BiPol

Classifying Political Bias from Tweets

Index

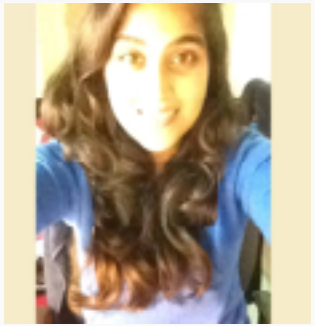
1. Team
2. Motivation
3. Use Case & Problem Statement
4. Challenges
5. Implementation
    - Architecture
    - Tools
    - Attributes
6. Demo + UI

Team

Saloni

Shivangi

Akbar



Anavi

Motivation for Study

- Number of major celebrities using twitter
- These have large follower base
- Most used tool by politicians for political communication
- Important for voters to know - what is their behavior like?

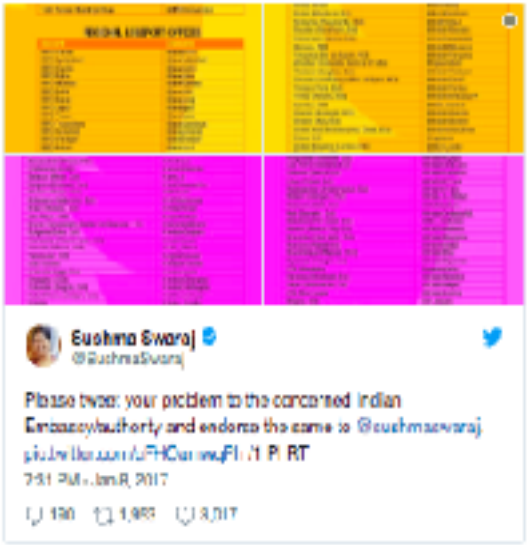
BiPol: Tool for Use by News Media

- Most news articles quote politician's Twitter handles
- Good tool for journalists/ researchers who want to study polarization on social media
- Pun on bye-polls ☺
- Goal: Detect Political Bias in Tweets

External affairs minister Sushma Swaraj on Sunday said Indians abroad facing problems can tweet to the respective embassy and tag her in the tweet, adding that she will personally monitor the missions' responses.

The minister, who has been known for reaching out to citizens through Twitter, said a 'new pattern' is being adopted.

"We are adopting a new pattern here. Pl (please) tweet your problem to the concerned Indian Embassy/authority and endorse the same to @sushmaswaraj.



Detecting Political Bias: Major Challenges

Is this bias?

Major Challenges

How to remove links?

Handling non-text media?

