A

Project Report On

**Natural Language Processing**

Mini Project

**“Election Results Prediction by analyzing Tweets”**

SUBMITED TOWARDS THE

PARTIAL FULFILLMENT OF THE REQUIREMENTS OF

Bachelor of Computer Engineering

By

**Mr. Bhosale Bhushan Sunil Roll No - 408**

**Mr. Kunjir Shubham Ganesh Roll No - 457**

**Miss. Markad Prajakta Dhananjay Roll No - 460**

**Miss. Salvi Neha Lilachand Roll No - 478**

**Under The Guidance Of**

**Prof. Jagtap V.G**



Department Of Computer Engineering

SVPM’s College Of Engineering,Malegaon(Bk.),Baramati,Pune-413115.

SAVITRIBAI PHULE PUNE UNIVERSITY

2022-23

### 

### SVPM’s COLLEGE OF ENGINEERING,

### DEPARTMENT OF COMPUTER ENGINEERING

**CERTIFICATE**

This is to certify that the Project Entitled

### “Election Results Prediction by analyzing Tweets”

Submitted by

Mr. Bhosale Bhushan Sunil Roll No - 408

Mr. Kunjir Shubham Ganesh Roll No - 457

Miss. Markad Prajakta Dhananjay Roll No - 460

Miss. Salvi Neha Lilachand Roll No - 478

is a bonafide work carried out by Students under the supervision of Prof Jagtap V.G and it is submitted towards the partial fulfillment of the requirement of Bachelor of Computer Engineering Project.

**Prof .Jagtap V.G Dr.Sinkar Y.D**

**Internal Guide H.O.D**

**Abstract**

The Election Results Prediction project focuses on leveraging Natural Language Processing (NLP) techniques and machine learning algorithms to predict election outcomes by analysing tweets. With the increasing influence of social media platforms like Twitter in shaping public opinion, it becomes essential to harness the power of NLP to extract valuable insights from large volumes of user-generated content.

This report presents a comprehensive overview of the project, starting with the motivation behind the research and the need to predict election results using tweet analysis. The project's problem statement is defined, followed by the objectives and scope of the study. The system architecture is described, highlighting the different modules involved in collecting, pre-processing, and analysing the tweets.

To implement the Election Results Prediction system, several tools and datasets are required. Python, along with libraries like NLTK and Scikit-learn, is used for development. Access to the Twitter API enables the collection of relevant tweets, while an election-related dataset with labelled data is utilized for training and testing the sentiment analysis model.

The report also discusses the various features employed in the project. Text pre-processing features such as noise removal, tokenization, stop-word removal, and stemming/lemmatization are utilized to enhance data quality. Sentiment analysis features like Bag-of-Words, TF-IDF, N-grams, and sentiment lexicons aid in determining the sentiment of the tweets. Machine learning features, including feature selection, feature engineering, and model-based features, contribute to accurate sentiment classification and prediction.

An implementation screenshot showcasing the project's progress, such as a user interface displaying sentiment analysis results or the prediction module generating election result predictions, is included. Finally, the report concludes by highlighting the significance of the project's contributions to political analysis and decision-making based on real-time social media data.

Overall, the Election Results Prediction project demonstrates the potential of NLP techniques and machine learning algorithms in analysing tweets to predict election outcomes. By harnessing the power of sentiment analysis and feature extraction, this research aims to provide valuable insights into public sentiment and assist in making informed decisions in the context of elections.

Top of Form

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

customers. Chatbots utilize methods and algorithms from two Artificial Intelligence domains:

Natural Language Processing. However, there are many challenges and limitations in their

application. In this survey we review recent advances on chatbots, Natural Language

processing are used. . Chatbots are also known as smart bots, interactive agents, digital assistants,

or artificial conversation entities. Chatbots can mimic human conversation and entertain users but

they are not built only for this. They are useful in applications such as education, information retrieval,

business, and e-commerce. They became so popular because there are many advantages of chatbots for

users and developers too. Most implementations are platform-independent and instantly available

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

customers. Chatbots utilize methods and algorithms from two Artificial Intelligence domains:

Natural Language Processing. However, there are many challenges and limitations in their

application. In this survey we review recent advances on chatbots, Natural Language

processing are used. . Chatbots are also known as smart bots, interactive agents, digital assistants,

or artificial conversation entities. Chatbots can mimic human conversation and entertain users but

they are not built only for this. They are useful in applications such as education, information retrieval,

business, and e-commerce. They became so popular because there are many advantages of chatbots for

users and developers too. Most implementations are platform-independent and instantly available

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

customers. Chatbots utilize methods and algorithms from two Artificial Intelligence domains:

