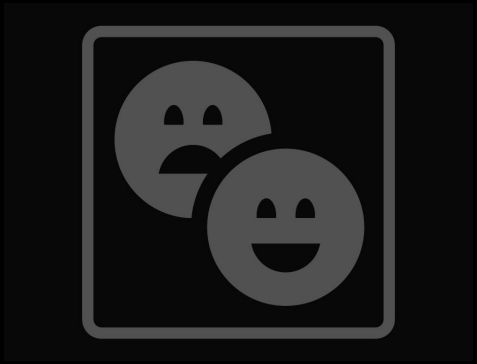


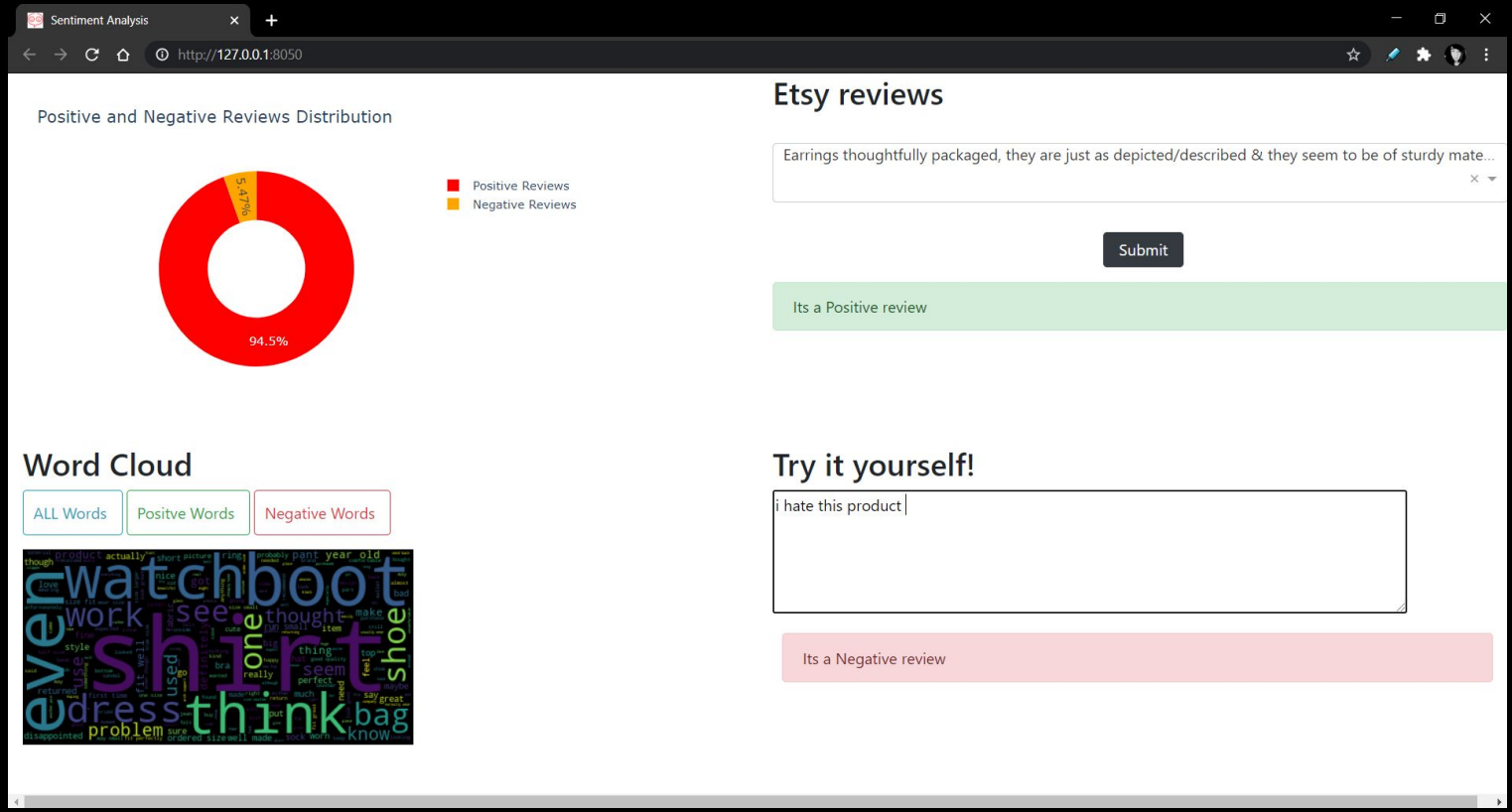
Sentiments Analysis

By Prajwal Mani

Problem Statement

The client who has online jewelry business receives from websites and TV channels .So the client wants us to build a dashboard through which they can Analyze the reviews of their products and do a sentiment Analysis of it.

