finally tunes its parameters dynamically. Model selection is carried out through calculations of initial values for SVR parameters, ε, and C, using the empirical rules proposed by [12]; and utilizing grid search algorithms around those values. To conclude, the proposed algorithm is compared with a standard neural network and a standard ARMAX approach. The final results depicted a slightly better performance from the proposed algorithm. [13] introduced a localized support vector regression for time series predictions to overcome the issues of the standard SVR algorithm. One such shortcoming of the SVR algorithm is that it fails to consider data in any other format aside from a global fashion. Which could cause an inflexibility to capture the local trend of data. This could be detrimental in financial time series analytics due to the volatile nature of financial data. The proposed LSVR method is a systematic and automatic scheme for adapting the margin locally and flexibly as opposed to the globally fixed SVR margin. One other advantage of the LSVR algorithm is that the performance remains almost unchanged against different C's. [14] used a hybrid support vector regression with a mixed kernel function (HSVRIA) to predict time series data. Parameters of the previously mentioned model were determined using an immune algorithm (IA). This showed an appropriate method to do time series forecasting when compared to the original SVR model which would not give the best results in the presence of complicated data patterns. In [15], SVR was used to forecast the demand and supply of Pulpwood which has an integral influence on the socio-economic development in India, as the steady increase in paper demand requires better forecasting abilities. In this day and age, one important consideration is the use of the internet for analytics purposes. [16] used web search data and time series analysis to predict the sales of an e-commerce website. This study explored the correlation of costumers' search and purchase behaviors, eliminated any seasonality and trends in the sales data, and reached a near 5% error when predicting 7-days sales volume. [17] used a hybrid sales forecasting model through the combination of a variable selection method and support vector regression in order to forecast the sales of computer products. The results of this study proved that not only the hybrid forecasting model

gave better more accurate predictions with lower errors, but also displayed the ability to identify the most important predictor variables.

[18] compared the strength of the Autoregressive Moving Average (ARMA), Neural Network (NN) and Support Vector Regression (SVR) to predict the PM 2.5 values. The results proved that both the neural network and the support vector regression give more accurate predictions than ARMA. In addition, NN and SVR can go as far as predicting the exact fluctuations of the data up to 400 days as opposed to only 100 days for ARMA. In other words, ARMA is best used if the task at hand is the prediction of a short time range as it's very simple to implement. However, both NN and SVR are better models for long term prediction. The paper concluded that neither of these is better as it heavily depends of the properties of the dataset as well as the objective of the study. [19] studied the performance of state space and ARIMA models on consumer retail sales forecasting on five different categories of women's footwear. The overall performance of the state space and ARIMA models were evaluated via RMSE, MAPE and MAE. It was concluded that both of these models showed results fairly similar on both one-step and multi-step forecasts. [20] proposed a novel NRS-GA-SVM algorithm for demand forecasting of retail supply chain emergency logistics. The sample attribute index reduction was reached using the NRS algorithm, dynamic demand forecasting model based on non-linear support vector machine regression theory and finally, to optimize the parameters of the machine learning algorithm, a Genetics Algorithm was utilized. This study found that in addition to more accurate forecasting results, the execution time of the forecasting model is also greatly reduced. [21] studied a sales forecasting model based on time series analytics and data mining methods in the German automobile industry. Additive components such as trend, seasonality, calendar and errors were considered in the time series model. The trend component was multivariate estimated through Multiple Linear Regression as we all Support Vector Machine. The other additive components were all estimated univariately. Other influences on the sales such as macro-economic and market-specific factors were also taken into account. The results proved that quarterly data analytics using non-linear models had the superior outcome. The contribution of this study was the use of data mining methods alongside the time series analytics. Hence, proving the importance of realizing if in most scenarios the data mining methods outperform traditional time series analysis. One objective of this study is to fill in the gap of such comparative studies which investigate traditional methods versus machine learning methods in sales forecasting literature.

