

PRODUCT DEMAND PREDICTION WITH MACHINE LEARNING

PROJECT TITLE: PRODUCT DEMAND ANALYSIS

PROBLEM STATEMENT:

Create a machine learning model that forecasts product demand based on historical sales and external factors, helping businesses optimize inventory management and production planning to meet customer needs efficiently.

AIM:

The project aims to build a machine learning model that forecasts product demand based on historical sales data and external factors. The primary goal is to assist businesses in optimizing inventory management and production planning to efficiently meet customer needs. The project encompasses data collection, data preprocessing, feature engineering, model selection, training, and evaluation.

DATASET: Product demand dataset

link: <https://www.kaggle.com/datasets/chakradharmattapali/product-demand-prediction-with-machine-learning>

some of the screenshot of product demand Exel sheet of dataset are:

ID	Store ID	Total Price	Base Price	Units Sold
1	8091	99.0375	111.8625	20
2	8091	99.0375	99.0375	28
3	8091	133.95	133.95	19
4	8091	133.95	133.95	44
5	8091	141.075	141.075	52
9	8091	227.2875	227.2875	18
10	8091	327.0375	327.0375	47
13	8091	210.9	210.9	50
14	8091	190.2375	234.4125	82
17	8095	99.0375	99.0375	99
18	8095	97.6125	97.6125	120
19	8095	98.325	98.325	40
22	8095	133.2375	133.2375	68
23	8095	133.95	133.95	87
24	8095	139.65	139.65	186
27	8095	236.55	260.0125	54
28	8095	214.4625	214.4625	74
29	8095	266.475	296.4	102
30	8095	173.85	192.375	214
31	8095	205.9125	205.9125	28
32	8095	205.9125	205.9125	7
33	8095	248.6625	248.6625	48
34	8095	200.925	200.925	78
35	8095	190.2375	240.825	57
37	8095	427.5	448.1625	50

ID	Store ID	Total Price	Base Price	Units Sold
35	8095	190.2375	240.825	57
37	8095	427.5	448.1625	50
38	8095	426.6375	458.1375	62
39	8095	177.4125	177.4125	22
42	8094	87.6375	87.6375	109
43	8094	88.35	88.35	133
44	8094	85.5	85.5	11
45	8094	128.25	180.975	9
47	8094	127.5375	127.5375	19
48	8094	123.975	123.975	33
49	8094	139.65	164.5875	69
50	8094	235.8375	235.8375	32
51	8094	234.4125	234.4125	47
52	8094	235.125	235.125	27
53	8094	227.2875	227.2875	69
54	8094	312.7875	312.7875	49
55	8094	210.9	230.9	60
56	8094	177.4125	177.4125	27
57	8094	177.4125	177.4125	33
58	8094	240.825	240.825	18
59	8094	213.0375	213.0375	72
60	8094	190.95	213.0375	81
61	8094	426.7875	448.1625	11
62	8094	426.7875	448.875	13
63	8094	426.7875	448.1625	28
65	8094	170.7875	170.7875	16

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
20127	212809	9961	213.0375	213.0375	23																		
20128	212810	9961	213.0375	213.0375	16																		
20129	212812	9961	252.9375	252.9375	97																		
20130	212813	9961	232.275	232.275	107																		
20131	212814	9961	236.55	236.55	26																		
20132	212816	9991	355.5375	489.5375	91																		
20133	212817	9961	856.9625	469.5875	127																		
20134	212819	9961	141.7875	177.4125	13																		
20135	212822	9984	87.6375	87.6375	52																		
20136	212823	9984	86.925	86.925	27																		
20137	212824	9981	86.2125	86.2125	10																		
20138	212826	9984	131.1	131.1	37																		
20139	212827	9981	130.3875	130.3875	26																		
20140	212828	9984	170.2875	170.2875	30																		
20141	212832	9984	194.5125	194.5125	85																		
20142	212833	9981	327.0375	327.0375	15																		
20143	212834	9984	163.875	210.9	204																		
20144	212835	9981	205.9125	205.9125	20																		
20145	212836	9984	205.9125	205.9125	12																		
20146	212837	9984	239.4	239.4	23																		
20147	212838	9981	235.8375	235.8375	28																		
20148	212839	9984	285.8375	285.8375	80																		
20149	212842	9981	357.675	483.7875	31																		
20150	212843	9984	141.7875	191.6625	12																		
20151	212844	9984	234.4125	234.4125	15																		

Project steps:

- 1.problem definition
- 2.Design Thinking

Step-1: Problem definition:

The problem is to create a machine learning model that forecasts product demand based on historical sales data and external factors. The goal is to help businesses optimize inventory management and production planning to efficiently meet customer needs. This project involves data collection, data preprocessing, feature engineering, model selection, training, and evaluation.

Step-2: Design Thinking:

(1).Data Collection:

Data collection is a systematic process of gathering observations or measurements. Whether you are performing research for business, governmental or academic purposes, data collection allows you to gain first-hand knowledge and original insights into your [research problem](#).

While methods and aims may differ between fields, the overall process of data collection remains largely the same.

Before you begin collecting data, you need to consider:

The aim of the research

The type of data that you will collect

The methods and procedures you will use to collect, store, and process the data.

(2).Data Preprocessing :

Data preprocessing is an important step in the data mining process. It refers to the cleaning, transforming and integration of data in order to make it ready for analysis. The goal of data preprocessing is to improve the quality of the data and to make it more suitable for the specific data mining task.

Some common steps in data preprocessing are:

(a).Data cleaning

(b).Data Integration

(c).Data Transformation

(d).Data Reduction

(e).Data Discretization

(f).Data Normalization

(3).Feature Engineering :

Feature engineering involves creating relevant features from the raw data. For instance:

- Lag features: Include past sales data (e.g., sales from the previous week or month) as features.
- Date-related features: Extract features like day of the week, month, quarter, or year.
- External factors: Incorporate external data such as holidays, economic indicators, or weather forecasts.

(4).Model Selection:

Choose an appropriate machine learning algorithm for your demand forecasting task. Time series models like ARIMA or machine learning models like Random Forest, XGBoost, or LSTM (if you have a significant amount of data) are common choices.

For this example, we'll use a Random Forest regressor.

```
from sklearn.ensemble import RandomForestRegressor
```

```
model = RandomForestRegressor(n_estimators=100,  
random_state=42)
```

(5).Model Training:

Data Splitting: Split the dataset into training, validation, and test sets.

Model Training: Train the selected regression model using the preprocessed training data.

Example:

```
model.fit(X_train, y_train)
```

(6).Evaluation:

Evaluate your model's performance on the testing dataset using appropriate metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), or Mean Absolute Percentage Error (MAPE).

Example:

```
from sklearn.metrics import mean_absolute_error
```

```
y_pred = model.predict(X_test)
```

```
mae = mean_absolute_error(y_test, y_pred)
```

```
print(f"Mean Absolute Error: {mae}")
```

Conclusion:

This document outlines the project's objectives, design thinking considerations, and the workflow for building a demand forecasting machine learning model. By following these steps, we aim to empower businesses with the ability to make data-driven decisions regarding inventory management and production planning, ultimately enhancing operational efficiency and customer experience.