Natural Language Processing. However, there are many challenges and limitations in their

application. In this survey we review recent advances on chatbots, Natural Language

processing are used. . Chatbots are also known as smart bots, interactive agents, digital assistants,

or artificial conversation entities. Chatbots can mimic human conversation and entertain users but

they are not built only for this. They are useful in applications such as education, information retrieval,

business, and e-commerce. They became so popular because there are many advantages of chatbots for

users and developers too. Most implementations are platform-independent and instantly available

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

customers. Chatbots utilize methods and algorithms from two Artificial Intelligence domains:

Natural Language Processing. However, there are many challenges and limitations in their

application. In this survey we review recent advances on chatbots, Natural Language

processing are used. . Chatbots are also known as smart bots, interactive agents, digital assistants,

or artificial conversation entities. Chatbots can mimic human conversation and entertain users but

they are not built only for this. They are useful in applications such as education, information retrieval,

business, and e-commerce. They became so popular because there are many advantages of chatbots for

users and developers too. Most implementations are platform-independent and instantly available

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

customers. Chatbots utilize methods and algorithms from two Artificial Intelligence domains:

Natural Language Processing. However, there are many challenges and limitations in their

application. In this survey we review recent advances on chatbots, Natural Language

processing are used. . Chatbots are also known as smart bots, interactive agents, digital assistants,

or artificial conversation entities. Chatbots can mimic human conversation and entertain users but

they are not built only for this. They are useful in applications such as education, information retrieval,

business, and e-commerce. They became so popular because there are many advantages of chatbots for

users and developers too. Most implementations are platform-independent and instantly available

Chatbots are intelligent conversational computer systems designed to mimic human

conversation to enable automated online guidance and support. The increased benefits of

chatbots led to their wide adoption by many industries in order to provide virtual assistance to

customers. Chatbots utilize methods and algorithms from two Artificial Intelligence domains:

Natural Language Processing. However, there are many challenges and limitations in their

application. In this survey we review recent advances on chatbots, Natural Language

processing are used. . Chatbots are also known as smart bots, interactive agents, digital assistants,

or artificial conversation entities. Chatbots can mimic human conversation and entertain users but

they are not built only for this. They are useful in applications such as education, information retrieval,

business, and e-commerce. They became so popular because there are many advantages of chatbots for

users and developers too. Most implementations are platform-independent and instantly available

**Contents:**

1.Introduction

2.Problem Statement

3.Objectives and scope

4.System Architecture

5.Tools and Dataset required

6.Implementation screenshot

7.Conclusion

**Introduction**

**1.1 Motivation**

Chapter 1

Introduction

1

Chapter 2

Literature Survey

2-4

Chapter 3

Problem Statement

5

Chapter 4

Objectives and scope

6

Chapter 5

Proposed methodology

7

Chapter 6

System Architecture

8

Chapter 7

Tools and Dataset required

9

Chapter 8

Implementation screenshot

10-15

Chapter 9

Conclusion

Chapter 1

Introduction

1

Chapter 2

Literature Survey

2-4

Chapter 3

Problem Statement

5

Chapter 4

Objectives and scope

6

Chapter 5

Proposed methodology

7

Chapter 6

System Architecture

8

Chapter 7

Tools and Dataset required

9

Chapter 8

Implementation screenshot

10-15

Chapter 9

Conclusion

The motivation behind this project is to leverage Natural Language Processing (NLP) techniques and machine learning algorithms to predict election results by analysing tweets. With the growing influence of social media platforms, such as Twitter, in shaping public opinion, it becomes crucial to harness the power of NLP to extract valuable insights from large volumes of user-generated content.

**1.2 Need of Problem**

The need for predicting election results using tweet analysis arises from the desire to gain real-time insights into public sentiment and gauge the popularity of political candidates. By analysing the sentiments expressed in tweets, we can identify trends, potential winners, and understand the factors influencing the outcome of an election.

**Problem Statement**

**2.1 Problem Statement**

The objective of this project is to develop a system that can predict election results by analyzing tweets related to the candidates. The system will utilize NLP techniques and machine learning algorithms to classify tweets into sentiment categories and aggregate the results to generate predictions.

**2.2 Features**

1. In the Election Results Prediction project, various features play a crucial role in extracting valuable insights from tweets and predicting election outcomes. The following features are utilized in the project:

**2.2.1 Text Preprocessing Features:**

* Noise Removal: Removing irrelevant characters, symbols, URLs, and hashtags from the tweets to reduce noise and improve the quality of the data.
* Tokenization: Splitting the text into individual words or tokens to facilitate further analysis and feature extraction.
* Stop-word Removal: Eliminating common and non-informative words (e.g., "and," "the," "is") that do not contribute much to the sentiment analysis process.
* Stemming and Lemmatization: Reducing words to their base or root form (e.g., "running" to "run") to handle different word variations and improve feature representation.