This paper is organized as follows: Section 2 is a review of the literature that has used time series analytics and support vector regression in forecasting, specifically focusing on sales forecasting. Section 3 is devoted to discussions about time series analytics and machine learning techniques employed in this paper. Section 4 discusses the application of the methods being studied as well as a comprehensive comparison of results. In section 5 the results obtained from the previous section are further investigated and finally, section 6 concludes this research.

3. Methodology

3.1. Time series analytics

Sales forecasting of e-commerce businesses heavily depends on the data available. Sales data is gathered over time, hence making the dataset a time series data. Therefore, time series analytics are applied to do such forecasting. The one main issue of such data is the prominent seasonality aspect, which in traditional approaches would have been completely removed from the data. The process of removing seasonality is an approximation method, and therefore forecasts based on seasonally adjusted data are more prone to errors. There are other adjustment processes available such as the natural log transform [9].

A stochastic process is a series of random variables such as $\{Y_t : t = 0, \pm1, \pm2, \pm3, \ldots.\}$ which are the observations of a time series. One well-known stochastic process is the random walk; a series of independent random variables e_t with the same distribution, a zero mean and variance of σ_e^2 formulized as such:

$$Y_t = Y_{t-1} + e_t \tag{1}$$

This is interpreted as being in the Y_t place at time t, only dependent on where we were 1 iteration previously. Such a model where data are correlated is a time series. Time series analytics holds one strong assumption for the stability of data. [22] Stability is concerned with the unchanging probability rules governing the process behavior, in other words, the series distribution of $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}$ is the same as $Y_{t_1-K}, Y_{t_2-K}, \ldots, Y_{t_n-K}$. In this analysis, stability, stationarity and seasonality of the data are studied. Stable data has little to no change in standard deviation, and a simple method for finding stable data is through visualizations. In the case of unstable data, a natural log transform would stabilize data, and the new dataset is used for time series analysis. The stationarity of the data is the quality of unchanging statistics such as mean and variance over time. Statistical tests such as Kwiatkowski-Phillips-Schmidt-Shin (KPSS) are utilized to measure if a dataset has stationary quality or not. In the case of non-stationary data, calculating the diff and repeating the previous process would ensure that a new stationary dataset is reached. Finally, the seasonality of the data is addressed through the use of a $SARIMA(p, d, q) \times (P, D, Q)^s$ model. Assigning values to the *p,d,q,P,D,Q* and *s* parameters of the SARIMA model is an intricate process. One method for calculating these parameters is through the ACF and PACF graphs. The ACF is a bar graph depicting the correlation coefficients between a lag and a time series. The PACF is the ACF of the sample population since the true value of the coefficient is not available. In the case of ARMA modeling, the ACF shows the MA parameter and PACF the AR parameter.

3.2. Support vector machine

Machine learning techniques have been used in forecasting endeavors for many years. Machine learning is broken down into 3 main categories: (1) Supervised Learning Algorithms, (2) Unsupervised Learning Algorithms and (3) Reinforcement Learning Algorithms. A Support Vector Machine algorithm is of the first category and is one of the classification algorithms. Classification algorithms break data into classes based on their labels, and then analyze new, unseen and unlabeled data and apply a label to those. Since forecasting is a continuous measure, classification algorithms are very limited in giving exact reports, which is why the regression-based support vector algorithm is utilized.

3.3. Regression analysis

Why not use a simple regression algorithm? Regression analysis is a form of machine learning for predictive analytics that follows four main assumptions: (1) Linearity, (2) Multivariate normality, (3) No or little multicollinearity, (4) No auto-correlation, and (5) Homoscedasticity (Assumptions of Linear Regression, n.d.). If any of these assumptions are invalidated then regression analytics cannot be performed on a time series dataset. This is one of the methods used in this paper to make certain a simple regression analysis is not an option.

