A Project Report on

Intermittent Demand Classification using Machine Learning

Submitted in partial fulfilment for award of degree of

MBA

In **Business Analytics**

Submitted by

Ashish Kumar Singh

R19MBA12

Under the Guidance of

Dr. J.B. Simha

CTO, Abiba Technologies/Chief Mentor, RACE

REVA Academy for Corporate Excellence

REVA University

Rukmini Knowledge Park, Kattigenahalli,

Yelahanka, Bangalore – 560064

February 2021

REVA UNIVERSITY
Bengaluru, India

Candidate's Declaration

I, Ashish K. Singh hereby declare that I have completed the project work towards the first year of Master of Business Administration in Business Analytics at, REVA University on the topic entitled "**Intermittent Demand Classification using Machine Learning"** under the supervision of Dr. J.B.Simha, CTO, Abiba Technologies and Chief Mentor, RACE. This report embodies the original work done by me in partial fulfilment of the requirements for the award of the degree for the academic year 2019-21.

Place: Bengaluru
Date: 06-Mar-2021

Name of the Student: Ashish K. Singh
Signature of Student

REVA UNIVERSITY
Bengaluru, India

Certificate

This is to Certify that the Project work entitled "**Intermittent Demand Classification using Machine Learning"** carried out by Ashish K. Singh with R19MBA12, is a bonafide student of REVA University, is submitting the first-year project report in fulfilment for the award of Master of Business Administration in Business Analytics during the academic year 2019-21. The Project report has been tested for plagiarism and has passed the plagiarism test with a similarity score less than 15%. The project report has been approved as it satisfies the academic requirements in respect of PROJECT work prescribed for the said Degree.

Signature of the Guide	Signature of the Director
Dr. J.B.Simha	Dr. Shinu Abhi
Guide	Director

External Viva

Names of the Examiners

1. Indrajit Kar, Head of AI-Chief Architect & Data Scientists, Siemens
2. Pradeepta Mishra, Associate Principal & Head of AI, LTI -Larsen and Toubro Infotech.

Place: Bengaluru

Date: 06-Mar-2021

REVA UNIVERSITY
Bengaluru, India

Acknowledgement

Having been finally able to complete my work on this project, there are a bunch of people that I would like to thank for helping me through this important part of my study. Firstly, I am ineffably indebted to my mentor Dr. J.B.Simha for his unconditional support and conscientious guidance that helped my project find a structure. His neverending trust on me and my capabilities encouraged me to outperform my abilities and encouraged me on the topic of forecasting. I would also like to thank Dr. Shinu for reviewing my report with the utmost patience and providing me with useful insights for improvement. I would also grateful to my friends at RACE who supported me during this project.

Furthermore, I want to thank my family who always showed unconditional pride and support during the entire course of my study, especially my Father, Late Sri. Sanjay Singh who always encouraged me to stay positive and motivated in the most difficult phases of life and taught me that there is always light at the end of the tunnel.

I am thankful for and would like to acknowledge Hon'ble Chancellor, Dr. P Shayma Raju, Vice-Chancellor, Dr. K. Mallikharjuna Babu, and Pro-Vice Chancellor, Dr. M. Dhanamjaya, for giving me a chance to enrol in this program with which I have gained immense knowledge that cannot be quantified.

Place: Bengaluru
Date: 06-Mar-2021

REVA UNIVERSITY
Bengaluru, India

Similarity Index Report

This is to certify that this project report titled "**Intermittent Demand Classification using Machine Learning"** was scanned for similarity detection. Process and outcome are given below.

Software Used: Turnitin
Date of Report Generation: 06-Mar-2021
Similarity Index in %: 13%
Total word count: 7088
Name of the Guide: Dr. J.B.Simha

Place: Bengaluru
Date: 06-Mar-2021

Name of the Student:Ashish K. Singh
Signature of Student

Verified by: Andrea Brian Churchill

Signature
Dr. Shinu Abhi,
Director, Corporate Training

List of Abbreviations

Sl. No	Abbreviation	Long Form
1	SCM	Supply Chain Management
2	JIT	just-in-time
3	SMA	Simple Moving Average
4	SES	Simple Exponential Smoothing
5	ARIMA	Autoregressive Integrated Moving Average
6	IIF	International Institute of Forecasters
7	TP	True Positive
8	TN	True Negative
9	FP	False Positive
10	FN	False Negative
11	SVM	Support Vector Machine
12	NN	Neural Network
13	MLP	Multi-Layer Perceptron

List of Figures

List of Tables

Abstract

Supply Chain Management (SCM) is the most important aspect for any organization building a competitive advantage, irrespective of the size. However big organizations with large-scale inventory and multiple products to offer often experience huge proportions of items that have small irregular demand with long periods of zero demand. The problem with these products with low irregular demand is that these items need to be stocked and replenished at regular intervals irrespective of the demand cycle, hence adding to the cost of holding the inventory. To be efficient in the above situation, Supply Chain and Inventory Management require accurate and reliable demand forecasting. Accurate forecasting of demand helps an organization to maximize customer value and achieve sustainable competitive advantage.

The above phenomenon of low irregular demand in between long periods of zero demand is known as **Intermittent Demand** and can be defined as "Intermittent demand or ID (also known as sporadic demand) occurs when a product experiences several periods of zero demand. Often in these situations, when demand occurs is small, and sometimes highly variable in size (Waller, 1956)". Intermittent Demand is a common business problem in automotive, aerospace and defence, maritime and heavy machinery industries as these industries rely heavily on high-value speciality components.

This project considers the demand forecasting and replenishment process of low volume intermittent demand items to provide actionable recommendations to create a dynamic re-order point to improve the inventory performance of these items. The intermittent demand forecasting entails the forecasting of demand series whith its peculiar characterization of significant large time interval between demand tipically larger than the unit of time used used for the forecast period. This leads to a large unit of time intervals with no demand occurring on the demand time series. (Varghese & Rossetti, 2008).

Keywords: Inventory Management, Demand Forecasting, Intermittent Demand, Time Series, Machine Learning, Naïve Bayes Classification

Contents

Chapter 1: Introduction

"Prediction is very difficult, especially if it's about the future." - Nils Bohr

The supply chain function is a necessary aspect of any organization to fulfil all the basic requirements from food and water to medicine, electricity and such. However, the supply chain activities come with their own set of complex challenges such as packaging, transportation, logistics, warehousing and inventory management and the costs associated with each of these.

Good SCM can improve customer service, reduce operating and inventory costs, and bring significant gains to a company's financial position. It is, therefore, necessary for companies to devise and implement Modern SCM plans and strategies to get a competitive edge over rivals. Companies today are facing increasing customer expectations and with the everyday increasing competition, products need to be delivered to the customer at the right cost, right place, and the right time and in the required quantity.

Inventory management is the most critical cornerstone of SCM and is the process of controlling the products an organisation keeps. "Inventory carrying costs are expressed as a percentage number and typically accounts for 15% to 30% of the value of a merchant's inventory (Flowspace, n.d.)". Thus, the performance of the SCM and the Organization as a whole is highly related to the performance of inventory management and thus, forecasting of the demand has become an essential component of inventory control. The size of the safety stock required in an inventory is directly proportional to the uncertainty of the forecasting model.

Many manufacturers are moving to just-in-time (JIT) manufacturing strategies for optimal utilization of resources and also to reduce the cost of holding high levels of inventory. However JIT requires the manufacturing to be done and the items delivered when the demand is expected as increasing competition has made customers more demanding and they expect their products whenever they demand and if the same is not available they will move on to the next seller, causing lost sale and dented reputation. Forecasting, therefore, is the indispensable pillar of the inventory control system and requires accurate and reliable prediction of the demand.

To overcome the challenges of ever-increasing inventory and labour costs, organizations are increasingly adopting innovative SCM technologies that utilize advanced planning and scheduling techniques for creating the optimized schedule to increase efficiency and obtain maximum output with limited resources by synchronizing the supply with the demand period and reducing inventory. “Finding patterns in the demand should give valuable insights in future demand (Visser, 2017)”. To achieve the objective of efficient inventory management it is necessary to find the demand probability based on historic demand to get a reliable demand estimate. Various forecasting techniques are available to predict the demand which can then be used to replenish the inventory right before the demand is predicted.

Items that see regular sales can be considered as fast-moving products and needs to be reordered and replenished at regular intervals. The requirement for Intermittent demand items (also known as ‘lumpy demand’ or ‘erratic demand’) on the other hand, generally appears sporadically and are usually considered as slower movers in any organization. “However, despite the comparatively low contribution to the total turnover, these slow-moving Stock Keeping Units (SKUs) may constitute up to 60% of the total stock value (Babai et al., 2014)”. These components with stochastic demand pattern create hindrance in accurate forecasting and planning.

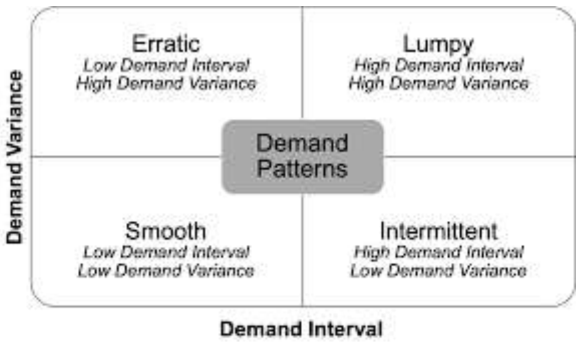

Figure 1.1 – Demand Patterns

Statistical Forecasting methods come in handy to discover the underlying pattern in historic data and use them to predict the future assuming that the historical pattern will continue in the future. Currently, most of the demand forecasting is done by using time series forecasting methodologies such as Simple Moving Average (SMA), Exponential Smoothing and Simple Exponential Smoothing (SES), Autoregressive Integrated Moving Average (ARIMA), etc. However these methods are not suited to predict Intermittent Demand as time series models do not take into account the underlying structure of intermittency in the demand and give more weightage to recent demand, also the variation in-demand size and period makes it difficult to forecast this type of demand and results in a severe error, rendering classical statistical models infeasible.

In 1972, J.D. Croston published "**Forecasting and Stock Control for Intermittent Demands**" (Croston, 1972) that introduced a new methodology to forecast products with intermittent demand that have a univariate forecast profile that came to be known as Croston's method. It is a modification of the exponential smoothing for intermittent demand products and is very useful in certain circumstances. However, even this method comes with its own bias and a meaningful alternative is the need of the hour.