Major Challenges

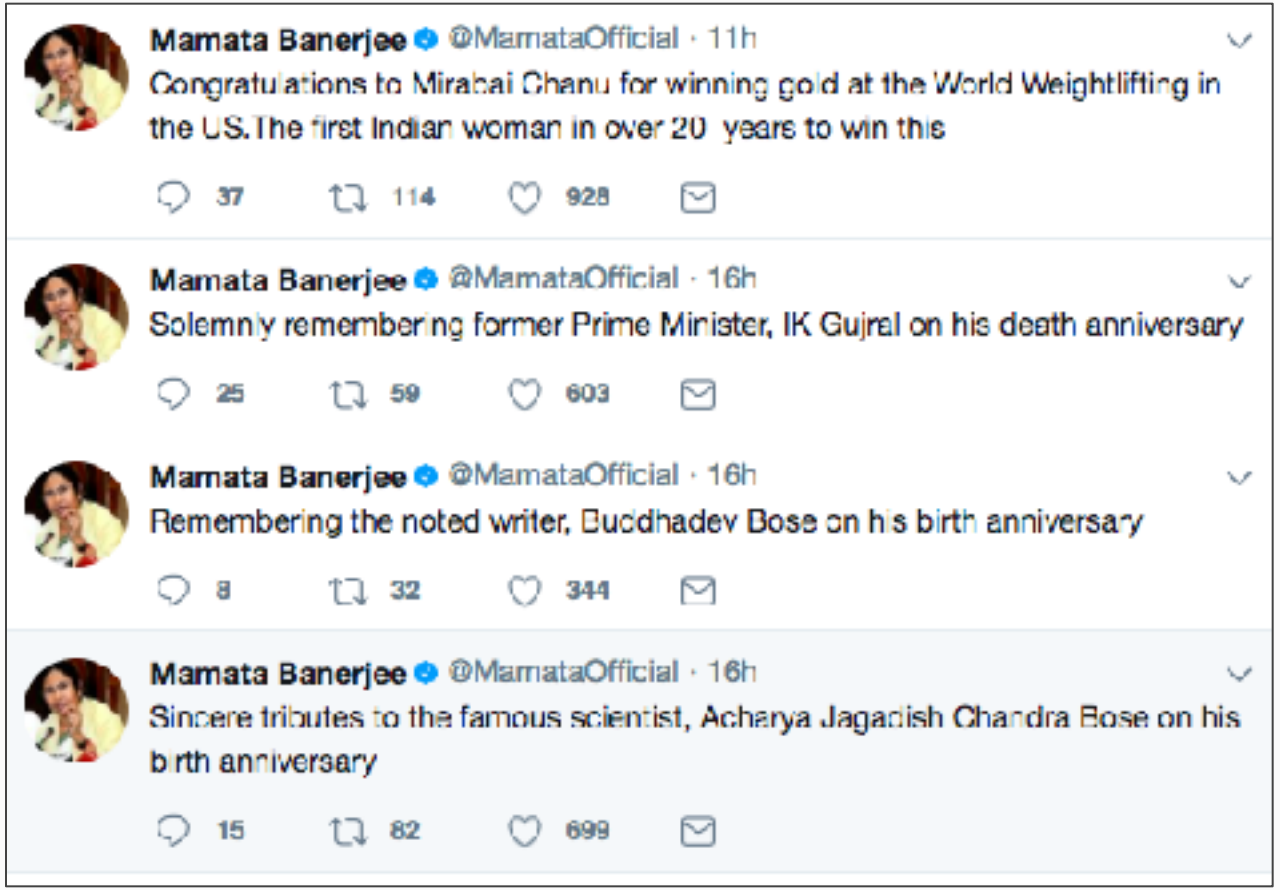
How to understand his ‘Mann ki baat’?

Findings – Politician's Behavior on Twitter

Sushma Swaraj replies actively to queries

Findings - Politician's Behavior on Twitter

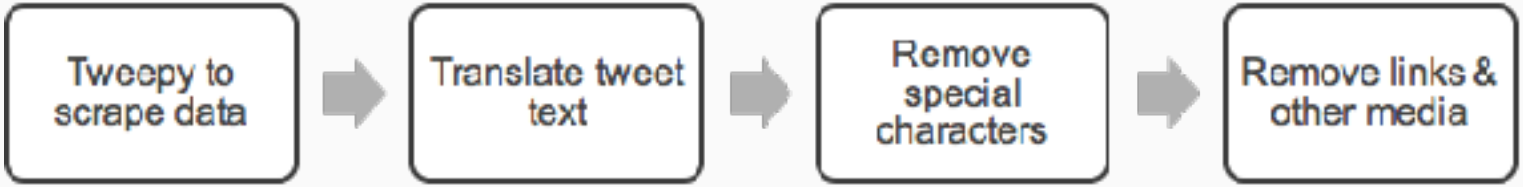
Mamata Banerjee typically tweets about anniversaries & birthday greetings



