True Chain

An Engineering Project in Community Service

Phase – I Report

Submitted by

1. **20BCE10015 Akhil**
2. **20BCE10082 Harsh Soni**
3. **20BCE10227 Shivam Saurabh**
4. **20BCE10128 Gandhi Monil**
5. **20BCE11104 Shivansh Rastogi**
6. **20BOE10072 Yashi Goswami**
7. **20MEI10039 Narendra Kumar**
8. **20MIM10100 Harshvardhan Saxena**

in partial fulfillment of the requirements for the degree of

Bachelor of Engineering and Technology

VIT Bhopal University
Bhopal
Madhyapradhesh

December, 2022

Bonafide Certificate

Certified that this project report titled **"True Chain"** is the bonafide work of "20BCE10015 Akhil, 20BCE10082 Harsh Soni, 20BCE10227 Shivam Saurabh, 20BCE11104 Shivansh Rastogi, 20BCE10128 Gandhi Monil, 20BOE10072 Yashi Goswami, 20MEI10039 Narendra Kumar , 20MIM10100 Harshvardhan Saxena" who carried out the project work under my supervision.

This project report (Phase I) is submitted for the Project Viva-Voce examination held on 26-12-2022.

Supervisor

Dr. Soumya Sankar Ghosh

1. INTRODUCTION -

Supply chain forecasting refers to the process of predicting demand, supply or pricing for a product or a range of products in a particular industry.
For example, the algorithms behind a forecasting model can look at data from suppliers and customers and forecast the price of a product. The algorithm can also examine external factors, such as weather or other disruptive events, to further increase the precision of the pricing forecast.

True Chain will help user to have clear visibility into the supply chain during peak demand seasons like Diwali and Christmas, to ensure that products were delivered to retailers on time. It helps user to predict events that could cause bottlenecks and accordingly plan for mitigating the impact of such situations

1.1 Motivation

During the VIT Bhopal experiential lerning industrial visit we came across various problem statements related to suppy chain forecasting. Correct forecasting is important from a business perspective. From cutting costs to keeping consumers happy, forecasting is a vital component of supply chain management, helping companies fill orders on time, avoid unnecessary inventory expenses and plan for price fluctuations. If our forecast assumed fewer products sold than the actual demand then we lose profit. On the other hand, if our forecast was overestimated, so we actually sell fewer products, we also lose, we incur the cost of product stock, or worse, our products expire. One bad forecast has a business, production and logistical consequences.

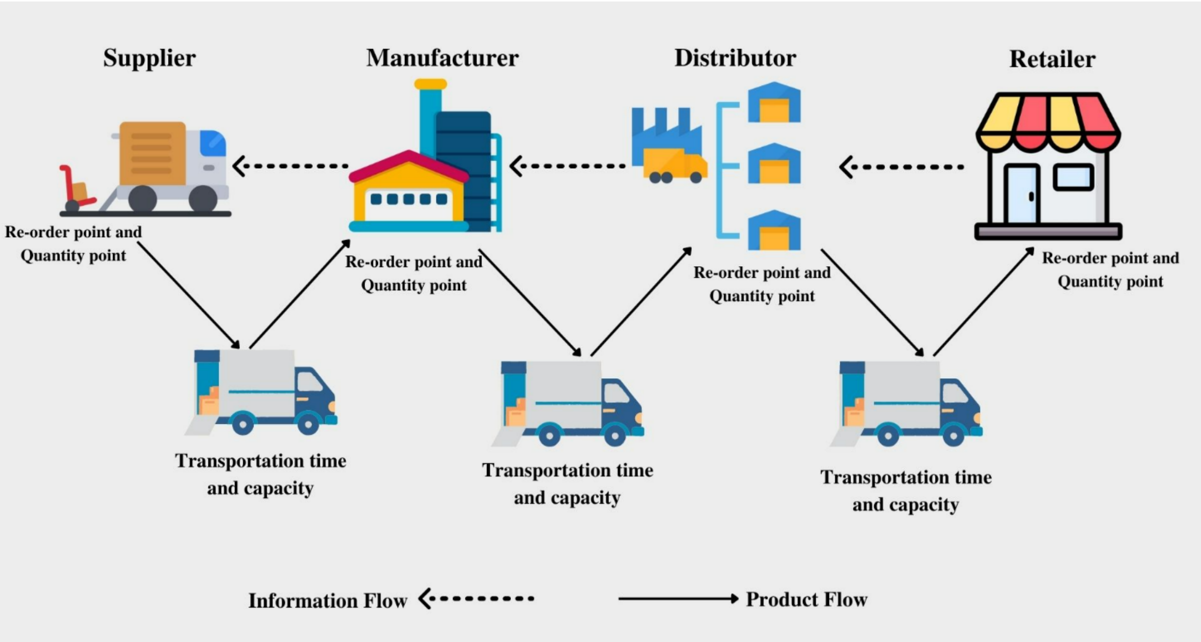
.

