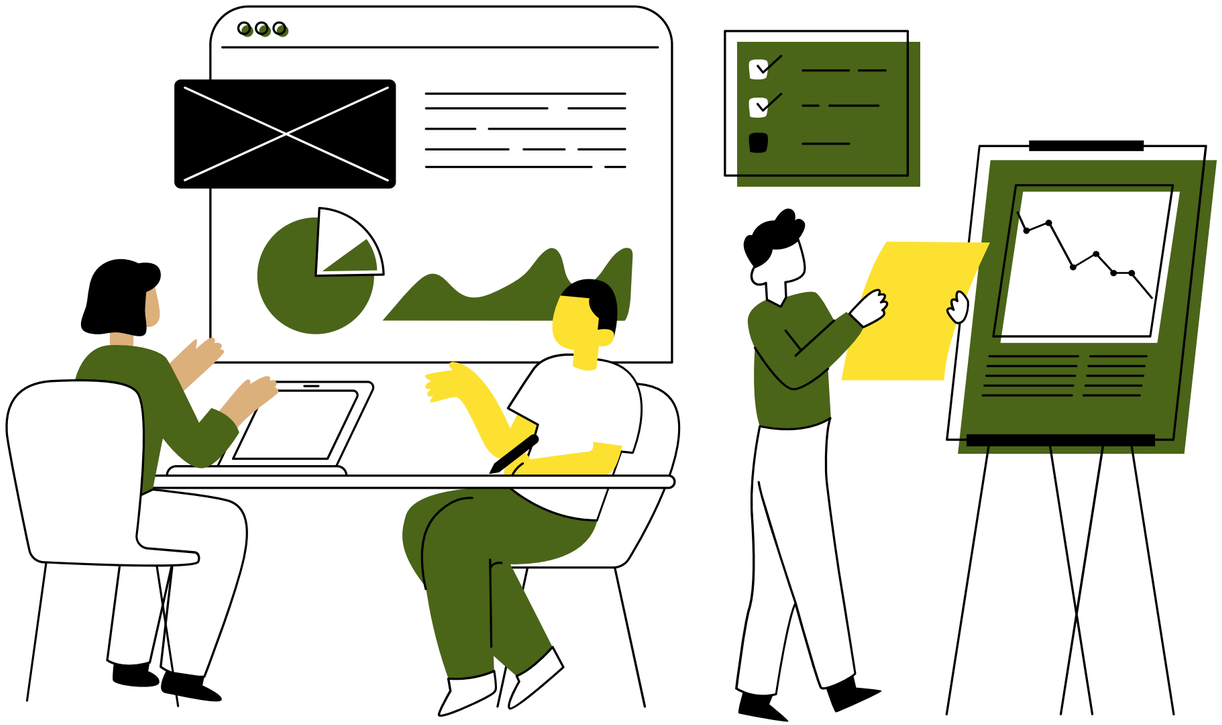
IBM Naan MUdhalvan

Applied Data Science

PROJECT 10:PRODUCT DEMAND PREDICTIONS WITH MACHINE LEARNING

AGENDA

- PROBLEM DEFINITION
- DESIGN THINKING

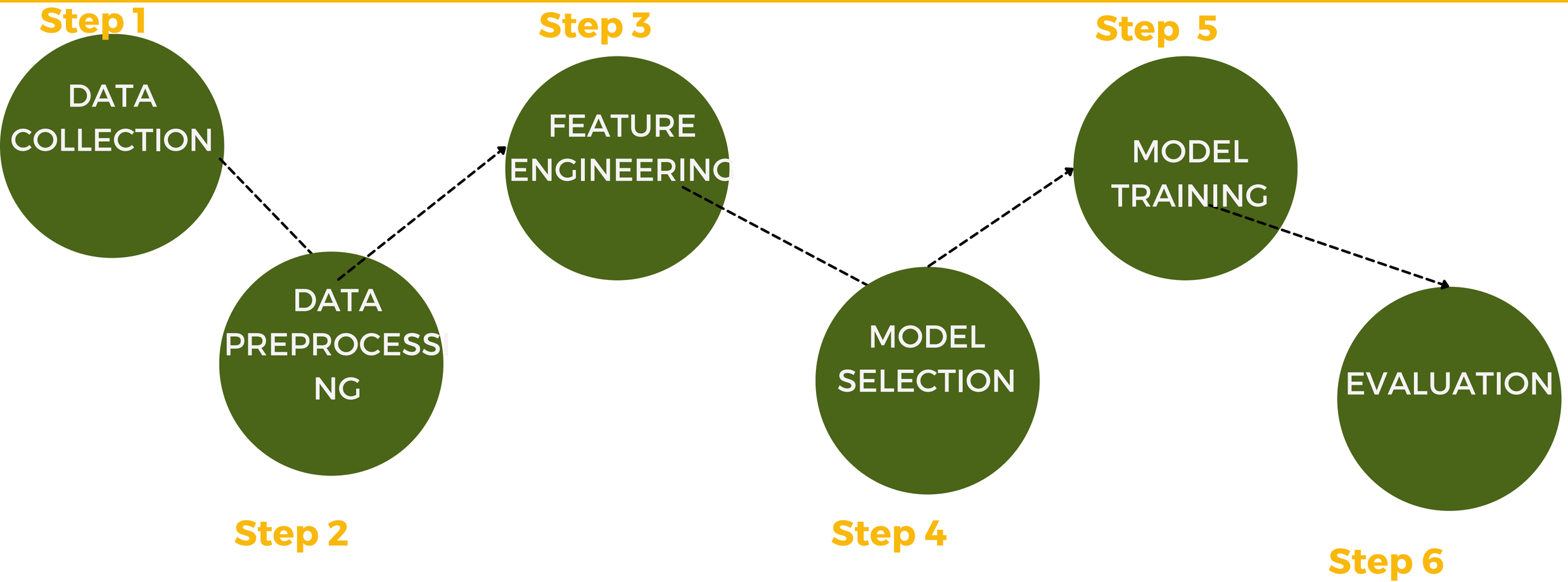


PROBLEM DEFINITION

- 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.

DESIGN THINKING

Design thinking is a non-linear, iterative process that teams use to understand users, challenge assumptions, redefine problems and create innovative solutions to prototype and test.



DATA COLLECTION

Here the datas for Product Demand Prediction are provided from the Kaggle website

Dataset Link:
https://www.kaggle.com/datasets/chakradharmattapalli/product-demand-prediction-with-machine-learning

The collected dataset contains

- product id
- store id
- Total price
- Base price