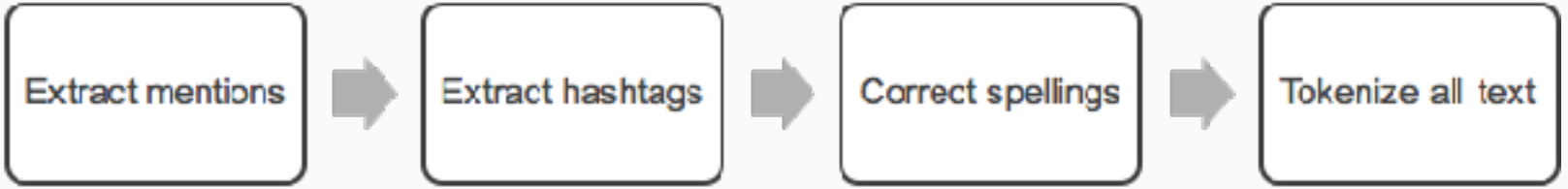
Implementation

Data Cleaning

Data Sourcing



Data Cleaning & Collating



Attributes

Attributes Used

Attribute Name	Description	Example
Politician Handles (y/n)	Presence of the handle of a politician https://surendran.info/twitter_handles_politicians_india/	Eg. “wonderful step by shri @narendramodi #DeMo” Presence → Bias
Controlled Vocab (y/n)	Presence of list of words (stop list) that could be indicators of political bias	Eg. “corruption”, “black money”, “communal”, “RSS” Presence → Bias
Party Mentions (y/n)	Presence of political party	Eg. ‘BJP’, ‘NDA’ Presence → Bias
Policy Mentions (y/n)	Presence of government policies	Eg. “Swachh Bharat”, “Jan Dhan” Presence → Bias

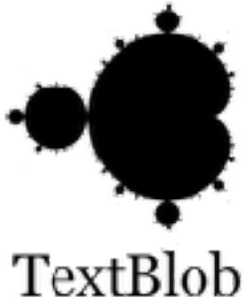
Attributes Used

Attribute Name	Description	Example
PM Mention (y/n)	Presence of variations of 'PM'	Eg. "namo", "modi", "Prime Minister" Presence → Bias
Politician Names	Presence of just names (not handles) of politicians Scraped data of Lok Sabha and Rajya Sabha	Eg. "Rahul Gandhi" Presence → Bias
Greetings (y/n)	Presence of any kind of greeting (congratulatory, death, condolence)	Presence → Low (or no) bias
News Bias (y/n)	Presence of news handles, interpolated with their bias (0,1)	Eg. "@IndiaToday", "@TimesNow" Presence → Bias
Polarity Score (−1,1)	> 0 → Positive > 0 → Negative 0 → Neutral	Positive/Negative → Bias

Data Scraping Sources– For Attribute Creation

1. List of twitter handles for politicians
2. List of all politician names
3. List of all policy names (bigrams & trigrams used)
4. Names of all Political Parties - Election Commission Website

Tools/APIs


TextBlob

What is Watson
Natural Language
Understanding?
NLP for advanced
text analysis
IBM

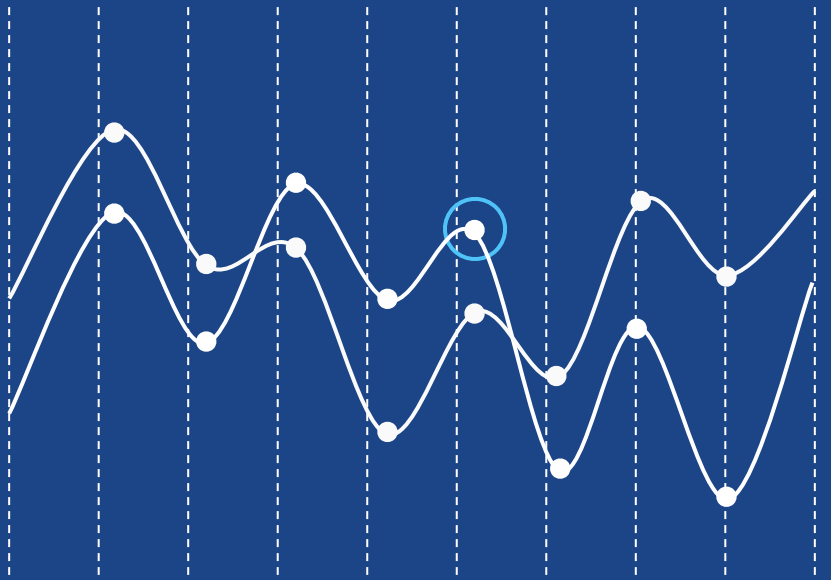
Natural Language Analysis
with Python NLTK