1.2 Objective

- Create new features based on the existing ones that enable machine learning algorithms to work. Getting the data right increases the quality of the data a lot because we transform the raw data into information that the model will transform into knowledge. Then the model can understand the pattern better and make better predictions.
- The user will be able to get both monthly and annual predictions regarding supply and demand.
- The objective of this project is to perform a comparative analysis of forecasting the distorted demand signals in the extended supply chain using non-linear machine learning techniques. More specifically, the work focuses on forecasting the demand at the upstream end of the supply chain to avoid the bullwhip effect.

2. Existing Work / Literature Review

We have studied the research papers and understood the steps involved in the supply chain (SC) management process, as shown below:



Here we can see the information and material flow between different phases. We will predict the product demand between retailer-distributor or distributor-producer or producer-supplier.

Demand Signal - A demand signal is a message issued within business operations or within a supply chain to notify a supplier that goods are required, and is, therefore, a key item of information for demand planners within a business.

Factors affecting supply chain performance

• Supply chain structure -

The number of facilities, the number of stages, and the structure of the material and information flow contribute to the complexity of the chain.

• Inventory control policy -

Inventory control means ensuring that the business has the right goods on hand to avoid stock-outs, prevent shrinkage, and provide proper accounting. There must be an economic balance between the costs incurred and the costs saved by holding the material in stock. The inventory control mechanism involves decisions regarding 'when' and 'how much to order.

• Information Sharing -

Information connects various SC partners and allows them to coordinate activities. Information is crucial to the daily operations at each stage of the SC. An information system can enable a firm to get a high variety of customized products to customers rapidly and to understand the changing customer's tastes and preferences.

• Customer demand -

Customer demand pattern is one of the environmental factors affecting the performance of the SC. In most industrial contexts, demand is uncertain and hard to forecast. When customer demand is wildly fluctuating, the member in an SC sends a highly variable order pattern to the associated member in the upper stage, which may cause amplification of order variance (which is termed as bullwhip effect) through the SC. Customer demand volatility also results in high capacity and inventory costs for the manufacturer.

• Forecasting method -

In a supply chain, the members need to forecast future demand, and it is impossible to predict demand with certainty. This uncertainty will result in distorted order quantity and via order variance amplification. The accuracy of the forecast highly influences the supply chain performance measures such as inventory cost, backorder cost, lost sales cost, and customer goodwill.

An inaccurate forecast results in the underutilization of the factory's capacity. It is identified as one of the main causes of the bullwhip effect in the supply chain is the use of demand forecasting.

• Lead time -

The time gap between the receipts of the order and to delivery of the product is referred to as lead time, which is the sum of the order lead time and delivery lead time

Under long lead time, the fluctuation in orders is more, which may result in a bullwhip effect.

