

election-prediction

December 24, 2023

1 *ELECTION SENTIMENT ANALYSIS*

```
[ ]: import numpy as np
import pandas as pd
from textblob import TextBlob
import plotly.graph_objects as go
import plotly.express as xp
```

```
[ ]: modi=pd.read_csv("/content/modi_reviews.csv")
rahul=pd.read_csv("/content/rahul_reviews.csv")
```

```
[ ]: modi.head()
```

```
[ ]:      Unnamed: 0      User \
0      0  advosushildixit
1      1      jiaeur
2      2  PVenkatGandhi
3      3  TheNirbhay1
4      4  ShakeChilli

                                Tweet
0  @anjanaomkashyap I am seeing you as future #bj...
1  #LokSabhaElections2019 \n23rd May 2019 will re...
2  #LokSabhaElections2019 \n23rd May 2019 will re...
3  PM Modi creates a new record of being the only...
4  @abhijitmajumder Appointment of Successor! \n\...
```

```
[ ]: rahul.head()
```

```
[ ]:      Unnamed: 0      User \
0      0  Sunnysweet16
1      1  drnitinchaube
2      2  mrvivek07
3      3  JosephPravinP
4      4  VandanaMegastar

                                Tweet
```

```

0 Wonder why no academic or journalist asks INC ...
1 Congrats for the change #australiavotes2019 an...
2 Peopel Say "Govt Ne 70 Years Kya kiya".\nUnse ...
3 @ajaymaken @RahulGandhi And as a final touch, ...
4 #LokSabhaElections2019 Anyone not having mass ...