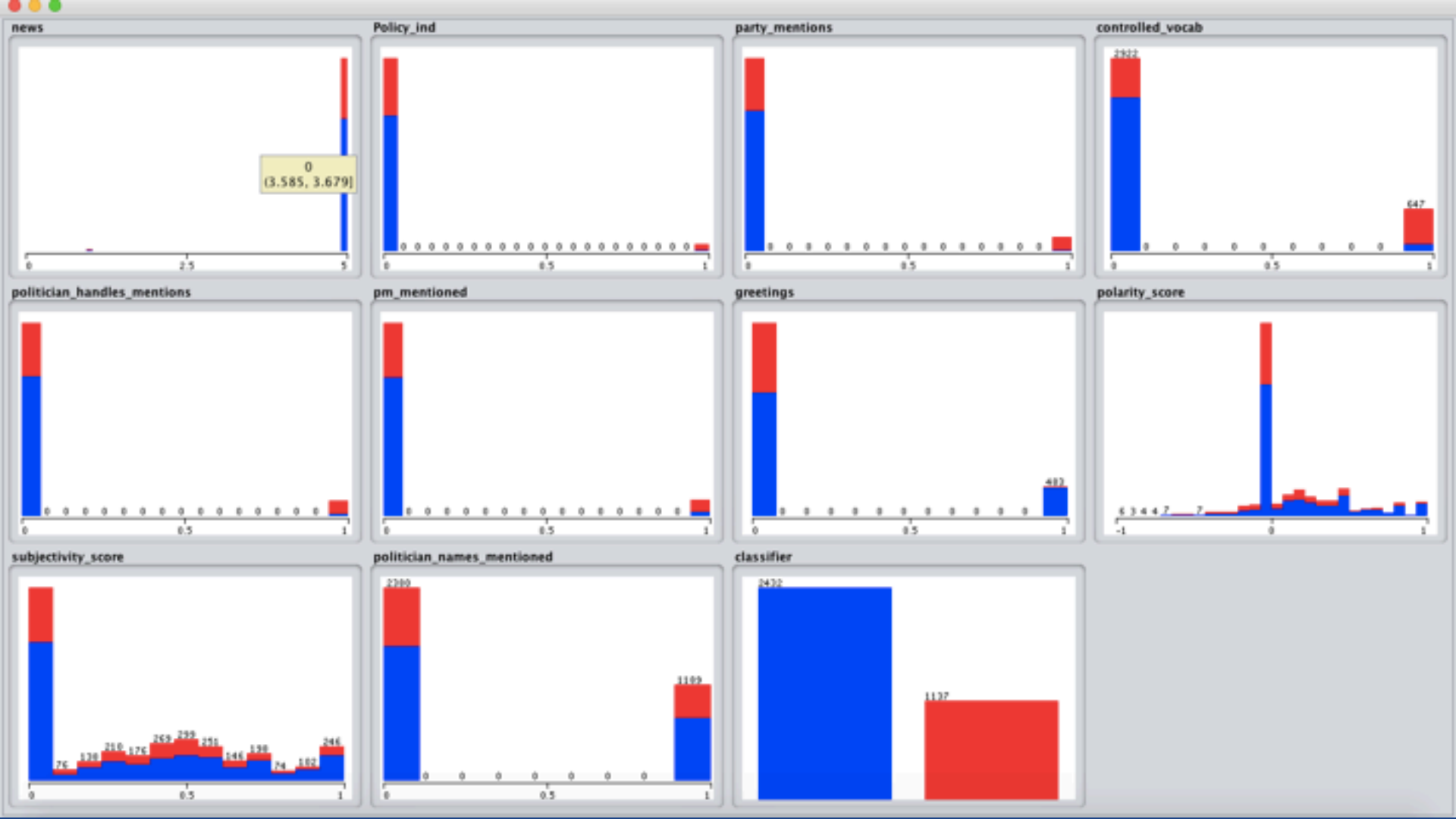
langdetect
build passing
Port of Google's language detection library (version from 03/03/2014) to Python.

