NATIONAL UNIVERSITY OF COMPUTER & EMERGING SCIENCES
DEPARTMENT OF COMPUTER SCIENCE
LAHORE, PAKISTAN

MACHINE LEARNING BASED DEMAND FORECASTING FOR SUPPLY CHAINS USING ENSEMBLE LEARNING

MUHAMMAD BILAL TOOR L19-1823
MS (DS)

DECEMBER 2022

SUPERVISOR: Dr. Tahir Ejaz
CO-SUPERVISOR: Dr. Mubashir Baig

This Thesis is submitted to the DEPARTMENT OF COMPUTER SCIENCE in partial fulfillment of the requirements for the degree of Master of Science in DATA SCIENCE.

Plagiarism Undertaking

I take full responsibility of the research work conducted during the MS Thesis titled MACHINE LEARNING BASED DEMAND FORECASTING FOR SUPPLY CHAINS USING ENSEMBLE LEARNING. I solemnly declare that the research work presented in the thesis is done solely by me with no significant help from any other person; however, small help wherever taken is duly acknowledged. I have also written the complete thesis by myself. Moreover, I have not presented this thesis (or substantially similar research work) or any part of the thesis previously to any other degree awarding institution within Pakistan or abroad.

I understand that the management of National University of Computer and Emerging Sciences has a zero tolerance policy towards plagiarism. Therefore, I as an author of the above-mentioned thesis, solemnly declare that no portion of my thesis has been plagiarized and any material used in the thesis from other sources is properly referenced. Moreover, the thesis does not contain any literal citing of more than 70 words (total) even by giving a reference unless I have the written permission of the publisher to do so. Furthermore, the work presented in the thesis is my own original work and I have positively cited the related work of the other researchers by clearly differentiating my work from their relevant work.

I further understand that if I am found guilty of any form of plagiarism in my thesis work even after my graduation, the University reserves the right to revoke my MS degree. Moreover, the University will also have the right to publish my name on its website that keeps a record of the students who plagiarized in their thesis work.

Bilal Toor

MUHAMMAD BILAL TOOR

Date: Dec 30th, 2022

Author's Declaration

I, MUHAMMAD BILAL TOOR, hereby state that my MS thesis titled MACHINE LEARNING BASED DEMAND FORECASTING FOR SUPPLY CHAINS USING ENSEMBLE LEARNING is my own work and it has not been previously submitted by me for taking partial or full credit for the award of any degree at this University or anywhere else in the world. If my statement is found to be incorrect, at any time even after my graduation, the University has the right to revoke my MS degree.

Date: Dec 30th, 2022

Certificate of Approval

It is certified that the research work presented in this thesis, entitled titled *MACHINE LEARNING BASED DEMAND FORECASTING FOR SUPPLY CHAINS USING ENSEMBLE LEARNING was conducted by MUHAMMAD BILAL TOOR under the supervision of Dr. Tahir Ejaz*

No part of this thesis has been submitted anywhere else for any other degree.

This thesis is submitted to the DEPARTMENT OF COMPUTER SCIENCE in partial fulfilment of the requirements for the degree of Master of Science in DATA SCIENCE

At the

National University of Computer and Emerging Sciences

Lahore, PAKISTAN

December 2022

Candidate Name: MUHAMMAD BILAL TOOR Signature: Bilal Toor ____________________

Examination Committee:

a) Name: __________________________________ Signature: ______________________

b) Name: __________________________________ Signature: ______________________

c) Name: __________________________________ Signature: ______________________

Dr. Irfan Younas
Graduate Program Coordinator, National University of Computer and Emerging Sciences, Lahore

Dr. Kashif Zafar
HoD of Computer Science, National University of Computer and Emerging Sciences, Lahore

Abstract

In this modern day and age, one of the most valuable sector of any industry or business is generating forecast. Creating accurate forecast not only aids the demand planner of the organization but also helps in cutting down the costs and optimize inventory. A faulty prediction might have a negative influence on the company's supply chain, resulting in high losses. Whereas a good forecast can do wonders for the company both in terms of growth and profit. All this depends upon the forecasting system that the company is using. For a successful business the demand planners requires consistently accurate results over a period of time. For this, the stability of the algorithms used for forecasting is very important. A good forecasting model is the one which not only gives good accuracy but should also be adaptable to changing context. In this study, we propose a customized method which will generate output by combining the product category level forecasts for Retail industry and E-commerce. For each category of products, we shall separately train a variety of machine learning algorithms for forecasting. Then we will create a customized model by selecting only the highest performing algorithms based on their score for each feature set. To get the final forecast of each category, the output of each algorithm is fed to a multilayer perception as input. This neural network will give the final forecast for the particular category as output. Forecasts will be generated by the ensemble of two kinds of algorithms; one are Time-Series algorithms and the other are Regression-based algorithms. Individual algorithm's performance is evaluated with MAE, RMSE and MAPE scores.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In today’s world of competitive environment, the most important aspect for any organization is cost reduction and inventory management especially for consumer packaged goods (CPG) industry. To gain a competitive advantage over others, industries focus on maintaining optimum inventory levels and try to cut their costs as much as possible. For Customers satisfaction maintaining the right amount of inventory is the key. To ensure this, Organizations use an approach known as demand forecasting. By using this technique, companies are able to produce accurate and efficient forecasts by anticipating demand and as a result distribute resources in such a way that will minimize stagnant inventory [1].

Demand Forecasting is a field of predictive analytics which predicts and tries to understand customer demand in advance to optimize supply decisions by using historical sales data. Forecasting demand and sales is one of the most crucial roles of manufacturers, distributors, and trading companies. By maintaining a balance between demand and supply, they eliminate excess and shortage of inventory and boost profitability. Surplus production results in additional stock holding, which locks up excess inventory when the manufacturer seeks to meet the overestimated demand. Underestimated demand, on the other hand, results in unmet orders, lost revenues, and missed opportunities due to an inefficient supply chain. As a result, accurate demand forecasting is a major challenge for supply chain [12].

The process of tracking goods, services and resources as they move from supplier all the way to the consumer is known as Supply chain management (SCM). Effective delivery of services and products, making and talking of orders and feedback of information are all part of part of this process. While there are numerous facets to SCM, we'd like to concentrate on the idea of demand forecasting. Each participant in the supply chain must deduce regarding the purchase of the products, when to buy the products and how much amount of purchase to be made depending upon the demand information from their respective consumers. Figure 1 [11] depicts the flow of products and information between the supplier, manufacturer, distribution center, supermarket, and consumer.

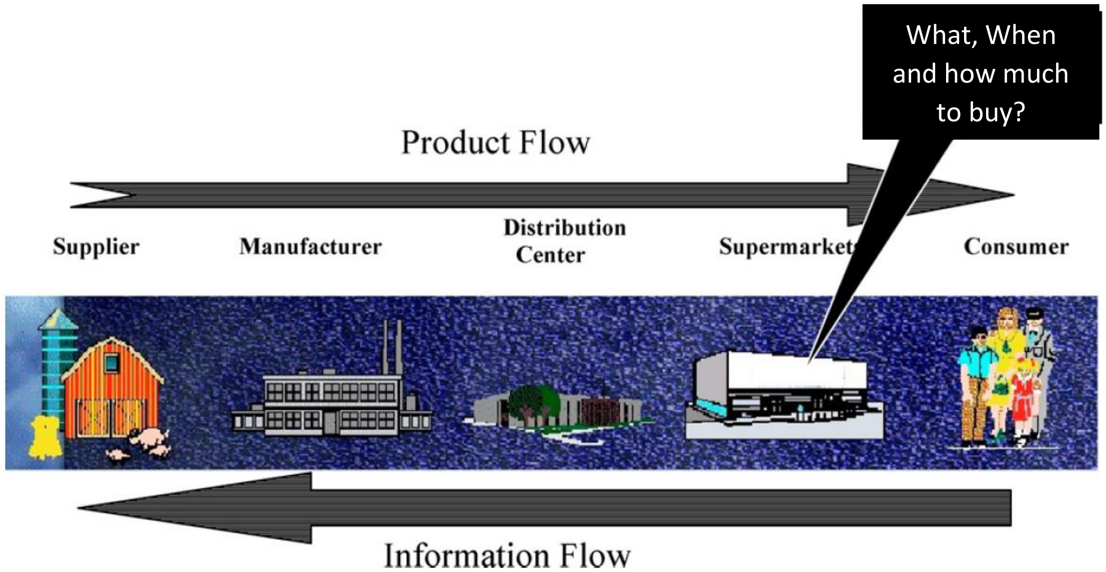

Fig.1. Components (Nodes) within the Supply chain [11].