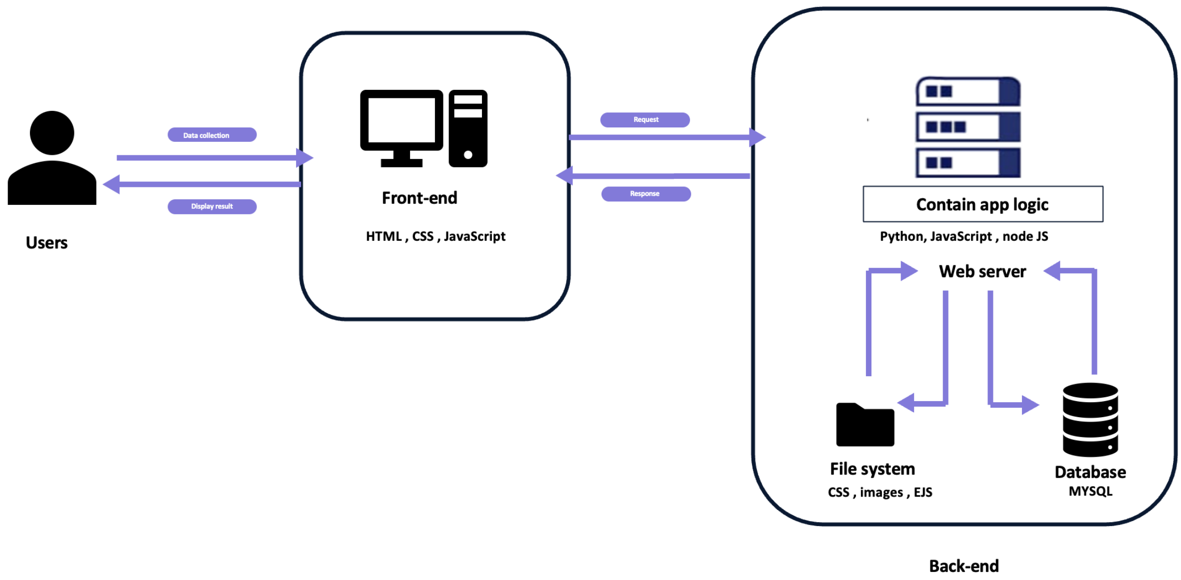
Sell-In / Sell-Out Model

Sales can be divided into Sell In and Sell Out.

Sell In is the number of products the manufacturer sells to the retailer, while Sell Out is the number of products sold by the retailer to the end customers.

Retailers will not sell products if we do not supply them. Moreover, if they want to sell more because they are planning a promotion for our product, they also have to order more products from us.

A) System Design / Architecture



B) Working Principle

We are trying to implement a machine learning model to examine the future statistics of any supply chain with the help of previous datasets.

We will calculate/predict the approximate requirement that is going to be needed by the supplier(upstream) before the stock outage. So, the supplier can place orders in advance to minimize the assumed demand variability. Suppose a retailer wants to order any product from a wholesaler, then the model will accordingly calculate the quantity depending upon previous records considering peak demands like festive or sale seasons.

How it will be implemented

To do so we are going to need a dataset of any supply chain.Then we train the data using different ML models and find the most accurate one. Some models are LSTM, Prophet, etc.

Approach

An accurate prediction for sales shortly can help managers to have a good plan for stocking, enhancing economic efficiency, and optimizing the business of the company.

Detecting accurate anomalies in sales enables the company to have an insight into its operating and marketing strategies. A negative anomaly in sales may correspond to not good strategies in marketing, leading to a decrease in sales. The strategies need to be reviewed and adjusted. By contrast, once a positive anomaly is detected from the model, it could be useful to investigate and explain the reason, thereby increasing sales and having appropriate strategies for the future.

We propose to solve the performance bottleneck in the SC on the Manufacturer's side by working on the following two aspects of the SC -

• SC Forecasting -

In this Part, we propose to work on the various traditional techniques of SC forecasting. And then extend our work towards using the new and current Industry leading Machine Learning based Methodologies of Forecasting.

We propose to provide a comprehensive Exploratory Data Analysis and Comparative Analysis of the above models.

• **Anomaly Detection in Sales and Demand Data -**

In this Part, we propose to use Machine Learning Techniques to build Anomaly Detection System over the Supply Chain data.