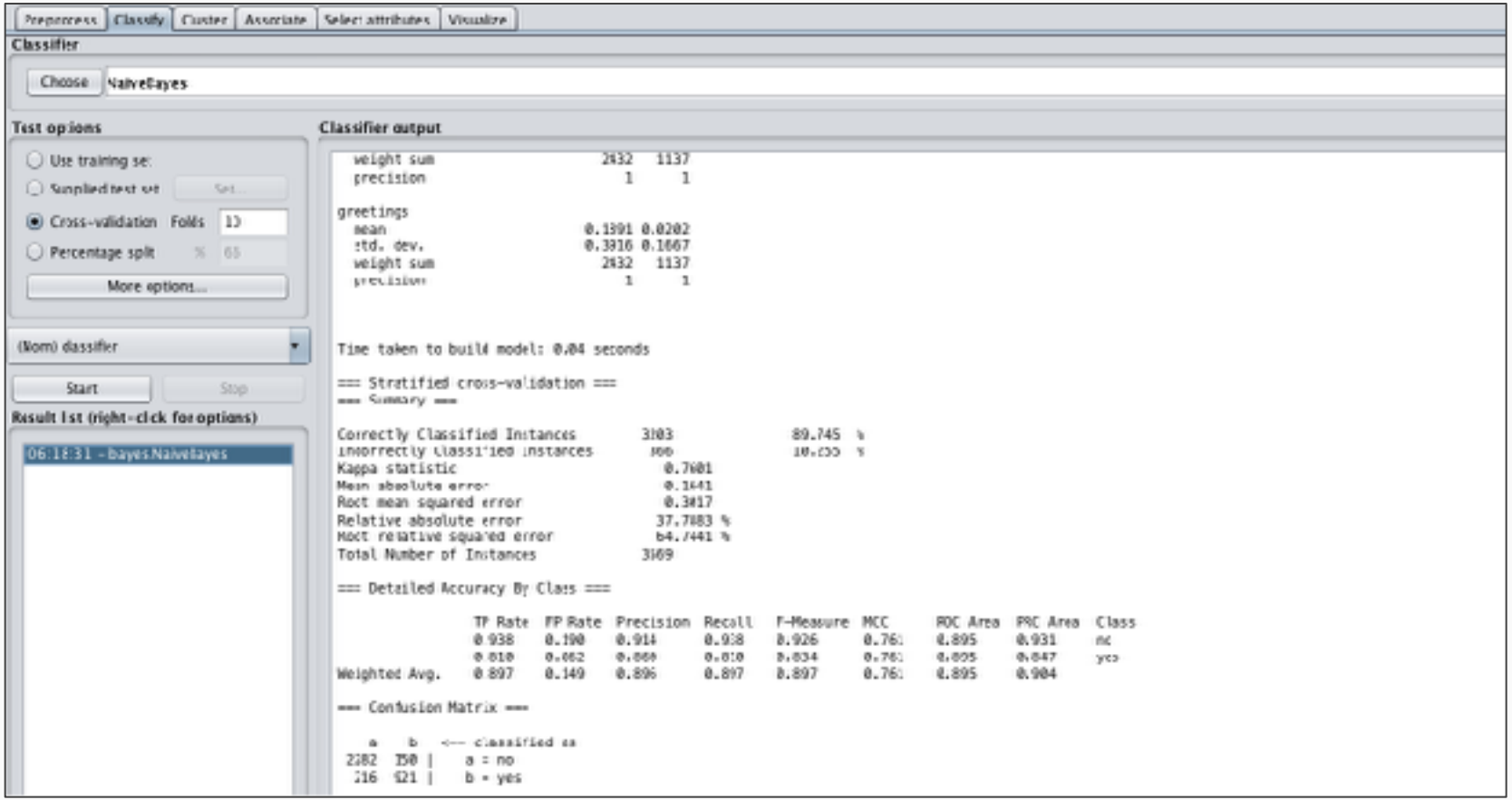
Dataset

Name	Tweet	News	Policy	Party Names	Controlled Vocabulary	Politician Handles	PM Mentions	Greetings	Polarity	Subjectivity	Politician Names	Classifier
Omar Abdullah	bFor once I find myself agreeing with @piersmorgan PIERS MORGAN: Endorsing Muslim-haters the WORST thing Trump has done\n https://t.co/h62m1f7Nuy'	none	0	0	0	0	0	0	-1.000	1.000	0	no
Omar Abdullah	b"Trumps anti-Muslim retweet fits a pattern https://t.co/d9H5Kux3Uu"	none	0	0	0	0	0	0	0.000	0.000	0	no
Omar Abdullah	bLet me correct this for you - Dineshwar Sharma & Home Ministry in Delhi ordered the withdrawal of cases. Ms Mufti was left to sign the file. https://t.co/3ec1IIx0va'	none	0	0	0	0	0	0	0.000	0.000	1	yes
Omar Abdullah	bIt SHOULD not matter but let\xe2\x80\x99s face it for a politician it does matter. It matters a lot & especially in this surcharged hyper-religious environment. https://t.co/lP3CDdxIbJ'	none	0	0	1	0	0	0	0.000	0.625	0	yes
Omar Abdullah	b@NigelBritto And mustard shipments at Sarsoon docks'	none	0	0	0	0	0	0	0.000	0.000	0	no
Omar	bMay be now some of our news channels & anchors will stop falling over themselves to											