3.4. Support vector regression

Support Vector Machine concepts can be generalized to fit regression problems, as in support vector regression (SVR) which is characterized by kernels, number of support vectors, sparse solution and control of the margin. SVR is an efficient supervised-learning approach for real-valued function estimation with asymmetrical loss function that penalizes high and low errors of estimation equally. It is a superior algorithm in that its computational complexity is not at all dependent on the dimensionality of its input space; hence its tremendously strong ability to generalize with high prediction accuracy. In simple terms, support vector regression works as follows: A tube with minimal radius is symmetrically placed around the function to be estimated. Any errors (the distance between the real and the estimated values) less than a certain threshold, ε, are ignored. This leaves only data points inside the tube while any datapoint outside the tube is penalized. The ε-insensitive region is called the ε-tube. Support vectors in SVR are what determine the shape of the tube, and the data is assumed to be independent and identically distributed (*iid*). [23]

In mathematical terms, the continues-valued function to be approximated using SVR is as follows, where M is the order of the polynomial used for function approximation:

$$y = f(x) = w^T x + b \quad x, w \in \mathbb{R}^{M+1} \tag{2}$$

This function is then treated by SVR as an optimization problem, in which the objective is to find the narrowest tube around the surface and minimizing the prediction error. $\|w\|$ In Eq(3) is the magnitude of the normal vector to the surface that is to be estimated in the objective function:

$$min_w \frac{1}{2} \|w\|^2 \tag{3}$$

The constraint is the minimization of the prediction error; the distance between the predicted and actual values. The value of ε depicts the width of the tube; naturally, a smaller tube indicates lower tolerance for error which affects the number of support vectors and eventually the solution sparsity. As with any regression analytics technique, a loss function should be employed; the choice of which is determined by a priori information about noise distribution of dataset, model sparsity desired, and computational complexity of training process. The most important feature of a loss function is that it should be convex to ensure a unique solution for the optimization problem. In this paper, the Huber ε loss function is chosen for its smoothness and the fact that it penalizes outliers with greater penalty as the error increases. One other method to guard against deviations is to employ slack variables ξ , ξ^* to determine how many points can be tolerated outside the tube. In this multi-objective optimization model, C is a regularization, a parameter that gives weight to the error and that can be tuned. A larger C gives more weight to minimizing the error. The Lagrangian method is used to solve this constrained quadratic optimization problem, with Lagrange multipliers: λ, λ^*, α, α^* which are all nonnegative real numbers. (Mariette Awad, 2015)

$$min_w \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i + {\xi^*}_i \tag{4}$$

Subject to

$$y_i - w^T x_i \leq \varepsilon + {\xi^*}_i \qquad i = 1, \dots, N$$

$$w^T x_i - y_i \leq \xi_i + \varepsilon \qquad i = 1, \dots, N$$

$$\xi_i, {\xi^*}_i \geq 0 \qquad i = 1, \dots, N$$

$$\mathcal{L}(w, \xi, \xi^*, \lambda, \lambda^*, \alpha, \alpha^*) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i + \xi^*{}_i + \sum_{i=1}^{N}\alpha^*{}_i\left(y_i - w^T x_i - \varepsilon - \xi^*{}_i\right)$$
$$+ \sum_{i=1}^{N}\alpha_i\left(-y_i + w^T x_i - \varepsilon - \xi_i\right) - \sum_{i=1}^{N}\lambda_i\xi_i + \lambda_i^*\xi^*{}_i \quad (5)$$

The minimum of the Lagrangian is calculated based on the KKT conditions. The Lagrange multipliers that are zero correspond to data points inside the tube, and the support vectors have nonzero Lagrange multipliers values. Since SVR is a Kernel-based algorithm, the appropriate Kernel Function is of great importance. The kernel is used as an extension of linear support vector regression into a non-linear format to better estimate the function. A kernel transforms feature vectors into high dimensional space, called the Kernel space, and executes the algorithm in the new high dimensional space; which allows the algorithm to separate complicated data through linear separation in a high dimensional space. It is easier to solve the non-linear transformed problem than the non-linear original problem. [5] In terms of mathematical representation, all instances of x in the previous equations could be replaced by $k(x_i, y_j)$ with φ(.) being the transformation from feature to kernel space.

$$k(x_i, x) = \varphi(x_i).\varphi(x) \quad (6)$$

In this paper the Gaussian Radial Basis Function is used as the kernel function where:

$$K(x, y) = exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) \quad (7)$$

Where σ is the global basis function width. Other famous kernel functions are Polynomial, exponential radial basis, multi-layer perceptron, Fourier series, Splines and more. In the case of Kernel selection, the GRBF was chosen based on its popularity and excellent prediction abilities especially when paired with the ε-insensitive and Huber loss function [24]; but other methods such as cross-validation and bootstrapping have also been used for kernel selection.

