EAI Endorsed Transactions
on Intelligent Systems and Machine Learning Applications EAI.EU

Influence of Promotion and Pricing on Purchase Incidence, Demand, and Sales Using Machine Learning

Rahul D Shanbhogue[1], Anwesh P R[2,*] and Narayana Darapaneni[3]

[1] PES University, Bangalore, Karnataka, 560050, India
[2]Great Learning, Hyderabad, Telangana, 500089, India
[3]Northwestern University, Evanston, IL 60208, United States

Abstract

The consumer goods industry is a dynamic and fast-paced sector that faces significant challenges in meeting the consumer's ever-evolving demands and preferences. Today's retail businesses focus on optimizing their supply and retail execution to maintain a competitive edge in the market and remain profitable. The most impactful method is to offer promotional events that stimulate large-scale purchases and attract new customers. Thus, it is vital to capture the influence of promotions on demand and sales to efficiently and effectively plan them. The study discusses the influence of different promotion patterns on normal sales days, promotion days, non-promotion days, and other promotion factors on demand and sales. This paper aims to unravel the intricacies of promotion and pricing strategies within FMCG retail, with further integration of customer segmentation, demand sensing, and sales prediction with Machine Learning (ML) and Deep Learning models to enhance strategic decision-making processes thereby enabling businesses to respond to more quickly to changes in the market and mitigate risks associated with the impact of disruptions and to ensure the continuity of the business. Machine Learning Models such as K-Means Clustering with PCA are used to identify customer segments based on features like – customer profile, behavior, and purchase pattern. A neural classification model is employed to obtain the incidence of purchase. Deep Learning models like the ensemble ELM model and the Bi-directional LSTM model are comparatively studied to predict demand and sales while simultaneously understanding the influence of promotion and pricing strategies.

Keywords: Promotion, Pricing, Demand, FMCG, Clustering, Machine Learning

Received on 09 February 2024, accepted on 04 April 2024, published on 10 April 2024

doi: 10.4108/eetismla.5082

1. Introduction

The consumer goods industry is a dynamic and fast paced sector that faces significant challenges in meeting the consumer's ever-evolving demands and preferences. Today's retail businesses focus on optimizing their supply and retail execution to maintain a competitive edge in the market and remain profitable. Technology has been playing a pivotal role in enabling businesses to optimize their operations and processes, as well as monitor and identify risks.

The advent of online retail and delivery has improved accessibility and convenience in purchasing goods. The popularity of social media has highly influenced customers and directly influenced their buying decisions. These factors have led to a rise of influencer and performance marketing which are effective ways for businesses and brands to reach consumers faster. This is evident in the way shoppers have begun basing their purchase decisions more on consumer-to-consumer interactions than commercials.

In the market, it is very important to remain competitive. The most impactful method is to offer promotional events that stimulate large-scale purchases and attract new customers. Thus, it is vital to capture the influence of promotions on demand and sales. However, large demand caused by promotions can create challenges in daily operations, especially for inventory control, delivery planning, and replenishment scheduling. Employing an accurate sales prediction method can relieve such problems.

Demand sensing monitors trends in the market, economic factors, trends in social media, promotional campaigns, pricing strategies, inventory levels, shipping history, logistics, and order transactions to estimate the demand ahead of time by predicting consumer demand, purchase incidence, and sales over a period of time.

Demand sensing can help win time by instantly responding to early demand insights and providing better input supply signals to replenish stocks, determine the stock threshold, and optimize inventory by factoring in expected and refreshed demand. With higher visibility in the supply chain

[*]Corresponding author. Email: anwesh@greatlearning.in

process, it is possible to optimize inventory when disruptions in product demand occur by making timely decisions to eliminate excess stocks, dead stocks, or obsolete stocks. To improve product positioning in the market, it is vital to understand customer profiles and demand by qualitative and quantitative factors to identify customer segments. The data from customer behavior, promotions, and transactions at point-of-sale are used for training the machine learning and deep learning models to sense demand and predict dynamic product prices.

Sales forecasting techniques were important in strategizing many supply chain activities. Traditional statistical forecasting techniques such as exponential smoothing, moving averages, ARIMA, and the SARIMA model were the most widely employed techniques. Over a period of time, more sophisticated methodologies have been employed in the scope of sales and demand forecasting. Many studies have discussed the use of non-linear models when dealing with modeling retail sales data have performed better than linear regression models.