With the advent of the millennium, a lot of emphases has been given to Industry 4.0 due to the increase in computational power and the decrease in the cost of the storage device. Essentially, the technologies that make Industry 4.0 possible is the utilization of past and current data. "To understand Industry 4.0, it is essential to see the full value chain which includes suppliers and the origins of the materials and components needed for various forms of smart manufacturing, the end-to-end digital supply chain and the final destination of all manufacturing/production, regardless of the number of intermediary steps and players: the end customer (*Industry 4.0: The Fourth Industrial Revolution – Guide to Industrie 4.0*, n.d.)."

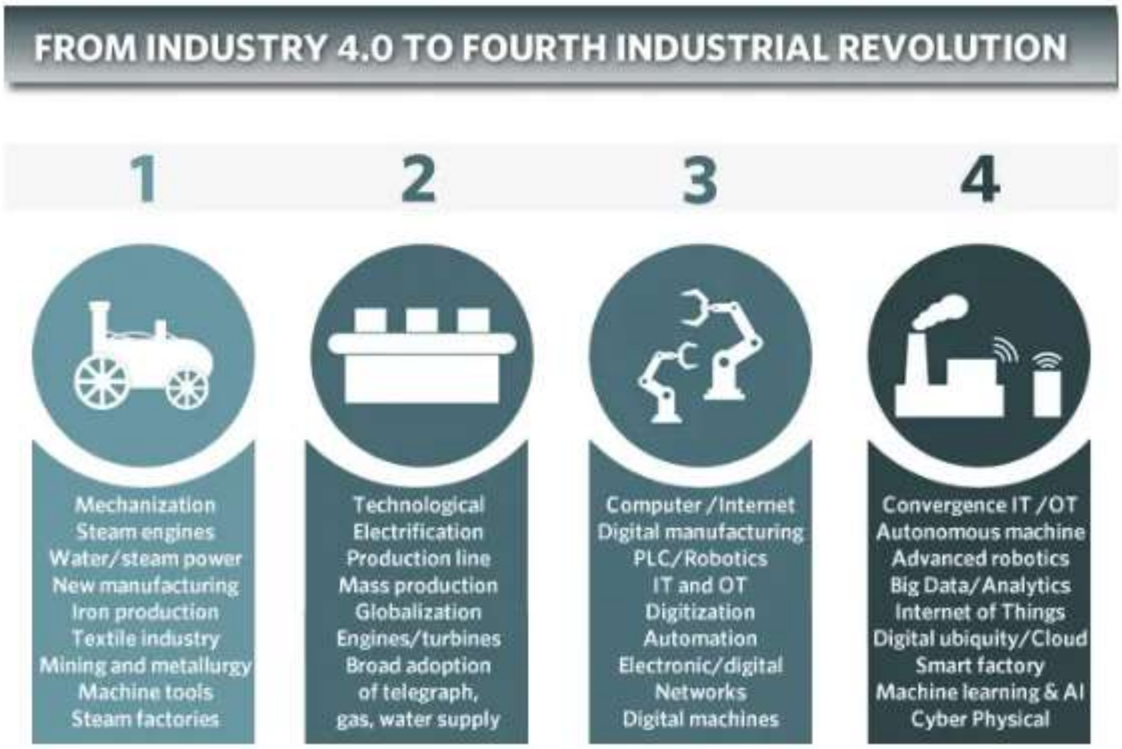

Figure 1.2 – Phases in the Industrial revolution

One of the cornerstones of the digital supply chain is inventory control and demand prediction. This problem can be solved using Machine learning and causal models that utilize historic data to predict meaningful demand forecast. Machine learning and causal models are powerful techniques that can be used to solve complicated problems in the varying domain and finding increasing utilization due to their ease of use and absence of complex mathematical formulations.

Figure 1.3 – Components in the 4th Industrial revolution

Chapter 2: Literature Review

"I have seen the future and it is very much like the present only longer." –K. Albran

Forecasting is the science of predicting future trends based on past and present data. Forecasting has become an integral part of decision making in today's world as these predictions support strategic planning and taking an informed decision, therefore these forecasts need to be as accurate as possible. No wonder it is a hotly contested domain due to its predictive nature and has received increasing attention since the 1980s with academic journals such as "**International Journal of Forecasting**" (Published by Elsevier on behalf of the International Institute of Forecasters, Volume 36, Issue 4) and the "**Journal of Forecasting**" (Published by John Wiley & Sons, Volume 39, Issue 7) dedicated to the cause. An "International Symposium on Forecasting" is also organized by the IIF on an annual basis.

Forecasting is used in a wide range of applications, from stock exchange prices to retail and electricity demand and from rainfall prediction to predicting the endangered animal extinction. It is, therefore, necessary to correctly identify the forecasting problem and applying the most appropriate methods based on the problem to obtain the best solution with further evaluation and refinement to reduce the forecast error.

There are two basic types—Qualitative techniques and Quantitative techniques. Quantitative techniques are further subdivided into two types of time series analysis and causal models.

The **Qualitative techniques** of forecast also sometimes referred to as judgmental of subjective techniques uses qualitative data (e.g. expert opinion, business knowledge, market and customer expectation) and may or may not consider experiences from previous incidents of similar nature. This method is generally used when little data is available for the forecast. Examples of qualitative techniques are the Delphi method, Market Research method, Panel Consensus techniques, Salesforce composite and Bayesian decision theory.

Quantitative techniques (also referred to as objective or mathematical techniques**)** utilize historical data and require a large number of calculations. Forecasting is done using statistical techniques and does not rely upon judgement. Therefore a good knowledge of underlying statistical procedures used for these techniques is needed before using it. As mentioned above there quantitative techniques are of two types– Time Series Analysis and Causal.

Time Series Analysis focuses on a sequence of observations that are ordered in time to extract its patterns and pattern changes, and thus relies completely on historical data. The publication of Box & Jenkins' (1970) book popularized time series methods by providing a practical approach for modelling linear non-stationary processes that included autoregressive and/or moving average components; their approach also extended to include modelling multiple time series (Robert & Makridakis Spyros, 1995). As defined by Granger and Newbold (1986), a time series is “a sequence of observations ordered by a time parameter”. Time series assumes that certain components like Trend, Seasonal Variations, Cyclic Variations etc. will repeat themselves over a while. SMA, SES, ARIMA and X11 Forecasting are some of the examples of Time series analysis.

Causal Forecasting, on the other hand, assumes that the forecast has a cause-effect relationship with one or more variables and uses a highly refined and specific information to determine relationships between these variables. Casual Forecasting is very robust and takes into consideration all possible factors and a special event that may impact the forecast, Therefore the past is important to causal models as also seen in time series analysis. Models used for Causal forecasting include Regression Model, Econometric Model, and Leading Indicator Models

The choice of the best forecasting model is based on the demand pattern which is used to select the demand estimation and forecast parameter optimization to obtain minimum forecast error. In 1984, T.M.Williams (Williams, 1984) proposed a method to categorize the demand patterns into “Smooth”, “Slow-Moving” or “Sporadic” using variance partitioning which he defined as “partitioning the variance of demand during a lead time into its constituent causal parts”. The purpose of this categorization was to have a separate inventory management policy and utilization of the best forecasting technique for each category of products.

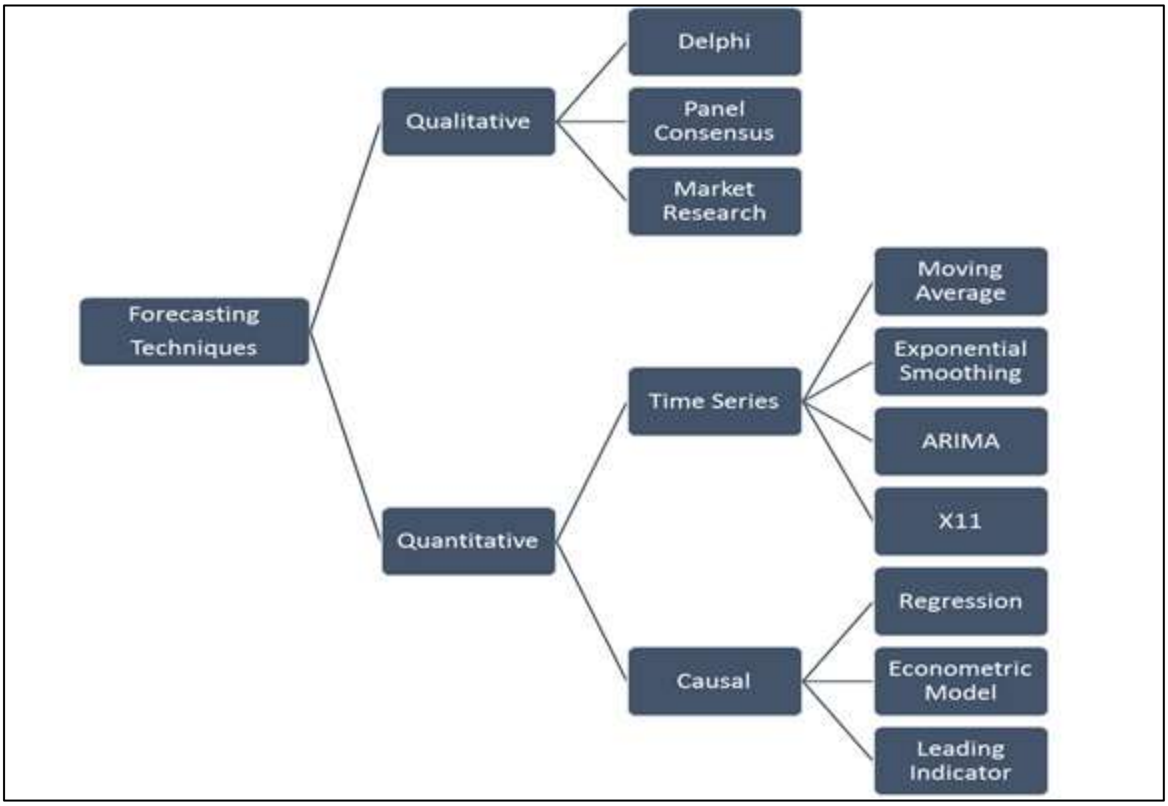

Figure 2.1 – Forecasting Techniques

To ensure that the most suitable forecasting method is used for demand with specific characteristics Demand patterns have been robustly classified into four categories by (Syntetos et al., 2005) using two metrics, Average Demand Interval (ADI) which is the average time between subsequent demands occurrence in the historical demand data (known as intermittence) and the squared coefficient of variation ($CV2$), which is variation in demand sizes when they occur ((known as lumpiness). Intermittent demand has a relatively high ADI.

