Implementation of Demand Forecasting – A Comparative Approach

Punit Gupta1, Harshit Ladia1, Kabir Kakkar1, Kriti Rai1, Yogesh Agrawal1, Rishika Mamgain1, Navaditya Gaur1

1Department of Computer and Communication Engineering,
Manipal University Jaipur, Jaipur, Rajasthan, India
punitg07@gmail.com

Abstract. Forecasting is a crucial factor in development of store head businesses either commercial or economical that require frontend sales to outgrow. Demand forecasting hence deals with providing our stores a significant amount of backup supply prediction which hence can deal with the supply to meet the ups and downs of our demand. Hence, a predictive model has been developed to maintain the supply of goods and featuring the predictions of upcoming demand for the next few selected time period. The model involves being highly trained in the field of Machine Learning using various predefined models such as Linear Regression, GBT Model and even using the time series analysis for the defined period and using the outcome from the factor of inputs provided by the user to finally predict the upcoming demand and provide (increase or decrease) the given supply in the specific field, thereby saving the businesses from any uncertain downhill and hence helping analytics to provide more reasonable theories for increasing the profit margin in the upcoming foreseeable future and even get a track record and deep insights of the previous demand using a business intelligence software, like PowerBI.

Keyword: Prediction, Sales, Demand Forecasting, Machine Learning

1. Introduction
The demand forecasts predict progress on the survival of business in a world where human society has traditionally undermined the most established structures. One of the perils of incentives is trade. The survival and endurance of this era of economic predators require strategy, expertise, and technology to effectively predict the future. Estimating or predicting is becoming a sign of existence and business language.

Economically speaking, demand is defined as the quantity of a product or service people are able to afford and buy in time or time. Consequently, demand estimation is fine art for estimating the amount of demand that may occur at some point or another in the future. Describes the strengths, weaknesses, limitations/responsibilities of major assessment methods, and their relevance in tourism and travel. It is important for a business to look at their future and prospects in terms of sales, cost, and profit. The value of future sales and demand is crucial because it affects cost savings, so forecasting future sales is a rational starting point for all business planning and the company.

Companies typically use three-step approaches to develop a sales forecast. They are environmental forecast, then industry forecast, then company sales forecast, inflation, unemployment, interest rates, consumer spending, profitability, business investment, government spending, net exports, additional environmental agency significant events, and events. To produce the necessary items in a timely manner and the various components, raw materials, of the future product. Equipment, machine tools, labour, etc. Prediction forecasting helps a company evaluate its product needs and plan its product accordingly. This allows management to reduce their dependence on luck or opportunity. Demand estimation is useful in better planning and national resource allocation. Demand estimation is favourable in industrially developed
countries, where demand conditions are always more uncertain than supply conditions. However, in developing countries, supply is often a limiting factor rather than demand. High prices and block markets indicate disruptions in supply [4].

Obviously, in a country like India, supply estimation is more important than the demand estimate. However, with the recent amendment of industrial licensing regulations and economic liberalization, competition for most sectors has increased except for the need for large investments. In such areas, supply is much higher than demand, and manufacturers are starting to struggle for the market. Therefore, demand estimation in India is important.

Demand assessment is a key business process that drives the company's strategic and operational plans [1-2]. Based on demand assessment, the business implements essential and long-term plans, such as budget and financial planning, sales and marketing plans, hierarchy planning, risk assessment, and mitigation plans. Small and medium-sized strategic strategies such as pre-building, make-to-stock, order-to-order, contract building, distribution outlines, and network balancing are all based on implementation. Demand forecasting further promotes key managerial functions such as decision making, performance appraisal, fair allocation of resources in a managed environment and business expansion planning.

2. Related Work

An organization named Frepple has been predicting the sales team’s needs for setting the expected sales volume and analyzing the demand trends. The forecasting algorithm produces forecast methods for Steady demand (single exponential smoothing), Trending demand (double exponential smoothing), Seasonal demand (Holt-Winters additive method), Intermittent demand (Croston method), Moving average. The system automatically chooses and tunes the most suitable procedure that gives the lowest forecast error.