This study aims to understand the effect of promotion and pricing strategies for retail businesses to maximize demand for each brand. Explore the use of ML models like Clustering and Neural Networks to identify various patterns in the demand signals and the influence of promotion and pricing on sensing the demand of each brand. Predictions made by demand sensing models and purchase incidence models will facilitate the prediction of the product purchase and fine-tune promotions and pricing of a product. Improved efficiency in sensing demand using Machine Learning can help businesses respond more quickly to changes in the market by enabling them to make better-informed decisions that can mitigate risks associated with the impact of disruptions and ensure the continuity of the business.

2. Related Work

Every business wants to ensure that its supply chain and retail execution are optimized to maintain a competitive edge in the market and remain profitable. The most significant way is to offer promotions that provide stimulus to large-scale purchases and attract new customers. It is crucial to understand the influence of promotions and pricing on demand and sales and plan them.

The promotion effect in the retail industry can be measured with the following performance indicators – the average number of quantities sold per day, the average number of transactions with the promoted product, the average value of a transaction containing the promoted product while ignoring the value of the promoted product, the average number of distinctive products in the transaction, and average number of transactions. The forecast on these indicators using the XGBoost model and evaluation based on promotion efficiency have given satisfactory results with changes in price or duration of promotion. The change in product price is a prime factor that affects the amount of sold units during promotions [1].

A method to forecast safe levels in retail for promotions using a cluster-based random forest model is discussed. Hierarchical clustering based on the location and sales volume of the store and further clustering based on price, promotion, sales, date, and product over daily, weekly, and monthly cycles have produced better clusters for modeling. The operation plans for promotions and customer behavior vary across regions. It is observed that price and promotion largely influence purchasing patterns. The clusters formed are classified as high, medium, and low sales groups. Weighted MAPE is used for the calculation of errors in the model [2]. Clustering techniques like KMeans are used to identify homogenous commodity groups into different purchase and demand groups, these segments are then fed into the demand prediction models. [3], [4]. The ensemble models have been observed to have performed better on these clusters [5]. The clustered segments are fed to Neural models to predict demand [6] [7] [8].

Along with appropriate feature engineering and scaling techniques the potential to obtain valuable patterns to predict sales of products using Neural models [9] [10]. Combinational models of time series forecasting techniques and neural network models like convoluted LSTM have good performance in demand prediction, although they are computationally more intensive [11] [12].

The joint effect of promotion and inventory stimulation on price discount is evaluated to find a promotional plan to maximize profit, optimal pricing, and optimal self-space allocation model is evaluated in a two-stage joint decision model [13].

The influence of promotion activities on sales predictions using the BP neural network model on different sales frequency cycles (daily, weekly, and monthly) of different products has been utilized to reduce supply chain costs while increasing efficiency. The sales plan forecasted can be tuned with the model. The prediction effect of the model is shown with a resultant MAPE less than 1 and R-square greater than 0.94 which indicates the performance of the model inference is very good. The sales promotion quantitative index is computed over the price discount index, promotion duration index, and promotion influence index as they major impact on sales [14].

Impact of product sales by relative price discount during promotions and demand forecasting models during a promotion for perishable products [15]. The interaction effects of demand with the weight, volume, and shelf life of the product have improved the performance of the model. Modeling of the ratio of promotional sales and baseline sales before and during promotions is also discussed. A non-linear relationship between the data has been identified. Five independent variables are considered to have the following discount intervals: 0 to 10 percent, 10 to 20 percent, 30 to 40 percent, 40 to 50 percent, and 50 to 60 percent.

The classification of sales periods into normal, promotional, and post-promotional periods [16] as a feature for forecasting sales using multiple Machine Learning models on product quantities sold weekly is evaluated. The baseline projection using a 4-day moving average model to estimate baseline demand during a normal sales period to recognize a dip in demand after the promotion period is recognized as the post-promotion period. The use of the final forecast of the baseline model and the Light Gradient Boosting Model (LGBM) is evaluated using the forecast improvement formula. Boosting algorithms like Gradient Tree and AdaBoost have also been observed to have performed better on the factors impacting sales such as promotions [17], branding, and packaging. It is essential to accurately forecast sales and predict demand to ensure optimal inventory levels and address consumer demand.

A replenishment, demand prediction, and inventory control regression method have been proposed for products [18] that have shorter lifecycles and the demand function for these products is highly dependent on the day of the week. The solution is also applicable for predicting demand for newly introduced products. Demand forecasting strategies employ LSTM and Light GBM models [19] using the lag function to take past data for the features to be trained for each observation.