Our Solution



Our Approach

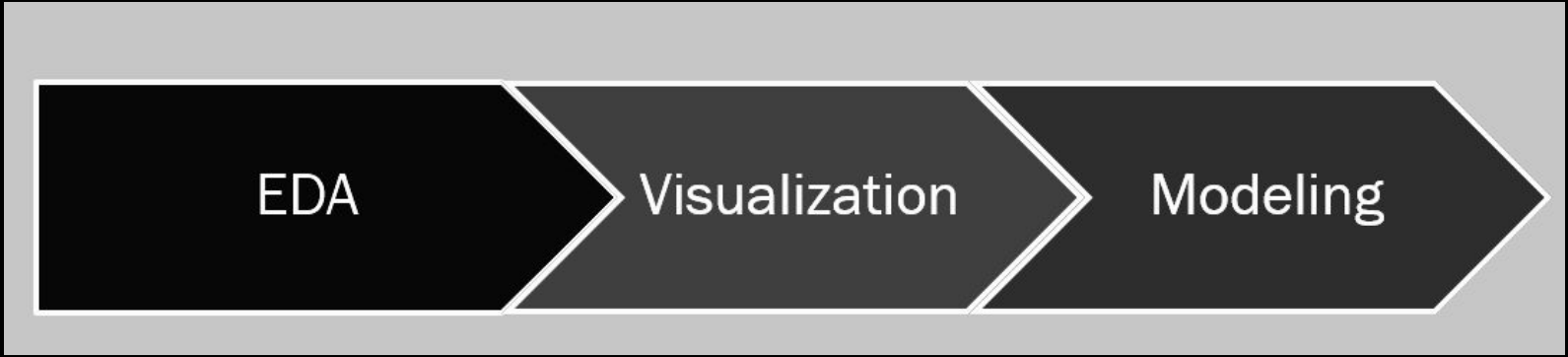
- Collect and Pre-Process the Dataset
- ML Life Cycle
- Web Scraping
- Building Dashboard UI and deploying the model

Python Modules

Collect and Pre-Process the Dataset

- We are Using Amazon Review Dataset (2018)
- We will be extracting just Jewellery category with around 750k samples
- With equal distribution of various reviews ranging from 1 to 5
- Convert that into CSV using pandas

ML Life Cycle



Exploratory Data Analysis

- In this particular stage we will be cleaning the data using pandas like checking null value ,dropping unwanted columns and handling data imbalance
- After all this we will be doing some NLP steps

Exploratory Data Analysis

- Some of the NLP steps are
  - Using regular expression to clean the text
  - Removing the stop words
  - Lemmatization
  - CountVectorizer and TfidfTransformer

Visualization

- Here we will be plotting some basic graphs like bar graph ,pie chart and word cloud to see the words which are more frequent in the given dataset
- There is not much plotting here..

Modeling

- I have tried out 4 modules
  - RandomForestClassifier
  - LogisticRegression
  - SVM
  - MultinomialNB
- We will use sklearn pipeline to combine CountVectorizer and TfidfTransformer with our model
- We will pickle it so that we can deploy it in dash

Web Scraping

- To check how our model works with new data we will be scraping Etsy using selenium and bs4
- Then convert that data into csv so that we can manipulate it later

User Interface

- Dash is python web framework for web analytic applications
- In my UI there are 4 components that are
  - Pie chart
  - Word Cloud
  - Etsy predict
  - Try it yourself!

User Interface

- All the above components where achieved using plotly ,dash_html_components,dash ,dash_core_components and dash_bootstrap_components
- The main role was played by app.callback() for handling events

Thank You