INNOMATİCS
RESEARCH LABS

INTERNSHIP PROJECT REPORT

ON

Streamlit Application Debugging

SUBMITTED BY

DIVYA MANOJ

BATCH

OCT – 21

GUIDED BY

CHIEF DATA SCIENTIST “KANAV BANSAL”

Nov 12, 2021

Debugging of Streamlit Application

Summary

A streamlit web based app "Airline Sentiment Analysis" displays visualizations on count of positive, negative and neutral reviews. It contains graphs of counts of different airlines and count of positive, negative and neutral tweets of each airline. Additionally, Logistic Regression model is used to classify entered tweet to either positive, negative, or, neutral tweet.

Purpose

The app had bugs because it was made with Python version 3.7. We needed to fix bugs.

Project Files -

The project consists of 4 python files:

1. project.py – It is the main file that runs the entire application.
2. data_app.py – It contains streamlit code to plot the graphs, visualize graphs and the dataframe.
3. ml_app.py – It contains ML code to predict sentiment of the entered tweet.
4. create_wordcloud.py – It contains code to generate wordcloud.

Technology Stack

1. Python

   Python is an interpreted high-level general-purpose programming language. It supports multiple programming paradigms,
   including structured (particularly, procedural), object-oriented and functional programming.

2. Streamlit

   Streamlit is an open-source python framework for building web apps for Machine Learning and Data Science. We can instantly develop web apps and

deploy them easily using Streamlit. Streamlit allows you to write an app the same way you write a python code. Streamlit makes it seamless to work on the interactive loop of coding and viewing results in the web app.

3. Matplotlib
   Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python.

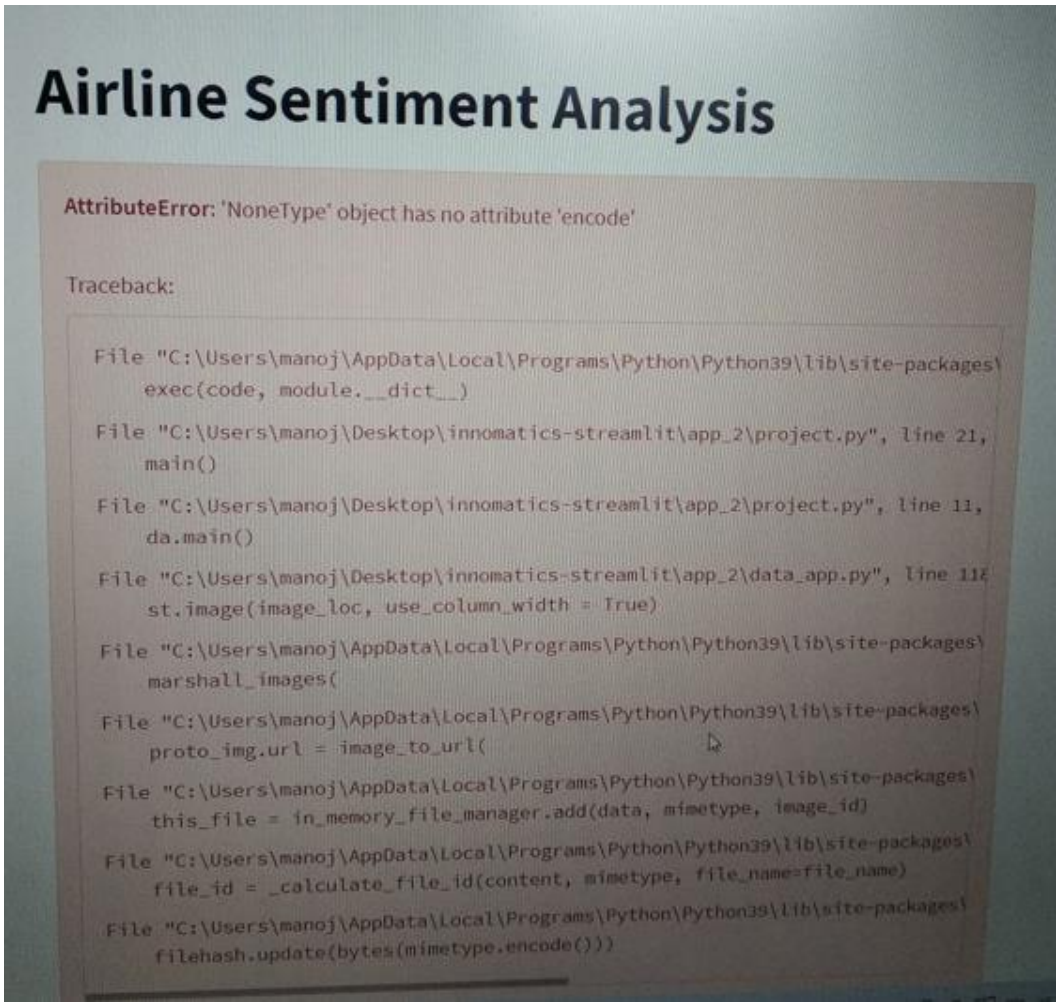
4. Seaborn
   Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

5. Wordcloud
   Word Clouds are visual representations of words that give greater prominence to words that appear more frequently.

Tool Stack

1. Visual Studio Code
   Visual Studio Code is a source-code editor made by Microsoft for Windows, Linux and macOS.[9] Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add additional functionality.