**2.2.2 Sentiment Analysis Features:**

* Bag-of-Words (BoW): Representing tweets as a collection of words, disregarding grammar and word order, to create a numerical feature vector for sentiment analysis.
* Term Frequency-Inverse Document Frequency (TF-IDF): Assigning weights to words based on their frequency in the tweet and rarity in the entire dataset, helping to identify important and distinctive terms.
* N-grams: Considering sequences of adjacent words as features (e.g., "good" and "not good") to capture contextual information and improve sentiment classification accuracy.
* Sentiment Lexicons: Using pre-built sentiment lexicons (e.g., AFINN, Sent WordNet) that associate words with sentiment scores to determine the overall sentiment of a tweet.

**2.2.3 Machine Learning Features:**

* Feature Selection: Identifying the most informative and relevant features from the dataset to reduce dimensionality and enhance the performance of the machine learning model.
* Feature Engineering: Creating new features based on domain knowledge or specific characteristics of the tweets that may contribute to sentiment analysis and election result prediction.
* Model-Based Features: Extracting features from pre-trained language models (e.g., BERT, GPT) that capture semantic information and contextual understanding of the text.

These features collectively contribute to the accuracy and effectiveness of sentiment analysis and prediction models in the Election Results Prediction project. Proper selection, pre-processing, and utilization of these features are crucial for obtaining meaningful insights from tweets and making accurate predictions regarding election outcome

**2.3 Objectives**

1. The main objectives of this project are as follows:

* Collect a large dataset of tweets related to the election.
* Pre-process the dataset by removing noise, handling text normalization, and feature extraction.
* Train a machine learning model using the pre-processed data to classify tweets into sentiment categories (positive, negative, neutral).
* Analyse the classified tweets to understand public sentiment towards candidates.
* Utilize the sentiment analysis results to predict election outcomes.

**2.4 Scope**

The scope of this project includes:

* Developing a Python-based system for tweet collection, pre-processing, and sentiment analysis.
* Exploring and implementing various NLP techniques for feature extraction and sentiment classification.
* Evaluating the accuracy and performance of the trained machine learning model.
* Generating predictions for election results based on sentiment analysis.

Problem statement

All the companies want their clients to know about them and they can get more projects or sell

their products. So the most primary way to do this is to have a good interactive website in the

advanced world of digital marketing. But in this advanced world, just having a website is not enough

as it does not cover all the details and can not solve the queries of the clients. Maybe a comment

section is an answer, but it's tedious work, and 'time is money' your client will not wait too long for

your response. And you cannot assign humans to answer everyone at the same time twenty-four by

seven. So for that, you need a chatbot, not a normal one "An

Intelligent Chatbot= with Artificial Intelligence technology. <Artificial intelligence chatbot is a

technology that makes interactions between man and machines using natural language possible. A

chatbot can give different responses from the same input given by the user according to the current

conversation issue". By using our "Intelligent Chatbot" you can overcome all the above-given issues,

you do not need humans to do manual work, your clients will be happy. A chatbot is a conventional

agent that is capable to communicate with operators by using natural languages. As numerous

chatbot platforms already exist, there are still some problems in building data-driven system because a

huge amount of data is required for their developmen

**System Architecture**

**Entity Classifier**

**Result**

**Joe Biden**

**Donald Trump**

**Compare**

**Sentiment Classifier 2**

**Sentiment Classifier 1**

**Implementation screenshot**

**Program Code:-**

import pandas as pd

import numpy as np

import seaborn as sns

import textblob

import wordcloud

import matplotlib.pyplot as plt

from textblob import TextBlob

from wordcloud import WordCloud

import plotly.graph\_objects as go

import plotly.express as px

trump\_reviews = pd.read\_csv("/Users/shubham/Desktop/NLP/Trumpall2.csv")

biden\_reviews = pd.read\_csv("/Users/shubham/Desktop/NLP/Bidenall2.csv")

print(trump\_reviews.head())

print(biden\_reviews.head())

textblob1 = TextBlob(trump\_reviews["text"][10])

print("Trump :",textblob1.sentiment)

textblob2 = TextBlob(biden\_reviews["text"][500])

print("Biden :",textblob2.sentiment)

def find\_pol(review):

return TextBlob(review).sentiment.polarity

trump\_reviews["Sentiment Polarity"] = trump\_reviews["text"].apply(find\_pol)

print(trump\_reviews.tail())

biden\_reviews["Sentiment Polarity"] = biden\_reviews["text"].apply(find\_pol)

print(biden\_reviews.tail())

trump\_reviews["Expression Label"] = np.where(trump\_reviews["Sentiment Polarity"]>0, "positive", "negative")

trump\_reviews["Expression Label"][trump\_reviews["Sentiment Polarity"]==0]="Neutral"

print(trump\_reviews.tail())

biden\_reviews["Expression Label"] = np.where(biden\_reviews["Sentiment Polarity"]>0, "positive", "negative")

biden\_reviews["Expression Label"][trump\_reviews["Sentiment Polarity"]==0]="Neutral"

print(biden\_reviews.tail())

