Title:

Product demand prediction with machine learning.

Problem statement:

The problem at hand is to analyse product demand data and discover product associations, enabling us to make informed decisions about product placements, bundling, and marketing strategies.

Problem Definition:

It is the process of using predictive analysis of historical data to estimate and predict customers' future demand for a product or service. Demand forecasting helps the business make better-informed supply decisions that estimate the total sales and revenue for a future period of time.

In demand forecasting, machine learning algorithms can analyze historical sales patterns and predict future trends. The first step is collecting data about past sales, such as product type, quantity sold, purchase frequency, seasonality, discounts, and more. This data is fed into the algorithm to create models that identify sales patterns over time. Once these patterns are identified, they can be used to create an accurate forecast for future trends.

The company wants to find the price at which its product can be a better deal compared to its competitors. For this task, the company provided a dataset of past changes in sales based on price changes. We need to train a model that can predict the demand for the product in the market with different price segments.

The dataset that we have for this task contains data about:

1. the product id;
2. store id;
3. total price at which product was sold;
4. base price at which product was sold;
5. Units sold (quantity demanded);

Importance of Product Demand:

Without demand, there is no business. And without a thorough understanding of demand, businesses aren't capable of making the right decisions about marketing spend, production, staffing, and more.

Demand forecast accuracy will never be 100%, but there are steps you can take to improve production lead times, increase operational efficiencies, save money, launch new products, and provide a better customer experience.

Preparing Budget:

It helps reduce risks and make efficient financial decisions that impact profit margins, cash flow, allocation of resources, opportunities for expansion, inventory accounting, operating costs, staffing, and overall spend. All strategic and operational plans are formulated around forecasting demand.

Demand planning and scheduling production:

Demand forecasting lets you provide the products your customers want when they want them. Forecasting demand requires that order fulfillment is synced up with your marketing prior to launching.

Nothing kills progress (or your reputation) faster than being sold out for weeks on end. Proper demand forecasting and inventory control can help ensure a business doesn't buy insufficient or excessive inventory.

Storing inventory:

It can help you spend less money on both inventory purchase orders and warehousing, as the more inventory you carry, the more expensive it is to store. Good inventory management involves having enough product on hand but not too much.

Developing a pricing strategy:

If you choose to slash prices or put an item on promotion, demand may temporarily increase for that product. Without that sale, you may not have experienced the boost.

If there is a limited supply of a high-demand product, you can use the scarcity principle to increase the price as an exclusive offer. You must keep an eye on new entrants though, as supply may increase.

Types of goods:

It will be very different for different products and services — from perishable goods that expire quickly to subscription boxes that come at the same time each month.

It's important to know the lifetime value of your customers (the total purchases they buy from you across channels over time), your average order value (how much they're

spending each time), and the combinations of products ordered to improve demand forecasting.

Design thinking:

Steps:

1.Set objectives
2.collect and record data
3.Measure and Analyze data
4.Budget Accordingly

1.**Set Objectives:**

-The goal of demand planning is to predict what,how much,and when customers will purchase in the next time period.

- Decide on parameters such as:
  - •time period
  - •product or product category
  - •Customer pool
- We will predict the sales for next three months. This will help the shopkeeper maintain an inventory and manage his finances. Also the shopkeeper can derive the marketing strategy.

2.**Collect and record data:**

-To get the most accurate demand forecasts, implement an omnichannel inventory,order,and warehouse management strategy.

-Pull data on:

- •order time and date
- •SKUs in each order
- •return rates
- •obsolete stock
- •Frequency of stockouts
- •Market conditions

-**SKU Velocity:** how frequently a SKU is picked over a certain period of time
Inventory Turnover Rate: how many times your entire inventory has been sold and then replaced within a given timeframe

-**Average Order Value (or AOV):** the average dollar amount a customer spends each time they place an order
Return rate: the frequency with which each SKU is returned

-**Stockout rate:** how often your business sells out of a particular SKU, depleting available units of the SKU to 0

-By tracking these inventory metrics over time, your business can forecast growth and trend projection on a more granular level and look back to see how your forecasts matched up to reality.

-In addition to your historical sales data, you may also need to pull in other pieces of data, like market conditions. Make sure any data you use is properly prepared to achieve the most reliable and accurate forecasts possible.