Multiple factors drive an incidence of purchase. A study discusses these factors that influence the willingness of a consumer to purchase a product and they are – price, availability, social proof, scarcity, novel features of the product, and social media activity [20]. They also provide empirical evidence to validate these factors to help markets and businesses improve their operational strategies to increase sales and consumer satisfaction. Employing the principal component analysis (PCA) technique and confirmatory factor analysis (CFA) technique has improved the performance of the demand prediction of the model.

To understand the effect of ever-changing consumer preferences and purchase behavior, employing a recommender system can learn from these effects on demand [21]. To evaluate the model performance, it is critical to understand the drivers and define KPIs that can measure a system's success. The McKinsey Global Institute identified major data providers that can elevate retail performance through significant improvements in demand prediction, inventory optimization, and operational efficiency [22].

Collaborative Planning, Forecasting, and Replenishment (CPFR) helps create a business model where each supply chain activity is efficiently planned to address demand at the lowest operational cost. Observed Unilever, Del Monte, and a few more major conglomerates who have implemented state-of-the-art sales forecasting and demand sensing models [23] that have significantly improved their supply chain processes at reduced operational costs, reduced instances of out-of-stock and maintained optimal inventory levels, and also improved their consumer satisfaction by reducing forecasting error greater than 30 percent and a 10 percent reduction in safety stock.

Employs Extreme Learning Machines (ELM) for demand sensing [24], a fast-learning feed-forward neural model that works on large real-time data with shorter turnaround time and has higher generalization performance over traditional gradient-based methods.

3. Implementation

After conducting a thorough literature survey, it is clear that engineering features related to promotions, pricing, and customers are crucial for understanding the influence of promotion and pricing on demand and sales. The learnings of each model that have been used to predict demand and sales in each reference paper were evaluated and it was observed that clustering data to identify the underlying patterns and employing non-linear models are observed to have performed better on this kind of data. Extreme Learning Machine (ELM) is a non-linear feed-forward neural network model (usually a single hidden layer network) that exhibits excellent generalization ability, fast and efficient learning speed with faster convergence, and lower runtime for the prediction which is very promising in predicting real-time refreshed demand and sales in a domain such as FMCG retail where predictions are required on-demand to take both strategic and tactical decision for the day-to-day operations of a business. We have also observed from the literature survey that the ELMs have superior performance and prediction over other machine learning and deep learning models in studies where the disruptions in data are frequent and factors impacting the target variable are many. Ensemble ELM has been chosen after research performed in the surveyed published works, to have alleviated the overfitting problem normally seen in the traditional ELM models.

The solution proposes to identify the customer segment the data point belongs to, and to predict the incidence of purchase, the demand for a product, and the sales in the store on a day. The customer segments are identified using an unsupervised learning method – KMeans clustering. The purchase incidence is predicted using a Neural Classification model. The proposed ELM model is the focus of this research study for the prediction of demand and sales. The approach of the bi-directional LSTM model offers an alternative solution to the proposed ELM model for the prediction of demand and sales, as a comparative study.

These ML techniques will provide the above predictions for a product by analyzing patterns from data sources such as – POS data, promotion occurrence, consumer profile, behavior, and purchase history data.

3.1. Overview of data sources

The demand and sales predictions have been performed on 5 distinct products purchased by 500 different customers over a span of 2 years of a leading retail chain spread across different states in the US. Data has close to 58000 incidences out of which 15000 are purchase incidence. Some of the important features from the dataset are tabulated below –

TABLE I. FEATURE DESCRIPTION

Feature Name	Feature Description
Day	Date Time of the sale day
ProductID	Identifier of the product
ProductPrice	Price of the corresponding product
ProductOnPromo	Indicates whether the product was on promotion or not
Quantity	Number of units sold for the product
CustomerID	Identifier of the customer
Sex	The gender of the customer
Marital Status	Indicates whether the customer is married or single.
Age	Age of the customer
Income	Income of the customer
Education	Possible values 0,1, 2, and 3 that correspond to unknown/illiterate, high school graduate, college graduate, and university graduate.
Occupation	Possible values 0,1 and 2 that correspond to unemployed, employed, and self-employed.
SettlementType	Possible values 0,1 and 2 correspond to town, small city, and metro city.
LastPurchasedProductID	Identifier of the product last purchased by the customer
LastPurchasedQuantity	Quantity of the corresponding last purchased product by the customer