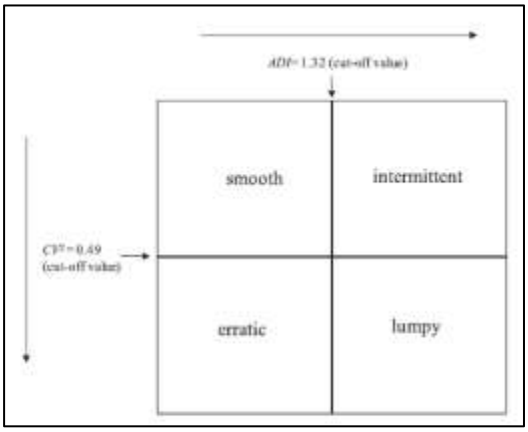

Figure 2.2 – Demand Classification

Ever since J.D.Croston came up with his forecasting approach known as Croston's method (Croston, 1972), this method is considered as the benchmark in intermittent demand forecasting as it is claimed to be unbiased. Forecasting reliable inventory size for intermittent demand is an arduous task and Croston's estimator that depends upon exponentially smoothing the demand sizes over the inter-demand interval is considered the benchmark in the industry. (Babai et al., 2014).

Chapter 3: Problem Statement

“Inventory is a current asset that should more than earn its keep; if inventory incurs more costs than benefits, it is a liability. Excess inventory is an operational liability.” - Toelle & Tersine

All organizations strive hard to decrease cost and increase profit. Optimizing SCM is the key to improve the bottom line. As discussed earlier inventory management is one of the building blocks of an efficient SCM and inappropriate inventory levels that require large investment is a liability to the organization. “An audit of a routine stock control system that used exponential smoothing for forecasting demand, revealed that for some low demand items the stock levels appeared to be excessive (Visser, 2017)”. Excess inventory is caused due to procuring higher quantity than required, therefore every organization must come up with an allowed inventory level to reduce inventory holding cost while holding sufficient inventory to fulfil all customer requirements to avoid losing business.

One of the major issues being faced by all organizations is insufficient forecast accuracy to predict the expected demand and fluctuation in actual vs. the predicted demand is a result of forecasting errors. This problem is further exaggerated with slow-moving and intermittent demand items. Standard forecasting techniques, like exponentially smoothing, moving averages do not predict intermittent demand accurately. Croston in his study (Croston, 1972) concluded that statistical methods like exponential smoothening is largely biased due to large smoothing constants and proposed possible improvements in the forecasting system that aimed to reduce this error that arises when there multiple periods of no demand.

The so-called “count data forecasting”, as proposed by Croston in 1972, focuses separately on estimating the demand size and period of the demand interval which was more intuitive and accurate for intermittent demand data. Separate single exponential smoothing estimates of the average demand size and the demand interval are calculated after the demand occurs, therefore, when no demand occurs; the forecast estimates before and after the demand period remain constant. The ratio of the above smoothing estimates of demand quantity and the demand interval then gives the average demand per period.

Let, D_t be the actual demand occured at time t,

Z_t be the approximate average of non-zero demand size for time t ,

P_t the approximate average of the interval size between non-zero demands, and,

q be the number of consecutive zero-demand periods, and,

F denote an estimate of mean demand size forecast

Then,

If $D_t > 0$,

$$Z_{t+1} = \alpha D_t + (1-\alpha) Z_t$$
$$P_{t+1} = \alpha q + (1-\alpha) P_t$$
$$F_{t+1} = Z_{t+1} / P_{t+1}$$

Where **α** is the learning Parameter that is used to allocate importance to either the most recent observations or the historical ones ($0<\alpha<1$)

If $D_t = 0$,

$$Z_{t+1} = Z_t$$
$$P_{t+1} = P_t$$
$$F_{t+1} = F_t$$

The most significant achievement of the Croston method over SES is its ability to estimate the period between the demand occurrences, which is of great value for inventory optimization. Syntetos & Boylan in their paper “On the bias of intermittent demand estimates” highlighted that the Croston's method is biased (Syntetos & Boylan, 2001). Various methodologies and adjustments of Croston’s method has been proposed ever since, like the Syntetos-Boylan Approximation (SBA), Shale-Boylan-Johnston (SBJ) method), Teunter- Syntetos - Babai (TSB) method.

Nevertheless, due to the underlying bias and inhibitions of the above methods, all these methods demonstrates mediocre accuracy across datasets. Also, since these techniques are built using simple exponential smoothing that contemplates predicting inter-demand interval, and demand size, which only gives an average demand over the range of the forecast period, which in turn burdens’ the organizations with excessive inventory and associated costs. Therefore, even the slightest improvement in the accuracy of intermittent demand prediction translates into remarkable savings.

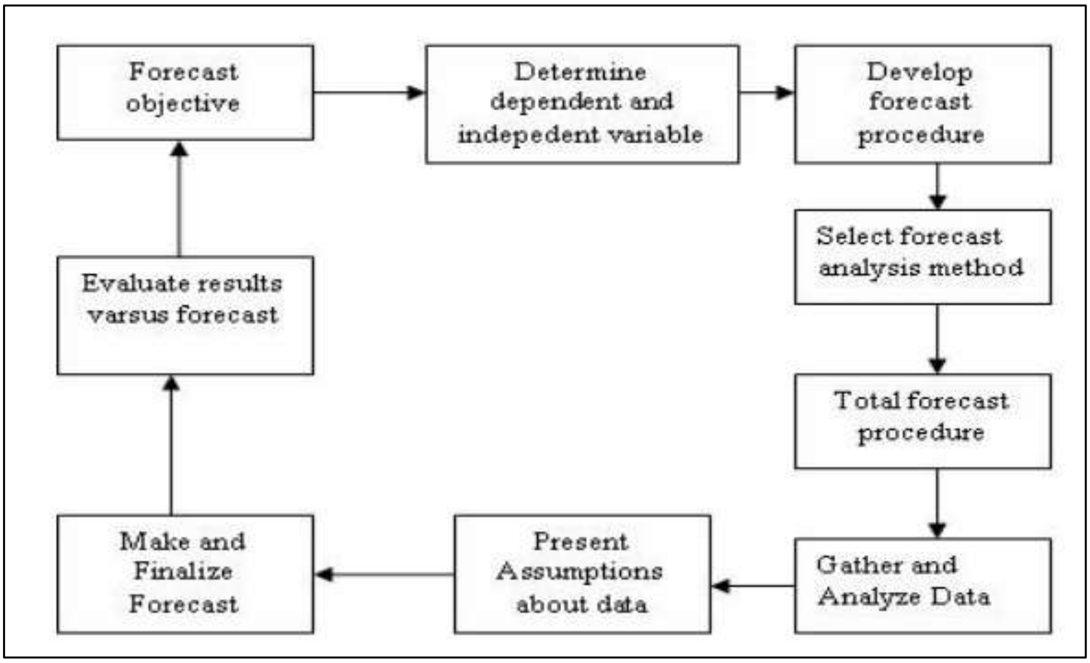

Figure 3.1 – Forecasting Methodology

Chapter 4: Objectives of the Study

“A lie gets halfway around the world before the truth has a chance to get its pants on.”
- Winston Churchill

There are very few forecasting methods that have been developed specifically to deal with the problem of intermittent demand. The special characteristics of Intermittent demand more than double the difficulty of the forecasting problem and hence it has received limited academic attention to further develop the available forecasting methods (Nikolopoulos, 2020). Intermittent Demand has two major uncertainties, variations in the demand size and the demand interval. Accurate and reliable forecasting techniques for intermittent demand can help up us deal with both the above uncertainties. Croston realized that applying traditional forecasting methods such as MA and SES to intermittent demand time series can lead to inaccurate forecast and unacceptable inventory size (Syntetos & Boylan, 2001; Teunter & Duncan, 2009). Croston’s approach focuses separately on both variables independently using a robust exponential smoothing approach and estimates a future mean demand size forecast using the ratio of the demand size and the demand interval forecasts.

Traditional forecasting models were compared with the Croston method and its variants, most results indicated that Croston methods and its variants gave inconsistent results for intermittent demand. It outperformed traditional methods with modest gains, and some studies even concluded that the Croston method and its variants had inferior performance (Syntetos & Boylan, 2001; Teunter & Duncan, 2009). “Croston method is based on the assumptions of independence (successive intervals are independent, successive demand sizes are independent and intervals and sizes are mutually independent) and normality of the demand size (Syntetos & Boylan, 2001)”. Also, the Croston method gives the average demand per period and hence, has a non–zero forecasts for all periods, which is not true in reality for intermittent demand.

Due to the limitation identified in the Croston method and the underlying bias in the estimated demand alternate methodologies such as Bootstrapping and Temporal Aggregation. “In all proposed applications so far in the literature, intermittent demand estimators were designed and tested to forecast spare part demand and respectively drive stock control decisions. There has not been a consideration, to the best of our knowledge, on how these methods can be used out of the core context of Operation Research and inventory management (Nikolopoulos, 2020)".

Machine learning and Causal models have gained a lot of attention in the recent past due to improvement in computational might. A causal forecasting method requires additional input data along with the historical demand values of the forecasted time series i.e. it works on a multivariate model and not univariate as employed by time series or Croston's Method. Causal forecasting methods are highly efficient it the dataset is non-stationary ("A time series non-stationary in the mean is where the data trend, or more generally, have a time-dependent mean" (Fildes et al., 2008)).

Classification of information is one of the most important components of decision-making tasks in a business. Classification methods are gaining increasing importance in business decision-making tasks and have become an integral part of organizational decision support systems (Kiang, 2003). The motivation for this project is to improve Croston's method of intermittent demand forecast by classifying the likelihood of the rare event demand occurrence.

The classification of demand occurrence will enable the business to control the inventory size and reduce safety stock by forecasting a dynamic reorder point just before the demand is expected which will overcome the limitations of Croston's method. This can help to save the cost of holding the inventory, as well as free up capital that is needed to procure the inventory; to be utilized for more pressing needs.