4. Numerical experiments

To provide a comprehensive comparison of the two forecasting methods of time series analytics and support vector regression, three datasets are used. The steps taken for analysis are different. As shown below, based on the steps of a time series analysis, we at first provide a look at some aspects of each dataset mainly seasonality trends and residuals. This way seasonality of data is readily realized and proper steps can be taken. However, such measures are not the best practices to check the stationarity of a dataset which is why statistical tests such as the oneway test or the Adfuller test are utilized. All the programming for both time series analytics and support vector regression is done using Python version 3.6.2 and R.

Flowchart of analysis:

Step 1 Plot data to observe any unusual behavior

Step 2 Use transformations to stabilize the variance

Step 3 Perform Adfuller test to find out if data is stationary

Step 4 Difference data if non-stationary

Step 5 Plot ACF/PACF to pick ARIMA parameters

Step 6 Test models based on minimum AIC/BIC values, and select best model

Step 7 Plot ACF of residuals

Step 8 Does the ACF plot of residuals look like a white noise term? If yes, forecast. If no, go to step 4

As for the performance measure in time series analytics, the Root Mean Square Error (RMSE) is chosen to measure the standard deviation of the errors. The equation for RMSE calculations is as follows:

$$RMSE(X,h) = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(h(x^{(i)}) - y^{(i)})^2} \quad (8)$$

And in the case of many outliers in the dataset, sometimes it is best to use other performance measures such as Mean Absolute Error (MAE), which also calculates the distance between the vector of predictions and target values.

Table 1. Comparison of model performances

	Model Performance			
	Model(ARIMA)	RMSE	Model (SVR)	RMSE
Dataset1	SARIMA(1,1,3)(1,0,1)12	355	Polynomial (degree = 3), C=10, ε= 0.2	0.88
Dataset2	ARIMA(3,1,1)	128	Polynomial (degree = 5), C=10, ε=0.02	1.23
Dataset3	SARIMA(1,1,2)(2,0,1)12	45	RBF, C=100, ε=0.1	0.36

Some of these results are not acceptable in terms of prediction accuracy, but for the sake of visualizing the difference in forecasting all models have been applied to datasets and can easily be compared with each other in [Figure 1] and [Figure 2].