- No.of.units sold.
- [150150 rows x 5 columns]

DATA PREPROCESSING

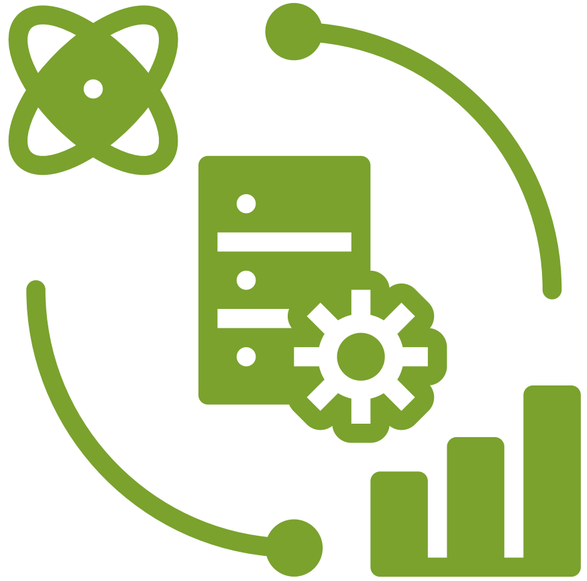
Here, our main goal is to remove the outliers, missing datas

PROBLEM

This data set contains only one missing values .

SOLUTION

As we have 150150 rows we can drop the row with has a missing value .This wont show a major difference while we are training the model



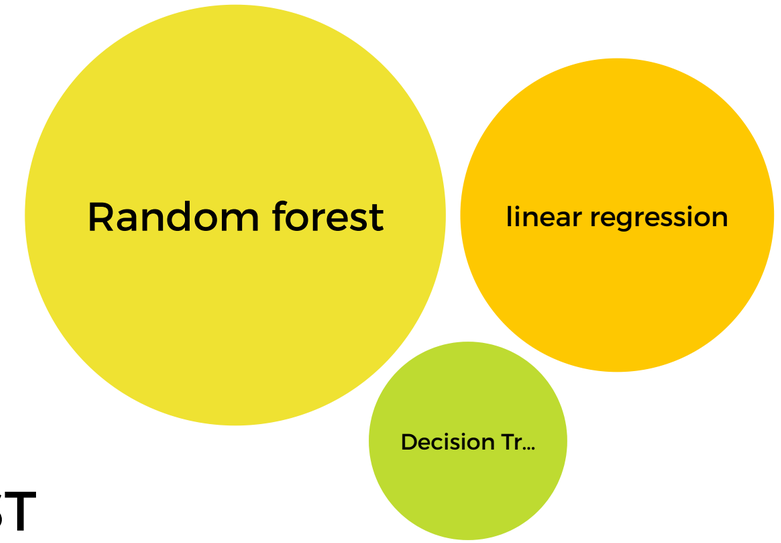
FEATURE ENGINEERING

Normalization: Rescaling the features to have a similar range, such as between 0 and 1, to prevent some features from dominating others.

.

Here, we convert all of our column data into the same scale(range) so that it wont create an impact while we are building a model

MODEL SELECTION



In this part we are choosing a suitable model RANDOM FOREST algorithm for demand forecasting.

