**Project:1** **Sentiment Analysis**

**Scope:**

* To find the sentiments from the reviews, whether it’s positive, negative, or neutral reviews.
* What words customers are using frequently during negative and positive reviews.
* We can use the model to find out the sentiments of the reviews.

**Libraries Used:**

**Pandas**: To clean and wrangle the data

**Matplotlib**: To see EDA

**TextBlob**: To count the sentiments.

**Word Cloud**: To create the word clouds for positive, negative and neutral groups.

**Langdetect**: To detect the language of the reviews.

**CountVectorizer, Tfidf Transformer**: To transform my text reviews into numerical columns.

**Word Tokenize**: To tokenize my reviews into separate words.

**Train test split:** To split my data into 80/20 split.

Logistic Regression

Random Forest

Grid Search CV: Hyperparameter Tuning

**Project:2** **Credit card default**

**Scope:**

* Credit card default is a very big problem for credit card industry now a days, just to figure out if the customer is going to default on their monthly payments.
* Increasing credit card debt becomes major problem for credit card companies to get their money back from the customers.
* Using this model, we can predict weather customer will default on their next payments or not on the basis of their education, marital status, last six months of their credit card payment details and their credit limit.

**Libraries Used:**

**Pandas**: To clean and wrangle the data

**Matplotlib**: To see EDA

**One hot encoding**: Convert categorical data into numerical data.

**Train test split**: To split the data in 80/20 split.

**Models:**

Logistic Regression, Random Forest, Gradient Boosting classifier, Logistic Regression

**Project:3 Credit card fraud**

**Scope:**

* Credit card fraud happens when someone — a fraudster or a thief — uses your stolen credit card or the information from that card to make unauthorized purchases in your name or take out cash advances using your account.
* we are on the express train to a cashless society. According to word payment report in 2016 total non-cash transactions increased by 10.1% for a total of 482.6 billion transactions!
* Now, while this might be exciting news, on the flip-side fraudulent transactions are on the rise as well. Even with EMV smart chips being implemented, we still have a very high amount of money lost from credit card fraud.
* What can we do to mitigate the risk? While there are a lot of methods to limit the loss and prevent fraud, and I’ll walk you through my process and model.

**Libraries Used:**

**Pandas**: To clean and wrangle the data

**Matplotlib**: To see EDA

**One hot encoding**: Convert categorical data into numerical data.

**Train test split**: To split the data in 80/20 split.

**Models:**

KMeans Clustering, Density based clustering (DBSCAN)