#reviews1

reviews1 = trump\_reviews[trump\_reviews['Sentiment Polarity'] == 0.0000]

print(reviews1.shape)

cond1=trump\_reviews['Sentiment Polarity'].isin(reviews1['Sentiment Polarity'])

trump\_reviews.drop(trump\_reviews[cond1].index, inplace = True)

print(trump\_reviews.shape)

#reviews1

reviews2 = biden\_reviews[biden\_reviews['Sentiment Polarity'] == 0.0000]

print(reviews2.shape)

cond2=biden\_reviews['Sentiment Polarity'].isin(reviews1['Sentiment Polarity'])

biden\_reviews.drop(biden\_reviews[cond2].index, inplace = True)

print(biden\_reviews.shape)

# Donald Trump

np.random.seed(10)

remove\_n =324

drop\_indices = np.random.choice(trump\_reviews.index, remove\_n, replace=False)

df\_subset\_trump = trump\_reviews.drop(drop\_indices)

print(df\_subset\_trump.shape)

# Joe Biden

np.random.seed(10)

remove\_n =31

drop\_indices = np.random.choice(biden\_reviews.index, remove\_n, replace=False)

df\_subset\_biden = biden\_reviews.drop(drop\_indices)

print(df\_subset\_biden.shape)

count\_1 = df\_subset\_trump.groupby('Expression Label').count()

print(count\_1)

negative\_per1 = (count\_1['Sentiment Polarity'][0]/1000)\*10

positive\_per1 = (count\_1['Sentiment Polarity'][1]/1000)\*100

count\_2 = df\_subset\_biden.groupby('Expression Label').count()

print(count\_2)

negative\_per2 = (count\_2['Sentiment Polarity'][0]/1000)\*100

positive\_per2 = (count\_2['Sentiment Polarity'][1]/1000)\*100

Politicians = ['Joe Biden', 'Donald Trump']

lis\_pos = [positive\_per1, positive\_per2]

lis\_neg = [negative\_per1, negative\_per2]