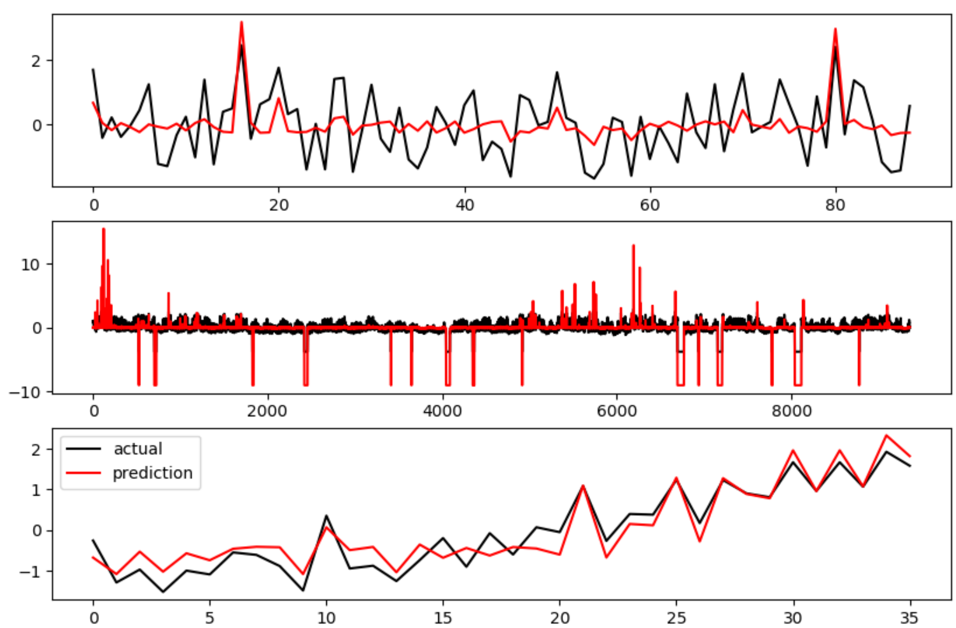

Figure 1. Comparison of support vector regression models on the 3 different datasets, red lines indicate predicted values

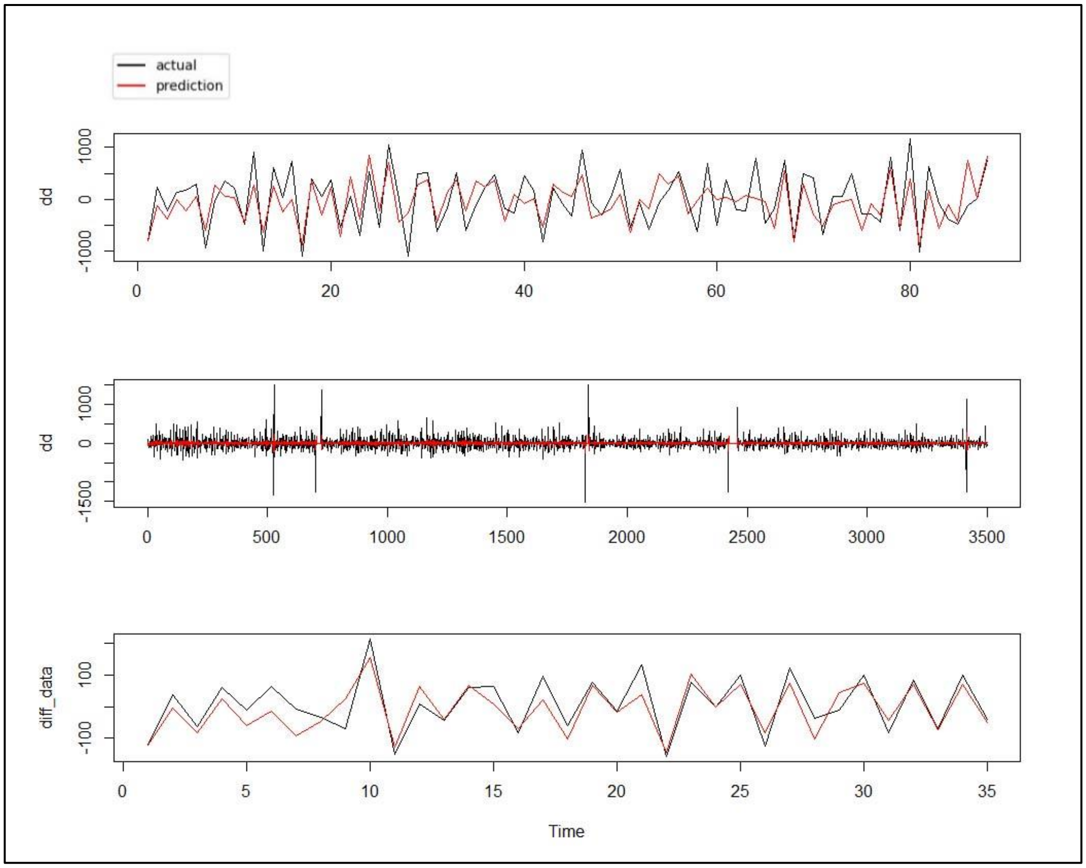

Figure 2. Comparison of Time Series ARIMA models on the 3 different Datasets, red indicates the predicted values

5. Results and discussion

Based on the results obtained from section 4, the most obvious takeaway from this study is the importance of data.

