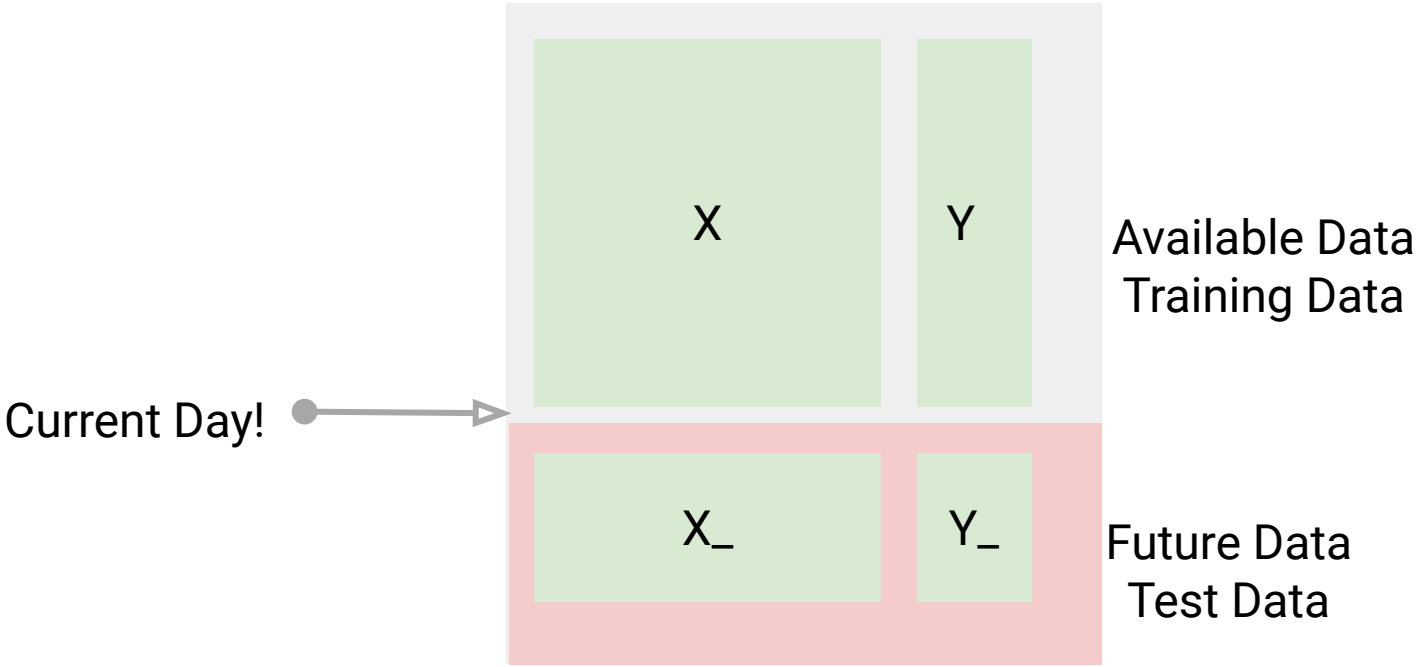
Building First Predictive Model

Creating the Dataset



Analytics Vidhya
Learn everything about analytics

Problem Types

Problem Types

Regression

Classification

Problem Types

Regression

Big Mart Sales

Classification

Titanic

Analytics Vidhya
Learn everything about analytics

Big Mart Sales

Analytics Vidhya
Learn everything about analytics

Big Mart Sales

To build a predictive model and find out the sales of each product at a particular store

Big Mart Sales

Variable	Description
Item_Identifier	Unique product ID
Item_Weight	Weight of product
Item_Fat_Content	Whether the product is low fat or not
Item_Visibility	The % of total display area of all products in a store allocated to the particular product
Item_Type	The category to which the product belongs
Item_MRP	Maximum Retail Price (list price) of the product
Outlet_Identifier	Unique store ID
Outlet_Establishment_Year	The year in which store was established
Outlet_Size	The size of the store in terms of ground area covered
Outlet_Location_Type	The type of city in which the store is located
Outlet_Type	Whether the outlet is just a grocery store or some sort of supermarket
Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

Big Mart Sales

Variable	Description
Item_Identifier	Unique product ID
Item_Weight	Weight of product
Item_Fat_Content	Whether the product is low fat or not
Item_Visibility	The % of total display area of all products in a store allocated to the particular product
Item_Type	The category to which the product belongs
Item_MRP	Maximum Retail Price (list price) of the product
Outlet_Identifier	Unique store ID
Outlet_Establishment_Year	The year in which store was established
Outlet_Size	The size of the store in terms of ground area covered
Outlet_Location_Type	The type of city in which the store is located
Outlet_Type	Whether the outlet is just a grocery store or some sort of supermarket
Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