fig = go.Figure(data=[

go.Bar(name='Positive', x=Politicians, y=lis\_pos),

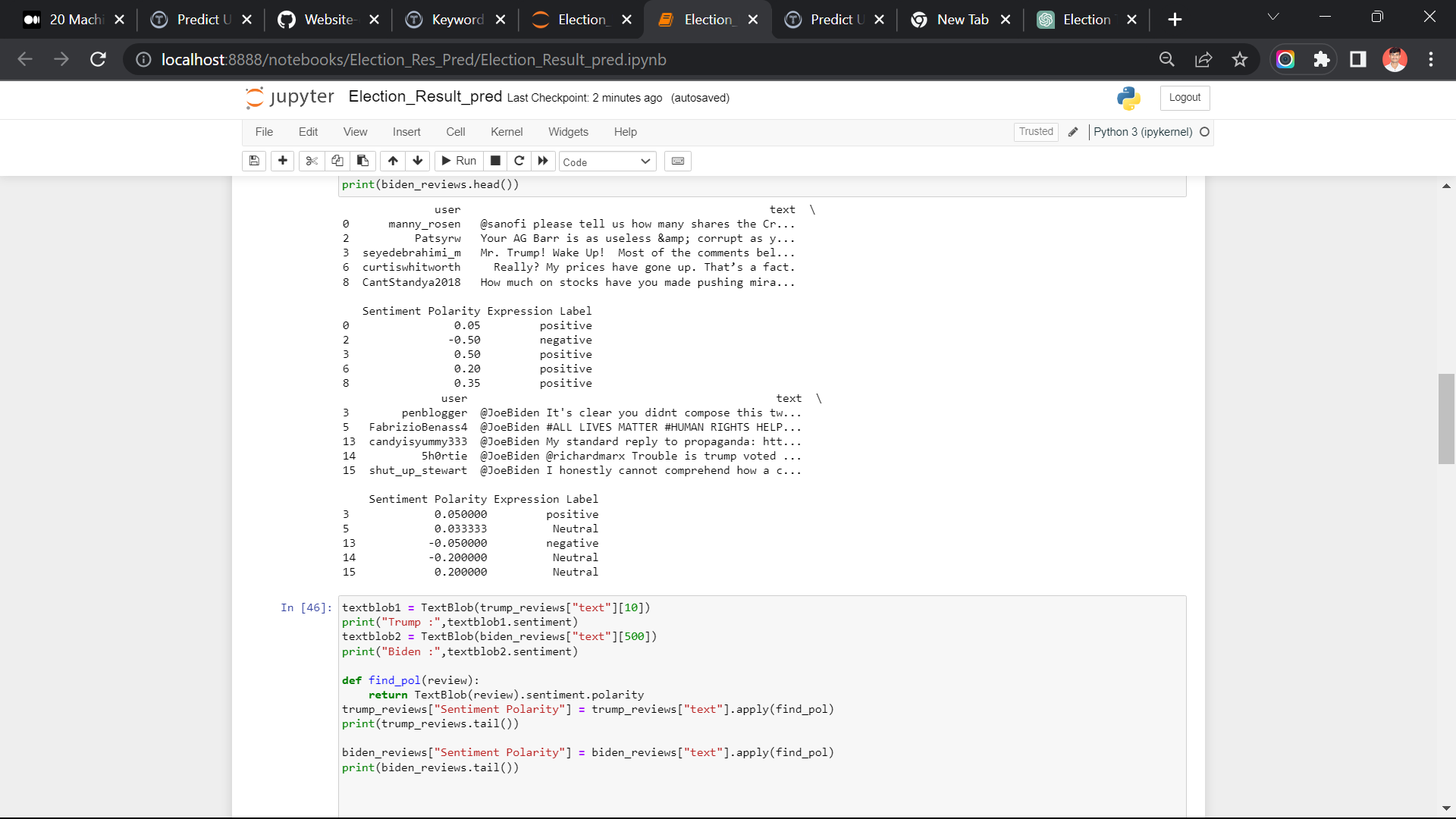
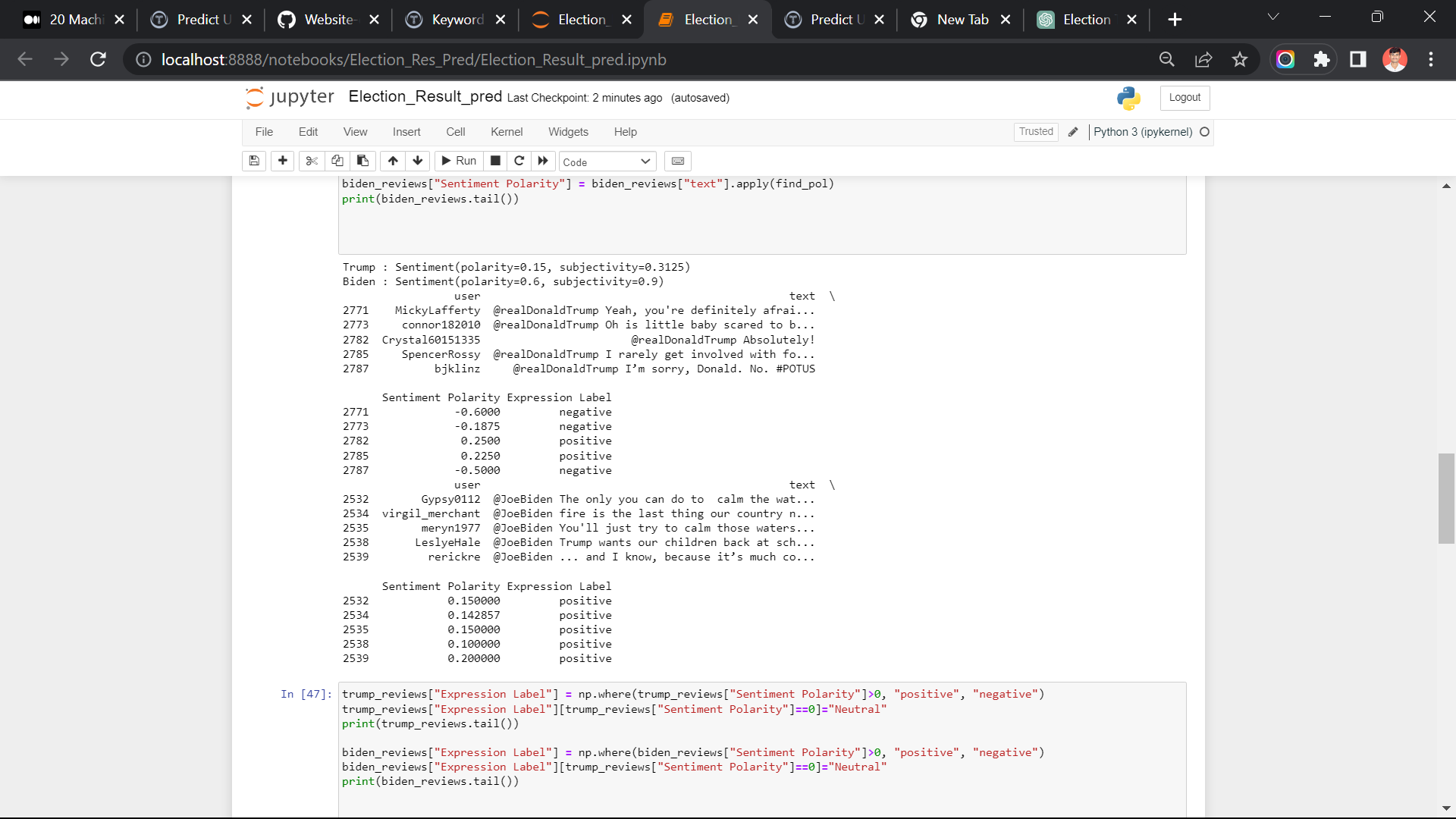
go.Bar(name='Negative', x=Politicians, y=lis\_neg)