The features that were engineered from the dataset are tabulated below –

TABLE II. ENGINEERED FEATURE DESCRIPTION

Feature Name	Feature Description
day_y, day_m, day_w	Day of the year, month, and week respectively was obtained by performing Repeated Basis Functions on the feature 'Day'.
PostPromotion	Indicates whether the product is in the post-promotion period. Identified by checking the dip in sales for 'n' days after a promotion has occurred.
LastPromotionInDays	Number of days when the last promotion was offered for the product.
ProductDiscount	Discount that was derived for the corresponding product by checking the dip in product price against the price before promotion.
NumOfAllPromotions	Total promotions offered on that day in the store.
5DayEMA (Sales)	5-day exponential moving average of sales
5DaySO (Sales)	5-day Stochastic Oscillator calculated on sales.
5DayROC (Sales)	5-day Rate of Change on sales
Sales Index	Derived by calculating the ratio of current sales to the previous day's sales.
Urban Household	Derived from the feature SettlementType
Earner	Derived from the feature Occupation
Literate	Derived from the feature Education

The preprocessing steps performed were - categorically encoding of data on the features ProductID, LastPurchasedProductID, Education, Occupation, and SettlementType and then scaling the data using the quantile transformer model.

3.2. ML Models

The study presented in this thesis describes a novel method of sensing demand, and sales and predicting the incidence of purchase. The overall approach to the solution illustrated in Figure 1, consists of the following steps – Preprocessing Layer, Clustering (PCA + K-Means, and Cluster classification Neural Network), Purchase Incidence Prediction Layer (Neural Classification Model) Demand and Sales Prediction Layer using - Bi-Directional LSTM (Base model) and Ensemble Extreme Learning Machine (Proposed solution).

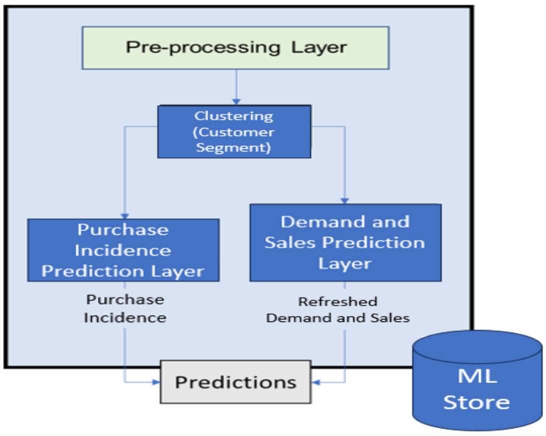

Figure 1 – End-to-End architecture

Clustering can help identify the homogenous groups of data points where each group belongs to a customer segment which will help understand the influence of promotions and customer patterns on demand and sales. The customer segments are identified in the data in the model by capturing unique underlying patterns of the customer profile and purchase history in the data. The unsupervised machine learning technique K-Means clustering model is used that identify patterns related to customer profile features such as – Sex, Married, Age, Income, Earned, Literate, and Urban Household; and purchase history features like – LastPurchasedProduct and LastPurchasedQuantity.

From the cumulative explained variance graph depicted in Figure 2, the PCA is performed on the above-mentioned features to obtain the number of principal components equal to 4 is observed to have cumulative explained variance greater than 0.8 and hence chosen as optimal number of principal components for further study.

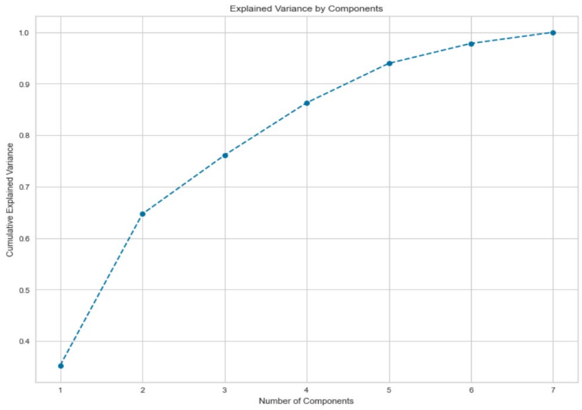

Figure 2 – Number of components to cumulative explained variance graph

The optimal number of clusters for the dataset was identified during hyperparametric tuning to be a K = 4 cluster using silhouette score and elbow plot. Figure 3, visualizes the number of clusters against the within-cluster sum of squares and the elbow is formed when the number of clusters = 4.

