Project Brief
Regression – Prediction of Store Sales

Store Sales Prediction

The success of any retail store depends upon its sales. More the sales made, more is the revenue. With a good customer service and care, the customer too enjoys a good shopping experience. This will lead to more in-flow of customers, opening more store branches across a city / country.

To enable this, store owners rely heavily on past data to predict future sales. This will help them in strategizing their business models and come up with innovative techniques and ideas to attract customers and set realistic goals.

Many medium to large stores implement this kind of analytics to understand trends like which products are getting sold faster, what are the slow moving products, impact of store types, etc. With new branches of the store, the analysis can get more complex and may also involve stiff competition from other competitors.

Business Objective

- Build a model to predict the Sales of a store

Data Dictionary: SalesStore Data

The dataset contains about 14000 records in the form of a CSV file. Click here to download the dataset for this project.

S.No	Feature	Data Type	Description
1	Item_identifier	Character	Unique Product ID
2	Item_weight	Numeric	Weight of the product
3	Item_Fat_Content	Numeric	Total fat content in the product
4	Item_Visibility	Numeric	How visible is the product in the store
5	Item_Type	Categorical	Product category of the selected product
6	Item_MRP	Numeric	Product cost
7	Outlet_Identifier	Categorical	The store ID
8	Outlet_Establishment_Year	Numeric	The year when the store was opened

Data Dictionary: SalesStore Data (contd.)

S.No	Feature	Data Type	Description
9	Outlet_Size	Categorical	Size of the store
10	Outlet_Size	Categorical	Store size type
11	Outlet_Location_Type	Categorical	Location type where the store is located
12	Outlet_Type	Categorical	The type of store
13	Item_Outlet_Sales	Numeric	Sales made by the store outlet

Technical Goals

- Understand the data very well. Do all transformations / data engineering / etc. wherever applicable

- Perform Exploratory Data Analysis (EDA)

- Carry out all the Data mining tasks

- Identify the salient features that will determine the best results

- Perform the model evaluation to select the best algorithms

General Objectives

- The objective of this mini project is to compare different regression models and proposed the best model through evaluation of models using evaluation metrices.

- Total time allocated for this sprint is 8 hours which include project work and presentation.

- Viva will be included during project presentation.