Chapter 5: Project Methodology

"Any intelligent fool can make things bigger and more complex. It takes a touch of genius –and a lot of courage – to move in the opposite direction." - Albert Einstein

Most Academicians and practitioners use advanced probabilistic models to forecast intermittent demand rather than forecasting unique incidence of "peaks-over-threshold time series". These methods typically aim to recreate the distribution of the underlying phenomena. (Nikolopoulos, 2020). This project aims to tackle the business problem from an analytic approach, utilizing new age analytical tools such as Machine Learning and Causal Modelling. The proposed method classifies the peaks in the demand to improve inventory control and eventually customer service levels as opposed to Croston's method that gives an average demand per period.

The intention is to highlight that the machine learning algorithm is more efficient and accurate than traditional forecasting method such as Croston's method for Intermittent Demand Forecast. The quantitative data of factors influencing the demand is taken and trained using Naive Bayes classifiers. The final objective is to see if machine learning methods can capture underlying factors in Intermittent Demand better than Statistical Methods.

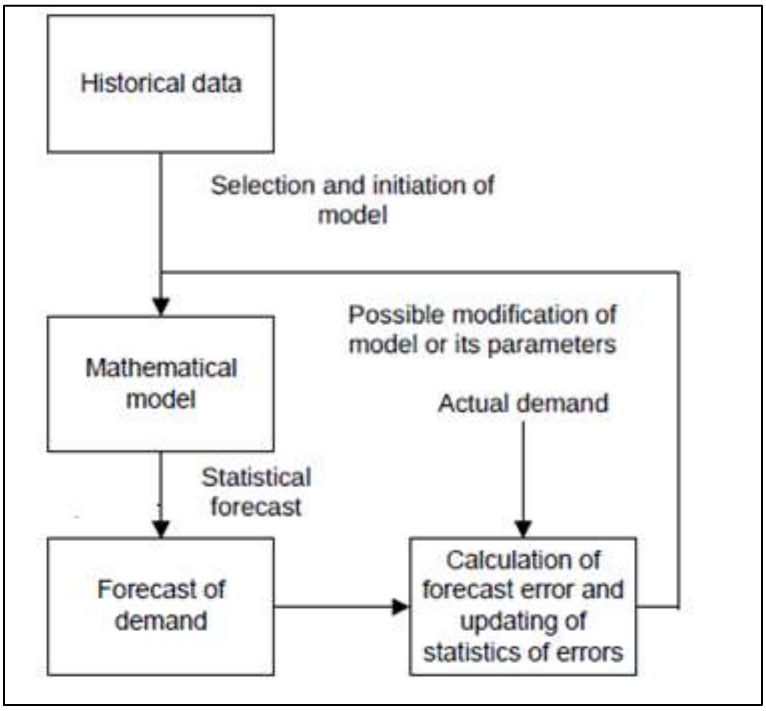

Figure 5.1 –Project Methodology

Chapter 6: Business Understanding

“A business that makes nothing but money is a poor business.” - Henry Ford

Any uncertainty in demand prediction raises the requirement of safety stock and thus ties up capital required to hold the inventory along with the cost of storage space. Most organizations can reduce their inventory size by using more efficient inventory control tools. The main purpose of an inventory control system is to determine the triggering point for stock replenishment (Axsater, 2006).

Inventory control becomes even more critical while dealing with intermittent demand components as the product would be lying for a very long duration before an order is received. Statistical time series has widely been used to predict intermittent demand forecast and multiple research has been done to improve its accuracy. Most of the progress made in this domain is also on statistical forecasting methods such as Croston Method and its variations that give the average demand per period and not the actual demand size or the demand period. However, the unique characteristics of intermittent demand (as compared to fast-moving or seasonal products) require special attention and also, these methods do not consider underlying factors. Since these methods cannot provide an accurate forecast quantity for the specific period organizations are required to holding the expected average demand quantity at all times and thus, entails to be capitally invested at all given time.

Businesses across the globe are targeting to maintain a high Customer Service Level (or just service level) to build reputation and customer loyalty. Customer Service Level a crutial Key Performance Indicators (KPIs) across industries and acts as a quantifiable measure of business effectively and performance. It is defined as the likelihood of being able to meet the customers’ demand and thus, it is can also be defined as the odds of avoiding lost sales (Simon & Joannes, 2014). Maintaining high levels of inventory is costly as buying or producing products require capital investment and also require space to store. In the current dynamically changing market robust solutions are needed to respond to customer demands; at the same time they effective techniques to decrease cost is also required.

Figure 6.1 – Inventory Control Tradeoff

All industries strive hard to achive high service levels, typically above 95%, however, achieving higher service levels is a classic case of diminishing returns where any additional inventory held, yields lesser returns, i.e. substantially lower stockouts being removed (Simon & Joannes, 2014). Optimizing Customer service levels to gain maximum returns is mostly convoluted. Inventory is replenished when in-hand inventory reaches a certain threshold known as reorder point. An effective and dynamic reorder point can help industries achieve higher service levels along with reducing time.

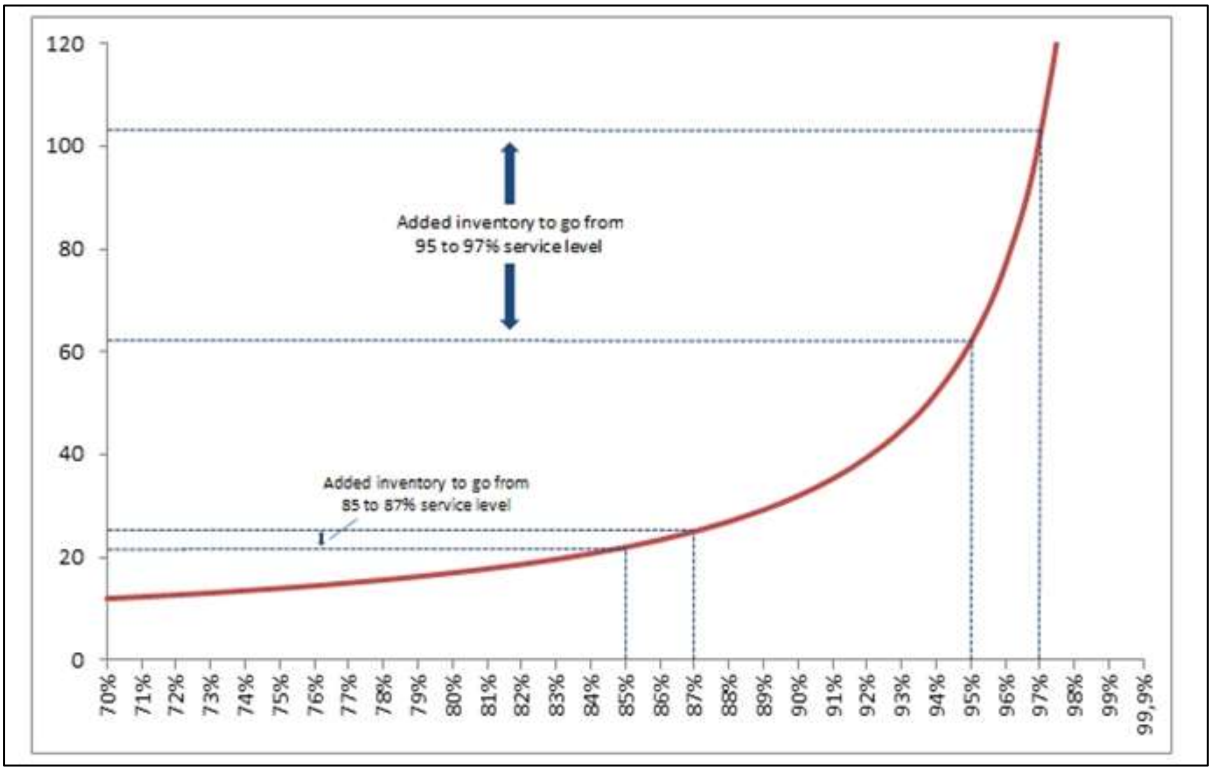

Figure 6.2 – Service Level Optimization Cost

Chapter 7: Data Understanding

“I deal in facts, not forecasting the future.” - Peter Lynch

The intermittent demand data consists for a period of 61 days of intermittent demand for unknown items and has 4 columns dayOfWeek, promo, marketing, Sales which are each explained in detail below.

Column	**Type**	**Explanation**
dayOfWeek	Numeric	The day of the week on which the sale is made
promo	Numeric	% of discount on the day of sale
marketing	Boolean	Marketing campaign performed or not
Sales	Numeric	Sale made on the day

Table 7.1 – Data Understanding

To gain more insight into the data we further explore the dataset
The intermittent demand shows the demand pattern as below.

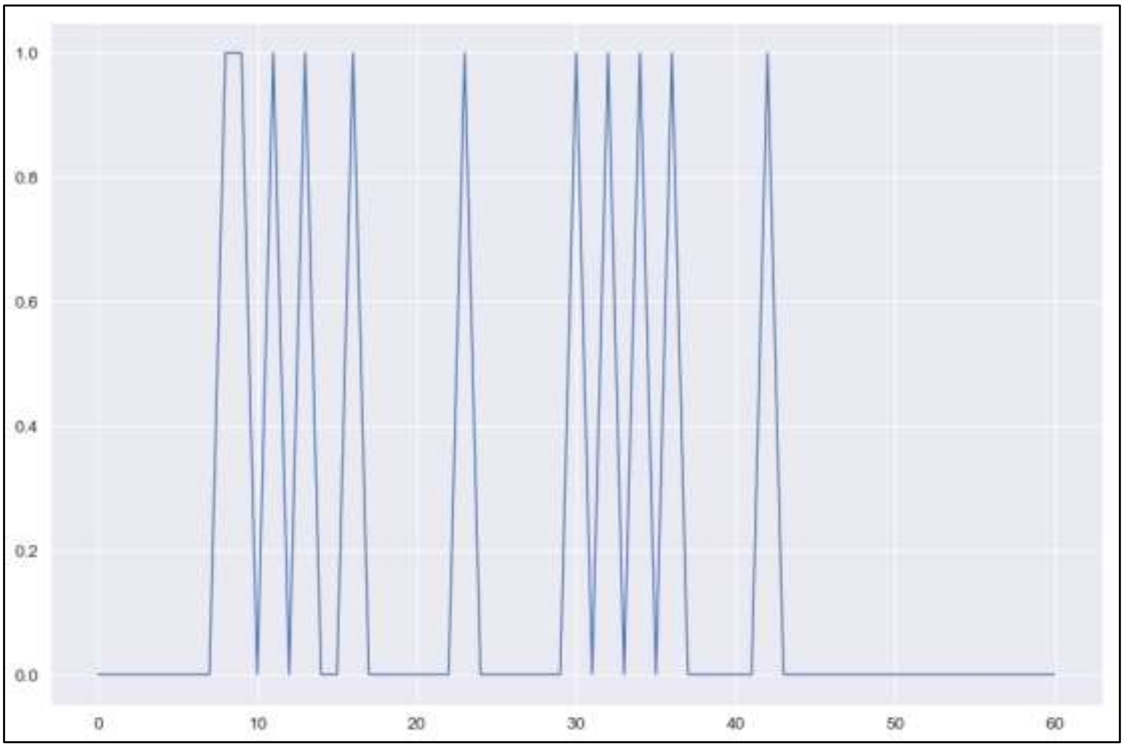

Figure 7.1 – Demand Pattern

The sales made on various days of the week are shown in the table below.