Forecast methods of the company are:

1. Moving average: This method solely calculates the average of the demand history for the last N time buckets where N is a configurable parameter.

2. Constant: This method is an implementation of the sole exponential smoothing method. It is practiced when the demand doesn’t evolve in time.

3. Trend: This process is an implementation of the double exponential smoothing method. It is used when a trend, either positive or negative, is observed in sales history.

4. Seasonal: This method is an implementation of the Holt-Winter’s process and is used when a seasonal pattern is detected in the demand history. A seasonal prediction has a recurring pattern in time such as ice creams being much more sold during summer related to winter.

5. Intermittent: This method is an implementation of the Croston design. This method is proper for sales history with intermittence, i.e. a sales history with many zeros.

Dynamic linear models [5] -

Dynamic linear models describe the different classes of models for time series forecasting. The concept is that at each time t these models correspond to a linear model, but the regression coefficients vary in time. An example of a dynamic linear model is given below.

\[ y(t) = \alpha(t) + t\beta(t) + w(t) \]  (1)
Dynamic linear models can be typically represented in a Bayesian framework; though supreme likelihood estimation procedures are still accessible. Similarly various models are proposed in past using LSTM for forecasting like energy model proposed by suganti and Samuel [3] for energy demand forecasting.

Some applications were LSTM is meant usefull for forecasting are language prediction [7], optimizing space odyssey [8], phoneme classification [9], precise timing [10] & analysis of time series [11] various other models which are considered in study are [12-15].

3. Proposed Model

This study aims to analyze the consumer mentality for purchasing a particular product and increase the profit by one company by simply going through the data of sales of previous years and recording its trend or pattern that can take place again in the future while analyzing the present pattern carefully.

Vast knowledge of Machine learning is applied here by carefully carrying out the general phases of analyzing averages applied under consideration that the data provided has been accurately worked upon.

The models are therefore divided into training and testing splits and thus implementing the model for obtaining an optimal accuracy while testing them amongst every different time series model. Transactional Sales data will be pulled into Azure SQL Database, using Azure DataFactory. The pulled data will then be read into Azure Databricks where a predictive machine learning model will be built. The predictions will be made into Azure Databricks using the trained Machine Learning model. The predictions will then be exported to the SQL database and will then be visualized on PowerBI.

First, we create a pipeline and the number of steps the data has to go to and fit the model via the pipeline. After completing the model configuration, we create the workspace and assign it to a resource group through which it can store and cooperate with different resources such as the deployment configuration in one single entity. Then we register the model to our ML Azure workspace. After registering, we set the required environment needed for the deployment. The required configuration of the CPU and memory desired is then set and finally, a container image containing our files is then uploaded to our Azure ML workspace. After getting the desired environment to work with, we capture a test data in a JSON file and send it to the workspace through which we get the input, and finally get the desired prediction needed.
Then our input data contains the features required by the model and the output we get is the predicted data sent by our model on the cloud.

3.1. Linear Regression Model
The Linear Regression model is a forecasting approach which has the potential to provide not only demand forecast for the dependent variable, but it also provides useful managerial information for adapting to the changes that cause the dependent variable to change.
In this model independent columns were named as “features” to predict the value of sales based on these independent variables

\[ y = a + bx, \quad (5) \]

where \( x \) is the independent variable and \( y \) is the dependent variable.

After fitting data using model predictions were made and RMSE and \( r^2 \) values were checked.
The predictions of this model were compared to the GBT tree model.

