International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 03 Issue: 08 | August -2019 ISSN: 2590-1892

BLACK FRIDAY SALES ANALYSIS FOR MACHINE LEARNING

VIVEK KUMAR SINGH , DHANUSHA , SHASHANK H S, BASAVARAJA J, ARUN KUMAR H, MAHENDRA KUMAR B
A1,A2,A3,A4,A5,A6 : Post Graduate
DEPT OF MCA , DAYANANDA SAGAR COLLEGE OF ENGINEERING

Abstract – Black Friday is considered as the most profitable shopping day in the U.S. the day after Thanksgiving- Black Friday. Our research is to gather data about the consumer's potential for shopping behavior using a combination of machine learning techniques. [1] This research identifies the relationships among different variables present in the data, which affects the sales prediction, which products are in high demand based on different variables. This helps the retailers to create a personalized offer for customers for various products which increase the sales. The approach develops a model to predict the above information. Using two theoretical approaches (Random Forrest Regression and Decision Tree Regression) which are used to investigate the factors, variables that affect consumer behavior on Black Friday.

1. INTRODUCTION

This research observes and analyses the consumer behaviour of the customer. On this Black Friday, the retailers in the United States signify the start of the shopping season as in Christmas season. [1] Black Friday affects the potential shoppers with an increase in the rate of shopping. The profitable season is highly dependent on this holiday season, as sales increases at the end of the year with a profitable range. So, with high type in Black Friday sales, the customers are entitled to shop in an exposed environment where are subjected to wait in long lines, regardless of the weather conditions. As a customer arrives in chaos to shop at the last moment, with a range of products which are highly in demand for higher prices [2] also, the products are likely to be available; lastly, sales advertisements are limited. Here the entitled project relies on a model that intends to find customer behaviour for shopping, which products are likely to be sold in demand, how much discount the retailer can withstand for the customized offer for various products. Based on the prediction from the machine the retailer can understand which products are likely to be sold.

2. Methodology

A machine learning model, developed to determine the accuracy rate of sales on various products. The model is trained with data given by retailers. Data is extracted from the client database which has different variables like age, marital status, product price, product id, product name, gender, etc. Prediction of sales depends on many factors like how the model is trained with appropriate datasets, variables that are needed to train the model the datasets to be cleansed for the training of the model, remove the anomalies, reduce the values to low dimension, which appropriate model is used. The algorithms used for Black Friday sales prediction are decision tree and random forest which predicts the accuracy of the sale.

- Decision Tree
- Random Forest

DECISION TREE

The decision tree algorithm falls under supervised learning which is used to solve both regression and classification problems [5]. It uses tree representation in which each leaf node corresponds to a class label and its attributes as internals node of the tree. Here whole training is set as root. The ordering attributes as a root node is accomplished by some statistical methods.[4] A tree is split bye source dataset, which consists of root node into subsets. Each subset derived in a recursive repeatedly. The recursion is completed until each subset has the same values of the output variable.

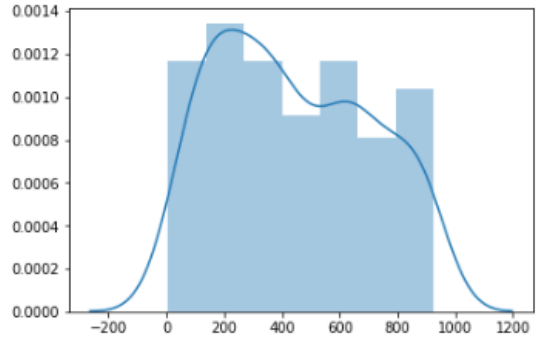
Once the model is trained with the test dataset, the performance of the model is evaluated.[5] The sklearn metric used in the decision tree here is an accuracy score. A classification metric used for evaluating the accuracy of the model which takes inputs as actual labels and the predicted labels. Accuracy Score of Decision Tree on the train set 98.14377383303507.

Decision trees are helpful visualizing for analysing a series of predicted output for this model. In this specific model, we examine the effect of illustrative factors Product ID, Gender, City Category, Product Category and Purchase.

Fig 1 – Output of Decision Model

```
1 seaborn.distplot(y_pred)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x2202e5edb38>
```



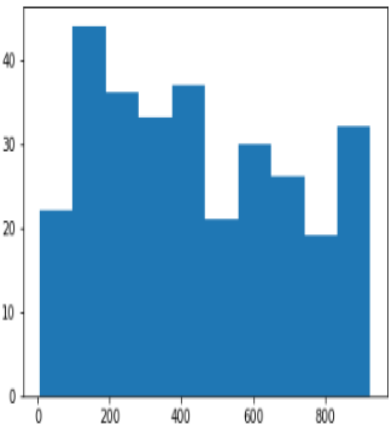
RANDOM FOREST

A Random Forest is a technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap Aggregation, commonly known as bagging [6]. The random forest consists of the large number of one individual decision tree which spits out a prediction. Model's prediction depends on the class which has most votes. Here a large number of unrelated models operate as an entity which outstand any other individual constituent models. The uncorrelated models produce ensemble predictions with more accuracy than individual predictions.[4] Random forest ensures that each individual tree behavior is uncorrelated with the help of Bagging. This model allows the individual tree to input random samples from a dataset with a replacement which results in different trees. [1] The algorithm has learned 96% on the training data and the value is 92 % on the test data.

Fig 2 – Output of Random Forest Model

```
1 plt.hist(y_pred)
```

```
(array([22., 44., 36., 33., 37., 21., 30., 26., 19., 32.]),
 array([  6.,  98., 190., 282., 374., 466., 558., 650., 742., 834., 926.]),
 <a list of 10 Patch objects>)
```



1. CONCLUSIONS

Black Friday Sales Prediction relies on the behaviour of the customer. The analysis clearly exposes that gender, product id, city category, age variables plays a vital role in the sales prediction. Also, the experience of the customer and demand driving of products, offerings tame the customer to shop more. Customer's interest in shopping has a promising effect on increases in sales for retailers. The retailers are expected to provide profitable bargains to the customers who intend them to shop increasingly. Black Friday is an effective utilizing offline key trade to ensure it is a profitable event.

REFERENCES

1. 1 Black Friday and Cyber Monday: Understanding consumer intentions on two major shopping days, Esther Swilleya, Ronald E. Goldsmith.
2. An Analysis of Consumer Behavior on Black Friday, Linda Simpson, Ph.D.
3. Black Friday Sales: Analysis and Prediction Kaggle
4. towardsdatascience.com.
5. Decision tree for classification

https://www.xoriant.com/blog/product-engineering/decision-trees-machine-learning-algorithm.html

6. Random forest algorithm works in machine learning https://dataaspirant.com/2017/05/22/random-forest-algorithm-machine-learing/