Big Mart Sales

Variable	Description
Item_Identifier	Unique product ID
Item_Weight	Weight of product
Item_Fat_Content	Whether the product is low fat or not
Item_Visibility	The % of total display area of all products in a store allocated to the particular product
Item_Type	The category to which the product belongs
Item_MRP	Maximum Retail Price (list price) of the product
Outlet_Identifier	Unique store ID
Outlet_Establishment_Year	The year in which store was established
Outlet_Size	The size of the store in terms of ground area covered
Outlet_Location_Type	The type of city in which the store is located
Outlet_Type	Whether the outlet is just a grocery store or some sort of supermarket
Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

Big Mart : First Predictive Model

Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

Big Mart : First Predictive Model

Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

1. Mean

Big Mart : First Predictive Model

Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

Variable
Item_Identifier
Item_Weight
Item_Fat_Content
Item_Visibility
Item_Type
Item_MRP
Outlet_Identifier
Outlet_Establishment_Year
Outlet_Size
Outlet_Location_Type
Outlet_Type

1. Mean
2. Mean Respect to another Variable

Analytics Vidhya
Learn everything about analytics

Big Mart : First Predictive Model

Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

Variable
Item_Identifier
Item_Weight
Item_Fat_Content
Item_Visibility
Item_Type
Item_MRP
Outlet_Identifier
Outlet_Establishment_Year
Outlet_Size
Outlet_Location_Type
Outlet_Type

1. Mean
2. Mean Respect to another Variable
3. Mean With Respect to two other variables

Analytics Vidhya
Learn everything about analytics

Big Mart : First Predictive Model

Item_Outlet_Sales	Sales of the product in the particular store. This is the outcome variable to be predicted.

Variable
Item_Identifier
Item_Weight
Item_Fat_Content
Item_Visibility
Item_Type
Item_MRP
Outlet_Identifier
Outlet_Establishment_Year
Outlet_Size
Outlet_Location_Type
Outlet_Type

1. Mean
2. Mean Respect to another Variable
3. Mean With Respect to two other variables

...or even more variables

Analytics Vidhya
Learn everything about analytics

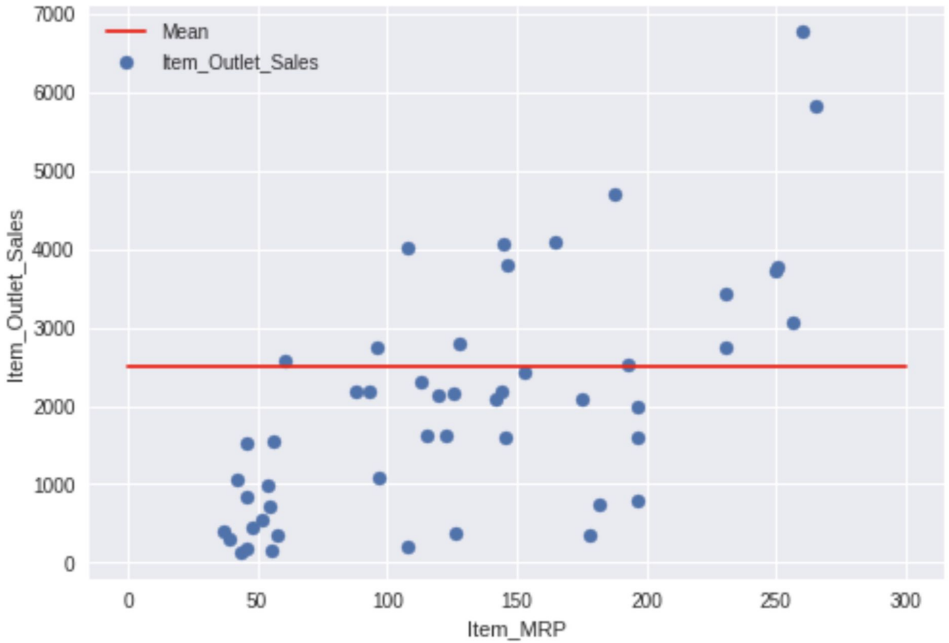
Big Mart : Evaluation

Big Mart : Evaluation

Mean Absolute Error: Sum of absolute difference between every observation, divided by the number of observations

Big Mart : Evaluation

Mean Absolute Error: Sum of absolute difference between every observation, divided by the number of observations



Analytics Vidhya
Learn everything about analytics