1.1 Background

The supply chain compromises of various nodes. These components (i.e. Supplier, Manufacturer, Distribution centre, Retailer and Consumer) need to make good synchronization within the supply chain. This coordination is necessary to ensure success and take competitive edge in the market. One of the more common issues that supply chain faces due to lack of coordination and information sharing is the bullwhip effect [12][14]. This phenomena shown in figure 2 explains the significance of change in demand at customer level and how it affects the demand at the wholesaler to Supplier level.
Important role is therefore played by demand forecasting in minimizing the extent of the bullwhip effect by making accurate predictions for demand. Great attention must be given to demand forecasting of the finished products as it is an essential stage in the supply planning process. If done incorrectly either it can lead to surplus in stock or shortage at the customer's end. Both ways it has an adverse impact on the business as well as customer satisfaction. Demand forecasting not only helps in inventory management but also optimizes warehousing, procurement, shipping and pricing. It can even be used for anomaly detection by comparing the actual sales with the one forecasted.

Demand forecasting is performed either by traditional statistical forecasting approaches or by modern machine learning based approaches. In the past, traditional forecasting methods were useful in making demand predictions according to the market needs. These approaches are still handy and works well with stable markets. Traditional methods relies heavily on the historical performance data. With the emergence of big data and technological advancements, the traditional statistical approach is vulnerable to the ever changing complex market environment. Apart from just historical data, companies now take into account a variety of demand drivers such as demographic, historical, behavioural and microeconomic factors to make proactive data-driven decision.

Fig.2. Bullwhip Effect in supply chain [14]

Modern industry is known for its randomness and demand volatility. Increased customer choice, product personalization, fast technical advancements, global rivalry, and upstream supply variations all contribute to volatility. Traditional methods are not agile enough to cater such dynamic behaviour. As a result forecasting accuracies are compromised with traditional statistical approaches. Modernization of demand forecasting is inevitable. Increased Market competition and unpredictability in demand will boost attempts to employ machine learning to enhance forecasting accuracy.

1.2 Scope

The scope of demand forecasting can vary depending upon the scale at which the forecast is made. Area of functioning of the organization is a key aspect in determining the scope of the demand forecast. It could be done at a local, national or global level. It could be limited to a small setup in a local area or large enough to cater products and services at global level. The present area of operation of an organization and also the future projection will determine the extent of the demand forecasting task.
Much would rely on the time and effort spent in comparison to the utility of the information obtained through demand analysis. The required trade-off, depicted in Figure 3, must be made between the expense of predicting and the advantages derived from such forecasting. Sophisticated statistical models that lie in the optimal region include Simple Moving Average (SMA), Exponential Smoothing (SES), Autoregressive Integration Moving Average (ARIMA) and Neural Network (NN). These models gives fair accuracy at a nominal cost. On the other hand, regression and causal model gives higher forecast accuracy but are relatively expensive.

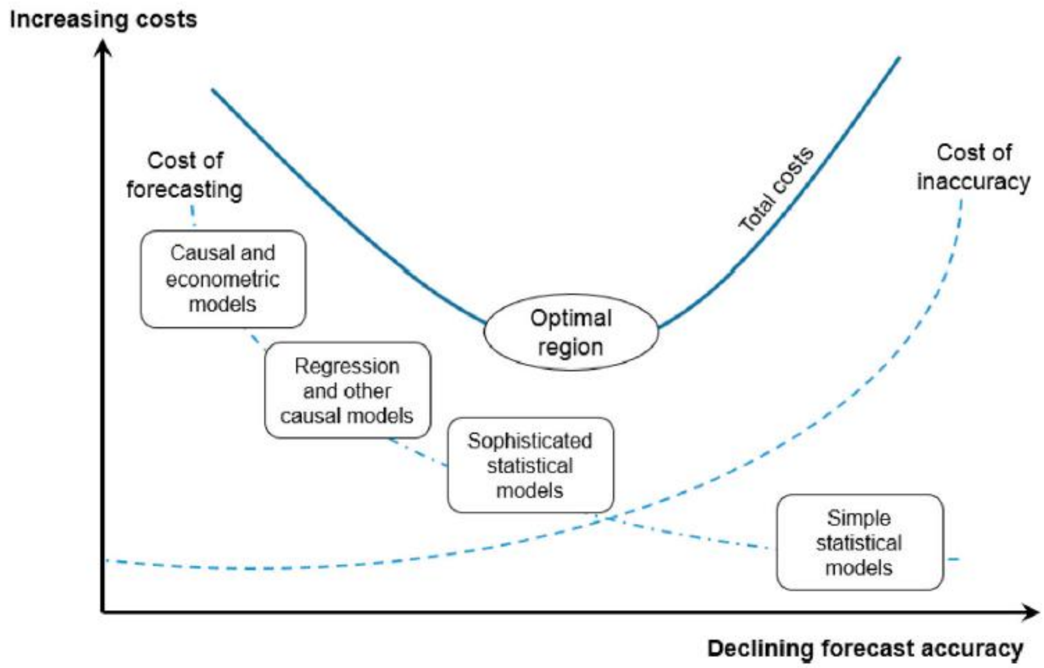

Fig.3. Trade-off between Accuracy and Cost [31]

1.3 Goals and Objectives

The fundamental purpose of this research is to thoroughly understand the performance of various machine learning techniques for demand forecasting. The focus of this research is on two alternative forecasting methods. Time Series approaches are the first, while regression-based models are the second. Based on the demands of the business, a customized demand forecasting model is created. The model will be prepared for each product category and it will be a combination of several time series and regression models. Depending on the amount of previous data, this model can produce short and long term projections. Each model can then be evaluated based on its accuracy and performance. Top models with highest accuracies are selected and their predictions are summed up to get the overall forecast for the business. The goal is to identify key demand drivers and attributes which contributes in making accurate forecasts for future demand and sales.

These wide range of external drivers and relationships that impact demand can only be incorporated in the system with the help of machine learning. Machine learning approaches allows us to make more accurate forecasts. It helps us to analyze more data and speeds up data preprocessing. It automates forecasting and uncovers data trends. With the application of machine learning, a more resilient system is built that is more adaptable to change. The goal of machine learning-based demand forecasting is to provide enterprises and businesses with a wide variety of benefits. Table 1 below outlines the benefits of how machine learning can capture demand drivers. It also assists a company in reducing risks associated with business activities and making critical business decisions. Aside from that, demand forecasting aims to improve the processes listed below [32].

System Benefits

• **Supplier relationship management**. You can compute how many items to purchase once you have a handful of client demand predictions, making it easier to assess if you really need additional supply chains or to minimize the number of providers.
• Customer relationship management. Customers who want to buy something expect it to be accessible right away. Demand forecasting predicts which consumer products will need to be purchased from a given store in the near future. This increases client happiness and loyalty to your brand.
• **Logistics and order Attainment**. Supply chain optimization is a component of demand forecasting. This implies that there are more chances of product been in stock at the time of order, and unsold products will not take up valuable shop space.
• **Marketing campaigns.** Marketing and advertisement strategy are greatly influenced by demand forecasting and it has the potential to affect sales. Marketing data may be used into advanced machine learning forecasting models.
• **Manufacturing flow management**. Time series-based demand forecasting, which is part of the ERP, predicts manufacturing requirements depending on the number of items sold in the end.