3.2. GBT Model
A Gradient Boosted Tree (short for GBT Model) helps in the making of a tree for the demand produced by the supplier being built in a greedy fashion, which can be incredibly good for getting the details of the company enlisted in a single regression tree.
The procedure or structure which can be represented in a pipeline is fitted into an iterative process which is found a bit less identical than the optimization of numerical methods via the gradient descent method.

\[
\text{grad}(\mathcal{L}(F)) = \left( \frac{\partial L(y_1, F(x_1))}{\partial F(x_1)}, \frac{\partial L(y_2, F(x_2))}{\partial F(x_2)}, \ldots, \frac{\partial L(y_N, F(x_N))}{\partial F(x_N)} \right) \quad (6)
\]

In each step, a singular regression tree is built so that an anti-gradient vector component can be predicted via the bottom down approach of the regression tree.

The length of the step is even calculated corresponding to the loss function and every region is separated which is like the loss function’s region determined by the key. It can be removed by separating the values and thereby changing it from the leaves directly.

Since this is part of the regression problem, the main steps followed by our model will be to first determine the constant model and then iterate over the size of the model.

1. Compute and calculate the values to be plugged in for the anti-gradient.
2. Build the regression tree which will be used to determine the anti-gradient factors and thereby its components.
3. Change the values in the tree leaves.
4. Plug in the tree for regression inside the model.

Thereafter, a loss function will be implemented for the tree built by our regression problems which is going to be used in determining the model.

The following features will be displayed in our loss function:

Huber Loss

\[
L(y, f(x)) = \begin{cases} 
\delta \cdot \left( |y - f(x)| - \frac{\delta}{2} \right) & : |y - f(x)| > \delta \\
\frac{1}{2} \cdot (y - f(x))^2 & : |y - f(x)| \leq \delta 
\end{cases} \quad (7)
\]
Squared Loss

\[ L(y, f(x)) = \frac{1}{2}(y - f(x))^2 \]  (8)

Absolute loss

\[ L(y, f(x)) = |y - f(x)| \]  (9)

4. Results Section

In this section we have shown the various outcomes of the work and the accuracy of the work using various measures like Accuracy of prediction, RMSE value and many more. The data set of the medical store is used to analyse and predict the future sales and demands. The product uses Azure cloud platform for the data analysis and prediction using machine learning models.

Predictions were made by fitting the Linear Regression model as in the figure there is a column named prediction which carries the predicted values corresponding to the features it contains:

The RMSE and r2 value of the LinearRegression model was checked. It can be seen that r2 value is 0.76 approximately and RMSE is 1511.30:

The GBT model is fitted into the pipeline which contains our regression tree and then predictions were made as the figure depicts the predictions corresponding to the sales and features columns:
The RMSE (Root mean squared error) of GBT tree model is checked and it came down from 1511.30 to 704.91 using the GBT model:

```python
from pyspark.ml.feature import VectorAssembler
headers={"ContentType":"application/json"}
input_data = """
| "data": [{
  "Store": 299.0,
  "DayOfWeek": 2.0,
  "Customers": 746.0,
  "Promo": 1.0,
  "SalePerCustomer": 9.953083109919572,
  "WeekOfYear": 21.0,
  "StoreType": 4.0,
  "Assortment": 3.0
}
]"
```

http_res = requests.post(
    service.scoring_uri, input_data,
    headers = headers)
print("Predicted : ", http_res.text)
### Actual Sales: 7425.0

The data is plugged in our json file and prediction is hence fetched from the servers:
Prediction graph of GBT:

![Prediction graph](image)

PowerBI report: Figure 3–9 showcases various components of power BI report which helps the client to take a look at the current status of inventory. Once the predictions were done, we used Power BI to make a dashboard of the predicted data so that the user can understand it better. Some of the visualizations are – **We made a card of the total sales of the 41088 stores from our prediction.**
Figure 3. Total Sales
A card of the total customers across 41088 stores.

Figure 4. Customer count
A clustered column chart of the sum of stores open per week of the year. The axis is the WeekOfYear and the value is sum_of_store.

Figure 5. sum of stores by week of year
Figure 6 pie chart of the sum of stores vs day of week. Here we can see how many stores are open per day of the week.

