A Comprehensive Study on Demand Forecasting Methods and Algorithms for Retail Industries

Udbhav Vikas[1], Karthik Sunil[2], Pattem Deeksha[3], Rohini S.Hallikar[4] and Dr Ramakanth Kumar P[5]

[1,2]Student, Department of Electronics and Communication, RV College of Engineering, Bangalore, India.

[3]Student, Department of Computer Science, RV College of Engineering, Bangalore, India.

[4]Assistant Professor, Department of Electronics and Communication, RV College of Engineering, Bangalore, India.

[5]Professor & HOD, Department of Computer Science, RV College of Engineering, Bangalore, India.

Abstract

Without a doubt, demand forecasting is an essential part of a company's supply chain. It predicts future demand and specifies the level of supply-side readiness needed to satisfy the demand. It is imperative that if a company's forecasting isn't reasonably reliable, the entire supply chain suffers. Over or under forecasted demand would have a debilitating impact on the operation of the supply chain, along with planning and logistics. Having acknowledged the importance of demand forecasting, one must look into the techniques and algorithms commonly employed to predict demand. Data mining, statistical modelling, and machine learning approaches are used to extract insights from existing datasets and are used to anticipate unobserved or unknown occurrences in statistical forecasting. In this paper, the performance comparison of various forecasting techniques, time series, regression and machine learning approaches are discussed, and the suitability of algorithms for different data patterns is examined.

Keywords: **Supply Chain, Demand Forecasting, Time Series, Statistical Modelling, Data Pattern, Retail, Machine Learning, ARIMA.**

1. Introduction

Demand forecasting, to a large extent, controls almost all the activities of Supply Chain Management. It is a vital force behind Supply Chain Management, Decision-making and Enterprise-level planning in any business establishment. To make meaningful choices like capacity building, resource allocation, expansion, and forward or backward integration, all big enterprises rely on demand forecasting accuracy. Forecasting is the process of predicting or estimating the actual value of something in the future. Today's rapidly changing, volatile, uncertain, complex and ambiguous markets require accurate demand forecasting for the efficient and agile supply chain management. There are different ways to forecast the demand. The forecast may also differ based on the forecasting model we use. Passive Demand Forecasting is used for businesses that are stable and have relatively

conservative growth plans. This type of passive forecasting is generally applicable for local and small enterprises, where the historical data is extrapolated with the fewest assumptions possible.

On the other hand, active demand forecasting is used for scaling up the business of diversified organizations with aggressive growth plans in terms of product portfolio development, geographical expansion and exposure to the external economic climate. Active demand forecasting is also used in start-ups with insufficient historical data to base the predictions. The demand forecasting is also based on the time period for which the projections are made. When the forecasting is only for the next three to twelve months, to quickly respond to the changes in customer demand, it is known as short-term demand forecasting. Medium to long future forecasting is usually applicable if the forecasting is for the next 12 to 24 months. This type of forecasting helps plan business growth and various business functions—the long-term forecasting roles out a roadmap for designing the capital requirement and supply chain operations. The last two types of demand forecasting are external and internal. The external macro-level demand forecasting helps in analyzing how broad market trends influence the business goals. Consideration of external market factors is an essential ingredient of this forecasting to assess a company's strategic objectives. Lastly, the internal business level demand forecasting is concerned with internal business processes such as manufacturing, production, sales, finances, human resources, customer services etc.

Accurate demand forecasting is critical for production, inventory and distribution management and marketing, finance, investment design, research and development, and human resource management. Demand forecasting is a sub-set of demand planning to deliver products on time and keep up the customer satisfaction. Demand forecast determines the volume and placement of the products and the time horizon over which the products are required for marketing. Both the quantitative and qualitative aspects of demand are essential for demand forecasting. The quantitative elements include the volume of goods as determined by the demand, and the qualitative aspects include the type of customer needs. The demand forecasting process broadly depends on selecting the appropriate method, viz, survey methods and statistical methods. Survey methods involve directly asking the consumer about their preferences and their opinions on the future of the product. This method is often used for shorter forecasting horizons and includes consumer survey methods and opinion poll methods. Statistical methods involve very little subjectivity when it comes to forecasting and are often used for longer forecast horizons. These forecasts are also more reliable than survey methods and are less expensive.

