SALES FORECASTING USING MACHINE LEARNING

ANALYSING SALES OF WALMART

I.PRAVEEN KUMAR
S180221
CSE-2E

K.HARI BABU
S180115
CSE-2E

ABOUT WALMART

Walmart Inc. helps people around the world save money and live better - anytime and anywhere in retail stores. online, and through their mobile devices. Each week, over 265 million customers and members visit approximately 11,500 stores under 56 banners in 27 countries and eCommerce websites. With fiscal year 2020 revenue of $524 billion. Walmart employs over 2.2 million associates worldwide. Walmart continues to be a leader in sustainability. corporate philanthropy opportunity. and employment.

Business Problem

In this project, we are provided with historical sales data for 45 Walmart stores located in different regions. Each store contains many departments, and participants must project the sales for each department in each store. To add to the challenge, selected holiday markdown events are included in the dataset. These markdowns are known to affect sales, but it is challenging to predict which departments are affected and the extent of the impact. You may only use the provided data to make your predictions.

Data

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, Labor Day, Thanksgiving, and Christmas. The weeks including these holidays are weighted five times higher in the evaluation than non-holiday weeks. Part of the challenge presented by this competition is modeling the effects of markdowns on these holiday weeks in the absence of complete/ideal historical data.

The basic idea of analyzing the Walmart Forecasting dataset is to get a fair idea about the factors affecting the Sales of the Walmart Store.

PROBLEM STATEMTENT

Forecasting means to predict the future. Forecasting is used to predict future conditions and making plans accordingly. In our daily life, we are using a weather forecast and plan our day activity accordingly. Forecasting is used in many businesses. Here we will learn Sales Forecasting using Walmart Dataset using Machine Learning in Python

By using these data we have to Predict the Walmart sales forecasting based on different parameters.

Data Description

- ***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 2012-11-01. Within this file you will find the following fields:

  Store - the store number

  Dept - the department number

  Date - the week

  Weekly_Sales - sales for the 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

MarkDown1-5 - anonymized data related to promotional markdowns that Walm art 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-Nov13 Christmas: 31-Dec-10, 30-Dec-11, 28-Dec-12, 27-Dec-13.

Business objectives and constraints :-

1. *The cost of a mis-classification can be very high.*
2. *There is some latency concerns.*

Libraries used :-

import numpy as np import pandas as pd import

matplotlib.pyplot as plt import matplotlib.patches as

patches import seaborn as sns import plotly.express as px

import plotly.graph_objs as go from plotly.offline import

iplot from sklearn.model_selection import train_test_split

from math import sqrt from sklearn.linear_model import

Ridge from sklearn.linear_model import Lasso from

sklearn.metrics import mean_squared_error as mse from

sklearn.metrics import r2_score from

sklearn.model_selection import GridSearchCV from

sklearn.model_selection import RandomizedSearchCV

import warnings

Step-1 First, importing libraries of Python.

Step-2 Now, we preparing data. Here we perform four tasks:

1. *Get the data from .csv file.*
2. *Merging the data.*
3. *Analyze the data.*
4. *Manipulating the data.*

Step-3 Now, we perform learning tasks on this data in four steps.

1. *Splitting the train and test data.*
2. *Applying linear regression.*
3. *Predicting the value*
4. *Evaluate the model*

Conclusion :-

Forecasting sales is a common and essential use of machine learning (ML). Sales forecasts can be used to identify benchmarks and determine incremental impacts of new initiatives, plan resources in response to expected demand, and project future budgets.

In conclusion, the forecast should use for better plans and get more benefit from it. Here we get the following topics.

- *Sale Forecasting*
- *Sale forecasting using machine learning in Python.*

It is a projection of future sales revenue and a prediction of which deals will move through the sales cycle. Sales forecasts drive short-term spending decisions and impact decisions on key deals.