Figure 6. Number of days the store is open
An area chart of sales per month. The axis is month and the value is sales. This shows that the sales were highest in which month and vice versa.
A card of the total stores on which the predictions were done.

A line chart of the total sales by customer. Here the axis is customers and the value is sales.

Figure 7. Monthly sales

Figure 8. Total stores

Figure 9. Total sales

Figure 10 & 11 showcases the complete live dashboard to the user with live prediction and feed based on the status of inventory.
5. Conclusion

In the retail industry, one of the main demands is to advance the supply chains, lessen costs, and expand sales, profits, and client supporters. To master this problem, there are many different approaches, such as statistical interpretation and machine learning strategies, to visualize and determine complicated communications and models from traditional data.

The model we have developed uses all the advanced demographic schemes utilized in demand estimation areas. In the effects, we have seen that the design also regression-based model yields a more reliable result,
thanks to the production of higher estimation and lower estimation and edge reference values. Special. These models are much better than the specific algorithms used in the 2 models. Separate assessment demand is very essential for the outlining team. The information and model established in this research are conducting sales data for different markets. As scheduled work, we decide to accelerate the model by collecting data from additional origins, such as economic readings, buying drifts, social media, social functions, moreover location-based demographic data. Another factor is often used to determine the parameters of deep learning techniques. In detail, we decide to use intensive learning techniques such as Confidential Neural Network, Recursive Neural Network, and training algorithms such as Deep Neural Network.

REFERENCES

[1] Wooldridge, J. M.. Introductory econometrics: A modern approach. Nelson Education., 2002, pp.191-193.

[2] Campbell, J. Y., Champbell, J. J., Campbell, J. W., Lo, A. W., Lo, A. W., & MacKinlay, A. C. The econometrics of financial markets. Princeton University Press. (1997)

[3] Suganithi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. Renewable and sustainable energy reviews, 16 (2), pp.1223-1240.

[4] Daganzo, C. (2014). Multinomial probit: the theory and its application to demand forecasting. Elsevier.

[5] Michie, D., Spiegelhalter, D. J., & Taylor, C. C. (1994). Machine learning. Neural and Statistical Classification , 13 (1994), pp. 1-298.

[6] Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM. IET journal, pp. 850-855.

[7] Sundermeyer, M., Schlüter, R., & Ney, H. (2012). LSTM neural networks for language modelling. In Thirteenth annual conference of the international speech communication association.

[8] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: A search space odyssey. IEEE transactions on neural networks and learning systems, 28 (10), pp.2222-2232.

[9] Graves, A., Fernández, S., & Schmidhuber, J. (2005, September). Bidirectional LSTM networks for improved phoneme classification and recognition. In International Conference on Artificial Neural Networks, pp. 799-804. Springer, Berlin, Heidelberg.

[10] Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. Journal of machine learning research, 3, pp.115-143.

[11] Quenouille, M. H., & Quenouille, M. H. (1957). The analysis of multiple time-series (Vol. 1). (London: Griffin).

[12] Bjoern  H  Menze,  B  Michael  Kelm,  Ralf  Masuch,  Uwe  Himmelreich, Peter Bachert, Wolfgang Petrich, and Fred A Hamprecht. A comparison of random forest and its gini importance with standard chemometric methods for the feature selection and classification of spectral data. BMC bioinformatics, 10(1), pp.213, 2009.

[13] K-R Muller, Alexander J Smola, Gunnar Ratsch, Bernhard Scholkopf, Jens Kohlmorgen, and Vladimir Vapnik. Predicting time series with support vector machines. In International Conference on Artificial Neural Networks, pp 999–1004. Springer, 1997.

[14] Gregory Daniel Noble. Application of modern principles to demand forecasting for electronics, domestic appliances and accessories. 2009.

[15] David W Opitz and Richard Maclin. Popular ensemble methods: An empirical study. J. Artif. Intell. Res.(JAIR), 1999,11,pp.169–198.