Value-added benefit

• **Reveal seasonal trends.** You can notice seasonal variations by looking at previous sales data.
• **Rationalize your cash flow.** You can trace sales expenses and income by looking at the previous balance sheet. It will help you deciding how much goods to buy with the available cash on hand.
• **Supply chain planning.** Demand forecasting will assist you in planning ahead of time so that you have goods on hand when customer demand surges.
• **Recognize how external forces will affect your sales.** Beside historical data these outside factors also affect your sales projections such as current market trend, condition of economy and data about competitors.
• **Prepare for the future.** Demand forecasting can help you prepare your supply chain and business for anticipated setbacks.

Table 1: Demand Impact Machine Learning can capture

Recurring demand pattern	Holidays, weekdays and seasons can cause recurring variation in demand.
Internal business decisions	Price changes, changes in how products are displayed and promotion are served by Machine learning.
External factor	The influence of external elements such as weather, local events, and customer footfall that are outside the retailer's control.
Unknown factor	Demand changes for which the influencing reason has not been identified.

1.4 Thesis Structure

This is the chapter wise structural breakdown of the report.

Abstract summarizes the overall process and benefit of Machine learning based demand forecasting for supply chains, brief describe the details of the proposed methodology, background of the problem and research directions.

Introduction gives a thorough understanding of the problem at hand. Background is discussed to understand the importance of the need of demand forecasting. The scope of the research is defined and finally the objectives and goals of the study are made clear.

Literature Review summarizes the scholarly articles and related work being carried out in traditional statistical and modern machine learning forecasting methods.

Proposed method outline the details of proposed methodology that is machine learning based demand forecasting for supply chain using ensemble learning.

Methodology presents two kinds of algorithms, one are regression based and the other are time series based models. Initial Experimentation is presented along with data exploration, data pre-processing, and feature engineering, model building and evaluation performance. In the end outlined the implementation steps in a form of flowchart.

Results, Discussion and Conclusion Examines the work done on the Thesis, its possibilities, and the implementation details in general. In addition, it displays the overall plan for Thesis, with each step outlined separately and in the sequence in which it will be completed.

1.5 Summary

It gives a thorough understanding of the problem at hand. Background is discussed to understand the importance of the need of demand forecasting. The scope of the research is defined and finally the objectives and goals of the study are made clear.

Chapter 2

Literature Review

This section examines the related research projects that several researchers are working on. Different forecasting models are used to carry out the predictions. With historical time series data, forecasting models are created using a variety of statistical and machine learning approaches. The forecasting model is used in a variety of applications, including commercial sales and economic forecasting, meteorological weather forecasting, electrical energy and power consumption forecasting, and so on. Sales forecasting is vital in business because it helps managers plan and make data-driven decisions by understanding feature demand. The forecasting model is often built using past sales data. Demand forecasting problem is either tackled with time series models or by regression based approaches.

This paper by Niu, Yiyang [2] recommends the XGboost sale prediction model, which combines the XGboost algorithm with thorough feature engineering processing, to solve Walmart's sales problem. The approach used in this study can efficiently mine attributes from several dimensions to create accurate predictions. The experimental findings show that the XGboost-based sale forecast model in this research outperforms existing machine learning models like the Logistic Regression and Ridge algorithms. Experiments show that the XGboost algorithm, as proposed in this study, is successful for sales forecasting tasks.

In this research by Wenxiang, et al. [3] Using previous data from the e-commerce firm, a model is built to anticipate sales. Depending on the qualities of the data, three types of prediction models are available: Long Short-Term Memory (LSTM), Artificial Neural Network (ANN) and Incentive Auto Regressive Integrated Moving Average (I-ARIMA). These three approaches can be used to address issues with varied degrees of precision and data formats. This research evaluates the advantages and disadvantages of three types of models for estimating item sales using various data sources. The amount of data volume is proportional to the model’s accuracy, and the model may be selected based on the data type and accuracy requirements. To attain maximum accuracy, a variety of forecasting methodologies, including Ridge Regression, Lasso Regression, Multiple Regression, and Polynomial Regression, as well as numerous boosting algorithms, including GradientTree Boosting and AdaBoost[4], are investigated. Because the combination of numerous forecasts can enhance forecast accuracy, a broad range of forecasting approaches are addressed.

Ding, Jingyi, et al. [5] proposed effective feature engineering strategy for providing sufficient information to our model in order for it to generate accurate predictions. CatBoosting is the algorithm we're utilizing, which was first announced in 2017. It's been proven to work in page ranking, click prediction, and recommendation algorithms. However, we haven't seen this strategy used extensively in sales forecasting. Its key benefit over previous Boosting algorithms is that it includes categorical features, eliminating the requirement for one-hot encoding during training and inference. This minimizes the number of input parameters and avoids the model from becoming unnecessarily complicated.

In this work [6], we look at how machine-learning models may be used to predict sales. The major purpose of this study is to investigate different techniques to applying machine learning for time series sales forecasting. This competition's aim was to forecast inventory demand. The impact of machine-learning generalisation has been studied. A stacking technique for building regression ensembles from single models was examined. In this method, predictions on the validation set are used as input variable for the next level models.

This study [7] investigates the viability and comparative comparability of Deep Learning algorithms for forecasting demand with application to a public dataset. By comparing Deep Learning performance to RMSE performance metrics, we find that it outperforms other model techniques namely Support Vector Machine and random forest. Demand forecasting is critical for business decision-making. The proposed Deep Learning model was used to anticipate the load needs of a supply chain. In terms of performance, it was also compared to Support Vector Machine, Gradient Boosted Trees, and random forest.

Bandara, k. [8] implement and predicts sales and it correlation in e-commerce sector. In order to account for various external sales demand factors, he examines and evaluates two unique LSTM learning algorithms with varying error after back-propagation, as well as a combination of dynamic and static characteristics. This technique puts the framework to the test using sales data of a retail Walmart.com, which is then compared to our proposed framework utilising cutting-edge forecasting algorithms. According to the findings, our strategy outperformed state of the art uni-variate forecasting systems. The findings also suggest that E-commerce product hierarchies include a range of correlation between the products and similar such patterns in data are used to better the forecasting accuracy of the sales.

Javad Feizabadi [9] suggested an approach in which he uses ARIMAX and Neural Network to create a hybrid forecast for the demand. In this technique he considered time series and varies other components to develop the forecast. The method is performed on a steel manufacturing firm data and some functional product. The supply chain performance is substantially improved by the use of machine learning forecasting techniques compared to the statistical traditional approaches. The use of the ML-based forecasting strategy (ARIMAX and NN) might enhance both operational and financial indicators, according to these research findings. Our findings show that inventory performance has improved statistically significantly. Inventory reduction means cheaper cost for data storage as well as providing quality services with lower costs of transportation.

Another study done by Cheriyan, Sunitha, et al. [10] to assess and examine the usage of data mining techniques for sales forecasting in order to develop comprehensive and trustworthy models. In this work, we used the training dataset to develop three machine learning methods, namely

Generalized Linear Model, Decision Tree and Gradient Boosted Trees. The models are then assessed for performance. The optimal algorithm for prediction is chosen based on performance accuracy. According to the results, the Gradient Boost Algorithm has a 98 percent overall accuracy, followed by Decision Tree Algorithms with a roughly 71 percent overall accuracy, and Generalized Linear Model with a 64 percent overall accuracy.

Another study [11] uses a mixture of two model to produce an intelligent system. The algorithms used are ARIMAX and neural network to generate effective demand forecasts. There is a significant increase in accuracy after the implementation of this strategy. We have targeted Chilean supermarket for this project and proposed a replenishing system for it. After deployment of this model in the supermarket, it is observed to have less sales failure and inventory levels are also balanced. If we compare both models, neural network outperforms ARIMAX in terms of accuracy. Out of all, our proposed hybrid model gives the best accuracy. Retailers can improves their service quality as the inventory levels are optimal and less failures of sales. This method is undoubtedly successful and have edge over other proposed strategies in the retail industry.

The goal of this study [12] is to apply hybrid approaches to demand forecasting problems where certain conditions are met: customers' orders are requested with a reasonable amount of time between them and on a regular basis. Experimenting with real demand data with the attributes in mind allows for the correct technique to be selected. These hybrid methods will be compared to a set of traditional procedures in order to demonstrate that hybrid methods outperform traditional methods. The results shown, the hybrid model obtained better results than every simple regression model.