Statistical demand forecasting methods are of three types. The Trend Projection approach takes the factors responsible for a variable's past trends into consideration to project the future directions of the variable. This approach is based on the assumption that past trends of a variable influence its future trends in an identical manner. These forecasting methods thus require the data to be in the form of a time series. Barometric forecasting has its application in meteorology and is not relevant to our case. The Econometric Methods employ a mixture of statistical methods and economic theories to forecast future trends. The models used in the econometric models include a single-equation regression model or a set of simultaneous equations. The single-equation regression model is enough for most commodities. Thus, the two broadly used quantitative approaches are based on time series and regression.

As an extension of the time series approach, a Long Short Term Memory (LSTM) based Recurrent Neural Network (RNN) is used for some applications. An LSTM is a type of neural network architecture that deals with sequential data exceptionally well. The output of one cell is dependent on the output of the previous cell and the input of the previous cells. Simply put, an LSTM remembers the previous inputs that have been provided to it along with the previous outputs, thus being able to establish and learn a relationship between the output of the current cell and the previous data. In recent years

LSTM has been used extensively for multivariate time series forecasting and, in some applications, in conjunction with regression-based models. In this paper, we have discussed various time series and regression-based algorithms and introduced specific LSTM architectures that have proved successful in demand forecasting.

2. Time Series Forecasting

Time series is a chronological sequence of observations. Time series forecasting is used to predict events based on a sequence of time and has numerous applications. This technique forecasts future events by examining historical trends and assuming that future trends will follow prior trends. This forecasting method leverages information about historical values and associated patterns to predict future activity. This usually has to do with trend analysis, cyclical fluctuation analysis, and seasonality difficulties. According to the type of sales data and expected output, there are many time series forecasting methods used in the retail industry. Some of them are:

Simple Naive and Seasonal Naive Method – If we forecast for the future period using the prior period's data, without making any adjustments or seeking to identify causal causes., it is called the simple naïve method of forecasting. Seasonal Naive is a comparable strategy that works well with data that is highly seasonal. In the seasonal naive forecasting method, the most current observed value from the same year's season is used for the forecast.

Average Methods - Mainly, three types of average method are used for forecasting, viz., (i) **Simple Average**, (ii) **Simple Moving Average (SMA)** and (iii) **Weighted Moving Average (WMA)**. In the first average method, the forecast of all the future numbers is equal to the average or mean of the past data. The second method, i.e., SMA, is an average price over a specified period. More recent data points are given a higher weighting in the WMA method as they are more relevant than those from the past. The total weighting should either be equal to one or 100 per cent. All these methods require saving lots of past data points, and it ignores complex relationships in data.

Exponential Smoothing Methods – The three methods under this category are:

1. **Single Exponential Smoothing (SES) -** This strategy works well when there is no evidence of trend or seasonal pattern in the data and estimates only level component. It's a type of weighted moving average that uses historical data with decreasing weights. The forecast is given by New Forecast = Last Forecast - α (Last Forecast Error). Here, $0 < α < 1$ and is generally small for the stability of forecasts.

2. **Double Exponential Smoothing (DES) -** For univariate time series data, Double Exponential Smoothing can model trend components and level components. It consists of a forecast equation and a couple of smoothing equations, one for the level and the other for the trend. It has an additional parameter beta (β*) which is the trend factor.

3. **Triple Exponential Smoothing (TES) -** Also known as the Winters seasonal method, this method consists of the forecast equation and three smoothing equations – respectively for level, trend, and seasonal component - each with its own smoothing set parameters α, β∗ and γ.

