STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

MIS 637

Prediction of Walmart Sales Using Random Forest Regression

Data Analytics and Machine Learning Project

By-
Soham Sanjay Shinde
MS in Engineering Management
Guided By-
Mahmoud Daneshmand

What is the problem for Walmart?

Walmart is trying to find way to predict the sales for their stores. To predict the sales they have gathered the data for predicting sales. They have CPI, unemployment rate, fuel price, holiday flag and weekly sales in the dataset for 45 stores.

How are we approaching the problem?

- Business Understanding
- Data Understanding
- Data Preparation
- Modeling Phase
- Evaluation Phase
- Deployment Phase
- Conclusion

Business Understanding

- When you buy grocery, home accesory or any other food product, it might be bought from Walmart. As walmart is very much spreaded all over the world.
- In USA alone, Walmart has **4692** locations and **662** Sam‘s club locations.
- Walmart serves more than **100 million customers**. From this number we can guess, how much the sales of Walmart will impact by people and their buying trends.
- In fiscal year **2019, revenue** of Walmart was **$514.4 billion**. For this much big profit, Walmart should make sure, that they are not out of stock and their inventory is up to date.
- If they predict the sales of the week, they can maintain the stock and if inventory is minimum, they can refill at specific store.

Data Understanding

- Dataset consist of 8 columns and 6435 rows. Each row represents the first date of week.
- Each row has the weekly sales, CPI, unemployment rate, fuel price and holiday flag.
- This can be the 2

Data Understanding

Data Understanding

Data Understanding

Use Mixed Integer Quadratic Programming