Stores Sales Prediction

Objective:

Development of a predictive model for Item Outlet Sales . The model will predict the sales of the specified items in stores and Big Marts

Benefits:

- Estimates of Future Sales.
- Provides Production Planning.
- Helps in Purchasing the items.
- Helps in Sales Strategy .

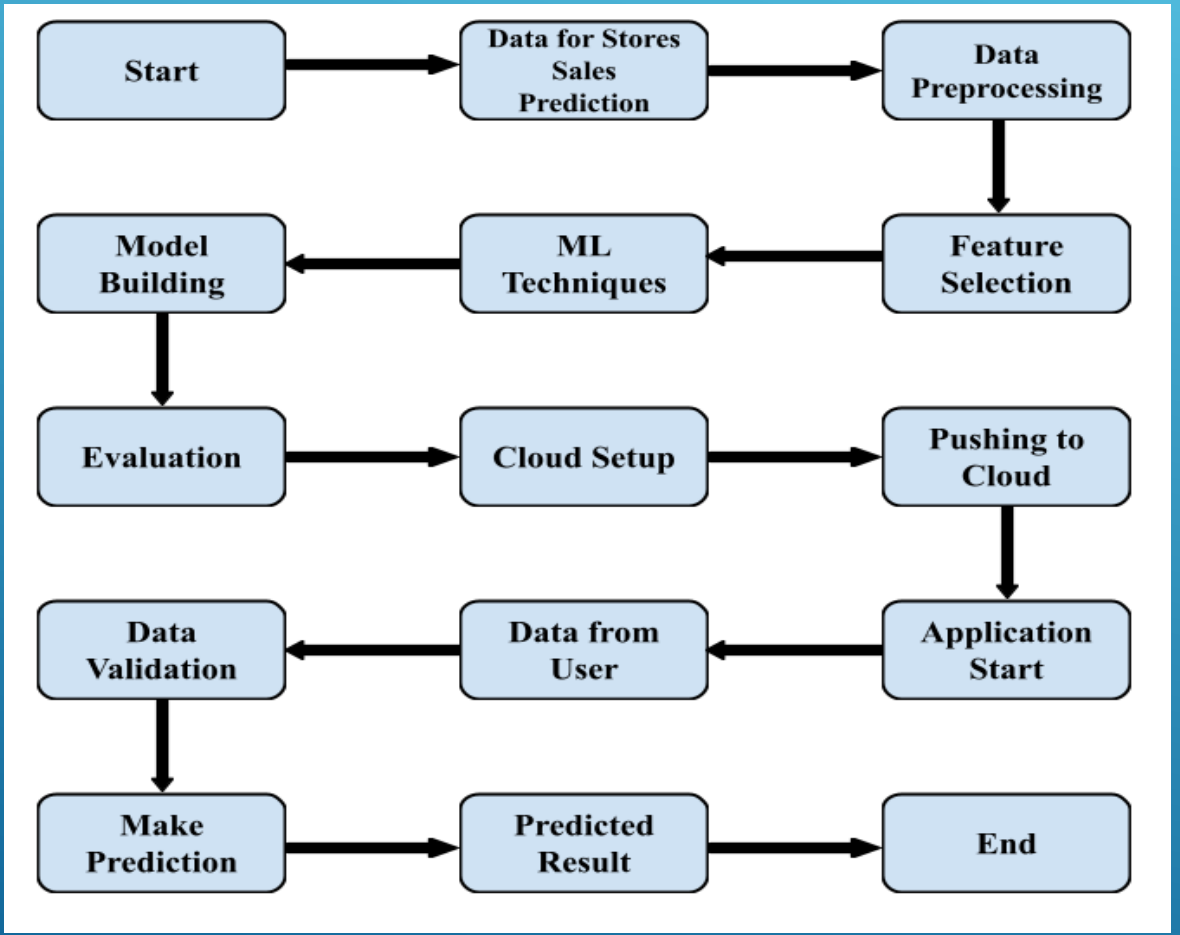
Data Sharing Agreement :

- Sample file name (Train.csv)
- Number of Columns
- Column names
- Column data type

Data Schema

Feature name	Datatype	Null/Required
Item Weight	float	Required
Item Fat Content	object	Required
Item Visibility	float	Required
Item Type	object	Required
Item MRP	float	Required
Outlet Identifier	object	Required
Outlet Establishment Year	int	Required
Outlet Size	object	Required
Outlet Location Type	object	Required
Outlet Type	object	Required
Item Outlet Sales	float	Required

Architecture



Architecture Description

Data Description

The dataset used in this project is collected from Kaggle. The dataset is divided into two sets of data. One is the test (5681) data and the other is the train (8523) data. The train dataset has both input and output variables.

Data Insertion into Database

a. Database Creation and connection - Create a database with name passed. If the database is already created, open the connection to the database.
b. Table creation in the database.
c. Insertion of files in the table

Data Pre-processing

Data Pre-processing steps we could use are data cleaning, data integration, data reduction and data transformation.

Feature Selection

Feature Selection helps us to find the best set of features that allows us to build the necessary model for the project. This helps in selecting a subset of features from an initially large volume of features.

Machine Learning Techniques

Based on the problem statement and requirements we can use supervised or unsupervised technique which fits the project.

.

Model Building

Depending on the data type of the target variable we are either going to be building a classification or regression model. The main aspect of machine learning model building is to obtain actionable insights and in order to achieve that it is important to be able to select a subset of important features from the vast number.

Evaluation

The Evaluation of accuracy can be done using the test data. Mean Absolute error can be found using test data and prediction data.

Cloud Setup

Using Heroku as the cloud deployment platform, the platform is setup for deploying the virtual app.

Pushing to Cloud

Once the cloud is setup, the virtual app created will be pushed to the cloud and will finally be deployed into the cloud

Application Start

Once the virtual app is deployed in to the cloud we can open the web application using any web browser.

Data from user

Using a web browser we open the web application and provide the necessary information as the input for prediction.

Data Validation

Once the input is provided and we click on the submit button, the system will provide the output based on its requirements.

Result Prediction

Once the data validation is completed the prediction will be done for the type of product in Stores and Big Marts provided in the input.