ARIMA Models – 'Auto-Regressive Integrated Moving Average', ARIMA in short, is a linear modelling technique. It is a class of models used to forecast future values by explaining a given time series based on its own past values. ARIMA models can be used to model any non-seasonal time series with patterns and without random white noise. The present data used in ARIMA models is assumed to be a linear function of previous data

points and mistakes.

 ARIMA model can be explained as under:

Predicted Forecast = Constant + Linear combination Lags of Y (upto p lags) + Linear Combination of Lagged forecast errors (upto q lags).

An ARIMA model is characterized by three terms: *p, q,d,* where
- *p* is the order of the AR (auto-regressive) part of the model term.
- *q* is the order of the MA (moving average) term.
- *d* is the number of differencing, i.e., the number of previous time points that must be subtracted from the current value for the time series to become stationary. If the statistical features of a Time Series, such as mean and variance, remain constant across time, it is considered stationary.

Short Term Load Forecasting (STLF) - STLF modelling presumes that a time series may be divided into three parts: error, trend, and seasonality. As against discovering patterns in the original time series, decomposition might make the process easier, and repeating the decomposed pattern in the future, makes the forecasting simply recombining the components. This method is beneficial for complex seasonality.

TBATS – TBATS is an abbreviation for Trigonometric seasonality, Box-Cox transformation, ARMA errors, and Trend and Seasonal components. The TBATS model can handle complex seasonalities such as non-integer seasonality, non-nested seasonality, and large-period seasonality with no seasonality limits, allowing for precise, long-term forecasting. This model is preferred when seasonality varies across time.

Theta Model – This model is based on the idea of changing the time-series local curvature using a coefficient called 'Theta' (the Greek letter, θ), which is applied directly to the data's second differences. The mean and slope of the original data are preserved in the produced series, but not their curvatures. Theta-lines are the names given to this new time series. The suggested technique divides the original time series into two or more Theta-lines.

Croston Model – The model suitable for the products with intermittent demand is Croston [1]. It is a modification of exponential smoothing for sporadic demand product time series. Firstly, the average magnitude of demand is estimated using exponential smoothing, and then the intermittent needs are determined. This information is then employed in a constant model to forecast the future demand.

3. Regression Based Forecasting

Regression-based demand forecasting methods are among the more popular techniques chosen for forecasting demand. The regression methodology approximates the demand function for a product, with demand as the dependent variable and the variables that determine the demand as the independent variables. A demand function with only one variable influencing demand is a 'single variable demand function'. The term 'multi-variable demand function' refers to a demand function that is modified by multiple variables. As a result, multiple regression is applied in this situation. The various regression models are listed below:

Linear Regression - The linear regression model provides a linear relationship between the forecast variable y and a single predictor variable x in the simplest example. This technique involves identifying the independent variable *x* and the dependent variable *y*. With this, a trend line $y = a + bx$ can be constructed to predict y, where b and a are given in (1) and (2) :

$$b = \frac{\sum XY - n\overline{XY}}{\sum X^2 - n\overline{X}^2} \qquad (1)$$

$$a = Y - bX \qquad (2)$$

Random Forest - Random Forest (RF) is one of many machine learning algorithms for supervised learning or learning from labelled data and making predictions based on the patterns learned. Both classification and regression tasks can be performed with RF. Use factors that drive the sales and use random forest modelling to see the relationship between these factors and the sales. Using the factor information in future, forecast the sales in future. This algorithm is capable of solving both classification and regression related issues, as well as providing reasonable estimation in both cases. RF is quite powerful when used on large volumes of higher-dimensional data.

Support Vector Machine - Small sample size, nonlinearity, high dimensionality, and local minima have all been proven to be helpful in forecasting issues using the Support Vector Machine (SVM). The primary idea behind SVM is to use a hyper-plane as a decision surface to maximize the separation between classes. Once trained with sample data, the Support Vector Machine learns a function called the kernel function that transforms the data such that a decision plane is formed between the different classes. A unique hyper-plane called optimize hyper-plane also exists to separate the data at its best.