dayOfWeek	1	2	3	4	5	6	7
Cumulative Sales	0	2	0	2	1	0	4

Table 7.2 – Sales Pattern

The correlation matrix for various components is as below:-

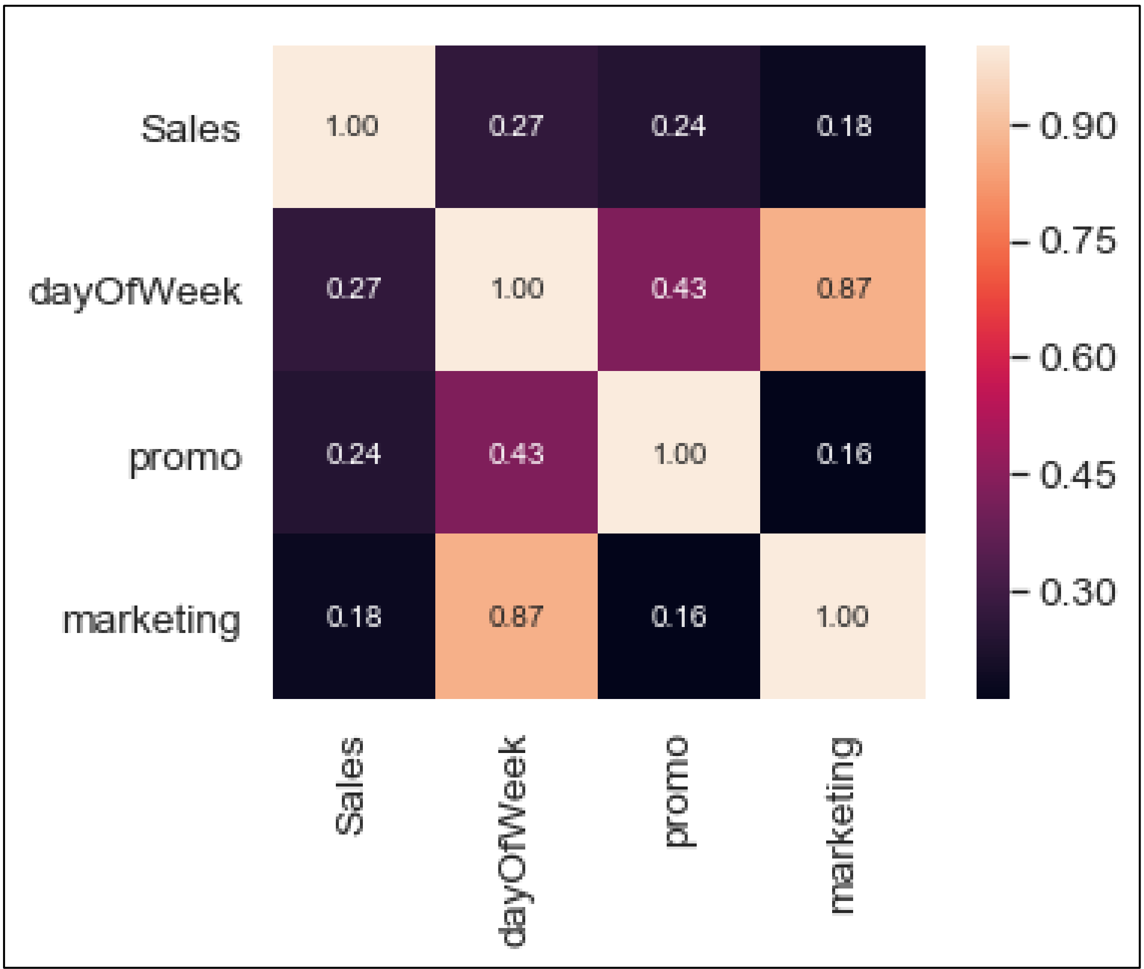

Figure 7.2 – Co-relation Matrix

Chapter 8: Data Preparation

“Forecasts usually tell us more about the forecaster than of the future.” - Warren Buffett

To improve forecast accuracy, additional features were created to capture the distinctive attribute of the intermittent demand forecasting problem, to help the algorithmic approaches learn the patterns better.

The following are features were added:

- Rolling window:- A rolling window with a window size of 7 days was created.
- Lag :- 7 lags of the demand were created (lag1 = Dt-1, lag 2 = Dt-2, lag 3 = Dt-3...and so on).
- Zero Cumulative:- The count of periods in between the non-zero demand periods (Hong et al., 2018).
- Rolling Mean:- Average for a window of data for 3 and 6 months.
- Demand Days: - Cumulative sum of demand rolling window of 7 days.
- No Demand Days:- 7 - Demand Days.

Note:-Naive Bayes, Linear Discriminant Analysis, and Tree-Based models are not affected by feature scaling as they are Not Distance-based, hence no scaling was used for the features.

Chapter 9: Data Modeling

"Most of us have roughly the same ability to predict the future. The trouble is, being right as often as the average forecaster won't produce superior results." - Howard Marks

Once additional features are added to the data we move ahead with the data modelling. The use of artificial intelligence in forecasting intermittent demand is because of its capability to deduce a non-linear process without requiring any distributional assumptions (Lolli et al., 2019). Machine learning can be used to provide dynamic forecast of the demand quantity, without presuming a demand rate to be constant in the future and are capable of encapsulating the relation between the non-zero demand and the inter-arrival rate of demand period, surpassing the limitations of Croston's method. Three Machine learning techniques are used to create models:- Naive Bayes classifiers, Support Vector Machine (SVM) Classifiers and Neural Network (NN) – Multi-Layer Perceptron (MLP).

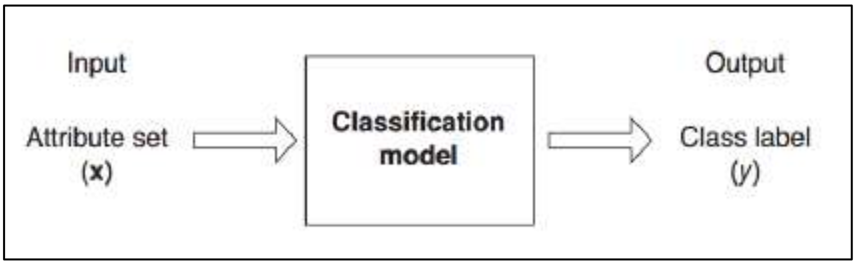

Figure 9.1 – Classification Model

Naive Bayes classifiers is a simple supervised learning algorithms using "probabilistic classifiers" that applies Bayes' theorem which has a very strong assumption of independence(naïve) between the features. Is is one of the most simple Bayesian network models (Wikipedia, n.d.)". Naive Bayes is very well liked due to its simplicity and ease of use and is one of the most extensively used machine learning algorithms. It sometimes, outperforms even highly sophisticated classification methods especially with less training data. It makes learning easy by presuming that the features are independent for a given class. Although assuming the features to be independeny is a very a poor assumption, in real life Naïve Bayes frequently comes up with better results than more sophisticated classifiers The beauty of this method lies in its incredible speed as compared to other classification algorithms.

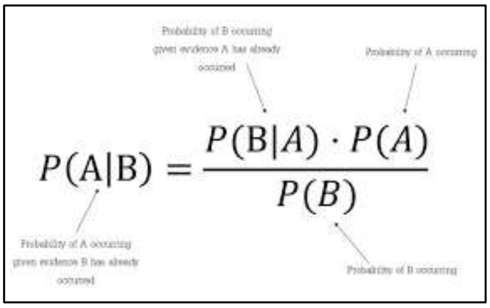

Figure 9.2 – Naïve Bayes Model

Support-vector machine or SVM was developed at AT&T Bell Laboratories by Vladimir Vapnik with colleagues. It is linear classifier that divides the a two-class features space using a hyperplane. The hyperplane differentiates the two-classes and distinctly classifies the data points. The hyperplane is designed in such a way that it has maximum margin i.e maximum distance from the data points of either classes. To create the hyperplane SVM uses a function called SVM kernel that converts an input space of low dimension, to a higher dimensional space which helps in converting a problem from non-separable to separable by addition of more dimensions. This improves SVM's very performance and accuracy.

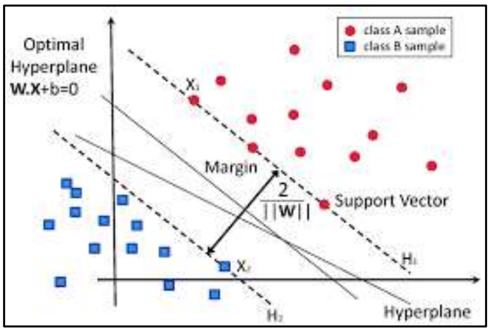

Figure 9.3 – SVM Model

"Multi-Layer Perceptron (MLP) is a class of feed-forward artificial neural network (ANN)". MLP consists of multiple perceptrons that are organized into layers. An MLP contains a mimimum of three layers that are nonlinearly-activating nodes: an input , hidden and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP works on a supervised learning technique called backpropagation that adjusts the weight and bias to reduce error during training. (Wikipedia, n.d.-b)

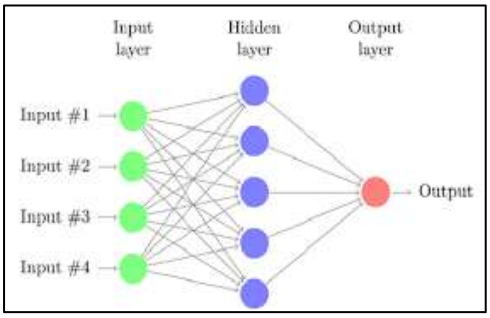

Figure 9.4 – NN-MLP Model

The dataset is split into train and test data sample. The training sample is used for training the model and the test sample is used to determine the reliability of the forecast. An 80-20 split is performed on the data set i.e 80% for training data and 20% for testing data.