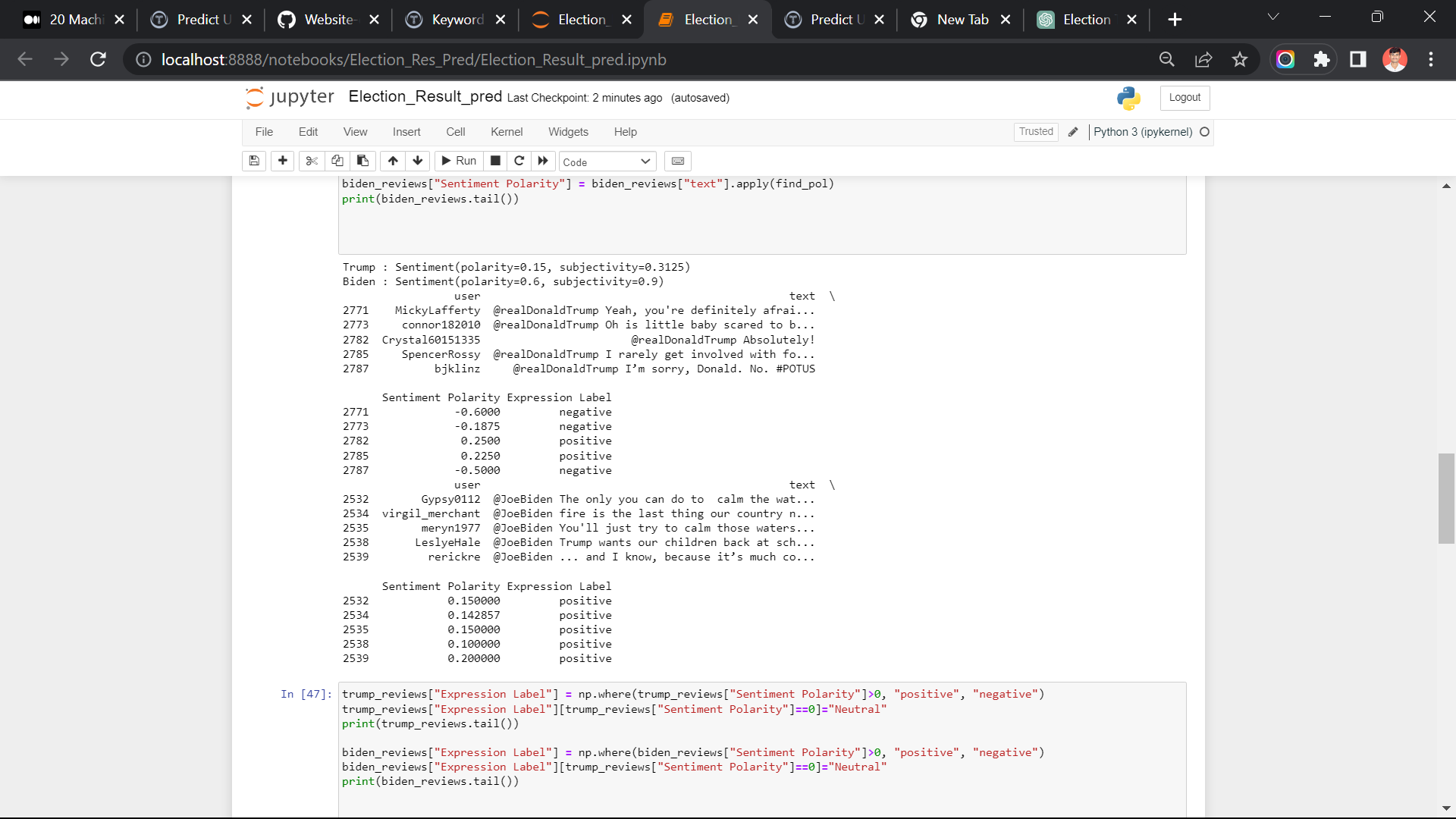
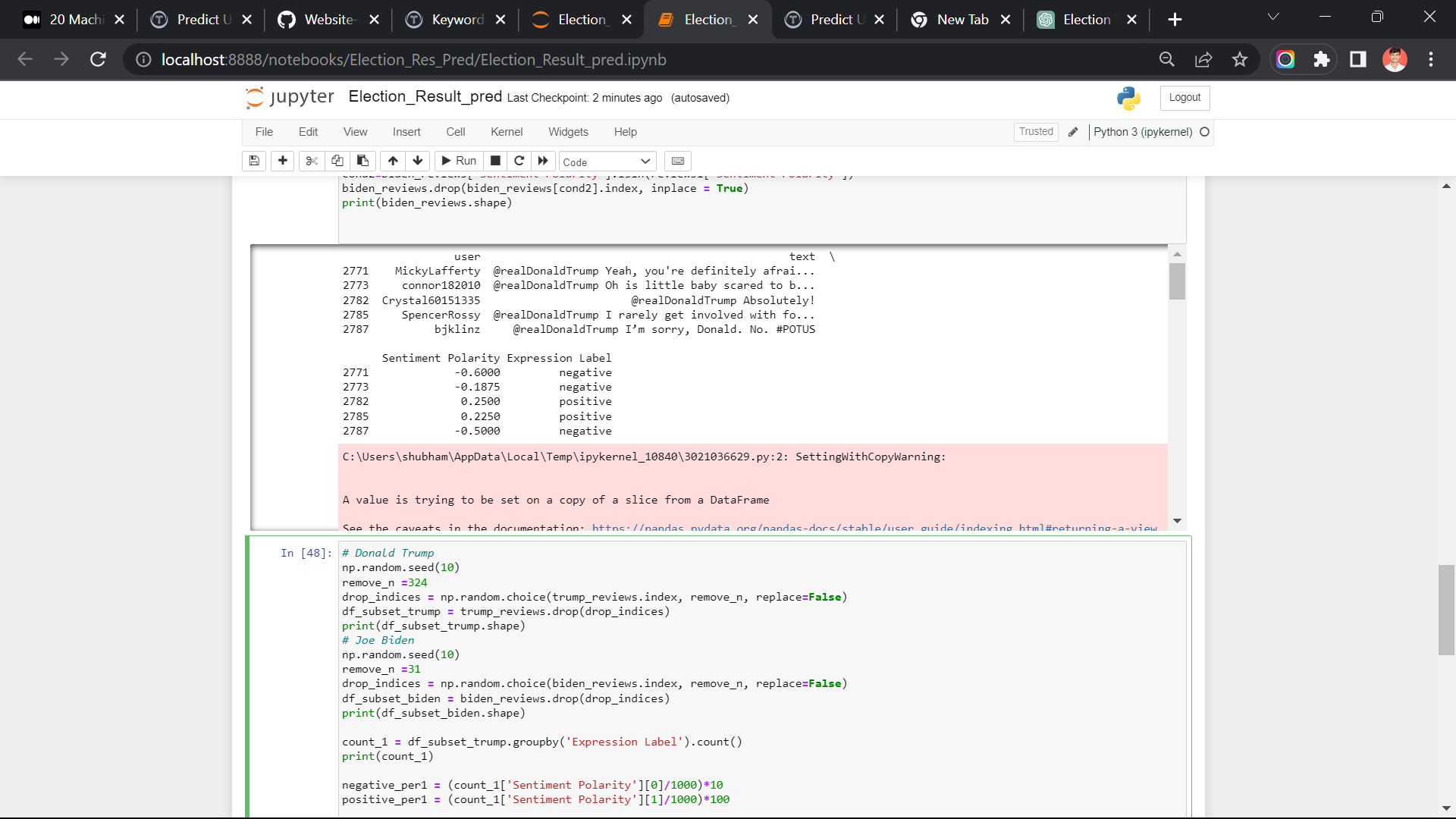