4. Neural Networks for Forecasting

Historically, linear methods dominate the time series forecasting as they are well understood and work well in simpler use cases. In recent times, neural networks are being employed to make time series-based predictions. They could learn arbitrary complex mappings from inputs to outputs and support multiple inputs and outputs. One disadvantage of using neural networks that we can see is unpredictability at times. This disadvantage is mitigated to a large extent with the use of specific neural network architectures than others. Various architectures under this category that have seen success in demand forecasting include:

Multi-layer Perceptrons - Multi-layer perceptrons (MLP) are the simplest type of neural networks. They are robust to noise in the mapping function and can support predictions even when there are missing values or outliers. They are nonlinear and hence do not lean heavily on the mapping function while making predictions and can learn linear and nonlinear trends. They support multi-variable input as any number of input nodes can be defined while varying the output dimension, which enables multi-variable forecasting. These capabilities make feed-forward networks useful in demand forecasting, where the demand associated with a particular SKU will depend on several variables, including past demand.

Convolutional Neural Networks - Convolutional neural networks (CNN) are used mainly to classify image data. However, their application can be extended to time series forecasting due to their ability to learn features from an image. A set of observations can be considered an image. A convolutional network could extract, distil and learn the features needed to predict the following observation from a set of previous observations. For time series forecasting, CNNs offer all of the advantages of Multilayer Perceptrons, including multivariate input and output and the ability to learn arbitrary yet complex functional relationships. The model can be taught a most relevant representation to the prediction issue from a vast input sequence, and as such, need not learn directly from lag observations.

Long Short Term Memory - Long Short-Term Memory network or LSTM is a special kind of recurrent neural networks (RNN) capable of learning long-term dependencies.

LSTM networks have gained a lot of importance in recent times with their application in many fields. Recurrent neural networks like LSTM incorporate the explicit handling of order between observations, which MLPs and CNNs do not. LSTM networks include native support for input data that are made up of a series of observations. This capability of LSTM networks has been put to excellent use in challenging natural language processing applications such as neural machine translation. The model has to learn the complex interrelationships between words both within and across languages when translating from one language to another.

5. Hybrid Models

Hybrid models used for demand forecasting employ both the time series approach as well as the regression model approach. These models generate the forecast using time series analysis and then perform regression analysis on the resulting data. One such model [3] employs an LSTM network to generate the forecast for the given data, and then computes the residual as:

$$r(t) = x(t) - x^f(t)$$

Where $x(t)$ is the input data, $x^f(t)$ is the generated forecast and $r(t)$ is the residual. The residual signal is then fed to a Random Forest regressor and the result $r^r(t)$ is added to $x^f(t)$ to generate the final forecast.

6. Literature Review of Different Forecasting Methods and Algorithm

There are various techniques and algorithms available today for carrying out demand forecasting. Still, it is essential to review the latest research papers to identify which algorithm or method is best for different data patterns and make the best use of it. Adhikari et al. (2019) discussed various demand forecasting algorithms like Moving Average, SES Model, Croston Model, Seasonal Linear Regression and Double Exponential Smoothing [1]. They have reported that combining the results of the time-series model and the regression-based model produces a superior outcome by eliminating over-forecasting and under-forecasting and bringing forecast values closer to the actual. These results are considerably superior to the individual algorithms used in the two models.

Punia et al. (2020) proposed a novel forecasting approach that blends deep learning – long short-term memory (LSTM) networks and random forest (RF) that was tested on a real-world multivariate dataset from a multi-channel retailer [3]. Their work suggested that the hybrid model method can handle complicated temporal and regression relationships, giving it an accuracy advantage over current forecasting approaches like neural networks, multiple regression, ARIMAX, LSTM networks, and RF. Similar to this, Babu and Reddy (2014) presented a new hybrid ARIMA–ANN model, which first characterizes the given data based on the nature of its volatility [13]. The time-series data considered by them were sunspot data, electricity price data from the Australian National Electricity Market, and the close prices of stocks from the New York Stock Exchange. The technique suggested by Babu and Reddy (ibid) can provide superior accuracy as compared to existing hybrid models that fit an ARIMA model to the input data directly.