```

```
[ ]: modi.shape
```

```
[ ]: (25688, 3)
```

```
[ ]: rahul.shape
```

```
[ ]: (14261, 3)
```

```
[ ]: modi['Tweet'][0]
```

```
[ ]: '@anjanaomkashyap I am seeing you as future #bjp spokesperson..\nGood luck.
Anjana Om Modi oops Kashyap.\nJournalists like you changed the meaning of
journalism. Janta maaf nai karege'
```

```
[ ]: TextBlob(modi['Tweet'][0]).sentiment
```

```
[ ]: Sentiment(polarity=0.35, subjectivity=0.36250000000000004)
```

```
[ ]: TextBlob(rahul['Tweet'][0]).sentiment
```

```
[ ]: Sentiment(polarity=0.21875, subjectivity=0.4017857142857143)
```

```
[ ]: def find_polarity(review):
      return TextBlob(review).sentiment.polarity
```

```
[ ]: find_polarity('I hate living without coding')
```

```
[ ]: -0.8
```

```
[ ]: modi['Tweet'] = modi['Tweet'].astype(str)
      rahul['Tweet'] = rahul['Tweet'].astype(str)

      def find_polarity(review):
          return TextBlob(review).sentiment.polarity
```

```
[ ]: """modi['Tweet'] = modi['Tweet'].astype(str)
      rahul['Tweet'] = rahul['Tweet'].astype(str)
      These lines ensure that the 'Tweet' column in both the modi and rahul_
      ↪DataFrames is converted to strings using the astype method.
      This step is helpful to ensure that all values in the 'Tweet' column are_
      ↪treated as strings, even if they were originally of a
```

```
different data type."""
```

```
[ ]: "modi['Tweet'] = modi['Tweet'].astype(str)\nrahul['Tweet'] = rahul['Tweet'].astype(str)\nThese lines ensure that the 'Tweet' column in both the modi and rahul DataFrames is converted to strings using the astype method.\nThis step is helpful to ensure that all values in the 'Tweet' column are treated as strings, even if they were originally of a\n different data type."
```

```
[ ]: modi['Polarity'] = modi['Tweet'].apply(find_polarity)\nrahul['Polarity'] = rahul['Tweet'].apply(find_polarity)
```

```
[ ]: modi
```

```
[ ]:      Unnamed: 0      User \
0          0  advosushildixit
1          1      jiaeur
2          2  PVenkatGandhi
3          3  TheNirbhay1
4          4  ShakeChilli
...
25683      25683      shv_indian
25684      25684      SONUPINKOO
25685      25685      SONUPINKOO
25686      25686      SONUPINKOO
25687      25687  parthajit0101

      Tweet  Polarity
0  @anjanaomkashyap I am seeing you as future #bj...  0.350000
1  #LokSabhaElections2019 \n23rd May 2019 will re...  0.800000
2  #LokSabhaElections2019 \n23rd May 2019 will re...  0.800000
3  PM Modi creates a new record of being the only...  0.312121
4  @abhijitmajumder Appointment of Successor! \n\...  0.098788
...
25683  #1DDrive #ModiSpeaksToNews18 #LokSabhaElection...  0.000000
25684  I am seriously worried about future of Indian ...  0.209524
25685  If @narendramodi is so Bad then Why does @INCI... -0.700000
25686  On this 23rd March (Pakistan Day), I Urge all ...  0.000000
25687  @timesofindia #ArvindKejriwal, #AAP sirji pls ... -0.213542

[25688 rows x 4 columns]
```

```
[ ]: modi['Polarity'].head()
```

```
[ ]: 0    0.350000
1    0.800000
2    0.800000
3    0.312121
```

```
4    0.098788
Name: Polarity, dtype: float64
```

```
[ ]: rahul
```

```
[ ]:      Unnamed: 0      User \
0          0      Sunnysweet16
1          1      drnitinchaube
2          2      mrvivek07
3          3      JosephPravinP
4          4      VandanaMegastar
...
14256      14256      SunjayJK
14257      14257      SunjayJK
14258      14258      RAMANKAIRA
14259      14259      RAMANKAIRA
14260      14260      RAMANKAIRA

      Tweet Polarity
0      Wonder why no academic or journalist asks INC ...  0.218750
1      Congrats for the change #australiavotes2019 an...  0.000000
2      Peopel Say "Govt Ne 70 Years Kya kiya".\nUnse ...  0.000000
3      @ajaymaken @RahulGandhi And as a final touch, ...  0.320000
4      #LokSabhaElections2019 Anyone not having mass ... -0.166667
...
14256      @quizzicalguy In this #LokSabhaElections2019, ...  0.000000
14257      @AnumaVidisha @RahulGandhi @ArvindKejriwal whe... -0.500000
14258      @sherryontopp Early morning #political move af...  0.080000
14259      @mayankgandhi04 @mallesh_2004 Early morning #p...  0.080000
14260      @DrKumarVishwas Early morning #political move ...  0.080000

[14261 rows x 4 columns]
```

```
[ ]: rahul['Polarity'].head()
```

```
[ ]: 0    0.218750
1    0.000000
2    0.000000
3    0.320000
4   -0.166667
Name: Polarity, dtype: float64
```

```
[ ]: import numpy as np

modi['Label'] = np.where(modi['Polarity'] > 0, 'positive', 'negative')
modi.loc[modi['Polarity'] == 0, 'Label'] = 'Neutral'
```

```
rahul['Label'] = np.where(rahul['Polarity'] > 0, 'positive', 'negative')
rahul.loc[rahul['Polarity'] == 0, 'Label'] = 'Neutral'
```

```
[ ]: rahul.head()
```

```
[ ]:      Unnamed: 0      User \
0          0      Sunnysweet16
1          1      drnitinchaube
2          2          mrvivek07
3          3      JosephPravinP
4          4      VandanaMegastar
```

	Tweet	Polarity	Label
0	Wonder why no academic or journalist asks INC ...	0.218750	positive
1	Congrats for the change #australiavotes2019 an...	0.000000	Neutral
2	Peopel Say "Govt Ne 70 Years Kya kiya".\nUnse ...	0.000000	Neutral
3	@ajaymaken @RahulGandhi And as a final touch, ...	0.320000	positive
4	#LokSabhaElections2019 Anyone not having mass ...	-0.166667	negative

```
[ ]: modi.head()
```

```
[ ]:      Unnamed: 0      User \
0          0  advosushildixit
1          1          jiaeur
2          2      PVenkatGandhi
3          3      TheNirbhay1
4          4      ShakeChilli
```

	Tweet	Polarity	Label
0	@anjanaomkashyap I am seeing you as future #bj...	0.350000	positive
1	#LokSabhaElections2019 \n23rd May 2019 will re...	0.800000	positive
2	#LokSabhaElections2019 \n23rd May 2019 will re...	0.800000	positive
3	PM Modi creates a new record of being the only...	0.312121	positive
4	@abhijitmajumder Appointment of Successor! \n\...	0.098788	positive

2 Removing neutral coloumn

```
[ ]: neutral_modi = modi[modi['Polarity']==0.0000]
neutral_modi
```

```
[ ]:      Unnamed: 0      User \
6          6      Mdsr20351488
9          9      TheShobhitAzad
10         10          sheikh541
20         20          UdayiCODE
21         21      SandipGhose
```

```

...
25674      25674      CrazySatire
25676      25676      SunjayJK
25682      25682      CrazySatire
25683      25683      shv_indian
25686      25686      SONUPINKOO

