**Project 3) - End to End Case Study-Retail Wallmart Store Sales Forecasting**

**Objective** –The objective is predicting store sales using historical markdown data.

One challenge of modelling retail data is the need to make decisions based on limited

history. If Christmas comes but once a year, so does the chance to see how strategic

decisions impacted the bottom line.

**Data Availability & Business Problem:**

You are provided with historical sales data for 45 Walmart stores located in different

regions. Each store contains a number of departments, and you are tasked with predicting

the department-wide sales for each store.

In addition, Walmart runs several promotional markdown events throughout the year.

These markdowns precede prominent holidays, the four largest of which are the Super

Bowl, Labour Day, Thanksgiving, and Christmas. The weeks including these holidays are

weighted five times higherin the evaluation than non-holiday weeks. Part of the challenge

presented by this competition is modelling the effects of markdowns on these holiday

weeks in the absence of complete/ideal historical data.

 **stores.csv:** This file contains anonymized information about the 45 stores, indicating

the type and size of store.

 **train.csv:** This is the historical training data, which covers to 2010-02-05 to 2012-11-

1. Within this file you will find the following fields:

 Store - the store number

 Dept - the department number

 Date - the week

 Weekly\_Sales - sales forthe given department in the given store

 IsHoliday - whether the week is a special holiday week

 **test.csv:** This file is identical to train.csv, except we have withheld the weekly sales.

You must predict the sales for each triplet of store, department, and date in this file.

 **features.csv:** This file contains additional data related to the store, department, and

regional activity for the given dates. It contains the following fields:

 Store - the store number

 Date - the week

 Temperature - average temperature in the region

 Fuel\_Price - cost of fuel in the region

AnalytixLabs, Website:www.analytixlabs.co.in Email:info@analytixlabs.co.in phone: +91-88021-73069 MarkDown1-5 - anonymized data related to promotional markdowns that

Walmart is running. MarkDown data is only available after Nov 2011, and is not

available for all stores all the time. Any missing value is marked with an NA.

 CPI - the consumer price index

 Unemployment - the unemployment rate

 IsHoliday - whether the week is a special holiday week

For convenience, the four holidays fall within the following weeks in the dataset (not all

holidays are in the data):

 Super Bowl: 12-Feb-10, 11-Feb-11, 10-Feb-12, 8-Feb-13

 Labor Day: 10-Sep-10, 9-Sep-11, 7-Sep-12, 6-Sep-13

 Thanksgiving:26-Nov-10, 25-Nov-11, 23-Nov-12, 29-Nov-13

 Christmas: 31-Dec-10, 30-Dec-11, 28-Dec-12, 27-Dec-13

**Historical Data** –

* Data of 45 Wallmart stores of different regions
* Each store contain a number of departments

**Tools Used –** Python (numpy, pandas, matplotlib, , sklearn,statsmodels,pickle , etc.), Excel

**Techniques –**

* Identified various factors that affect the total turnaround time for a successful ticket closure using **Random Forest Algorithm.**
* Used **Classification** to find out if ticket is delayed or not
* For dataset’s Graphical representation and insights used **Matplot lib and Seaborn**
* For exploratory analysis I used all **statistical methods like Univariate Analysis, Bivariate analysis Hypothetical Testing, Correlation.**

**Tuning Parameters –**

* Used machine learning models like random forest,linear regression and decision tress with without tuning parameters

**Validation –**

* It will train my model using train data set I will verify result against test data set. If we will get score value nearly then I will say my model is good model.

**Challenges** –

* Data Collection
* Feature Selection

------------------------------------------------------------------------------------------------------------------------------------------