It is difficult to apply the available traditional and advanced forecasting techniques to many customers (Murray et al., 2015) [2]. The authors of this study used data mining techniques to discover client categories with comparable demand patterns. The segmentation was calculated using Euclidean distance measurement and k-means based on the monthly volume of product provided.

The Root Mean Square Error (RMSE) and Mean Square Error (MSE) have long been

prominent in statistical modelling, owing to their theoretical relevance. Hyndman and Koehler (2006), who studied and contrasted univariate time series forecast accuracy measures, have proposed that in scenarios with a wide range of scales, including data close to zero or negative, scaled errors become the standard measure of prediction accuracy [8].

One of the most difficult challenges is forecasting demand for special days, which have different demand patterns than regular days [6]. Working in the retail domain, Huber and & Stuckenschmidt (2020) reported a large-scale demand forecasting scenario requiring daily projections at the store level. Compared to time series models with adjustments or a regularized linear regression model, the machine learning models incorporated special day-specific features that reduced error by more than 10% and up to 20% [6].

Like much other economic time series, aggregate retail sales in the United States display strong trend and seasonal patterns. Artificial Neural Networks (ANN) were compared against classic approaches such as Winters exponential smoothing, Box Jenkins ARIMA model, and multivariate regression for US retail sales by Alon et al. (2001) [5]. The overall finding is that ARIMA models produce more accurate forecasts than other econometric models for immediate and short-term forecasts. This study showed that, on average, ANNs outperform traditional statistical methods, followed by the BoxJenkins model.

Hybrid forecasting systems combine various methods to increase forecasting quality as compared to individual techniques. Aburto and Weber (2007) developed a sequential hybrid forecasting system (SHFS) in which SARIMAX is applied to the original time series and then the neural network to the output from the SARIMAX process [7]. The output from the neural network is the sequential hybrid forecast for the original time series. This system was used to forecast a supermarket's sales data. The same authors in another work reported in the same year [11] described a hybrid intelligent approach for demand forecasting that aids supply chain management in the retail sector. In a supermarket, providing advanced projections allow all agents in the chain to manage their inventory selections better. Neural networks surpassed ARIMA models in terms of forecast accuracy, and the proposed additive hybrid strategy, which combines SARIMAX and Multilayer Perceptrons-type neural networks, produced the best results.

Another highly competitive market where inventory control plays a crucial role in the business's profitability is, Fashion retail. Loureiro et al.. (2018) have shown the potential of using a deep learning approach, mainly to estimate the sales of future products for which no historical data exists [14]. This study demonstrated the efficacy of Deep Neural Network (DNN) and other data mining techniques for forecasting sales in the fashion retail industry, where there is no historical sales data.

The retail food industry's time series sales are characterized by significant volatility and skewness, which change with time (Arunraj and Ahrens, 2015) [10]. To anticipate daily sales of a perishable product, the authors of this paper have created a seasonal autoregressive integrated moving average with external variables (SARIMAX) model. They have built and used the SARIMAX with Multiple Linear Regression (SARIMA-MLR) and hybrid SARIMA with Quantile Regression (SARIMA-QR) models to anticipate daily banana sales in a German supermarket. In comparison to seasonal naive forecasting, both of the above models produce better forecasts for out-of-sample data models. However, the SARIMA-QR model has an advantage over the SARIMA-MLR model. It allows for direct and accurate forecasting of greater service levels without the need for extrapolation. It was further reported that the SARIMA-QR model could help businesses make correct and suitable judgments whether the focus of interest is on higher (for promotional activities) or lower (due to harsh weather conditions) sales.

A complete framework for generating nonlinear time series sales forecasting models was presented by Doganis et al. (2006) [9]. The GA-RBF technique suggested by them combines two powerful artificial intelligence technologies, namely the RBF neural network architecture and a specially constructed genetic algorithm for selecting acceptable

explanatory variables. This was verified using fresh milk sale data from a large dairy product manufacturer. All other setups, including neural network modelling, performed worse than the RBF model that used simply prior sales volume values.