```

		Tweet	Polarity	Label
6	After 5yrs of Modi \n\n-SC vs SC\n-EC vs EC\n...		0.0	Neutral
9	#AzadPrediction\n#LokSabhaElections2019 \n\nBJ...		0.0	Neutral
10	#BJP was renamed as Safroon Party then again i...		0.0	Neutral
20	I am seeing 3 Options for PM this time #Modi #...		0.0	Neutral
21	@thakkar_sameet @BDUTT @surjitbhallla 4/n smack...		0.0	Neutral
...				
25674	Poll:\nWhich province (s) of Pakistan after di...		0.0	Neutral
25676	@IASassociation MT:We stand by our @IPS_Associ...		0.0	Neutral
25682	No one calling that 'BJP is communal' in this ...		0.0	Neutral
25683	#1DDrive #ModiSpeaksToNews18 #LokSabhaElection...		0.0	Neutral
25686	On this 23rd March (Pakistan Day), I Urge all ...		0.0	Neutral

[8207 rows x 5 columns]

```

[ ]: neutral_modi = modi[modi['Polarity']==0.0000]

remove_neutral_modi = modi['Polarity'].isin(neutral_modi['Polarity'])
modi.drop(modi[remove_neutral_modi].index,inplace=True)
print(neutral_modi.shape)
print(modi.shape)

```

(8207, 5)

(17481, 5)

```

[ ]: neutral_rahul = rahul[rahul['Polarity']==0.0000]
remove_neutral_rahul = rahul['Polarity'].isin(neutral_rahul['Polarity'])
rahul.drop(rahul[remove_neutral_rahul].index,inplace=True)
print(neutral_rahul.shape)
print(rahul.shape)

```

(4894, 5)

(9367, 5)

```

[ ]: print(modi.shape)
print(rahul.shape)

```

(17481, 5)

(9367, 5)

```
[ ]: #modi
np.random.seed(10)
remove_n = 8481
drop_indices = np.random.choice(modi.index,remove_n,replace=False)
df_modi = modi.drop(drop_indices)
```

```
[ ]: #rahul
np.random.seed(10)
remove_n = 367
drop_indices1 = np.random.choice(rahul.index,remove_n,replace=False)
df_rahul = rahul.drop(drop_indices1)
```

```
[ ]: print(df_modi.shape)
print(df_rahul.shape)
```

```
(9000, 5)
(9000, 5)
```

3 Prediction about Election

```
[ ]: df_modi.groupby('Label').count()
```

```
[ ]:      Unnamed: 0  User  Tweet  Polarity
Label
negative      2600  2600   2600      2600
positive      6400  6400   6400      6400
```

```
[ ]: df_rahul.groupby('Label').count()
```

```
[ ]:      Unnamed: 0  User  Tweet  Polarity
Label
negative      3039  3039   3039      3039
positive      5961  5961   5961      5961
```

```
[ ]: modi_count = df_modi.groupby('Label').count()
neg_modi = (modi_count['Polarity'][0] / 1000) * 100
pos_modi = (modi_count['Polarity'][1] / 1000) * 100
```

```
[ ]: rahul_count = df_rahul.groupby('Label').count()
neg_rahul = (rahul_count['Polarity'][0] / 1000) * 100
pos_rahul = (rahul_count['Polarity'][1] / 1000) * 10
```

```
[ ]: modi_count
```

```
[ ]:      Unnamed: 0  User  Tweet  Polarity
Label
negative      2600  2600   2600      2600
```

positive	6400	6400	6400	6400
----------	------	------	------	------

```
[ ]: rahul_count
```

```
[ ]:      Unnamed: 0  User  Tweet  Polarity
Label
negative      3039  3039   3039     3039
positive      5961  5961   5961     5961
```

```
[ ]: import plotly.graph_objects as go

politicians = ['Modi', 'Rahul']

neg_list = [neg_modi, neg_rahul]
pos_list = [pos_modi, pos_rahul] # Assuming you have a pos_rahul variable

fig = go.Figure(
    data=[
        go.Bar(name='Negative', x=politicians, y=neg_list, marker_color='red'),
        go.Bar(name='Positive', x=politicians, y=pos_list, marker_color='green')
    ]
)

fig.update_layout(barmode='group')
fig.show()
```

3.1 THANKS (AMAN)

<https://www.linkedin.com/in/aman-6570871b5/>