**PRODUCT DEMAND PRODUCTION 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 business optimize inventory management and production planning to meet customer needs efficiently.**

****DATASET:**product demant dataset**

**Link:** <https://www.kaggle.com/datasets/prasoonkottarathil/microsoft-lifetime-stocks-dataset>

**Problem Definition and Design Thinking**

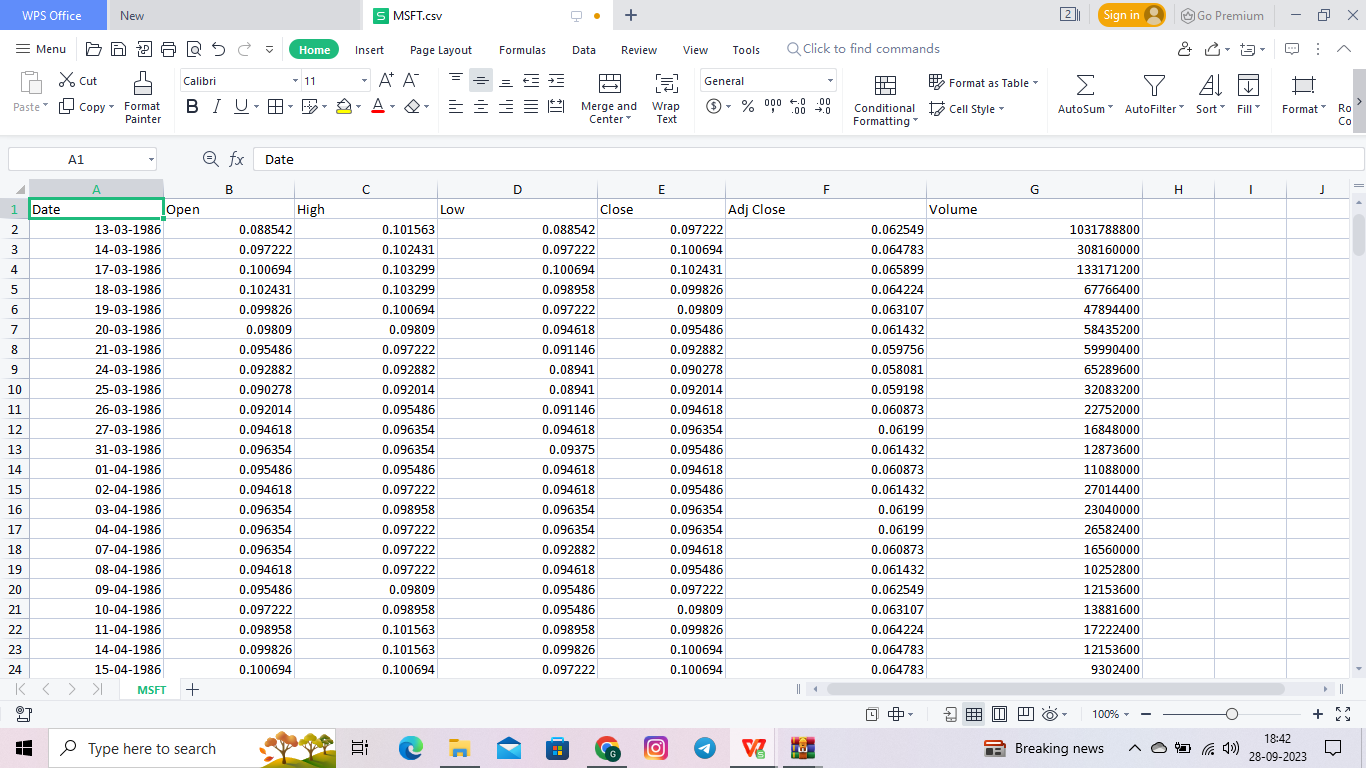
In this part we will need to understand the problem statement and create a document on what i have understood and how I will proceed ahead with solving the problem. Please think on a design and present in form of a document.

**Problem Definition:** The problem is to build a predictive model that forecasts stock prices based on historical market data. The goal is to create a tool that assists investors in making well-informed decisions and optimizing their investment strategies. This project involves data collection, data preprocessing, feature engineering, model selection, training, and evaluation.

**Design Thinking:**

* Data Collection
* Data Preprocessing
* Feature Engineering
* Model Selection
* Model Training
* Evaluation

GIVEN DATA



**Data Collection :**

Regardless of the field of study or preference for defining data (quantitative, qualitative), accurate data collection is essential to maintaining the integrity of research. Both the selection of appropriate data collection instruments (existing, modified, or newly developed) and clearly delineated instructions for their correct use reduce the likelihood of errors occurring.

**Data Preprocessing :**

Data Preprocessing can be defined as a process of converting raw data into a format that is understandable and usable for further analysis. It is an important step in the Data Preparation stage. It ensures that the outcome of the analysis is accurate, complete, and consistent. The main objective of Data Understanding is to gather general insights about the input dataset that will help to perform further steps to preprocess data. Let’s review two of the most common ways to understand input datasets

**Feature Engineering :**

Feature engineering refers to manipulation — addition, deletion, combination, mutation — of your data set to improve machine learning model training, leading to better performance and greater accuracy. Effective feature engineering is based on sound knowledge of the business problem and the available data sources.

Creating new features gives you a deeper understanding of your data and results in more valuable insights. When done correctly, feature engineering is one of the most valuable techniques of [data science](https://domino.ai/data-science-dictionary/data-science/), but it is also one of the most challenging. A common example of feature engineering is when your doctor uses your body mass index (BMI). BMI is calculated from both body weight and height, and serves as a surrogate for a characteristic that is very hard to accurately measure: the proportion of lean body mass.

* Model Selection

**Model Selection :**

Variable selection is the process of selecting the best subset of predictors for a given problem and predictive model, while model selection is done to select one specific model from the list of available predictive models for a given business problem.

The set of best variables might vary according to the change in the predictive model used as different types of predictive modeling algorithms works differently. A specific set of features might yield very different results with different predictive models

From sklearn.ensemble import randomforestregressor

Model = randomforestregressor(n\_estimators=100,random\_state=42)

**Model Training :**

Model training is at the heart of the data science development lifecycle where the data science team works to fit the best weights and biases to an algorithm to minimize the loss function over prediction range. Loss functions define how to optimize the ML algorithms. A data science team may use different types of loss functions depending on the project objectives, the type of data used and the type of algorithm.

When a supervised learning technique is used, model training creates a mathematical representation of the relationship between the data features and a target label. In unsupervised learning, it creates a mathematical representation among the data features themselves.

Train the selected model on your training data.

Example:

Model.fit(X\_train,y\_train)

**Evaluation :**

Model evaluation is the process of using different evaluation metrics to understand a machine learning model’s performance, as well as its strengths and weaknesses. Model evaluation is important to assess the efficacy of a model during initial research phases, and it also plays a role in model monitoring.

To understand if your model(s) is working well with new data, you can leverage a number of evaluation metrics.

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}”)