The authors of this study, Das Adhikari, Nimai Chand, and colleagues [13], seek to highlight a novel ensemble strategy based on the averaging method that not only prioritises the algorithm that consistently maintains a high level of accuracy but also reduces the divergence from real sales. In this investigation, we can see that two sorts of forecasts are being made:

• Using a cutting-edge Time-Series Algorithm
•Algorithms Based on Regression

The new method combines the two forecasts, gives each one a weighted average, and produces the outcome. It is intended to show how the Time-Series Model and the Regression-Based Model compare in terms of accuracy. The overall outcomes will be determined by a comparison of three-month consistent accuracy measures.

In this study [14] the distorted demand in supply chain is predicted using three machine learning algorithms namely support vector machine, neural network and recurrent neural network. By making this prediction we can better understand the effects of bullwhip. These Machine learning models are compared with traditional approaches such as naïve forecasting, moving averages and linear regression. Modern ML techniques yielded better results overall. Yet they do not show a substantial different compare to traditional methods when run on simulated dataset. On real world data the accuracies are better of machine learning models such recurrent neural network (RNN) and support vector machines (SVM) on foundries dataset. Overall it could be concluded that machine learning model and multiple linear regression produce good accuracies for the distorted demand signal than previous methods.

Kilimci, Zeynep Hilal [15] proposed an intelligent method for demand forecasting. Here he uses the historical sales data to run various machine learning algorithms such as time series, deep learning and support vector regression. This is the first study to have employed the mixture of these techniques to generate the forecasts. Deep learning models and support vector algorithms are used together for the first time to best of our knowledge. Time series models are also considered while generating the forecasts. Finally the results of these algorithms are combined using a intelligent integration strategy. A boosting ensemble technique is used to produce the demand forecast. A unique integration strategy is yielding a higher accuracy by the introduction of deep learning approach.

For the mentioned dataset, the product demand is predicted using Seasonal ARIMA and long short term memory (LSTM). The two approaches are judged on the following performances such as scalability, execution time and convenience. An e-commerce business is evaluated for demand forecasting for its sales. Purpose is to cut down their costs and to minimize stagnant inventory. The LSTM and SARIMA model are compared with each other in this study. These models are used to predict demand for the specified dataset. Based on the comparison of various aspects it is seen that SARIMA is most powerful models and trained better. SARIMA is also better at long term forecasts whereas LSTM provided good result for short term results. It is also observed that SARIMA gives on overall better and consistence.

Catal, Cagatay, et al. [18] used time series analysis techniques as well as regression approaches in machine learning to anticipate sales amounts based on many features in this investigation. Regression approaches included the following: Neural Network Regression, Decision Forest Regression, Bayesian Regression, Boosted Decision Tree Regression and Linear Regression are some of the most common types of regression. Drift Method, Non-Seasonal ARIMA, Seasonal ETS, Non-Seasonal ETS, Seasonal ARIMA, Naive Method and Average Method were applied in addition to these regression approaches. Boosted Decision Tree Regression is the best technique this time

Table 2: Literature Review Summary

Studies	Motivation for the Research	Result of the Research
Comparison of Statistical and Machine Learning Methods for Daily SKU Demand Forecasting (Spiliotis, Evangelos, et al.) [20] Statistical and Machine Learning-Based E-Commerce Sales Forecasting (Dong, Wenxiang, et al.) [3]	To study the traditional methods of demand forecasting and make comparison between traditional statistical models and modern machine learning methods	ARIMA models have a high level of accuracy. Because they are reliant on heuristic parameter selection, they may pose issues in the initial model selection. When analysing a large number of time series observations, it might be time intensive. Forecast performance is commonly measured using RMSE and MAPE.
A Sales Forecasting Model for New-Released and Nonlinear Sales Trend Products (Tanaka, Kenji.) [25] A comparative study of linear and nonlinear models for aggregate retail sales forecasting (Ching-Wu Chu, & Guoqiang Peter Zhang.) [19]	To see if newer nonlinear machine learning models perform better than traditional methods	Neural networks with deseasonalized data perform the best overall Experiments show that the suggested technique is more reliable in terms of accuracy and efficiency than current established methods.
Ensemble Methodology for Demand Forecasting (Das Adhikari, Nimai Chand, et al.) [13] Ensemble Deep Learning for Regression and Time Series Forecasting (Qiu, Xueheng, et al) [20]	Checking whether ensemble methods improves accuracy and performance for forecasting tasks.	Ensemble methods have outperformed various algorithms for both time series and regression datasets based on the RMSE and MAPE score. Methods proposed using ensemble techniques has the potential to deal with larger and more complex datasets.

Data Analytics in the Supply Chain Management: Review of Machine Learning Applications in Demand Forecasting (Aamer, Ammar, et al.) [23] Application of Machine Learning in Supply Chain Management: A Comprehensive Overview of the Main Areas. (Tirkolaee, Erfan Babaee, et al.) [1]	The study's goal is to assist other researchers in closing the research gap by extending machine learning's application to other relevant industries.	Machine learning methods, in comparison to traditional forecasting models, may deliver improved accuracy and lower computing costs for demand forecasting. Introduction of deep learning models for demand forecasting could be beneficial for supply chain management.
Deep Learning with Long Short-Term Memory for Time Series Prediction (Hua, Yuxiu & Zhifeng, Zhao & Li) [21] Highly Efficient Short Term Load Forecasting Scheme Using Long Short Term Memory Network (Rafi, Shafiul Hasan) [22]	To understand different Recurrent Neural Network models and Compare their performance with state of the art Machine learning models.	Long Short-Term Memory (LSTM) is a strong tool for sequence prediction. LSTM models can outperform significant machine learning models

Table 2 present a brief summary of the literature review by discussing the motivation of each study and observing the result and outcomes of the research. Following section will discuss the commonly used approaches to demand forecasting problem.

2.1 Time Series Approach

The historical data for time series forecasting is a set of historically organised raw data items. The natural arrangement of the data points distinguishes it from Causal forecasting. A time series prediction makes the assumption that components such as trends, seasonality, cycles, and so on will repeat themselves. Time series forecasts are frequently shown using line charts. Most business sectors, such as finance, sales, and operations, employ time series forecasting. Businesses can use time series to discover cyclical patterns, trends, growth rates, and any irregularities or variations in a data set. Figure 4 shows the overview of the entire prediction process of time series model.

The following are some of the most often used time series forecasting techniques:

Moving Average (MA): A moving average, often known as a simple moving average, is the most basic approach to forecast by taking the average of the previous 'n' periods. The average value is used to anticipate the value for the following period.

Exponential Smoothing (EA) is a widely used approach in which various weights are assigned to the observed data point, depending on how old the data is, to generate a smoothed time series. The Box Jenkins technique is a specific example of Exponential Smoothing, in which the model is used to identify the best fit of a time-series model to previous values of a time series. EA is well suited to datasets with no discernible trend and a wide range of values. Holt's approach and winter’s method are two EA improvements that may be used on datasets with varied trends.

ARIMA (Autoregressive integrated moving average) is a statistical approach for forecasting the future that uses time series data. The autoregressive, integrated, and moving parts of the dataset are the three components of an ARIMA model. ARIMA effectively auto-correlates its own past deviations from the mean. When creating predictions, it considers trends, seasonality, cycles, mistakes, and non-stationary features of a data collection.