Two different models are created each for Naive Bayes classifiers, SVM Classifiers and NN – MLP, one, without any feature engineering and a second with feature engineering.

Chapter 10: Data Evaluation

"Even the most serious efforts to make predictions can end up so far from the mark as to be more dangerous than useless." -Peter Bernstein

The forecast accuracy metrics of Demand Classification models are evaluated for the stock control implications. Customer Service level is one of the Key Performance Indicators for the Classification model i.e sufficient stock should be available to cover all demand. The following accuracy metrics are considered for evaluation of the model.

True-positive (TP) is the output in which the model correctly predicts the positive class i.e the model correctly predicts the demand period.

True-negative (TN) is the output in which the model correctly predicts the negative class i.e the model incorrectly predicts the no demand period.

False-positive (FP) is the output in which the model incorrectly predicts the positive class i.e the model incorrectly predicts the demand period to be no demand period.

False-negative (FN) is the output in which the model incorrectly predicts the negative class i.e the model incorrectly predicts the no demand period to be demand period.

Accuracy is the overall proportion of correct classifications i.e the fraction of the prediction that the classifier predicted accurately.

$$\text{Accuracy} = \frac{\text{Number of correct prediction}}{\text{Total Number of predictions}}$$

$$Accuracy = \frac{TP+TN}{\text{TP+FP+FN+TN}}$$

Precision is also known as Positive Predictive Value is the proportion of positive classification that were predicted correctly.

$$Precision = \frac{TP}{\text{TP+FP}}$$

Recall also known as the True Positive rate or Sensitivity is the proportion of actual positives that were predicted correctly.

$$Precision = \frac{TP}{\mathrm{TP+FN}}$$

F score sometimes called the F1 score is the harmonic mean of precision and recall and provides a more balanced measure of precision and recall

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

For Naive Bayes Classifier without feature engineering:-

```
13
Positive cases: 1
Negative cases: 12
Accuracy Score
0.7692307692307693
Precision/Recall Metrics
              precision    recall  f1-score   support

           0       0.91      0.83      0.87        12
           1       0.00      0.00      0.00         1

    accuracy                           0.77        13
   macro avg       0.45      0.42      0.43        13
weighted avg       0.84      0.77      0.80        13



[[10  2]
 [ 1  0]]
```

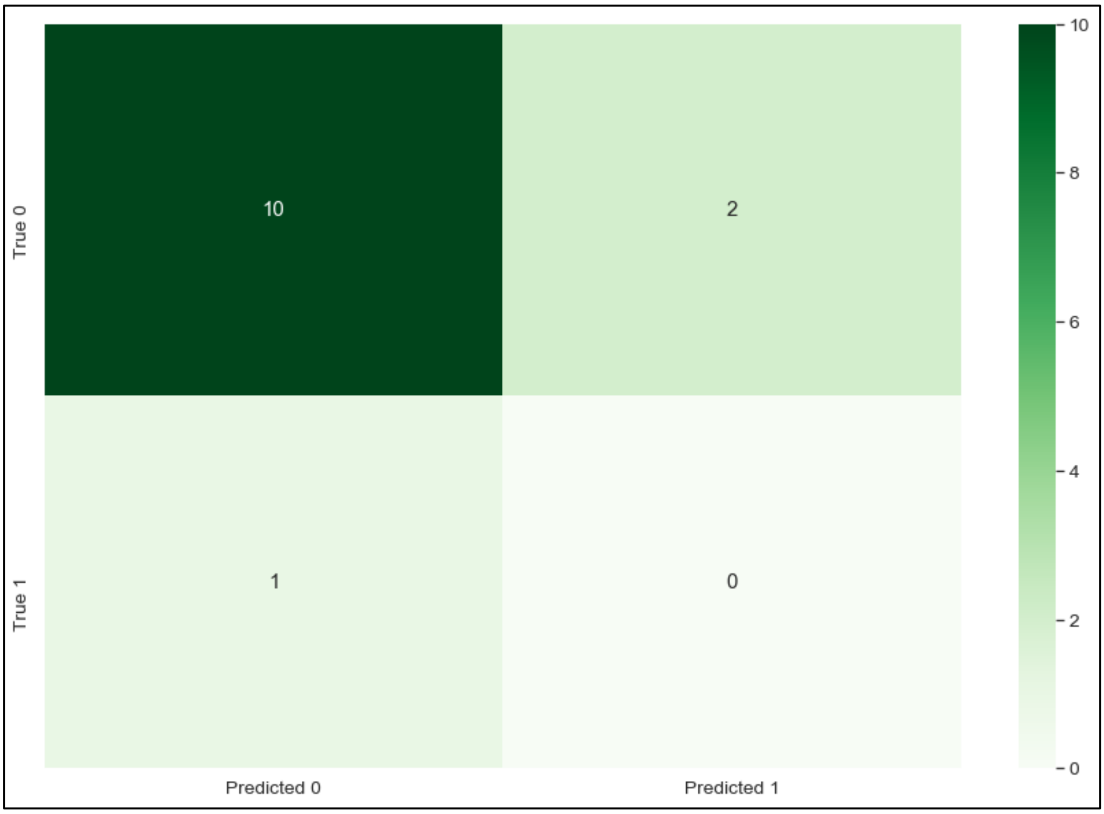

Figure 10.1 – Naïve Bayes Accuracy/Confusion Matrix without Feature Engineering

For Support Vector Classifier without feature engineering:-

```
13
Positive cases: 1
Negative cases: 12
Accuracy Score
0.9230769230769231
Precision/Recall Metrics
              precision    recall  f1-score   support

           0       0.92      1.00      0.96        12
           1       0.00      0.00      0.00         1

    accuracy                           0.92        13
   macro avg       0.46      0.50      0.48        13
weighted avg       0.85      0.92      0.89        13



[[12  0]
 [ 1  0]]
```

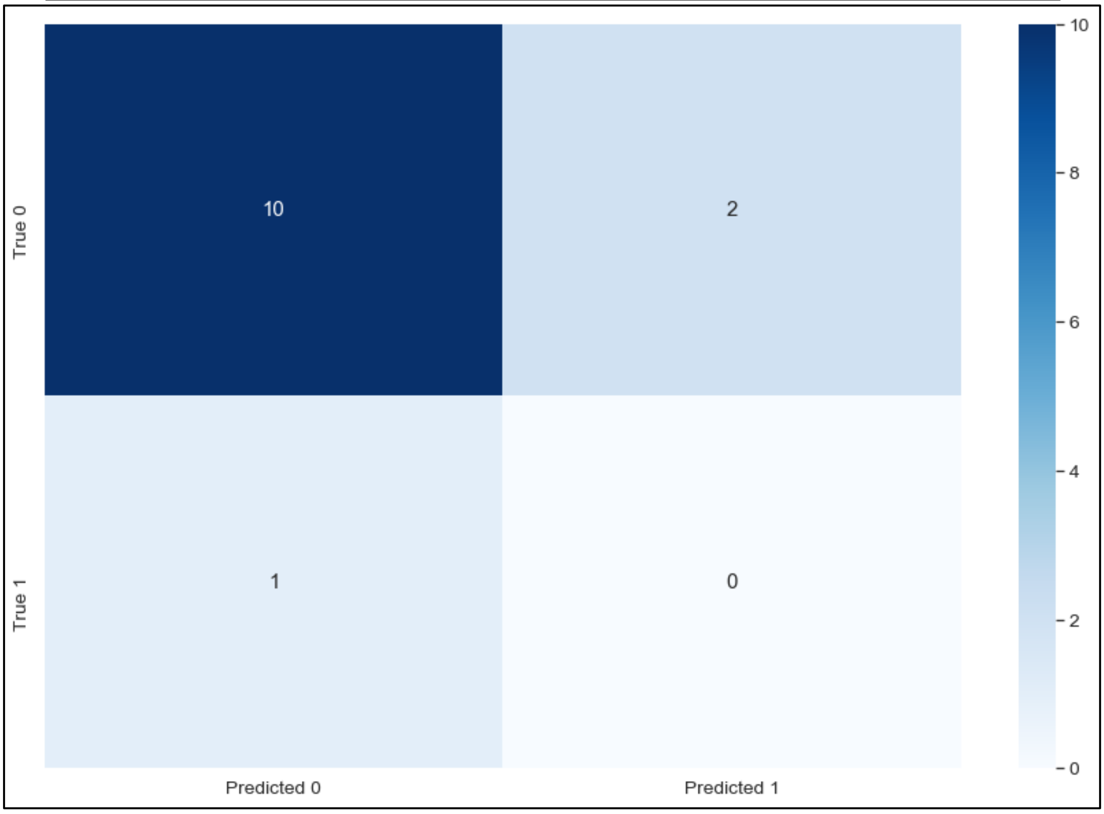

Figure 10.2 – SVM Accuracy/Confusion Matrix without Feature Engineering

For Neural Network – MLP without feature engineering:-

```
13
Positive cases: 1
Negative cases: 12
Accuracy Score
0.9230769230769231
Precision/Recall Metrics
              precision    recall  f1-score   support

           0       0.92      1.00      0.96        12
           1       0.00      0.00      0.00         1

    accuracy                           0.92        13
   macro avg       0.46      0.50      0.48        13
weighted avg       0.85      0.92      0.89        13



[[12  0]
 [ 1  0]]
```