For SC Forecasting we will use the historical data from a manufacturing company that produces a variety of products and ships them to 4 warehouses.

The warehouses are placed in different locations all over the world Thus, it normally takes months to ship products to the warehouses.

This Dataset is perfect to analyse the Supply chain event in the upper part of the Supply Chain Model by establishing forecasting models that can predict the Events well in advance to give the manufacturer insight into the future demand signal.

A reasonably accurate forecast of the demand signal of the coming months will be highly beneficial for the Manufacturer.

The demand signal to the warehouse and the average Lead Time that scales to months show a strong correlation to the scenario that can be affected by the Bullwhip effect & other Factors that affect the Supply Chain Performance.

Analysis of the Dataset

- The first dataset will be collected from various archives which will be analyzed as Time-series data for forecasting and the second one will be the generated datasets for detecting an anomaly

- The training data and validation data will be prepared and data from the training set will be fed to train the models. In the training process, the number of cells, dropouts, and the learning rate of the model will be optimized.

- Comparative analysis will be performed on various models from traditional techniques like LSTM model to Prophet model of Forecasting and . This will help the client make decisions by providing historical trends in the company's sales data.

- Prophet provides several tools and techniques for detecting anomalies in time series data. By combining these techniques, we will develop a robust anomaly detection system that can identify and alert you to potential issues in our data.

C) Results

Exploratory data analysis-
[link]

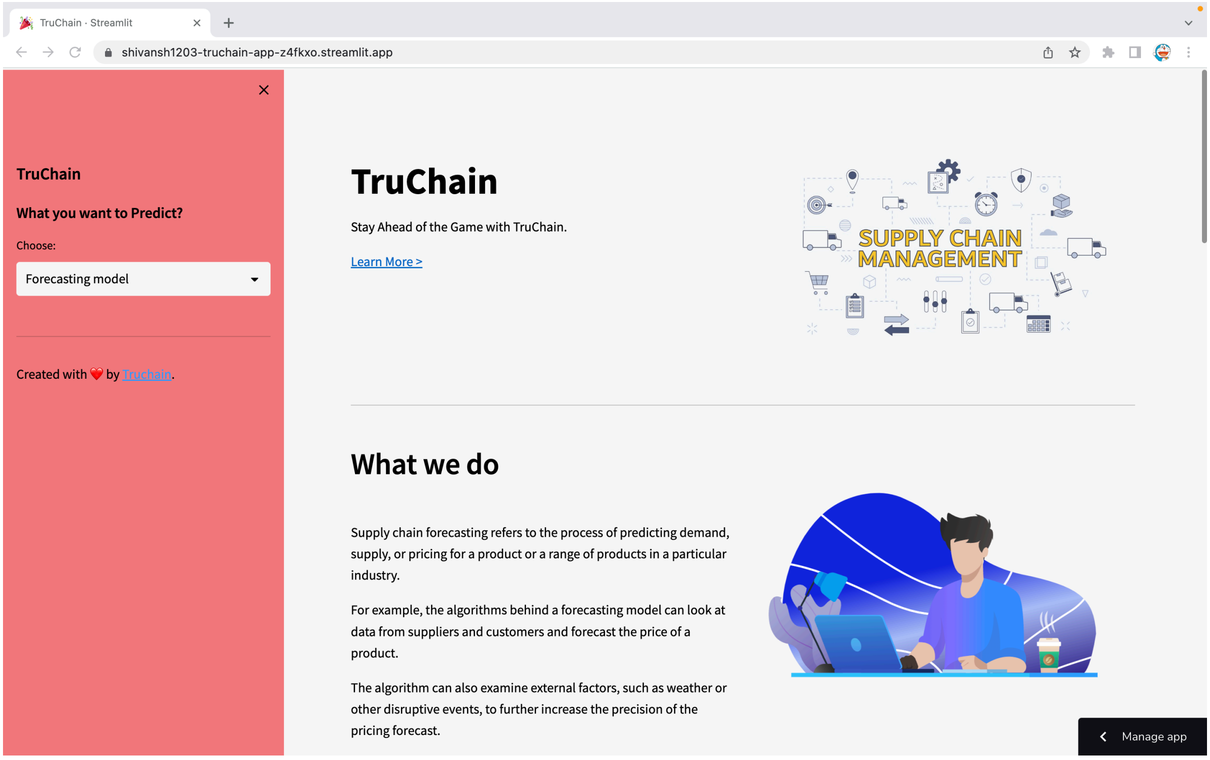
- There are a total of 4 categorical features, namely Season, is_holiday, is_weekend, and weather_code, and 7 numerical/continuous features along with 1 datetime object, resulting in a total of 12 independent features with 10,886 rows. No missing data or duplicated rows are present in the dataset.
- The data covers a time period of almost 2 years, ranging from 1st January 2011 to 19th December 2012. During the months of September and October, the maximum number of bikes are rented. However, the bike rental count is less during the cold seasons of winter (November, December, January, February) when people prefer not to ride bikes due to the cold.
- The month-wise bar plot shows that the demand for bikes at the beginning of the dataset is quite low compared to months from March 2012 onwards. There's a drop in the middle of the dataset due to the cold and winter season. Outliers are present in windspeed and casual users, indicating that windspeed is not uniform, while the count of casual users varies as they are not registered and not serious about riding bikes.
- The exponential decay curve for the count (registered and non-registered users combined) indicates that as the number of users renting bikes increases, the frequency decreases. For weather_code, in the fourth category, i.e., Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog, the number of users renting bikes is much lower. Hence it's better to drop this feature while conducting further tests.