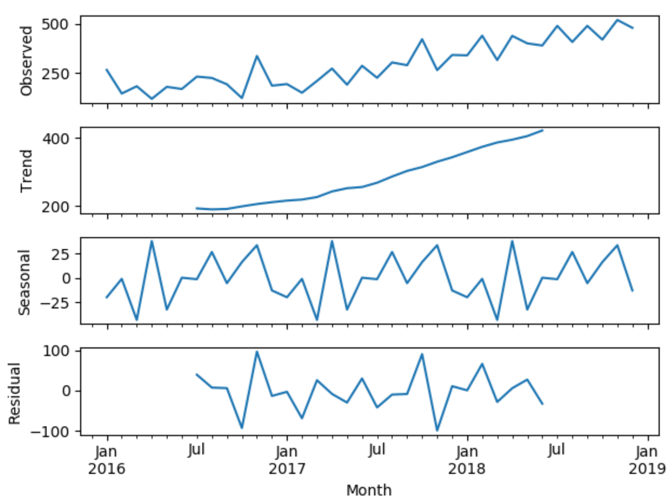

Figure 3. Dataset 3 decomposition

In time series analytics, a lower RMSE indicates better results which we've achieved with our model based on the monthly dataset 3. Taking a look at the observations, trends, seasonality and residuals of dataset 3 we have Dataset 3 decomposition [Figure 3].

This data follows an upward trend which is why predictive analytics provides very fine results. The dataset has some seasonality year by year, which is why the 12 months cycle was chosen in the modeling. Datasets 1 and 2 did not produce good enough results in either method. One reason for this would be the heavy presence of oscillation in the dataset which could result in overfitting. The presence of too many outlier datapoints in Dataset 2 is another reason for a lackluster performance as both SVR and ARIMA models try to make up for those data points.

6. Conclusion

Time series analytics methods in particular the ARIMA modeling was compared in its prediction capabilities against a machine learning technique, support vector regression, to find the best method for predictive analytics. The results of these algorithms were compared on three different and unrelated datasets. These results clearly emphasize the importance of data. Both methods show acceptable results on one of these datasets with the support vector machine producing slightly better results in almost all three comparisons but still a large error value. In conclusion, while neither method is superior to the other every single time since the use of each depends heavily on the dataset at hand, based on our analytics, it seems that support vector regression has produced better results when it comes to predictive analytics on time series data. The reason for this is its ability to deal with seasonality and outliers in a dataset and generalize well to unseen data. For future studies, such datasets with prominent oscillation could be studied using Fourier Transform.

References

[1] B. a. Hilliard, "A Cross-Spectral analysis of retail inventories and sales," The Journal of Business, vol.48, no.3, pp.365-382, **(1975)**

[2] T. Bowen, "How supermarkets are using big data & predictive analytics to win," [Online], Available: https://expert360.com/resources/articles/supermarket-retail-big-data

[3] G. P. Chin-wu Chu, "A comparative study of Linear and Nonlinear models for aggregate retail sales forecasting," International Journal of Production Economics, vol.86, no.3, pp.217-231, **(2003)** DOI: 10.1016/S0925-5273(03)00068-9

[4] A. J. Smola and B. Scholkopf, "A tutorial on support vector regression," Statistics and Computing, vol.14, no.3, pp.199-222, **(2004)** DOI: 10.1023/B:STCO.0000035301.49549.88

[5] H. Mori and D. Kanaoka, "Application of support vector regression to temperature forecasting for short-term load forecasting," Neural Networks, International Joint, **(2007)** DOI:10.1109/IJCNN.2007.4371109

[6] V. N. Vapnik, A. Y. Chervonenkis, and И. Телемех, "Math-Net . Ru," vol.25, no.6, pp.937-945, **(2019)**

[7] V. N. Vapnik, Statistical Learning Theory, **(1998)** DOI: 10.1007/978-1-4419-1428-6_5864