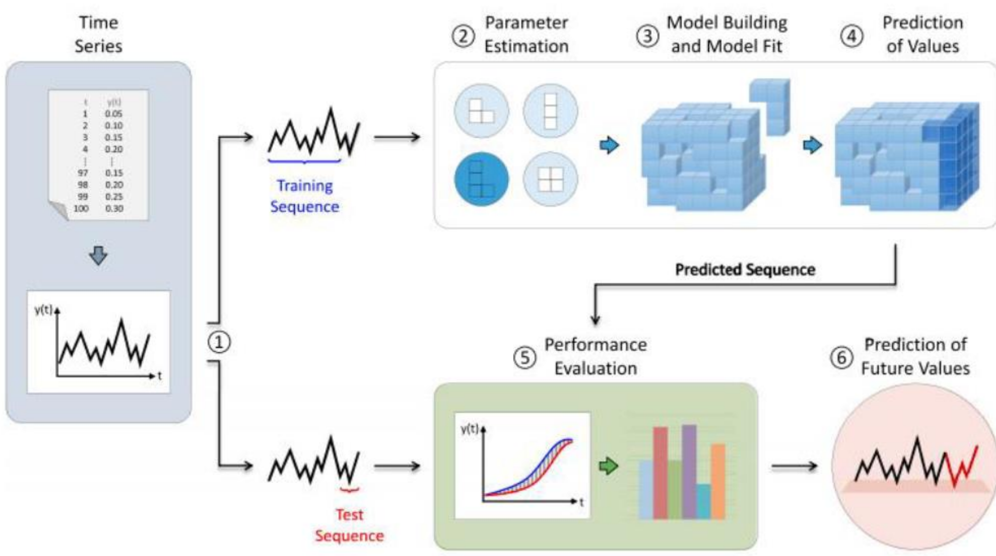

Fig.4. Time Series prediction process [33]

2.2 Causal (Regression) techniques

Causal forecasting is a strategy that presupposes a cause-and-effect relationship exists between the forecasted variable and one or more other independent variables. Causal approaches normally take into account all potential influences on the dependent variable. As a result, the data needed for forecasting should be converted into a supervised learning form. The flowchart depiction of a regularly used demand forecasting procedure is shown in Figure [5].

One of the most popular strategies used to understand a variable connection in a dataset is regression. The least square methodology is used to estimate a function between the dependent and independent variables, which characterises the interaction between them. A basic example would be estimating a business's margin (dependent variable) based on parameters such as cost of goods sold, inventory holding, and so on (independent variables).

Following are the three regressors that are used in this study.

Gradient boosting, an ensemble learning technique that combines the predictions of many weak models to produce a stronger, more accurate model, is implemented by XGBoost.

Linear regression is a statistical technique that fits a linear equation to observed data to represent the connection between a dependent variable and one or more independent variables. The purpose of linear regression is to identify the best fit line.

A **random forest regressor** is a form of ensemble learning approach that may be used for regression problems that require predicting a continuous numerical value. At training time, a number of decision trees are generated using a sample of the training data to build a random forest regressor. The resultant decision trees are integrated to form the final model, which can subsequently be used to generate predictions on new data.

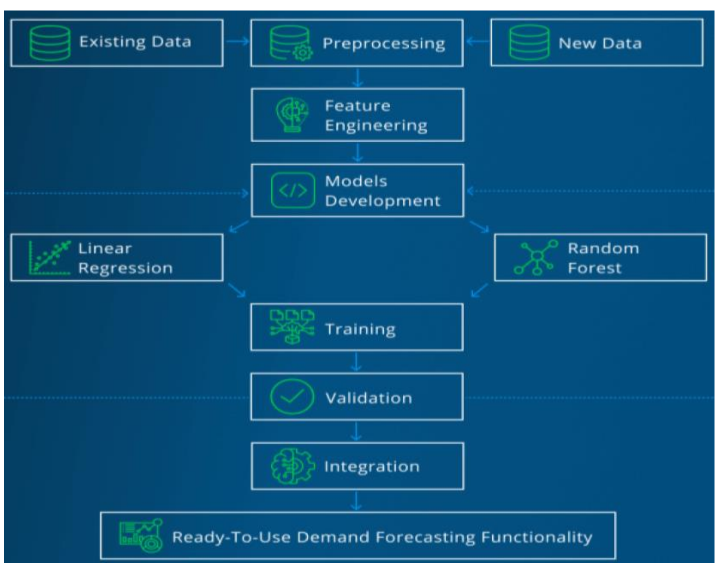

Fig.5. Flowchart representation of demand forecasting process [32]

Chapter 3

Details of the proposed method

An ensemble learning (EL) approach is a machine learning process that intelligently combines various machine learning algorithms to improve prediction performance [13]. Under different circumstances different algorithms has their own strengths and weaknesses. Hence combining the outputs of various models provides with more powerful and effective results for decision making processes.

Heterogeneous and homogeneous are the two approaches used in the literature for ensemble learning. Heterogeneous ensemble learning is when different types of learners are used as base algorithms. This work uses heterogeneous ensembles. Using different base learners a set of diverse models are produced in the ensemble generation phase. For this study five different regression based and time series algorithms are employed on each category.

Our proposed method in Fig [6] will assist businesses to generate product category level forecast and use that to predict the demand of a particular product in a store or supermarket. In this study, we propose a customized method which will generate output by combining the product category level forecasts for Retail industry and E-commerce. For each category of products, we shall separately train a variety of machine learning algorithms for forecasting. To get the final category forecast, the output of each algorithm are combined which is discussed later in Chapter 5.

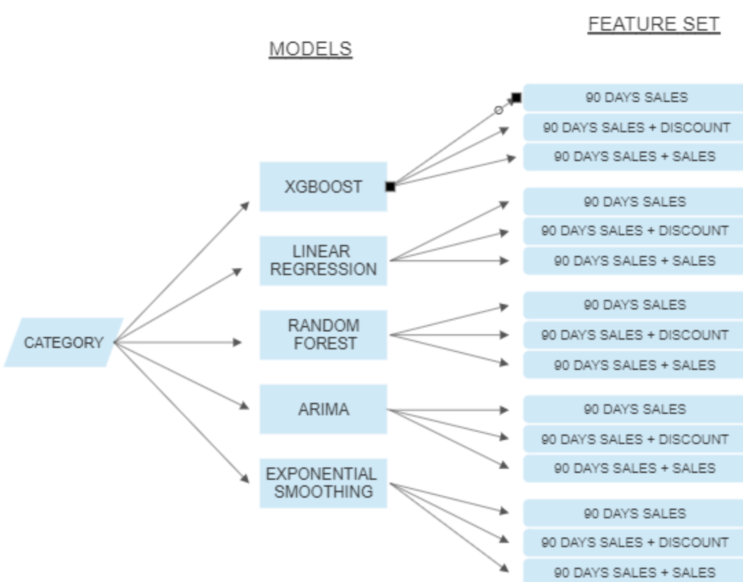

Fig.6. Proposed Methodology

The dataset used contains E-commerce Sales from 2015 to 2018. Each row represent a particular record of a single transaction. There are a total of 53 product categories. We are going to forecast the number of items sold for 5 such category. The selected categories are following:

- Cleats
- Women's Apparel
- Indoor/Outdoor Games
- Water Sports
- Hiking & Camping

For each Category we will filter out the transactions related to that category only. After data pre-processing and feature engineering, the data is prepared to be fed to the Machine Learning model. Each record of the prepared data shows daily Items sold in the previous 90 days as features and a target variable which represents aggregated sum of the monthly Items Sold following the 90 days Sales. This is depicted in Fig [7].

	Day0	Day1	Day2	Day3	Day4	Day5	Day6	Day7	Day8	Day9	...	Day80	Day81	Day82	Day83	Day84	Day85	Day86	Day87	Day88	Order Item Quantity
0	77	73	88	90	62	71	83	79	81	63	...	79	73	82	45	55	66	61	131	76	2187
1	74	47	58	63	77	88	82	53	92	101	...	66	72	58	97	82	67	63	59	105	2220
2	78	98	66	71	72	81	93	81	61	86	...	67	57	66	81	57	55	102	74	63	2142
3	64	65	86	69	42	65	73	94	73	48	...	46	52	87	78	84	80	102	57	61	2274
4	78	37	62	64	78	92	55	72	54	90	...	79	72	76	77	45	91	59	79	93	2411
5	71	57	37	88	98	37	46	47	65	89	...	77	80	91	57	78	120	92	69	94	2177
6	65	71	70	47	57	73	78	79	65	75	...	103	99	65	61	87	72	49	57	84	2171
7	102	68	71	83	56	71	85	54	97	84	...	86	68	64	57	63	58	93	59	84	2205
8	55	75	55	90	69	120	50	52	93	88	...	72	60	49	79	68	67	78	124	71	2218
9	58	76	93	80	80	81	61	74	73	74	...	92	92	72	86	95	72	57	63	66	2188

Fig.7. Prepared Data

For the mentioned Product Categories, we've trained 5 different Algorithms. These algorithms consists of regression based models as well as time series models. Following are the models used:

- Xgboost
- Linear Regression
- Random Forest
- Arima
- Simple Exponential Smoothing

For each model, we have performed 3 experimentations using different feature sets. The features set for regression based models are 90 days sales, 90 days sales + 90 Days Aggregate Discount and 90 days sales + 90 Days Aggregate Sales. The feature for Arima model are Previous Monthly Sales, Previous Monthly Sales + discount and Previous Monthly Sales + Monthly Sales. For Simple Exponential Smoothing the model is trained on Previous Monthly sales + learning rate=0.3, Previous Monthly sales + learning rate=0.5 and Previous Monthly sales + learning rate=0.7.