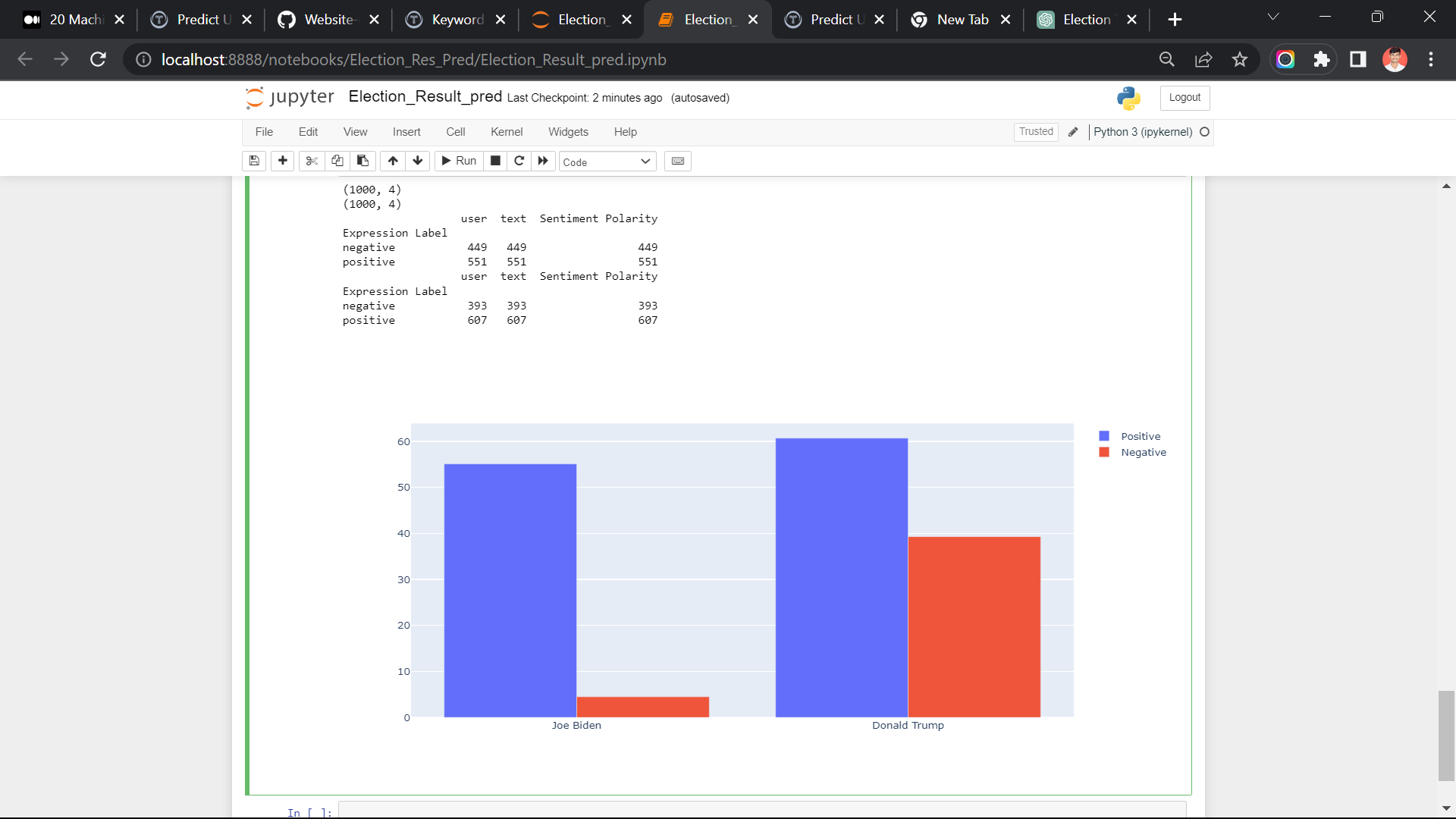
])