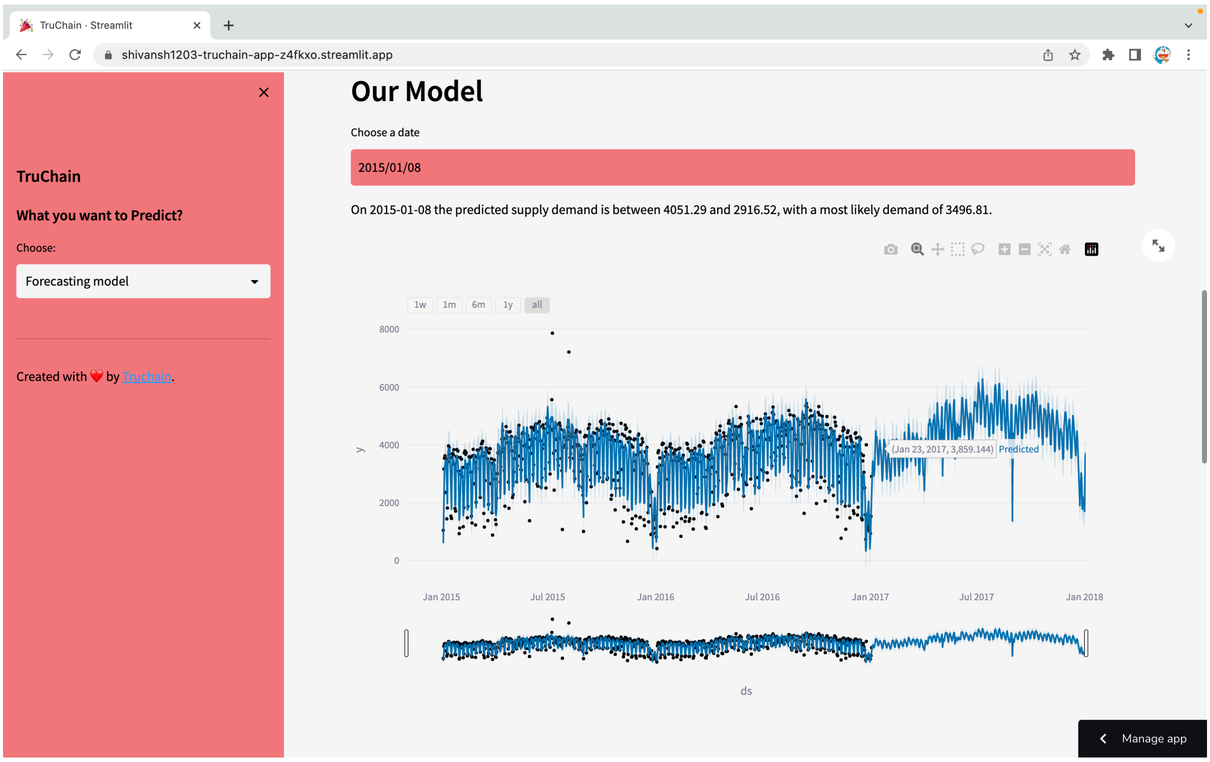
- The count for bikes rented on working days is much higher than on non-working days. During holidays, people do not prefer to ride bikes. However, when the weather_code is Clear with Few clouds, people tend to rent more bikes for their commute.
- During spring, summer, fall, and winter, the count is more or less equal for the users renting bikes. The registered user count has a higher correlation with the count as compared to the casual user count. The windspeed and season have a very low (near zero) positive correlation with the count, indicating that windspeed and season do not have a significant effect on the demand for rented bikes.
- The temperature and user-specific feeling of heat/cold have a moderate correlation (0.3) with the count. People tend to go out on bright sunny days when the temperature is normal. However, during harsh conditions such as too hot or too cold, the demand for bikes has seen a considerable dip. Casual users who rent bikes prefer to ride when the temperature is suitable.
- During holidays, the user count has seen a considerable dip, whereas, on working days, the count is normal.