Chapter 4

Research Methodology (Materials and Methods)

The methods used to carry out this thesis work will be described in the following section. It also focuses on the processes done to prepare data for predictive modelling. Following are the data analysis step that we performed: Data pre-processing, feature engineering and Model Building.

4.1 Dataset

The dataset used in the experimentation for Thesis belongs to Supply Chains used by the company DataCo Global [30]. The transaction record from Jan 2015 till Jan 2018 is available in the dataset. The dataset contain sales record of variety of products. Each product falls under a certain product category. The categories ranges from Cleats, Clothing, Sports and Electronics etc. This data set contains E-commerce transactions and contains 53 product categories that are ordered from all overall the world.

4.1.1 Dataset details

There are 180520 observations (rows) and 53 attributes in this data collection (columns). Each of these observations corresponds to a record of order placed with a distinct identifying number. Each row also specify the status of the order and the payment method. The dataset records also shows the country and city from where the order is placed. Each row contains the price of the product, the discount given on that product and the number of items sold. Table 3 shows the list of some attributes in the dataset.

Table 3: Some Attributes of the dataset

Category Name	Customer Country	Customer Id
Department Name	Order Country	order date (Date Orders)
Order Id	Order Item Discount	Order Item Quantity
Sales	Order Status	Product Name

4.2 Data Preparation

4.2.1 Data Pre-processing

Since we are to make product category level forecast. We have filtered the data and selected the records for the required category. After filtering the data for a particular category, we did feature extraction and selected those features that will help us in generating a forecast. Then data was sorted by date value. As this is a time-series data, we make the ‘order date (Date Orders)’ column as the index and change its name to ‘Date’. After performing aggregation and grouping operation the data looks as it is shown in the fig [8].

Date	Sales IN DOLLARS	Discount IN DOLLARS	Item Quantity
2015-01-01	1051590	106209	11854
2015-02-01	927009	93998	10438
2015-03-01	1051253	107204	12062
2015-04-01	1014463	102460	11287
2015-05-01	1050478	106539	11902
2015-06-01	1024006	103765	11203
2015-07-01	1038081	105145	11800
2015-08-01	1029494	104303	11612

Fig.8. Processed Data

4.2.2 Feature Engineering

Feature engineering is an important step in data preparation. Daily aggregated Sales will assist us in engineering new feature. New features are made by performing slicing and sub selection on the data. The target variable ‘Order Item Quantity’ is made by taking the month wise aggregate of the items sold. The data prepared consists of 30 records and 91 features. Fig [7] shows the features which are to be fed to a machine learning model.

One feature is the target variable and the rest of the 90 features are Item sales record for the past 90 days. Each feature represent a preceding day of Item sold at that day.

Each record represents the aggregated Monthly Sales value and 90 previous daily sales values.

To implement machine learning regression based models, we have converted the time series problem into supervised learning problem by doing feature engineering.

4.2.3 Segregation of Training and Testing Data

One of the most important steps in machine learning is to feed some data into the algorithm and train it to recognise patterns in the data. Once the algorithm has learned the pattern, it must be fed another dataset to determine the algorithm's level of knowledge. For the purposes of training and testing, it is common practise to divide the available data into two groups.

4.3 Model Building and k-Fold Validation

The training set is fed into the algorithm after the dataset has been separated into training and testing sets so that it may learn how to predict values. Multiple Algorithms, such as linear regression, xgboost, random forest, ARIMA and exponential smoothing have been used. For Accurate scores k-fold cross validation is apply with 10 splits on regression based models.

4.4 Performance Evaluation

The performance of any regression approach is measured by feeding test data into the learning algorithm. This method shows how well a model has learned the pattern in the data and can forecast the values of new data. To calculate the performance of any given regression model, the root mean squared error (RMSE) and Mean Absolute Error (MAE)

4.5 Summary

This section discusses the methodology and the stepwise details in the data preparation phase. After that we did the all-important feature engineering and choose the most relevant features. Before building model, we split the data into train and test set and performed k fold validation. Then we build different machine learning models and evaluated them using root mean squared error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

Chapter 5

Discussion and Future Directions

5.1 Results and Discussion

We are able to discuss the results of the implemented part done in this Thesis. The experimentations are performed on jupyter notebook python 3.8.3 and library used were numpy, pandas, matplotlib, scikit learn and pmdarima. We performed the experiments on dataset "DataCo Smart Supply Chains" from kaggle [30]. Data Pre-processing followed by the steps mentioned in chapter 4 are executed on the dataset. We experimented our proposed methodology on 5 different product categories "*Cleats", "Women's Apparel", "Indoor/Outdoor Games", Water Sports" and "Camping & Hiking"*. Five Different models are used for each category with each model performing three experiments using different set of features. The results of all the 5 category are depicted in table 4 below. The best result in each experiment are highlighted.

Table 4: Results

Algorithms	**[CLEATS]** Model Evaluation on different feature Sets								
	Previous 90 days Sales			Previous 90 days Sales + 90 Days Aggregate Discount			Previous 90 days Sales + 90 Days Aggregate Sales(dollars)		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Xgboost	**86.07**	**115.30**	**2.85**	86.15	115.37	2.85	86.29	115.44	2.85
Linear regression	**106.05**	**144.91**	**3.79**	106.36	146.64	3.80	106.29	146.07	3.80
Random Forest	85.75	112.04	2.88	**85.26**	**111.98**	**2.85**	85.72	112.43	2.87
ARIMA	Previous Monthly Sales			Previous Monthly Sales + Discount			Previous Monthly Sales + Sales		
	95.93	133.93	4.62	94.43	131.54	4.60	**93.43**	**132.54**	**4.56**
Exponential Smoothing	Learning Rate α = 0.3			Learning Rate α = 0.5			Learning Rate α = 0.7		
	88.11	79.66	2.61	83.43	73.54	2.34	**81.43**	**69.43**	**2.24**

Algorithms	**[WOMEN'S APPAREL]** Model Evaluation on different feature Sets								
	Previous 90 days Sales			Previous 90 days Sales + 90 Days Aggregate Discount			Previous 90 days Sales + 90 Days Aggregate Sales(dollars)		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Xgboost	101.69	119.80	4.31	**101.68**	**119.61**	**4.30**	101.73	119.67	4.31
Linear regression	**121.03**	**146.15**	**5.40**	127.26	151.55	5.72	127.48	151.60	5.74
Random Forest	**91.08**	**111.26**	**3.80**	91.25	110.49	3.82	91.25	110.50	3.82
ARIMA	Previous Monthly Sales			Previous Monthly Sales + Discount			Previous Monthly Sales + Sales		
	98.65	**109.34**	**4.03**	101.43	107.34	4.46	99.68	104.96	4.22
Exponential Smoothing	Learning Rate α = 0.3			Learning Rate α = 0.5			Learning Rate α = 0.7		
	92.73	**102.78**	**3.28**	97.86	104.09	3.50	101.87	112.06	4.67