3.**Manage and Analyze data:**

Patterns that repeat over several time periods are typically good indicators of future outcomes.

Inspect your historical sales data and look for trends across:

- •Seasons
- •Holidays
- •Geographies
- •Customer Demographics
- •product lines

4.**Budget Accordingly:**

Once you have a feedback loop, you can set your next forecast (hopefully more accurately) and update your budget to allocate funds where they should go based on growth goals. Demand forecasting helps you reduce inventory carrying costs, plan marketing spend, future headcount, production and inventory needs, and even new products.

Methodology:

Dataset Information:

• The dataset is taken from the Kaggle in the text form then it is further converted in the .csv form to make the data available in tabular form.

• The system is taking the data of past 5 years.

•Further we divide the data into parts, first training data i.e. 1 year of data. Second is testing data i.e. 4 years of data and the training and testing of the data is done with the cross- validation of five folds.

Solving the problem:

1. Different machine learning classification algorithms are used to solve the problem and we have chosen the best model for the dataset of course of 5 years.

2.Parameters used to predict the best model out of the some models are:

*Correlation coefficient
*Mean absolute error
*Root mean squared error
*Relative absolute error
*Root Relative squared error

Correlation coefficient: A correlation coefficient is a numerical measure of linear interdependence of given variables.
It is also defined for a set of data. The lower limit is -1 whereas upper limit is +1.It shows relative strength of give data sets or variables as well. A negative correlation coefficient means decrease in one variable leads to increase in value of other variable. A positive correlation coefficient means increase in one variable will lead to increase in other variable as well. The exact amount of change in one variable when other variable is increased or decrease is unknown but relative increase decrease can be specified.

Mean absolute error: It is difference between the original and predicted value averaged over all the points in datasets. The difference in mean absolute error and root mean squared error is that it take cares of the direction of the error. For every observation in our dataset, we take absolute difference from original observation and get the summation. This is divided by total no of observations to get the result.

Root Mean squared error: Root mean squared error is standard deviation of errors which occur when prediction is made. This is similar to mean squared error, only root is taken. Here errors are squared before getting average of them and followed by square root of them. Thus errors are not differenced and it accurately tells large errors in the prediction. This is why it doesn't take absolute error. It can tell model's accuracy if the value is low, model is good to go. Lesser the value, more accurate and close are values predicted to the original values in the future.

Relative Absolute error: Relative absolute error is another performance matrix used to determine if the predictive
model is accurate or not. It is often used as relative error but they are different. Relative error is just to measure the effectiveness of scales or clocks
Relative absolute error is measured by current orginal observation minus the predicted observations. This is summed up and divided by summation of original values. Thus, result is obtained. Absolute error is simple difference between prediction and actual observations whereas relative absolute error takes care of this difference by dividing by the original value.

Root Relative Squared error:
Root relative squared error is comparative with what it would have been if a basic indicator had been utilized. All the more explicitly, this basic indicator is only the normal of the real qualities. In this manner, the squared error is taken. This is further decided by square error of
the main predictor. By taking the square foundation of the relative squared error one lessens the mistake to indistinguishable measurements from the amount being anticipated.

***Tools used:**

1.jupyter notebook
2.Colab
3.github
4.Exel(.csv file)

Results and Interpretation:

Results will be presented through visualizations, including association rules, scatter plots, and heatmaps.
Key performance metrics for the analysis include support, confidence, and lift.

Visualization:

Visualization tools such as Matplotlib and Seaborn will be used to create visual representations of the discovered associations.

Recommendations:

*Provide actionable recommendations, including bundling related products together.
*Suggest optimizing inventory based on product associations.
*Segment customers for targeted marketing strategies.
*Quantify the potential impact of recommendations on sales.

Project Timeline:

A detailed project timeline will be created to outline milestones, tasks, and deadlines. This will
ensure the project stays on track and is completed within the specified timeframe.

Resources:

TEAM:

Project Manager :Meena
Data Analyst :Sneka
Data Engineer :Nagapriya
Data Scientist : Nandhini
Visualization specialist :Menaka

Conclusion:

This project design document outlines our approach to performing product demand prediction with machine learning. It provides a clear roadmap for addressing the problem and delivering actionable insights to the retail business.