Our Model-

- Based on the EDA analysis, it can be concluded that the demand for bike rentals varies significantly based on seasonal factors such as weather, holidays, and time of year. The data contains 12 independent features and 10,886 rows, with no missing data, null values, or duplicated rows. Outliers were observed in the windspeed and casual user features, indicating that these factors do not have a uniform effect on bike rental demand.
- Prophet, a time series forecasting model, is a suitable choice for predicting bike rental demand due to its ability to capture seasonality and holidays. It also allows for the inclusion of external regressors such as weather codes and temperature. The model is based on an additive regression model with four main components: trend, seasonality, holidays, and regressors.
- Compared to a neural model like LSTM, Prophet has several advantages. First, Prophet is much easier to implement and requires minimal tuning, whereas LSTM requires complex hyperparameter tuning and model selection. Second, Prophet is faster to train and provides better interpretability of results, whereas LSTM requires a lot of computational power and can be challenging to interpret. Finally, Prophet's built-in feature selection and handling of missing values make it more suitable for time series forecasting tasks like bike rental demand prediction.
- In summary, based on the EDA analysis, the use of Prophet with seasonal regressors is justified for predicting bike rental demand. Prophet's ability to capture seasonality and holidays and handle external regressors makes it a superior choice over neural models like LSTM for this particular task.

UI/UX (tenta)-


TruChain · Streamlit
shivansh1203-truchain-app-z4fkxo.streamlit.app
TruChain
What you want to Predict?
Choose:
Forecasting model
Created with ❤️ by Truchain.
TruChain
Stay Ahead of the Game with TruChain.
Learn More >
SUPPLY CHAIN MANAGEMENT
What we do
Supply chain forecasting refers to the process of predicting demand, supply, or pricing for a product or a range of products in a particular industry.
For example, the algorithms behind a forecasting model can look at data from suppliers and customers and forecast the price of a product.
The algorithm can also examine external factors, such as weather or other disruptive events, to further increase the precision of the pricing forecast.
Manage app


TruChain · Streamlit
shivansh1203-truchain-app-z4fkxo.streamlit.app
TruChain
What you want to Predict?
Choose:
Forecasting model
Created with ❤️ by Truchain.
Our Model
Choose a date
2015/01/08
On 2015-01-08 the predicted supply demand is between 4051.29 and 2916.52, with a most likely demand of 3496.81.
1w
1m
6m
1y
all
8000
6000
4000
2000
0
y
(Jan 23, 2017, 3,859.144) Predicted
Jan 2015
Jul 2015
Jan 2016
Jul 2016
Jan 2017
Jul 2017
Jan 2018
ds
Manage app

TruChain · Streamlit

shivansh1203-truchain-app-z4fkxo.streamlit.app

TruChain

What you want to Predict?