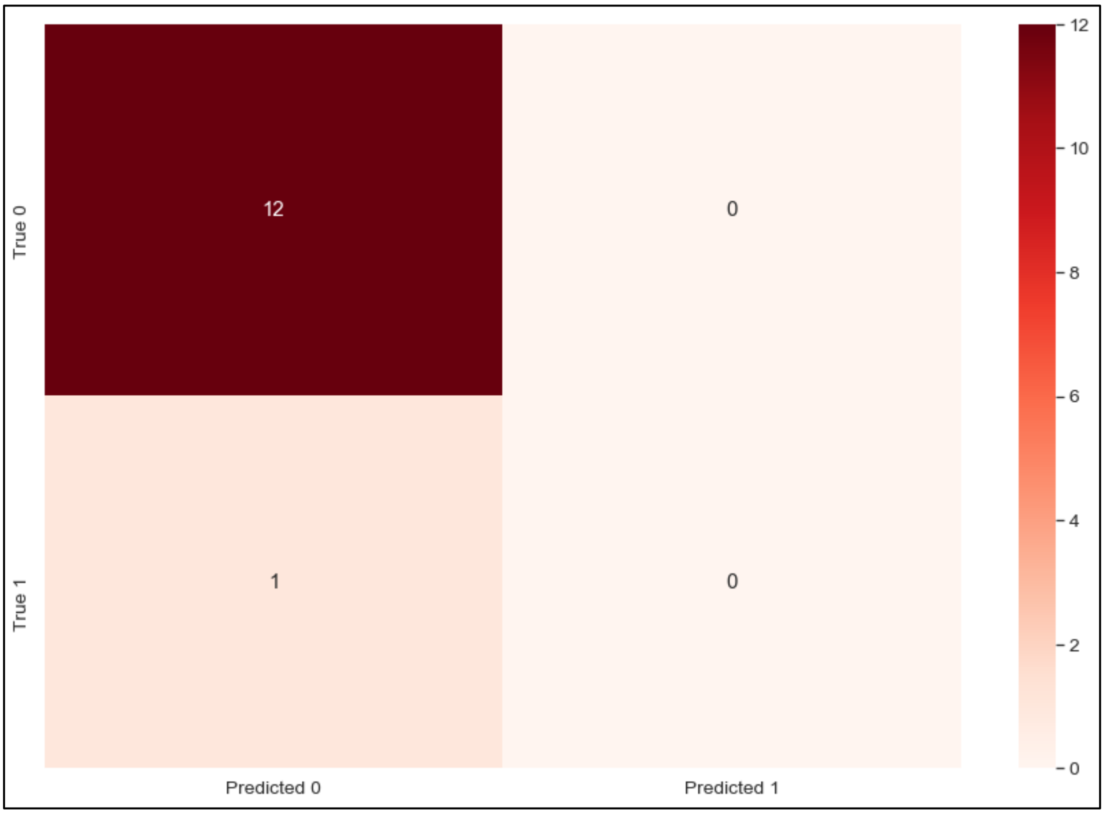

Figure 10.3 – NN-MLP Accuracy/Confusion Matrix without Feature Engineering

For Naive Bayes Classifier with feature engineering:-

```
13
Positive cases: 1
Negative cases: 12
Accuracy Score
0.9230769230769231
Precision/Recall Metrics
              precision    recall  f1-score   support

           0       1.00      0.92      0.96        12
           1       0.50      1.00      0.67         1

    accuracy                           0.92        13
   macro avg       0.75      0.96      0.81        13
weighted avg       0.96      0.92      0.93        13



[[11  1]
 [ 0  1]]
```

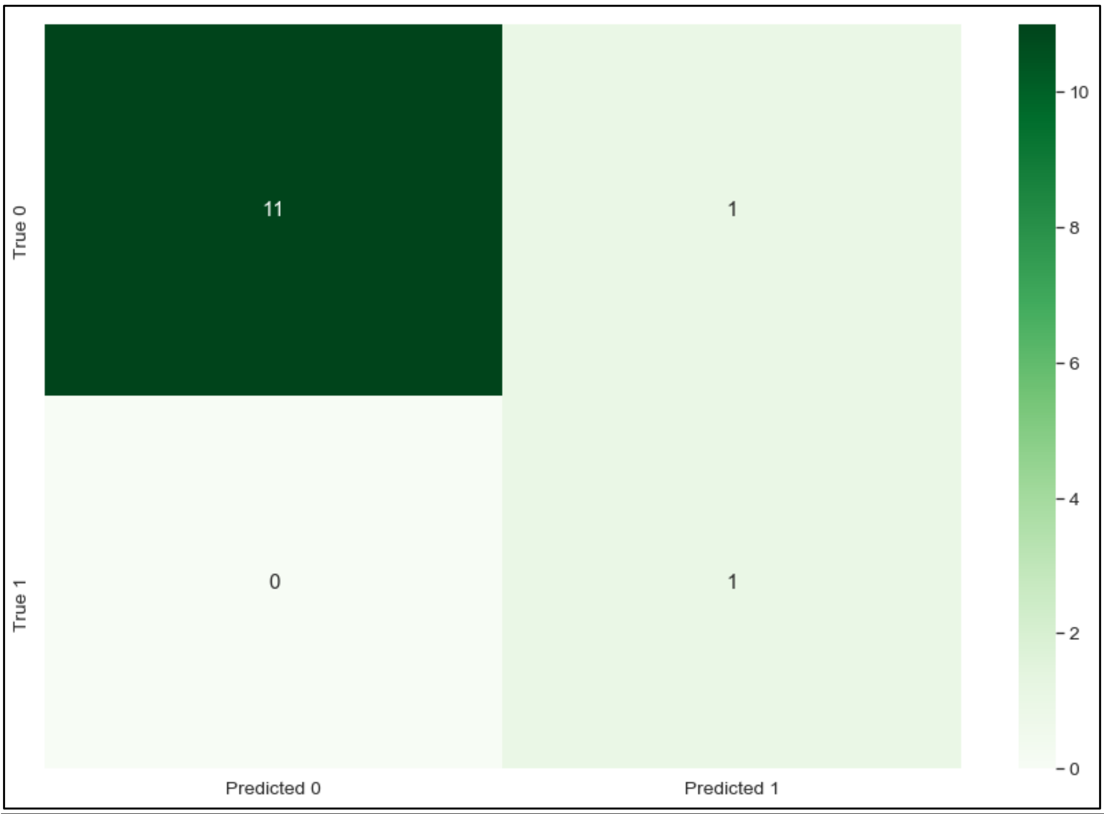

Figure 10.4 – Naïve Bayes Accuracy/Confusion Matrix with Feature Engineerin

For Support Vector Classifier with feature engineering:-

```
13
Positive cases: 1
Negative cases: 12
Accuracy Score
1.0
Precision/Recall Metrics
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        12
           1       1.00      1.00      1.00         1

    accuracy                           1.00        13
   macro avg       1.00      1.00      1.00        13
weighted avg       1.00      1.00      1.00        13



[[12  0]
 [ 0  1]]
```

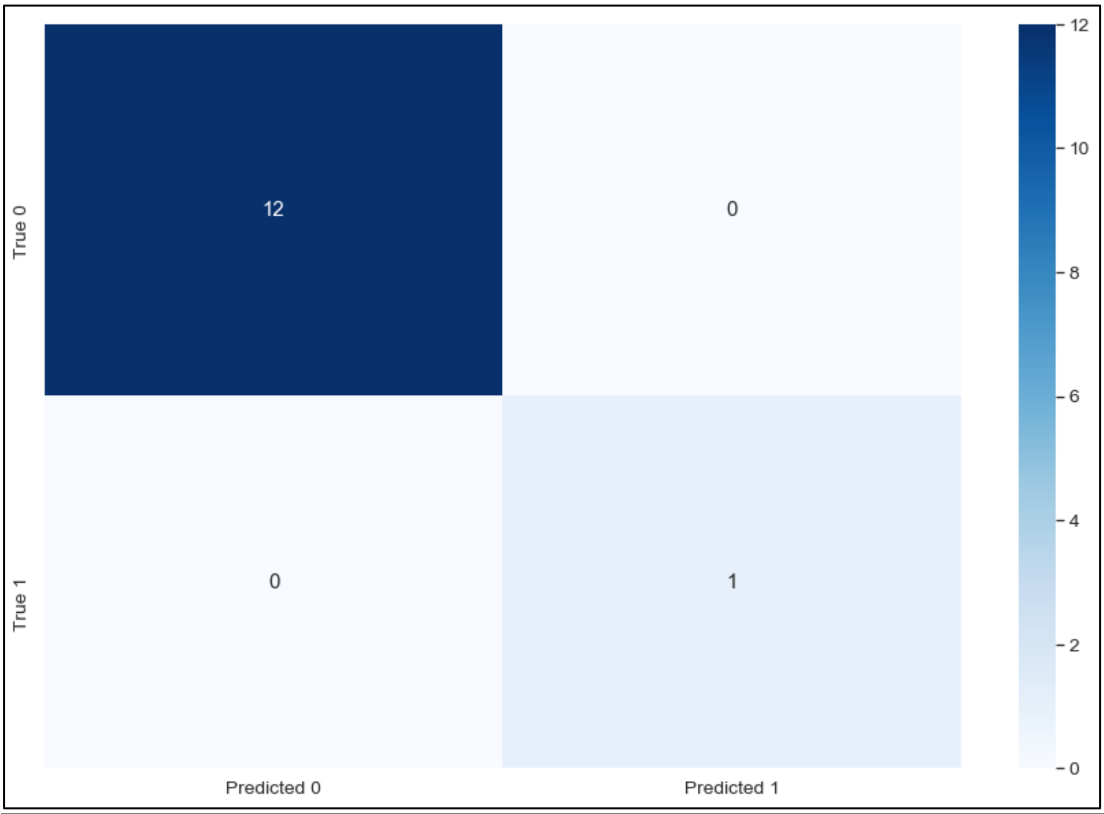

Figure 10.5 – SVM Accuracy/Confusion Matrix with Feature Engineering

For Neural Network – MLP with feature engineering:-

```
13
Positive cases: 1
Negative cases: 12
Accuracy Score
0.8461538461538461
Precision/Recall Metrics
              precision    recall  f1-score   support

           0       1.00      0.83      0.91        12
           1       0.33      1.00      0.50         1

    accuracy                           0.85        13
   macro avg       0.67      0.92      0.70        13
weighted avg       0.95      0.85      0.88        13



[[10  2]
 [ 0  1]]
```