To learn fuzzy IF-THEN rules for promotion gathered from marketing professionals, Kuo (2001) presented a fuzzy neural network with initial weights created using a genetic algorithm (GFNN) [4]. The GFNN output is then combined with an ANN forecast based on time series data and another ANN's promotion length. The results of combining ANN and GFNN models for a convenience store (CVS) company show that the suggested system outperforms the traditional statistical technique and a single ANN.

Grocery sales forecasting has become more challenging due to promotions and shorter life cycles necessitating a more complicated methodology, as indicated by Ali et al. (2009) [12]. For a more accurate model, Ali et al. (ibid) proposed a model using regression trees with explicit features constructed from sales and promotion time series of the focal and related SKU-store combinations, as well as large-scope models to exploit product and store similarity. Individual time series-based exponential smoothing to stepwise regression and SVR with three kernels are among the approaches used. At the same time, the scope of models ranges from single store-SKU models to those including multiple SKUs and stores. When given rich input data containing generated explicit features, the regression tree methodology significantly enhanced forecast accuracy. The findings of this study show that in case of promotional data, employing more specific input data is only helpful if more advanced methodologies are applied.

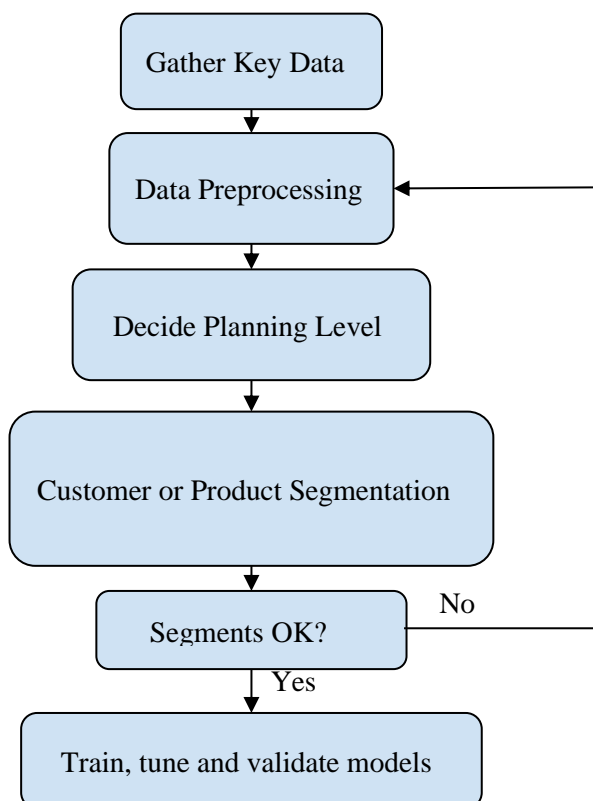## 7. Demand Forecasting Methodology for Retail Industries

Figure 1: Demand Planning Methodology

After researching various methodologies followed by the retail industries, we came to

the conclusion that the methodology presented in this paper is unique and effective when compared to the existing forecasting methodology adopted in the industry.

7.1. Gather Key Data

Data in today's world is very important as most of the supply chain management processes work on data. Gathering key data is the initial step for carrying out demand planning. It includes gathering the most important data like product, customer, retailer, time, region and other related details which are mostly static in nature. After gathering the master data, the next step is to gather all the fact data which are mostly dynamic in nature like sales history, shipments, orders, marketing, promotion and other details as per requirement.

7.2. Data Preprocessing

After gathering all the data, data preprocessing is required before proceeding to further steps. The data needs to be transformed into vectors so that it is easy to perform various cleaning operations on it. Unexpected events, stockouts and dumping etc., will have an impact on actual sales. They behave as outliers in data, and it is handled with outlier correction by using statistical language like R Programming. In some cases, there is a bullwhip effect, which can be reduced using data preprocessing methods.