Results & Conclusion

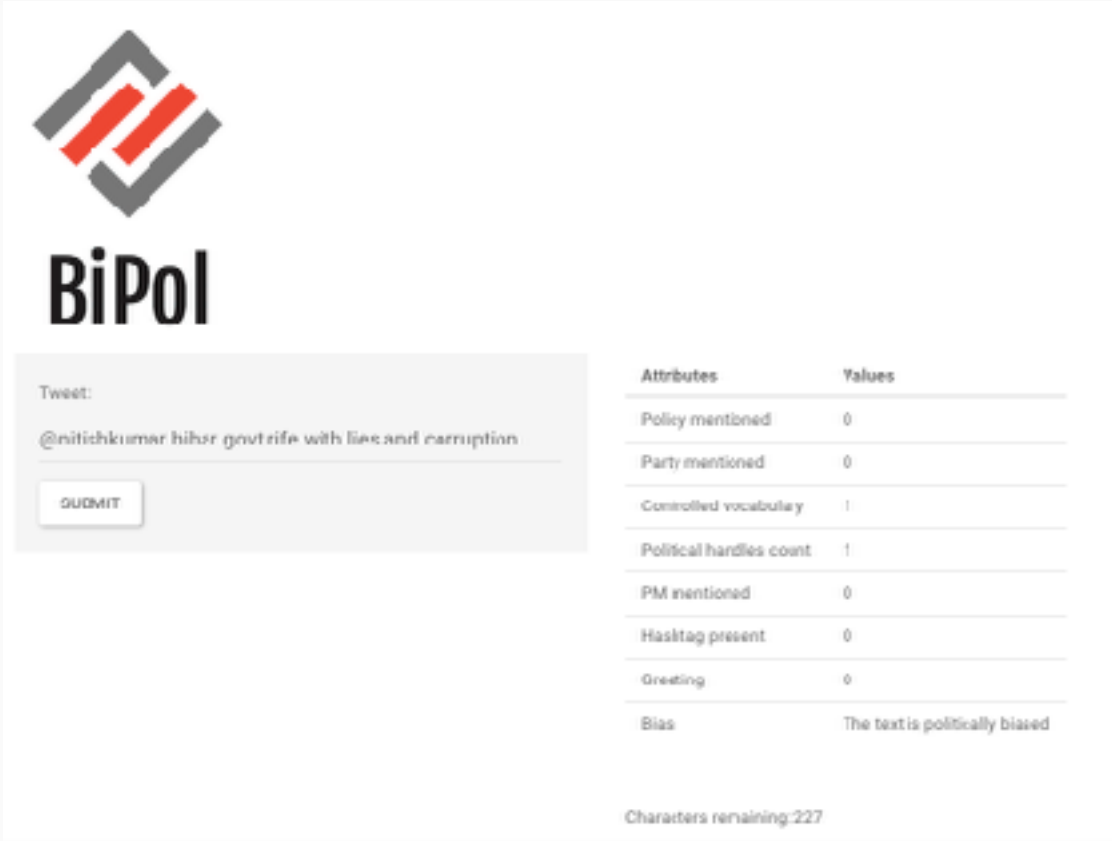



news
0
(3.585, 3.679]
Policy_ind
party_mentions
controlled_vocab
2922
647
politician_handles_mentions
pm_mentioned
greetings
403
polarity_score
subjectivity_score
76
138
210
176
269
293
251
146
198
74
182
246
politician_names_mentioned
2309
1109
classifier
2472
1137



Selecting top 6 attributes on info gain → applying Naïve Bayes using 10 cross-validation folds

UI in R Shiny


BiPol
Tweet:
@nitishkumar bihar govt rife with lies and corruption
SUBMIT
Attributes
Values
Policy mentioned
0
Party mentioned
0
Controlled vocabulary
1
Political handles count
1
PM mentioned
0
Hashtag present
0
Greeting
0
Bias
The text is politically biased
Characters remaining:227