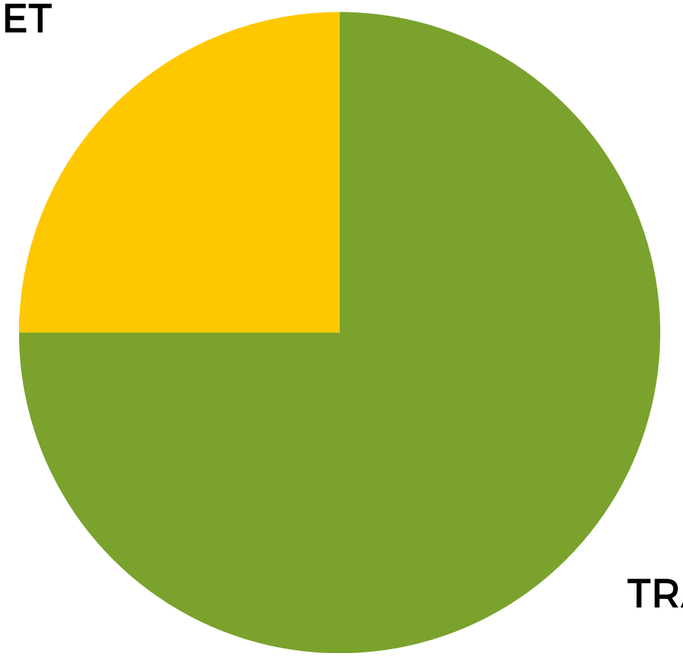
Model selection involves 3 steps.

Step 1: In the Random forest model, a subset of data points and a subset of features is selected for constructing each decision tree. Simply put, n random records and m features are taken from the data set having k number of records.

Step 2: Individual decision trees are constructed for each sample.

Step 3: results from each tree are aggregated to give a prediction for each observation

MODEL TRAINING



The preprocessed data is taken for training the data.

This preprocessed data is further splitted into # train data set

test data set

MODEL EVALUATION

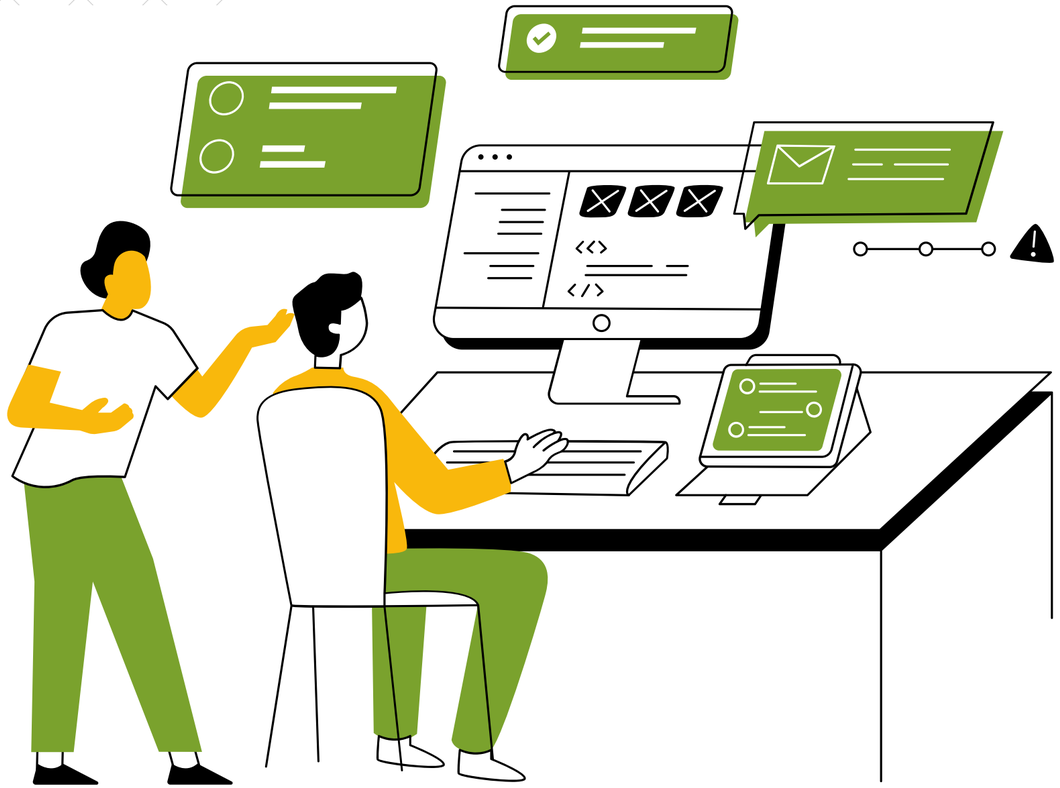
At this point, we have a trained Random Forest model, but we need to find out whether it is making accurate predictions.

- The simplest way to evaluate this model is using accuracy;
- we check the predictions against the actual values in the test set and count up how many the model got right.

CONCLUSION

These are the steps to be followed while we are working on the product demand prediction with MachineLearning.

Therefore we would create a successful and accurate model for Product Demand Prediction.



THANK YOU

P

BY
L PRIYA NANDHINI