Choose:

Forecasting model

Created with ❤️ by Truchain.

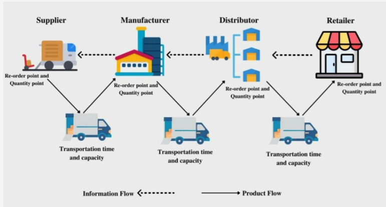
ds

Our Approach

A machine learning model will be used to predict future statistics in a supply chain using previous datasets

The goal is to forecast distorted demand signals in the supply chain using non-linear machine learning techniques, specifically focusing on forecasting demand at the upstream end to avoid the bullwhip effect

A dataset including product information, dates, locations, prices, and quantities will be used to train the model using various machine learning techniques, including linear regression, neural networks, support vector machines, and more



Our Vision

The proposed model can be implemented as a robust platform to aid in decision-making for the supply chain.

Thus, the following qualitative support from the data and concepts established in the background work is enough to prove that the proposed model can be implemented as a robust platform that can act as an aid to the client in making a better decision for the supply chain.

Manage app

TruChain · Streamlit

shivansh1203-truchain-app-z4fkxo.streamlit.app

Thus, the following qualitative support from the data and concepts established in the background work is enough to prove that the proposed model can be implemented as a robust platform that can act as an aid to the client in making a better decision for the supply chain.

TruChain

What you want to Predict?

Choose:

Forecasting model

Created with ❤️ by Truchain.

Get In Touch With Us

Name

Email

Message

Send

Manage app

CONCLUSION

- During the course of this project, we will perform a comparative analysis of various models of forecasting, from traditional techniques like LSTM model to Prophet model, to improve accuracy and efficiency using the dataset provided by the manager based on our research for this project, which will be displayed on the web platform and android platform (Hybrid application). We will implement anomaly detection for quickly determining anomalies or unexpected patterns in sales, which will help the company gain insight into its operating and marketing strategies to make more effective decisions.

- As mentioned in the first phase of our approach, we proposed to provide a comprehensive analysis using a variety of ML techniques to highlight the importance of the featured parameters and how they can prove to be beneficial as information to the client.

- The parameters are collected from various datasets and analyzed for their feasibility in supporting the hidden trends of the data.
  - Seasonality patterns
  - Sell in / sell out
  - Bullwhip
  - The effects of an anomaly in the data.

- We have outlined the issues in the problem statement and have matched them to the concepts introduced in the various existing types of research that are mentioned in the background work section.

- Further better SC visibility can be achieved by establishing contracts between manufacturers and clients to share their sales data which will increase the quality of the data fed to the proposed model and ultimately increase the accuracy of the predictions.

- Thus, the following qualitative support from the data and concepts established in the background work is enough to prove that the proposed model can be implemented as a robust platform that can act as an aid to the client in making a better decision for the supply chain.

3. Reference:

- Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management HD Nguyen, KP Tran, S Thomassey, M Hamad - International Journal of Information Management, 2021

- Hau L. Lee, V. Padmanabhan, Seungjin Whang, (1997) Information Distortion in a Supply Chain: The Bullwhip Effect. Management Science 43(4):546-558

- Nathalie Silva, Luís Miguel D.F. Ferreira, Cristóvão Silva, Vanessa Magalhães, Pedro Neto, Improving Supply Chain Visibility With Artificial Neural Networks, Procedia Manufacturing, Volume 11, 2017

- Real Carbonneau, Kevin Laframboise, Rustam Vahidov, Application of machine learning techniques for supply chain demand forecasting, European Journal of Operational Research, Volume 184, Issue 3, 2008