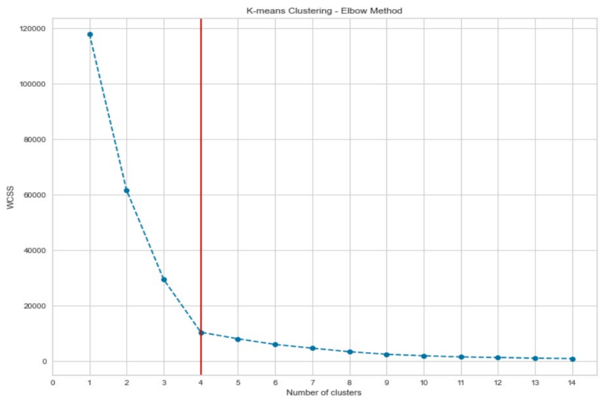

Figure 3 – Elbow Plot KMeans cluster model

The cluster labels obtained were used to analyze the data points in each segment and this helped in classifying them into the following customer segments –

- Cluster 0 is labeled as a "Low-Value" customer segment.
- Cluster 1 is labeled as a "Standard" customer segment.
- Cluster 2 is labeled as a "High-Value" customer segment.
- Cluster 3 is labeled as a "Need-Based" customer segment.

Cluster classification using a Multi-Class Classification Neural Model is performed for classifying customer segment labels after the formation of clusters. In the inference phase, data is pre-processed and fed to the model to obtain the customer segment for the customer sample in the data point. Precision, recall, and accuracy parameters are considered for comparing the models for applying the appropriate model from the ML engine.

The incidence of purchase at a store is predicted using the Binary Class Neural Classification model. The model is trained on the following input features – day component features; customer profile features - Sex, Married, Age, Income, Urban Household, Literate, Earner; customer purchase history features – LastPurchasedProduct, LastPurchasedQuantity; promotion and pricing features – ProductPrice, ProductOnPostPromotion, ProductOnPromo, ProductSales, and ProductDiscount. In the inference phase, this model will help in classifying during normal sales, promotion, or post-promotion period what would be the purchase incidences in the store.

The Long Short-Term Memory (LSTM) is a back-propagation network model that is capable of retaining memory of past inputs and has better generalization ability, unlike the traditional RNN. The LSTM models can contribute to demand sensing and sales forecasting as they capture complex temporal patterns, handle non-linear patterns, adapt to new information, easily incorporate external factors, handle missing data in time series data, and are very scalable. A Bi-directional LSTM for each customer segment is trained for each feature in the dataset containing observation from a period of 3-day lag.

The Ensemble Extreme Learning Machine (ELM) is a feed-forward network model where there is usually a single hidden layer that learns from a linear function and a non-linear function in every single neuron, the input layer captures the data features and performs no computation, and the output layer is linear with no kernel or bias. The weights and biases in the input layer are randomly assigned and are not adjusted throughout the learning process, improving the generalization ability of the linear output layer as it provides weakly correlated hidden layer features. The ELM model can handle large amounts of data in real-time. The first layer of the combinational neural network model contains multiple ELMs whose outputs are concatenated and fed to dense layers and the output layer employs linear function to predict the demand and sales. There are a total of 3 ELMs in the first layer that focuses on learning promotion pricing, customer profile and purchase history, and sales indicators respectively.

Here, the neural network is trained for each customer segment as a different model. In the inference phase, based on the identified customer segment for the sample data point, the workflow picks the right fit ELM model from the ML store. The model architecture of the ensemble ELM combinational neural model is depicted in Figure 4.

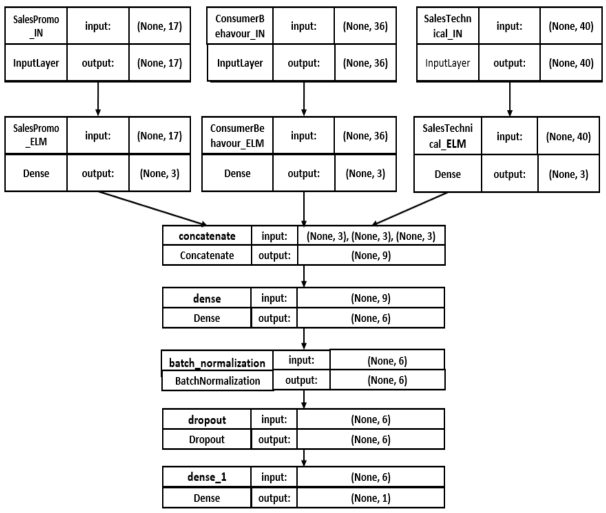

Figure 4 – Ensemble ELM combinational Neural Network Model Architecture

4. Model Results and Discussion

The classification models that were trained are evaluated based on – precision, recall, and accuracy. The bi-directional LSTM and ensemble ELM combinational neural models are evaluated based on – mean squared error (MSE) and mean absolute percentage error (MAPE).

