BLACK FRIDAY SALES PREDICTION ANALYSIS USING MACHINE LEARNING REGRESSION MODELS

Reeba Babu
Supervisor : Prof. Jason Barron
M.Sc. in Data Science

INSTITUTE of TECHNOLOGY CARLOW
Institiúid Teicneolaíochta Cheatharlach

Introduction

Black Friday is the biggest sales day of the year in the United States. It takes place the Friday after Thanksgiving. Every year brands and retailers run special offers, deals, and discounts on tons of products and, shoppers get pretty hyped about it all. In 2017 174 million Americans went shopping on Black Friday. The prediction of Black Friday sales would help to understand the customer experience and their demand products.

Research Objective

This research focus on below research question.

- How to predict the purchase amount using machine learning regression models?
- Which gender spends the most money for shopping?
- Do the jobs of people has any impact on sales?

The Dataset

- The dataset is taken from kaggle which contains about 550068 rows and 12 features
- The data points included are:

User_ID , Product_ID, Gender, Age , Occupation, City_Category,Stay_In_Current_City_Years,Marital_Status,Product_Category_1,Product_Category_2,Product_Category_3 , purchase

Literature Review

In previous analysis, the extreme gradient boosted trees algorithm was used to predict Black Friday sales. Bagging and boosting methods can produce excellent results that can be further improved by hyper parameter tuning (Trung et al., 2019).

Exploratory data analysis on the dataset is helpful to understand the customer behavior on various products (kalra et al.,2020).

There is a research focus on the comparison of different machine learning algorithms for multiple regression on Black Friday sales data (Wu et al., 2018).

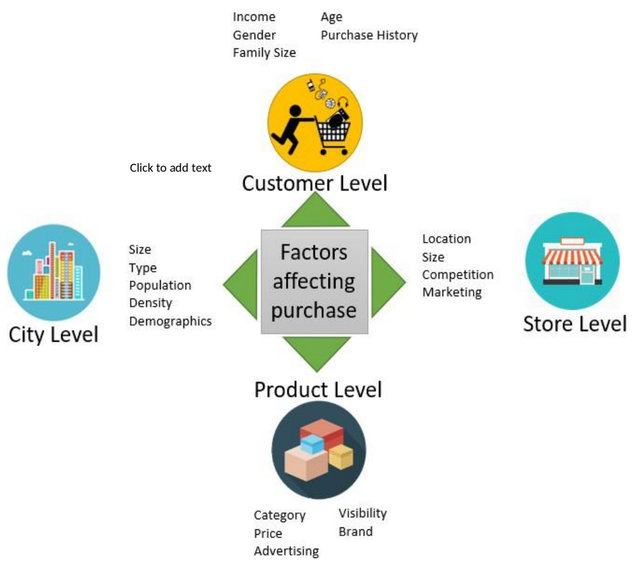

Figure1. Possible parameters that influence the customer's expenditure in Black Friday

Methodology

- Hyper parameter tuning will be applied to arrive at better accuracy.
- Exploratory data analysis including graph visualization will be used to investigate the dataset and to understand the customer behavior on various products.
- The Machine learning regression models would be use for predicting the purchase amount.
- The model was then integrated into a flask web application

Technologies

python, TensorFlow, matplotlib, jupyter, scikit learn, pandas

Next Steps

In the future, by this model, the brands and companies can understand customer behavior and their demand products. By this, they can achieve customer satisfaction, profit, and company growth.

References

1. Trung, N.D., Thien, T.D., Luu, T.D. and Huynh, H.X., Black Friday Sale Prediction Via Extreme Gradient Boosted Trees.

2. Kalra, S., Perumal, B., Yadav, S. and Narayanan, S.J., 2020, February. Analysing and Predicting the purchases done on the day of Black Friday. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (pp. 1-8). IEEE.

3. Wu, C.S.M., Patil, P. and Gunaseelan, S., 2018, November. Comparison of Different Machine Learning Algorithms for Multiple Regression on Black Friday Sales Data. In 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS) (pp. 16-20). IEEE.

Contact

Contact : Reeba Babu
Phone: +353 892084213
Email: reebababu1993@gmail.com