[8] A. J. Smola and B. S. C. H. Olkopf, "A tutorial on support vector regression∗," Statistics and Computing, vol.14, no.3, pp.199-222, **(2004)** DOI: 10.1023/B:STCO.0000035301.49549.88

[9] C. Chu and G. Peter, "A comparative study of linear and nonlinear models for aggregate retail sales forecasting," International journal of production economics, vol.86, no.3 pp.217-231, **(2003)** DOI: 10.1016/s0925-5273(03)00068-9

[10] A. A. Levis and L. G. Papageorgiou, "Customer demand forecasting via support vector regression analysis," Chemical Engineering Research & Design, vol.83, no.8, pp.1009-1018, **(2005)** DOI: 10.1205/cherd.04246

[11] J. Guajardo, R. Weber, and J. Miranda, “A forecasting methodology using support vector regression and dynamic feature selection,” Journal of Information & Knowledge Management, vol.5, no.4, pp.329-335, **(2007)** DOI: 10.1142/S021964920600158X
[12] V. Cherkassky and Y. Ma, “Practical selection of SVM parameters and noise estimation for SVM regression,” Neural Networks, vol.17, no.1, pp.113-126, **(2004)** DOI: 10.1016/s0893-6080(03)00169-2
[13] H. Y. Ã, K. Huang, I. King, and M. R. Lyu, “Neurocomputing localized support vector regression for time series prediction,” Neurocomputing, vol.72, pp.2659-2669, **(2009)** DOI: 10.1016/j.neucom.2008.09.014
[14] I. M. A. Lgorithms and M. O. Hsvria, “A hybrid support vector regression for time series prediction,” Third International Conference on Knowledge, pp.506-509, **(2010)** DOI: 10.1109/WKDD.2010.92
[15] V. Anandhi and R. M. Chezian, “Support vector regression to forecast the demand and supply of pulpwood,” International Journal of Future Computer and Communication, vol.2, no.3, pp.1-4, **(2013)**
[16] D. Wei, P. Geng, L. Ying, and L. Shuaipeng, “A prediction study on e-commerce sales based on structure time series model and web search data,” The 26th Chinese Control and Decision Conference (2014 CCDC), pp.5346-5351, **(2014)** DOI: 10.1109/CCDC.2014.6852219
[17] Chi-jie Lu, “Neurocomputing Sales forecasting of computer products based on variable selection scheme and support vector regression,” Neurocomputing, vol.128, pp.491-499, 27 March, **(2014)** DOI: 10.1016/j.neucom.2013.08.012
[18] M. Liu and X. Cao, Beijing PM 2.5 time series analysis and prediction using regression models, pp.1-20, **(2016)**
[19] P. Ramos, N. Santos, and R. Rebelo, “Robotics and Computer-Integrated Manufacturing Performance of state space and ARIMA models for consumer retail sales forecasting,” Robotics and Computer-Integrated Manufacturing, vol.34, pp.151-163, **(2015)** DOI: 10.1016/j.rcim.2014.12.015
[20] H. Xue, C. Jiang, B. Cai, and Y. Yuan, “Research on demand forecasting of retail supply chain emergency logistics based on NRS-GA-SVM,” 2018 Chinese Control Decis, Conf., pp.3647-3652, **(2018)**
[21] B. Brühl, M. Hülsmann, D. Borscheid, C. M. Friedrich, and S. Augustin, “A sales forecast model for the german automobile market based on time series analysis and data mining methods,” Advances in Data Mining. Applications and Theoretical, pp.146-147, **(2009)** DOI: 10.1007/978-3-642-03067-3_13
[22] J. D. Cryer, “Time Series Analysis with Applications in R,” Springer, **(2008)**
[23] R. K. Mariette Awad, “Efficient Learning Machines theories, concepts, and applications for engineers and system designers,” Effic. Learn. Mach., pp.263, **(2015)**
[24] S. Gunn, “Support vector machines for classification and regression,” ISIS technical report, May, **(1998)**

This page is empty by intention.