Bugs and their Solution



Added following lines of code to solve the bug –

data_app.py

Streamlit > data_app.py > ...

```python
import streamlit as st

import os
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
from wordcloud import WordCloud,STOPWORDS
import create_wordcloud

from PIL import Image

dataset_loc = "data/Tweets.csv"
image_loc = "img/airline.jpeg"
image = Image.open(image_loc)
pos_loc = "img/pos.png"
neg_loc = "img/neg.png"

st.set_option('deprecation.showPyplotGlobalUse', False)

```

data_app.py

Streamlit > data_app.py > ...

```
108
109
110
111 def main():
112
113     # sidebar
114     load_sidebar()
115
116     # Title/ text
117     st.title('Airline Sentiment Analysis')
118     st.image(image, use_column_width = True)
119     st.text('Analyze how travelers in February 2015 expressed their feelings on Twitter')
120
121     # loading the data
122     df = load_data(dataset_loc)
123
124     # display description
125     load_description(df)
126
127     # data viz
128     load_viz(df)
129
130
131
```

Final Result

Twitter US Airline Sentiment

Analyze how travelers in February 2015 expressed their feelings on Twitter

This data originally came from Crowdflower's Data for Everyone library.

Made with ❤️ 😎

Airline Sentiment Analysis

Analyze how travelers in February 2015 expressed their feelings on Twitter

Data Preview

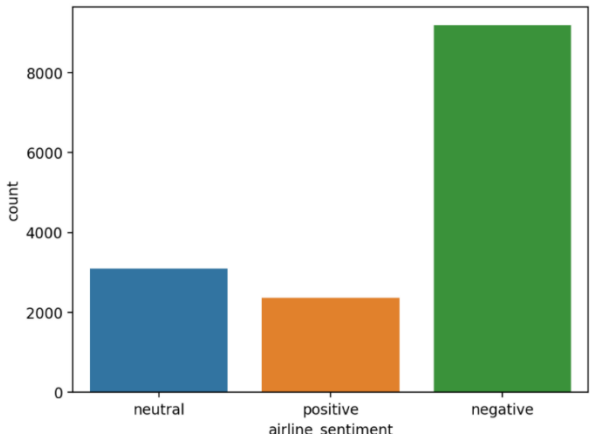
Choose Head/Tail?

Top

Data Visualisation

Seaborn - Tweet Sentiment Count

AxesSubplot(0.125,0.11;0.775x0.77)



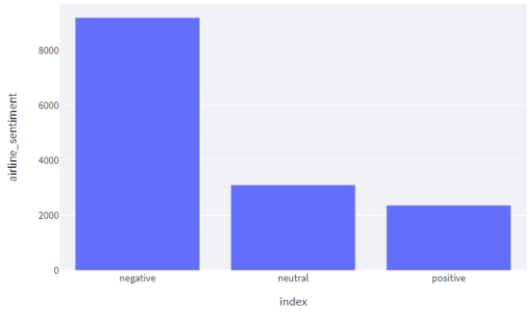
Twitter US Airline Sentiment

Analyze how travelers in February 2015 expressed their feelings on Twitter

This data originally came from Crowdflower's Data for Everyone library.

Made with ❤️ 😎

Plotly - Tweet Sentiment Count



Twitter US Airline Sentiment

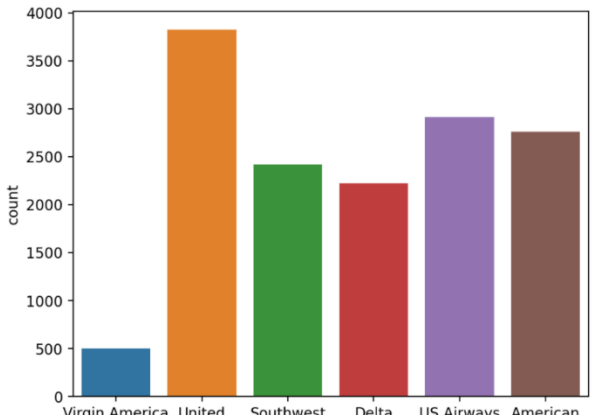
Analyze how travelers in February 2015 expressed their feelings on Twitter

This data originally came from Crowdflower's Data for Everyone library.

Made with ❤️ 😎

Airline Count

```
AxesSubplot(0.125,0.11;0.775x0.77)
```



Twitter US Airline Sentiment

Analyze how travelers in February 2015 expressed their feelings on Twitter

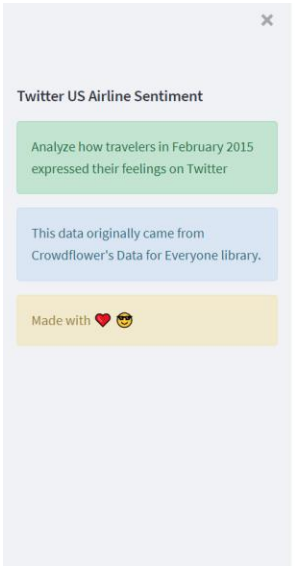
This data originally came from Crowdflower's Data for Everyone library.

Made with ❤️ 😎

Word Cloud

Choose the sentiment?

- (•) positive
- ( ) negative



Code Link - https://github.com/yamadivya/Innomatics-Research-Labs/tree/main/Streamlit