TABLE III. RESULTS FOR NEURAL CLASSIFICATION MODELS

Metrics	Customer Segment Model (Cluster Label)	Purchase Incidence Model
Precision	0.9801	0.9985
Recall	0.9762	0.9961
Accuracy	0.9782	0.9982

The performance of the classification model is evaluated based on the following metrics – precision, recall, and accuracy.

The classification of the customer segment is a multi-class model and the test evaluation is that the model is very precise with a high recall of 0.9762 which signifies that the model has been able to classify with very little miss in true positive, and accuracy of 0.9782 which signifies the total number of correct classification is very high in classifying the customer segment labels on the sample of the dataset that was retained for model testing and validation.

The classification of incidence of purchase is a binary classification model. The model has precisely identified the incidence of purchase with a high recall value of 0.9961 which signifies that the model has been able to classify with very little miss in true positive and accuracy of 0.9982 which signifies the total number of correct classifications is very high on the sample of the dataset that was retained for model testing and validation.

TABLE IV. MODEL RESULTS FOR DEMAND PREDICTION

Customer Segment	Metrics	Bi-Directional LSTM (3-lag)	Ensemble ELM Combinational Neural Model
Low-Value Segment	MSE	3.3007	0.2617
	MAPE	0.6388	0.0678
Standard Segment	MSE	3.3297	0.3671
	MAPE	0.5199	0.1332
High-Value Segment	MSE	2.5964	0.4146
	MAPE	0.4697	0.2009
Need-Based Segment	MSE	2.1620	1.4512
	MAPE	0.4715	0.3443

TABLE V. MODEL RESULTS FOR SALES PREDICTION

Customer Segment	Metrics	Bi-Directional LSTM (3-lag)	Ensemble ELM Combinational Neural Model
	MSE	16.4812	5.2444
Low-Value Segment	MAPE	0.5599	0.1238
Standard Segment	MSE	20.8582	4.8992
	MAPE	0.5442	0.1695
High-Value Segment	MSE	8.2346	3.8033
	MAPE	0.4696	0.2784
Need-Based Segment	MSE	6.7407	4.8813
	MAPE	0.4260	0.4232

The performance of the LSTM and ELM models are good and they can learn from the data progressively without much overfitting and underfitting.

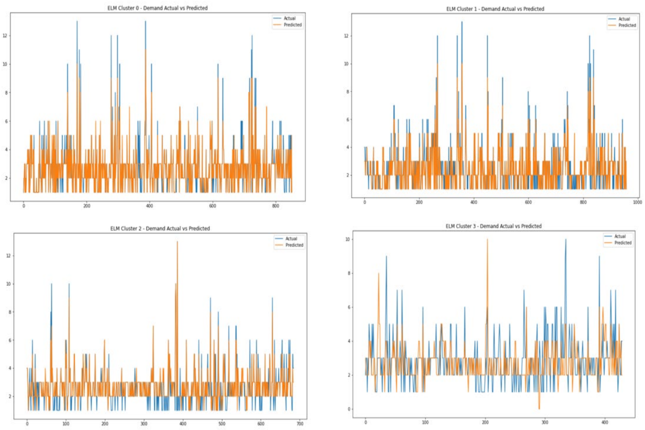

Figure 5 –Demand Actual vs Predicted Plot of Ensemble ELM combinational Neural Network Model per customer segment (Low-Value, Standard, High-Value, and Need-Based Segment *from left to right and top to bottom*)

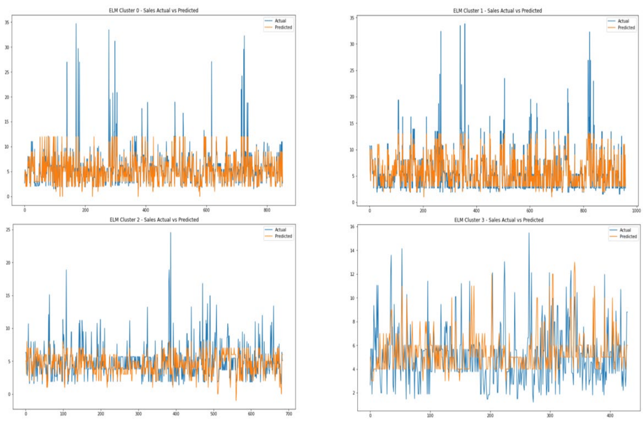

Figure 6 – Sales Actual vs Predicted Plot of Ensemble ELM combinational Neural Network Model per customer segment (Low-Value, Standard, High-Value, and Need-Based Segment *from left to right and top to bottom*)

The actual vs predicted plots in Figure 5 and Figure 6 for demand and sales respectively for the ensemble ELM combinational Neural Network model per each customer segment are observed to give better results as it progresses with learning and performs well even during demand or sales disruptions. However, the actual vs predicted plots of the LSTM model depict an average prediction line over the actual demand and sales value and have poorly captured the disruptions in demand and sales.