7.3. Decide Planning Level

All the data for various intersections cannot be worked upon. To get an appropriate forecast, level of forecast needs to be decided. After processing the data, demand planners are required to select the appropriate planning level to forecast demand. The planning level is decided by considering various factors like which level drives the business of company, where is the noise in actual data and which level would give better forecast accuracy. The quad plot and scatter can be used here to get a clear picture of the most ideal intersection of attributes for conducting the forecast in order to achieve a high level of accuracy. It is an iterative procedure that takes multiple iterations to achieve a precise level.

7.4. Segmentation

Often deciding the planning level is to be followed with segmentation. If the number of products are huge, then product segmentation is used and if the number of customers is large, then customer segmentation is used. It is also possible to have both the segmentation when working on a very large scale. Segmentation is responsible to divide the products or customers into categories where each category represents some specific characteristics. It is generally based on mathematical attributes like volume of the sales , product life cycle, intermittency, coefficient of variability etc.. If the company is satisfied with the segmentation technique they go to final step of forecasting else they go back to the data preprocessing step to prepare data from scratch.

7.5. Train, Tune and Validate Models

This is the final and most important step where various algorithms are tested on the prepared history data and the best fit model is chosen to predict the future demand of upcoming months or even years. Testing of algorithm has to be done for each Stock Keeping Unit (SKU) or product and best fit algorithm may differ for each product or each product segment. The validation of various model is done my using accuracy metrics such as MAPE, MSE, RMSE etc. If the best fit model accuracy of certain SKU is low, various tuning of parameters has to be performed or a method from the pool of algorithms is suggested to be force fitted on the product for increasing the accuracy. But the main issue is, a variety of algorithms has to be tested

to filter out the best model. It would be easy if some algorithms are predefined for a specific data pattern.

8. Statistical Algorithms Used For Different Data Patterns

We tested some algorithms on retail data having different data patterns. Generally, a variety of statistical procedures can be used to analyse data that has a large sales volume, a lengthy history, and does not deviate significantly from its mean over time (variability). Simpler methods, such as moving averages and naïve approaches, are recommended for low volume, low history, and high variability. New Products require a trend reversal because their sales are expected to skyrocket. A new product's forecast can be based on the forecast of goods with similar features until it reaches maturity. For End Of Life products or the product whose demand is going steeply less, dampening for decreasing sales is required. For Intermittent products (several periods of zero demand), specialized algorithms such as Croston's method is advisable, as it is very difficult to predict the frequency and quantity of demand using traditional statistical forecasting techniques. after seeing the outcomes of numerous statistical algorithms we created a table that can be used as a reference to know which algorithms to test on which sort of retail data pattern.

Table 1: Algorithms for Different Data Patterns

Description of Data	Algorithms
High volume, low variability	ARIMA, TES, TBATS, NNET, STLF
High volume, high variability	ARIMA, TES
Low volume, low variability with long history	ARIMA, TES Auto, TES Damped, STLF ,Seasonal Naive
Low volume, high variability, lesser history	DES, SES, Seasonal Naive, Moving Average
Intermittent	Croston, Seasonal Naive
New launches	DES, TES
End of Life Product	Simple Naive

Abbreviations of different algorithms shown in the Table 1 are :

ARIMA - Autoregressive Integrated Moving Average
SES - Single Exponential Smoothing
DES - Double Exponential Smoothing
TES - Triple Exponential Smoothing
STLF - Short Term Load Forecasting

7. Conclusion

In order to generate an accurate forecast, one must pick the appropriate technique based on the nature of the data as well as the specific use case. In some cases, it may be necessary to use deep learning or regression techniques, but in most of the simpler use cases, traditional methods will yield excellent results. In the retail industry, demand forecasting is of utmost importance as the demand for each SKU has to be accurately predicted in order to plan the inventory, as excess inventory would increase holding costs. However, on a larger scale, demand forecasting plays a very important role in planning the entire supply chain infrastructure of a company. Apart from the qualitative measures that have been explained in this paper, there is also a quantitative method of forecasting which involves taking into account insights from subject matter experts who formulate the demand plan. With the advance of neural networks and their capabilities, one could design a network that is capable of arriving at these intuitions and applying them to the forecast in the least subjective way possible so as to generate a forecast that is as accurate as can be. It is in an industry's best interest to endeavor to develop more advanced demand forecasting techniques, given how important it is for the development of business.

