Data Mining Project - CRISP-DM Framework

# 1. Business Understanding

Objective: The goal is to forecast whether an online shopper is likely to make a purchase or not by analyzing their online behavior.

Stakeholders: Stakeholders include marketers that want to modify their tactics in response to users behavior.

Criteria: The performance evaluation will be done by looking at the accuracy and precision.

# 2. Data Understanding

Data Available: The available data was subtracted from Kaggle.com and is a dataset of online attributes including page visits from online shoppers, time spent online and other online activities.

Initial Exploration: The missing values were verified as nonexistent and the data types were examined. The first data exploration was conducted using simple statistics summary.

Data Quality Issues: The dataset is neat, reliable and clean without any missing values.

# 3. Data Preparation

Cleaning Requirements: I normalized the features and took out the missing rows so I could prepare my data properly.

Formatting or Transformation: I encoded all the categorical variables like “Month”, “Vistiro Type”.

# 4. Modeling

Modeling Approach: I applied the classification models to my python script.

Justification: Classification is suitable for forecasting binary outcomes, such as whether someone will make or not a purchase.

Models Considered: The three models I chose to work with are Naïve Bayes, Logistic regresson and Random forest.

# 5. Evaluation

Criteria for Model Performance: Accuracy and precision.

Other Areas to Explore: Additional models and different evaluation techniques are two more areas to investigate.

# 6. Deployment

Deployment Strategy: Real world scenarios will be used to test the models and assist e-commerce companies in forecasting shopper’s behavior.

Monitoring: It will be monitored through real time assessment..

Future Work: Possible imporevments or action can improve the model, such as adding more data features to give even more clean data to use and give a more optimized prefromance,

# Conclusion

Summary: The goal of this assignment was to forecast purchases made by online shoppers. A thorough evaluation of this issue was made possible with the use of data preparation and the use of modelling proceses. The models will provide a better understanding of the online shoppers dataset.