Algorithms	**[Indoor/Outdoor Games]** Model Evaluation on different feature Sets								
	Previous 90 days Sales			Previous 90 days Sales + 90 Days Aggregate Discount			Previous 90 days Sales + 90 Days Aggregate Sales(dollars)		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Xgboost	**110.68**	**133.04**	**5.38**	110.69	133.17	5.38	110.80	133.14	5.39
Linear regression	**103.01**	**130.84**	**4.95**	105.73	134.88	5.12	105.02	134.01	5.07
Random Forest	88.08	108.19	4.09	85.92	106.92	3.97	**85.88**	**106.80**	**3.97**
ARIMA	Previous Monthly Sales			Previous Monthly Sales + Discount			Previous Monthly Sales + Sales		
	82.93	**106.52**	**4.06**	83.08	111.98	4.34	99.86	114.50	5.19
Exponential Smoothing	Learning Rate α = 0.3			Learning Rate α = 0.5			Learning Rate α = 0.7		
	74.19	73.57	3.30	**68.50**	**75.17**	**2.54**	68.50	75.63	2.64

Algorithms	[Water Sports] Model Evaluation on different feature Sets								
	Previous 90 days Sales			Previous 90 days Sales + 90 Days Aggregate Discount			Previous 90 days Sales + 90 Days Aggregate Sales(dollars)		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Xgboost	**33.33**	**38.53**	**6.13**	33.43	38.58	6.15	33.43	38.58	6.15
Linear regression	**30.55**	**34.76**	**5.53**	32.01	36.37	5.86	32.16	36.52	5.89
Random Forest	**25.51**	**31.82**	**4.44**	25.62	31.91	4.47	25.55	31.83	4.45
ARIMA	Previous Monthly Sales			Previous Monthly Sales + Discount			Previous Monthly Sales + Sales		
	23.79	**29.24**	**4.74**	25.49	28.24	4.87	24.86	28.34	4.77
Exponential Smoothing	Learning Rate α = 0.3			Learning Rate α = 0.5			Learning Rate α = 0.7		
	26.34	31.27	5.58	27.87	35.27	5.88	**24.34**	**27.27**	**4.58**

Algorithms	[Camping & Hiking] Model Evaluation on different feature Sets								
	Previous 90 days Sales			Previous 90 days Sales + 90 Days Aggregate Discount			Previous 90 days Sales + 90 Days Aggregate Sales(dollars)		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Xgboost	**16.38**	**20.71**	**2.97**	16.49	20.87	3.00	16.41	20.79	2.98
Linear regression	**20.88**	**26.81**	**4.07**	22.20	27.80	4.38	22.21	27.78	4.38
Random Forest	**16.99**	**20.98**	**3.12**	17.24	21.39	3.17	17.18	21.32	3.16
ARIMA	Previous Monthly Sales			Previous Monthly Sales + Discount			Previous Monthly Sales + Sales		
	22.44	29.10	5.44	19.32	24.68	4.76	**17.45**	**22.35**	**3.96**
Exponential Smoothing	Learning Rate α = 0.3			Learning Rate α = 0.5			Learning Rate α = 0.7		
	15.23	19.23	4.23	**14.45**	**15.43**	**3.65**	15.65	17.04	3.37

For Each category, the predicted value of each model is fed to a multilayer perceptron. This neural network shown in fig [9] will decide the Final Demand Forecast for the particular category.

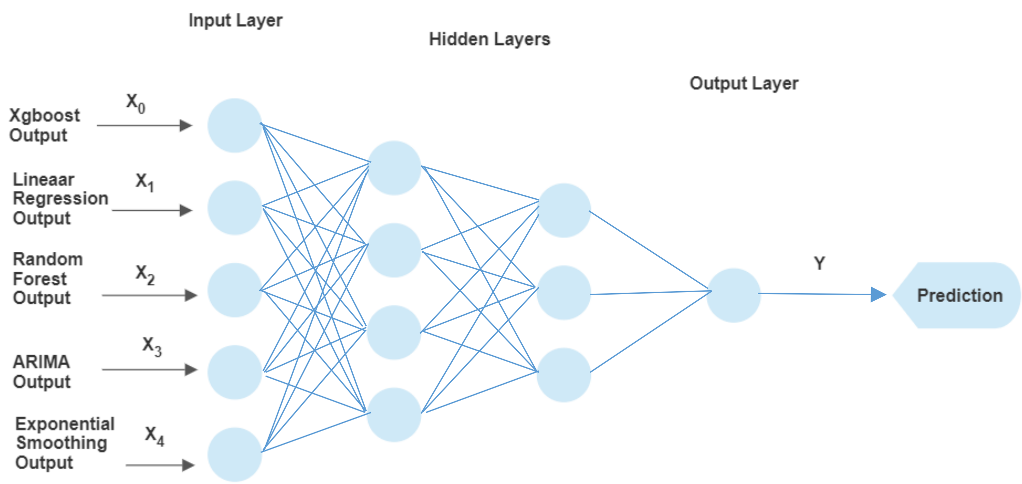

Fig.9. Final Forecast for Each Category

From the results in Table [4], it can be observed that majority of the models are performing better when the feature set with previous sales + discount and previous sales + Monthly sales are used as features. Considering the results we will use those features to predict an outcome. All the models are training on those features and their predictions are fed to the multilayer perceptron for the final prediction for each category.

The predictions of the five models are fed to a multilayer perceptron as input to the five neurons. There are two hidden layers with four and three neurons respectively. And there is one output layer which is the combined final prediction. The weights and biases are initialized between inputs Layer to hidden layer1 and between hidden layer1 to hidden layer2. The Output layer with only one neuron gives the final prediction. Table [5] shows the results of multilayer perceptron on the 5 categories to make the final demand prediction for each category.

Table 5: Multilayer perceptron output

CATEGORIES	MODELS USED Trained on best feature set respectively					
	MEAN ABSOLUTE ERROR (MAE)					
	Xgboost	Linear Regression	Random Forest	ARIMA	Exponential Smoothing	**Multilayer Perceptron**
Cleats	86.07	106.05	85.26	93.43	81.43	**83.45**
Women's Apparel	101.68	121.03	91.08	98.65	92.73	**95.23**
Indoor Outdoor Games	110.68	103.01	85.88	82.93	68.50	**70.23**
Water Sports	33.33	30.55	25.51	23.79	24.34	**25.48**
Camping & Hiking	16.38	20.88	16.99	17.45	14.45	**15.95**

Comparison between Multilayer Perceptron output and each individual Algorithm of each category with respect to Mean Absolute Error (MAE).

- Cleats: [MLP Outperforms 4 models]
- Women's Apparel: [MLP Outperforms 3 models].
- Indoor Outdoor Games: [MLP Outperforms 4 models].
- Water Sports: [MLP Outperforms 3 models].
- Camping & Hiking: [MLP Outperforms 4 models].

5.2 Thesis Conclusion

In this study, we highlighted the importance of demand forecasting for supply chains. Introduction of machine learning in demand forecasting has automated many processes within the Supply chain management (SCM) system. We have seen how ML-based Demand forecasting has help supply chains in inventory management and cutting down costs. A detailed literature review is presented related to this topic. Many State-of-the-arts algorithms are studied that can be used in make the predictions. With the introduction of Big Data, the potential for deep learning in demand forecasting has increased ever since. There are still many challenges and opportunities in forecasting domain. With the emergence of new technologies and E-commerce there will remain a room for improvement in the Supply chain management.

5.3 Future Directions

The data used in this study was stationary. It did not exhibit any trend or seasonality. The Causal regression based model used in this study are linear regression, xgboost and random forest. The Date time features such as hours, days and weeks and can be added as features to these model if there is seasonality in data. Similarly the time series model used in this study are ARIMA and Exponential Smoothing. For a data with Seasonality, SARIMA and Holt winter smoothing method tends to perform better. Feature set fed to the model could be varied by feature engineering.

5.4 Summary

This chapter summarises the results, discussion, and conclusion of Thesis as well as possible future directions. It also contributes to a better understanding of the problem statement at hand.

References

[1] Tirkolaee, Erfan Babaee, et al. "Application of Machine Learning in Supply Chain Management: A Comprehensive Overview of the Main Areas." *Mathematical Problems in Engineering*, edited by Ali Ahmadian, vol. 2021, June 2021, pp. 1–14. DOI.org (Crossref), https://doi.org/10.1155/2021/1476043.