5. Conclusions

The purpose of this research is to bring about an efficient approach and its respective architecture, along with each component, the modules, and their functionalities are clearly defined. Each has a specific role to play in the process and has been designed to be optimal in its functionalities. Computationally, the pipeline described in this project meets the requirements quite well, i.e., the ability of ML models to predict the customer segment, the incidence of purchase at the store, the demand for a product on a given day, and the sales on a given day.

The classification models for customer segment and purchase incidence are very precise with high recall of 0.9762 and 0.9961 respectively and high model accuracy of 0.9782 and 0.9982 respectively.

It was identified that the demand signals had a non-linear relationship and hence we explored the models – a multi-stage ensemble ELM neural network and a Bi-Directional LSTM model. When we compare the learnings and performance of the two models, we understand that ELM Neural Network is giving better results as it progresses with learning and it can perform well even during demand disruptions. The refreshed demand identified is a better representation of demand sensed for a product.

The ELM and Bi-LSTM model for prediction of demand is evaluated based on the metrics - MSE and MAPE on the test dataset for the following customer segments – for the low-value segment the ELM model is evaluated with MSE of 0.2617 and MAPE of 0.0678 compared to Bi-LSTM model evaluated with MSE of 3.3007 and MAPE of 0.6388; for the standard segment the ELM model is evaluated with MSE of 0.3671 and MAPE of 0.1332 compared to Bi-LSTM model evaluated with MSE of 3.3297 and MAPE of 0.5199; for the high-value segment the ELM model is evaluated with MSE of 0.4146 and MAPE of 0.2009 compared to Bi-LSTM model evaluated with MSE of 2.5964 and MAPE of 0.4697; and finally for the need-based segment the ELM model is evaluated with MSE of 1.4512 and MAPE of 0.3443 compared to Bi-LSTM model evaluated with MSE of 2.1620 and MAPE of 0.4715.

The ELM and Bi-LSTM model for prediction of sales is evaluated based on the metrics - MSE and MAPE on the test dataset for the following customer segments – for low-value segment, the ELM model is evaluated with MSE of 5.244 and MAPE of 0.1238 compared to Bi-LSTM model evaluated with MSE of 16.4812 and MAPE of 0.5599; for the standard segment the ELM model is evaluated with MSE of 4.8892 and MAPE of 0.1695 compared to Bi-LSTM model evaluated with MSE of 20.8582 and MAPE of 0.5442; for high-value segment, the ELM model is evaluated with MSE of 3.8033 and MAPE of 0.2784 compared to Bi-LSTM model evaluated with MSE of 8.2346 and MAPE of 0.4696; and finally for the need-based segment the ELM model is evaluated with MSE of 4.8813 and MAPE of 0.4232 compared to Bi-LSTM model evaluated with MSE of 6.7407 and MAPE of 0.4260.

The exploration and preparation of the final dataset for modeling have been challenging – a lot of the features for demand sensing such as supply logistics and inventory stock were not captured in the data, however, the primary focus of this study was the influence of promotion and pricing on demand sensing and the features related to customer profile, customer behavior (also the purchase incidence at the store), customer purchase pattern, promotion factors, holiday events, and pricing has been well captured in the dataset. Many features were engineered concerning promotion, customer profile, and sales technical indicators that have helped in better modeling the data. The step to interpreting the clustering of the dataset into the 4 segments and identification and understanding of these clusters were pivotal.

Future directions towards this work include incorporating –

1. Clustering based on sales and promotion period to understand more underlying patterns in the data related to these features.
2. Features related to human feedback like social media effects, NPS, and consumer reviews to understand their impact on the promotion scope and ultimately on the demand.

References

[1] Joanna Henzel, Marek Sikora. "Gradient Boosting Application in Forecasting of Performance Indicators Values for Measuring the Efficiency of Promotions in FMCG Retail." IEEE Conference, 2020.