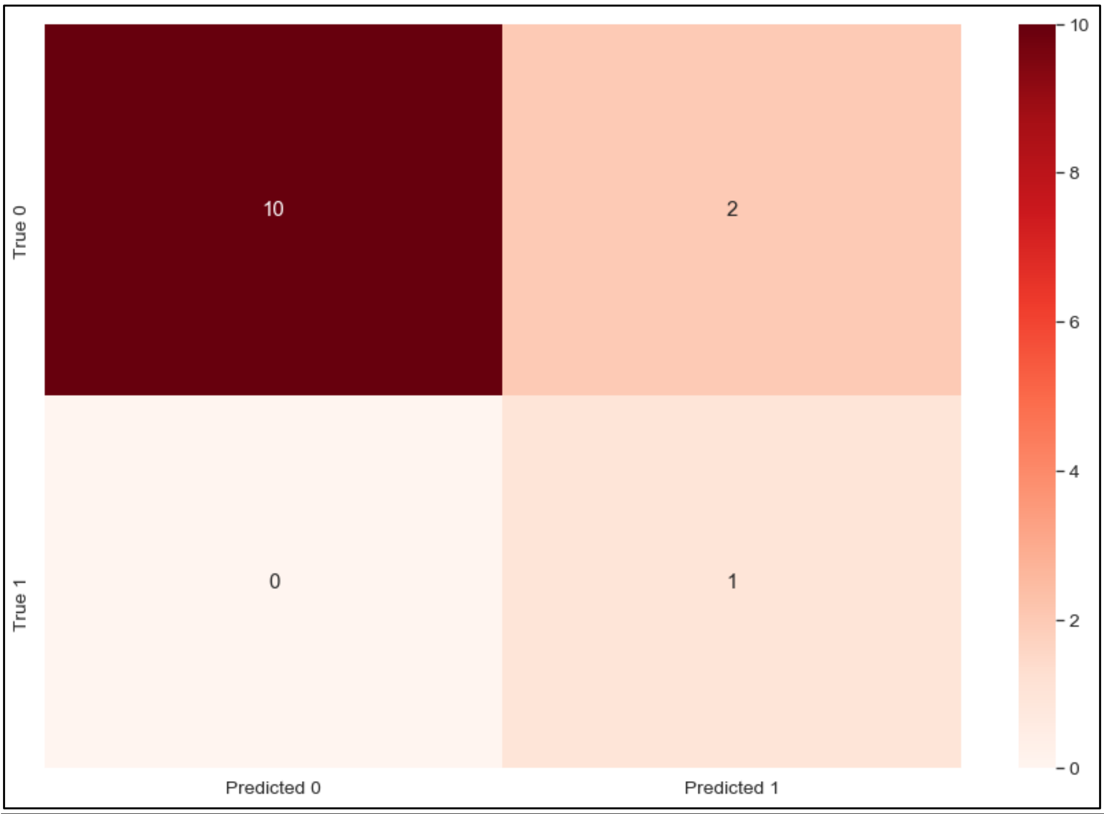

Figure 10.1 – NN-MLP Accuracy/Confusion Matrix without Feature Engineering

Chapter 11: Deployment

"Forecasts create the mirage that the future is knowable." - Peter L. Bernstein

The model has not been deployed yet and is currently being adapted to validate on a larger set of validation dataset to predict the classification of Intermittent demand components forecast for a large aerospace OEM engaged in design, manufacturing and MRO services for various systems and components for commercial, business and military aviation that include fixed and rotary wing aircrafts. The OEM also has presence in defence sector, space programes, airport infrastructure and several other industries in Road Transport, Marine and Critical Infrastructure.

Chapter 12: Analysis and Results

"Our expectations of the future are not unbiased and do not reflect all available information." -Peter Bernstein

Due to the randomness and nonlinearity of the intermittent demand distribution, traditional forecasting methods are not able to predict results with good accuracy. Traditional forecasting methods assume demand to be stationary whereas intermittent demand data is non-stationary due to periods with zero demand and variable demand sizes. Forecasting Intermittent Demand is both difficult as well as important for businesses to thrive. Machine Learning models such as Naïve Bayes, SVM, GBM, Neural Networks (NN's), etc.. have been widely used in the recent past to address the lack of effective forecasting methods in intermittent demand forecasting.

The goal of this project being to create a classification model to predict the inter-demand interval for intermittent demand more reliably and hence reduce average stock holding value whilst maintaining high Customer service levels. Most organizations are working tirelessly to reduce inventory costs by improving replenishment planning and to implementx methodologies like 'Just In Time'. The above classification can be used to predict the demand interval and replenishments can be ordered just before the demand is predicted.

Chapter 13: Conclusions and Recommendations for future work

"No prediction -- whether of a repetition of past patterns or a complete break with past patterns -- can be proved in advance to be right." - Benjamin Graham

This project focuses on the broader use of Machine Learning Models and Classification techniques precisely to address the intermittent demand forecasting problem for predict the special event of demand occurrence. The likelihood of demand occurrence is forecasted using classifiers which would be consecutively combined with the aggregated average for the demand period obtained from the conventional Croston's method, to give the demand size. The proposed method is easy, applicable, quick, and a robust substitute to more intricate Statistical Forecasting methods, that results in zero excessive stock and ensuring that the demand is met.

It is to the best of our knowledge that this proposition should intuitively appeal and extremely suitable for highly uncertain situations like intermittent demand forecasting. However, the limitation of this approach is that it relies on Croston's method for the 'demand per period' estimate to be aggregated for the actual demand size to be predicted.

.

Bibliography

Axsater, S. (2006). *Inventory Control - second edition*. https://link-springer-com.dianus.libr.tue.nl/content/pdf/10.1007%2F0-387-33331-2.pdf

Babai, M. Z., Syntetos, A., & Teunter, R. (2014). Intermittent demand forecasting: An empirical study on accuracy and the risk of obsolescence. *International Journal of Production Economics*, *157*(1), 212–219. https://doi.org/10.1016/j.ijpe.2014.08.019

Croston, J. D. (1972). *Forecasting and Stock Control of Intermittent Demand*.

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A. A. (2008). *Forecasting and operational research : a review*. 1150–1172. https://doi.org/10.1057/palgrave.jors.2602597

Flowspace. (n.d.). *No Title*. https://www.flow.space/blog/inventory-carrying-costs/

Hong, Y., Zhou, J., & Lanham, M. A. (2018). *Forecasting Intermittent Demand Patterns with Time Series and Machine Learning Methodologies*.

Industry 4.0: the fourth industrial revolution – guide to Industrie 4.0. (n.d.). https://www.i-scoop.eu/industry-4-0/#:~:text=Industry 4.0 is the current, called a “smart factory”.

Kiang, M. Y. (2003). A comparative assessment of classification methods. *Decision Support Systems*, *35*(4), 441–454. https://doi.org/10.1016/S0167-9236(02)00110-0

Lolli, F., Balugani, E., Ishizaka, A., Gamberini, R., Rimini, B., & Regattieri, A. (2019). Machine learning for multi-criteria inventory classification applied to intermittent demand. *Production Planning and Control*, *30*(1), 76–89. https://doi.org/10.1080/09537287.2018.1525506

Nikolopoulos, K. (2020). We need to talk about intermittent demand forecasting. *European Journal of Operational Research*, *xxxx*, 1–11. https://doi.org/10.1016/j.ejor.2019.12.046

Robert, F., & Makridakis Spyros. (1995). *The Impact of Empirical Accuracy Studies On Time Series Analysis and Forecasting*.

Simon, S., & Joannes, V. (2014). *No Title*. https://www.lokad.com/service-level-definition#:~:text=The service level can also, usually not a feasible option.

Syntetos, A. A., & Boylan, J. E. (2001). *On the bias of intermittent demand estimates*. *71*, 457–466.

Syntetos, A. A., Boylan, J. E., & Croston, J. D. (2005). On the categorization of demand patterns. *Journal of the Operational Research Society*, *56*(5), 495–503. https://doi.org/10.1057/palgrave.jors.2601841

Teunter, R. H., & Duncan, L. (2009). *Forecasting intermittent demand : a comparative study*. 321–329. https://doi.org/10.1057/palgrave.jors.2602569

Varghese, V., & Rossetti, M. (2008). A classification approach for selecting forecasting techniques for intermittent demand. *IIE Annual Conference and Expo 2008*, 863–868.

Visser, S. T. (2017). *Forecasting the intermittent and slow-moving demand of spare parts Forecasting the intermittent and slow-moving demand of spare parts Master of Science*.

Waller, D. (1956). *Methods for Intermittent Demand Forecasting*. 1–6.

Wikipedia. (n.d.-a). *No Title*. https://en.wikipedia.org/wiki/Naive_Bayes_classifier

Wikipedia. (n.d.-b). *No Title*. https://en.wikipedia.org/wiki/Multilayer_perceptron

Williams, T. M. (1984). Stock Control with Sporadic and Slow-Moving Demand. *Journal of the Operational Research Society*. https://link.springer.com/article/10.1057/jors.1984.185

Appendix

Plagiarism Report[1]

Intermittent Demand Classification using Machine Learning

ORIGINALITY REPORT

13%	10%	3%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	ukcatalogue.oup.com Internet Source	2%
2	Submitted to Sogang University Student Paper	1%
3	pure.tue.nl Internet Source	1%
4	bucks.collections.crest.ac.uk Internet Source	1%
5	matthewalanham.com Internet Source	1%
6	core.ac.uk Internet Source	1%
7	Kunal Rajput, Girija Chetty, Rachel Davey. "PHIs (Protected Health Information) identification from free text clinical records based on machine learning", 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2017 Publication	1%

[1] Turnitn report from the University is attached.

8	www.tandfonline.com Internet Source	<1%
9	en.m.wikipedia.org Internet Source	<1%
10	www.coursehero.com Internet Source	<1%
11	Submitted to University of Stellenbosch, South Africa Student Paper	<1%
12	etd.ohiolink.edu Internet Source	<1%
13	Submitted to City University Student Paper	<1%
14	aaltodoc.aalto.fi Internet Source	<1%
15	Submitted to Deakin University Student Paper	<1%
16	Konstantinos Nikolopoulos. "We need to talk about intermittent demand forecasting", European Journal of Operational Research, 2021 Publication	<1%
17	Submitted to Regional Centre for Biotechnology Student Paper	<1%

18	hdl.handle.net Internet Source	<1%
19	lutpub.lut.fi Internet Source	<1%
20	www.geeksforgeeks.org Internet Source	<1%
21	Submitted to CITY College, Affiliated Institute of the University of Sheffield Student Paper	<1%
22	Submitted to University of Westminster Student Paper	<1%
23	academic.oup.com Internet Source	<1%
24	orca.cf.ac.uk Internet Source	<1%
25	usir.salford.ac.uk Internet Source	<1%
26	www.inmybangalore.com Internet Source	<1%
27	Submitted to Institute of Management Technology Student Paper	<1%
28	Submitted to Visvesvaraya Technological University, Belagavi	<1%

Student Paper

29	Submitted to University of Leeds Student Paper	<1%
30	archive.org Internet Source	<1%
31	Submitted to Cardiff University Student Paper	<1%
32	essay.utwente.nl Internet Source	<1%
33	medium.com Internet Source	<1%
34	scholar.sun.ac.za Internet Source	<1%
35	www.pfu.co.jp Internet Source	<1%
36	res.mdpi.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		