REFERENCES

[1] Adhikari, N., Domakonda, N., Chandan, C., Gupta, G., Garg, R., Teja, S., Das, L., and Misra, A. (2019) An intelligent approach to demand forecasting, In International Conference on Computer Networks and Communication Technologies (pp. 167–183).

[2] Murray, P.W., Agard, B. and Barajas, M.A. (2015) Forecasting Supply Chain Demand by Clustering Customers. IFAC-PapersOnLine 48 (2015): pages 1834-1839.

 [3] Punia, S., Nikolopoulos, K., Singh, S.P., Madaan, J.K. and Litsiou, K. (2020) Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail, International Journal of Production Research, 58:16, pages 4964-4979, DOI: 10.1080/00207543.2020.1735666.

[4] Kuo, R.J. (2001) A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm, European Journal of Operational Research, Volume 129, Issue 3, 2001, pages 496-517.

[5] Alon, I., Qi, M. and Sadowski, R. (2001) Forecasting aggregate retail sales: A comparison of artificial neural networks and traditional methods. Journal of Retailing and Consumer Services. 8. 147-156. 10.1016/S0969-6989(00)00011-4.

[6] Huber, J. and Stuckenschmidt, H. (2020) Daily retail demand forecasting using machine learning with emphasis on calendric special days. International Journal of Forecasting, Elsevier, vol. 36(4), pages 1420-1438.

[7] Aburto L. and Weber R. (2007) A Sequential Hybrid Forecasting System for Demand Prediction. In: Perner P. (eds) Machine Learning and Data Mining in Pattern Recognition. MLDM 2007. Lecture Notes in Computer Science, vol 4571. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73499-4_39.

[8] Hyndman, R.J. and Koehler, A.B. (2006) Another look at measures of forecast accuracy, International Journal of Forecasting, Volume 22, Issue 4, 2006, Pages 679-688.

[9] Doganis, P., Alexandridis, A., Patrinos, P. and Sarimveis, H. (2006) Time series sales forecasting for short shelf-life food products based on artificial neural networks

and evolutionary computing, Journal of Food Engineering, Volume 75, Issue 2, 2006, Pages 196-204, ISSN 0260-8774, https://doi.org/10.1016/j.jfoodeng.2005.03.056.

[10] Arunraj, N.S. and Ahrens, D. (2015) A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting, International Journal of Production Economics, Volume 170, Part A, 2015, Pages 321-335, ISSN 0925-5273, https://doi.org/10.1016/j.ijpe.2015.09.039.

[11] Aburto, L. and Weber, R. (2007) Improved supply chain management based on hybrid demand forecasts, Applied Soft Computing, Volume 7, Issue 1, 2007, Pages 136-144, https://doi.org/10.1016/j.asoc.2005.06.001.

[12] Ali, Ö.G., Serpil Sayın, Woensel, T.V. and Jan Fransoo, SKU (2009) Demand forecasting in the presence of promotions, Expert Systems with Applications, Volume 36, Issue 10, 2009, Pages 12340-12348, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2009.04.052.

[13] Babu, C.N. and Reddy, B.E. (2014) A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data, Applied Soft Computing, Volume 23, 2014, Pages 27-38, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2014.05.028.

[14] Loureiro, A.L.D., Miguéis, V.L. and Lucas F.M. da Silva (2018) Exploring the use of deep neural networks for sales forecasting in fashion retail, Decision Support Systems, Volume 114, 2018, Pages 81-93, ISSN 0167-9236, https://doi.org/10.1016/j.dss.2018.08.010.