[2] Hongyan Dai, Haoyang Yu, Qing Xiao, Weihua Zhou. " A Clustering-based Sales Forecast Method for Big Promotion Days in O2O." IEEE International Conference, 2019.

[3] Yi Xiang Xingyu Lu Chunling Li Shasha Mo, Luqiao Zhang. Research on commodity sales forecasting based on combination model. 22 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA), 2022.

[4] Kao L.-J Lu C.-J. A clustering-based sales forecasting scheme by using an extreme learning machine and ensembling linkage methods with applications to a computer server. ScienceDirect, pages 231–238, 2016

[5] Langrock R. Pesch R. Senge R Ulrich M., Jahnke H. Classification-based model selection in retail demand forecasting. ScienceDirect, pages 209–223, 2022.

[6] Sahay K.B Singh A. Short-term demand forecasting by using ann algorithms. IEEE, pages 1–4, 2018.

[7] Uysal M. Akyokus S. Uysal M.O. Atak Bulbul B. Ekmis M.A Kilimci Z.H., Akyuz A.O. An improved demand forecasting model using deep learning approach and proposed decision integration strategy for the supply chain. Complexity, pages Article ID–9067367, 2019

[8] Janisha A Kamil R. Demand sensing using deep learning. IRJET, pages 1–5, 2018

[9] Houria Laaroussi, Fatima Guerouate, and Mohamed sbihi. Deep learning framework for forecasting tourism demand. ICTMOD, 12:666875, 2020.

[10] Haotian Zhu. A deep learning-based hybrid model for sales prediction of e-commerce with sentiment analysis. 21 2nd International Conference on Computing and Data Science (CDS), 2021.

[11] Fatima-Zohra Benhamida, Ouahiba Kaddouri, Tahar Ouhrouche, M. Benaichouche, D. Casado-Mansilla, D. López-de-Ipiña. " Stockbuy: A new demand forecasting tool for inventory control."5th International Conference on Smart and Sustainable Technologies (SpliTech), 2020.

[12] Zhang N Li Z. Short-term demand forecast of e-commerce platform based on convlstm network. Hindawi, pages Article ID 5227829–10, 2022

[13] Gao Jun-jun, Kang Ting. " A Joint Decision Model of Inventory Control and Promotion Optimization based on Demand Forecasting." IEEE International Conference, 2009.

[14] Qianwen Zhuang, Xiaodong Zhang, Pei Wang, Bin Deng, Hua Pan. " A Neural Network Model for China B2C E-commerce Sales Forecast Based on Promotional Factors and Historical Data." IEEE International Conference, 2019.

[15] K.H. van Donselaar, J. Peters, A. de Jong, R.A.C.M. Broekmeulen. " Analysis and forecasting of demand during promotions for perishable items." ScienceDirect, 2015.

[16] Harsha Chamara Hewage; H. Niles Perera. " Comparing Statistical and Machine Learning Methods for Sales Forecasting During the Post-promotional Period." IEEE International Conference 2021.

[17] Animikh Aich Chetana Hegde Akshay Krishna, Akhilesh V. Sales-forecasting of retail stores using machine learning techniques. 18 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), 2018.

[18] M. Kasahara H. Goto, M. Tanaka. Short-term replenishment method for short-lifecycle goods. 08 4th IEEE International Conference on Management of Innovation and Technology, 2008.

[19] Rony Mitra Dyutimoy Das Sushmita Narayana Manoj K. Tiwari Priyam Saha, Nitesh Gudheniya. Demand forecasting of a multinational retail company using deep learning frameworks. ScienceDirect, pages 395–399, 2022.

[20] Kakalejcík-L. Bucko, J. and M. Ferencová. Online shopping: Factors that affect consumer purchasing behavior. ˇ Cogent Business Management, pages 5(1), p.1535751, 2018.

[21] Karan Wanchoo. Retail demand forecasting: a comparison between CNN and gradient boosting method of univariate time series. IEEE, 2019.

[22] Chui-M. Brown B. Bughin J. Dobbs R. Roxburgh C. Manyika, J., and A. Hung Byers. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, pages pp.64–75, 2011.

[23] Samuel Rabi Dimitris Folinas. 1. Estimating benefits of demand sensing for consumer goods organizations. Springer, pages 245–261, 2012.

[24] Hongjiao Fu Mingming Wang Xun Liang Ming Gao, Wei Xu. A novel forecasting method for large-scale sales prediction using extreme learning machine. 14 Seventh International Joint Conference on Computational Sciences and Optimization, 2014