# Change the bar mode

fig.update\_layout(barmode='group')

fig.show()

**Output:-**

****

**Conclusion**

In conclusion, this project aims to predict election results by analyzing tweets using NLP techniques and machine learning algorithms. By harnessing the power of NLP and sentiment analysis, we can gain valuable insights into public sentiment and utilize that information for predicting election outcomes. The successful implementation of this project can contribute to the field of political analysis and assist in making informed decisions based on real-time social media data.

Minimal human interference in the use of devices is the goal of our world of

technology. Chatbots can reach out to a broad audience on messaging apps and be more

effective than humans are. At the same time, they may develop into a capable information

gathering tool. They provide significant savings in the operation of customer service

departments. With further development of AI and machine learning, somebody may not be

capable of understanding whether he talks to a chatbot or a real-life agent. The Chatbot must

be simple and conventional.

A chatbot is one of the simple ways to transport data from a computer without having

to think for proper keywords to look up in a search or browse several web pages to collect

information; users can easily type their query in natural language and retrieve information.

In this paper, information about the design, implementation of the chatbot has been

presented. From the survey above, it can be said that the development and improvement

of chatbot design grow at an unpredictable rate due to variety of methods and approaches used

to design a chatbot. Chatbot is a great tool for quick interaction with the user. They help

us by providing entertainment, saving time and answering the questions that are hard to

find.