[2] Niu, Yiyang. "Walmart Sales Forecasting Using XGBoost Algorithm and Feature Engineering." *2020 International Conference on Big Data & Artificial Intelligence & Software Engineering* (ICBASE), IEEE, 2020, pp. 458–61. DOI.org (Crossref), https://doi.org/10.1109/ICBASE51474.2020.00103.

[3] Dong, Wenxiang, et al. "Statistical and Machine Learning-Based E-Commerce Sales Forecasting." *Proceedings of the 4th International Conference on Crowd Science and Engineering*, ACM, 2019, pp. 110–17. DOI.org (Crossref), https://doi.org/10.1145/3371238.3371256.

[4] Krishna, Akshay, et al. "Sales-Forecasting of Retail Stores Using Machine Learning Techniques." *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions* (CSITSS), IEEE, 2018, pp. 160–66. DOI.org (Crossref),https://doi.org/10.1109/CSITSS.2018.8768765.

[5] Ding, Jingyi, et al. "Sales Forecasting Based on CatBoost." *2020 2nd International Conference on Information Technology and Computer Application* (ITCA), IEEE, 2020, pp. 636–39. DOI.org (Crossref), https://doi.org/10.1109/ITCA52113.2020.00138

[6] Pavlyshenko, Bohdan. "Machine-Learning Models for Sales Time Series Forecasting." Data, vol. 4, no. 1, Jan. 2019, p. 15. DOI.org (Crossref), https://doi.org/10.3390/data4010015.

[7] Wiyanti, Dt, et al. "Machine-Learning Algorithm for Demand Forecasting Problem." *Journal of Physics: Conference Series*, vol. 1918, no. 4, June 2021, p. 042012. DOI.org (Crossref), https://doi.org/10.1088/1742-6596/1918/4/042012.

[8]"Sales Demand Forecast in E-commerce using a long Short-Term Memory Neural Network Methodology" https://arxiv.org/pdf/1901.04028.pdf

[9] Feizabadi, Javad. "Machine Learning Demand Forecasting and Supply Chain Performance." *International Journal of Logistics Research and Applications*, Aug. 2020, pp. 1–24. DOI.org (Crossref), https://doi.org/10.1080/13675567.2020.1803246.

[10] Cheriyan, Sunitha, et al. "Intelligent Sales Prediction Using Machine Learning Techniques." *2018 International Conference on Computing, Electronics & Communications Engineering* (ICCECE), IEEE, 2018, pp. 53–58. DOI.org (Crossref), https://doi.org/10.1109/iCCECOME.2018.8659115.

[11] Aburto, Luis, and Richard Weber. "Improved Supply Chain Management Based on Hybrid Demand Forecasts." *Applied Soft Computing*, vol. 7, no. 1, Jan. 2007, pp. 136–44. DOI.org (Crossref), https://doi.org/10.1016/j.asoc.2005.06.001.

[12] Álvarez-López, Víctor, et al. "Hybrid Machine Learning Methods for Demand Forecasting: Consecutive Application of Classification and Regression Methods for the Forecasting of Periodic and Non-Continuous Demand." *Proceedings of the 2nd International Conference on Applications of Intelligent Systems*, ACM, 2019, pp. 1–6. DOI.org (Crossref), https://doi.org/10.1145/3309772.3309773.

[13] Das Adhikari, Nimai Chand, et al. "Ensemble Methodology for Demand Forecasting." 2017 *International Conference on Intelligent Sustainable Systems* (ICISS), IEEE, 2017, pp. 846–51. DOI.org (Crossref), https://doi.org/10.1109/ISS1.2017.8389297.

[14] Carbonneau, Real, et al. "Application of Machine Learning Techniques for Supply Chain Demand Forecasting." *European Journal of Operational Research*, vol. 184, no. 3, Feb. 2008, pp. 1140–54. DOI.org (Crossref), https://doi.org/10.1016/j.ejor.2006.12.004.

[15] Kilimci, Zeynep Hilal, et al. "An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain." *Complexity*, vol. 2019, Mar. 2019, pp. 1–15. DOI.org (Crossref), https://doi.org/10.1155/2019/9067367.

[16] Jain, Anupriya, et al. "Demand Forecasting for E-Commerce Platforms." 2020 IEEE *International Conference for Innovation in Technology* (INOCON), IEEE, 2020, pp. 1–4. DOI.org (Crossref), https://doi.org/10.1109/INOCON50539.2020.9298395.

[17] Tanaka, Kenji. "A Sales Forecasting Model for New-Released and Nonlinear Sales Trend Products." *Expert Systems with Applications*, vol. 37, no. 11, Nov. 2010, pp. 7387–93. DOI.org.

[18] Catal, Cagatay, et al. "Benchmarking of Regression Algorithms and Time Series Analysis Techniques for Sales Forecasting." *Balkan Journal of Electrical and Computer Engineering*, Jan. 2019, pp. 20–26. DOI.org (Crossref), https://doi.org/10.17694/bajece.494920.

[19] Spiliotis, Evangelos, et al. "Comparison of Statistical and Machine Learning Methods for Daily SKU Demand Forecasting." *Operational Research*, Sept. 2020. DOI.org (Crossref), https://doi.org/10.1007/s12351-020-00605-2.

[20] Qiu, Xueheng, et al. "Ensemble Deep Learning for Regression and Time Series Forecasting." *2014 IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL), IEEE*, 2014, pp. 1–6. DOI.org (Crossref),https://doi.org/10.1109/CIEL.2014.7015739.

[21] Hua, Yuxiu & Zhifeng, Zhao & Li, Rongpeng & Chen, Xianfu & Liu, Zhiming & Zhang, Honggang. (2019). Deep Learning with Long Short-Term Memory for Time Series Prediction. *IEEE Communications Magazine*. PP. 1-6. 10.1109/MCOM.2019.1800155.

[22] Rafi, Shafiul Hasan, and Nahid-Al-Masood Nahid-Al-Masood. "Highly Efficient Short Term Load Forecasting Scheme Using Long Short Term Memory Network." 2020 8th *International Electrical Engineering Congress* (IEECON), IEEE, 2020, pp. 1–4. DOI.org

[23] Aamer, Ammar, et al. “Data Analytics in the Supply Chain Management: Review of Machine Learning Applications in Demand Forecasting.” *Operations and Supply Chain Management: An International Journal*, Dec. 2020, pp. 1–13. DOI.org (Crossref), https://doi.org/10.31387/oscm0440281.

[24] Makridakis, Spyros, et al. “Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward.” *PLOS ONE*, edited by Alejandro Raul Hernandez Montoya, vol. 13, no. 3, Mar. 2018, p. e0194889. DOI.org (Crossref), https://doi.org/10.1371/journal.pone.0194889.

[25] Ching-Wu Chu, & Guoqiang Peter Zhang. A comparative study of linear and nonlinear models for aggregate retail sales forecasting. *International Journal of Production Economics*, 86(3), 217-231.

[26] L. Rokach, “Ensemble-based classifiers,” *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 1–39, 2010.

[27] R. Polikar, “Ensemble based systems in decision making,” *IEEE Circuits and Systems Magazine*, vol. 6, no. 3, pp. 21–45, 2006.

[28] Y. Ren, L. Zhang, and P. N. Suganthan, “Ensemble classification and regression-recent developments, applications and future directions,” *IEEE Computational Intelligence Magazine*, vol. 11 , no. 1, pp. 41–53, 2016.

[29] G. Tsoumakas, L. Angelis, and I. Vlahavas, “Selective fusion of heterogeneous classifiers,” *Intelligent Data Analysis*, vol. 9, no. 6 , pp. 511–525, 2005

[30] https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis

[31] https://hbr.org/1971/07/how-to-choose-the-right-forecasting-technique

[32] https://mobidev.biz/blog/machine-learning-methods-demand-forecasting-retail#:~:text=Ok-,What%20Is%20Demand%20Forecasting%20in%20Machine%20Learning%3F,from%20data%20for%20improved%20analysis.

[33] https://scholarworks.rit.edu/theses/10722/