REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

REVA Academy for Corporate Excellence (RACE)

Emotion recognition based Voice of Citizen using Twitter data

Surendra Tanniru

SRN: R20MBA09

Date: 22/10/2022

MBA in Business Analytics

Capstone Project Presentation
Year: II

race.reva.edu.in

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Agenda

01 Introduction
Back Ground | Current status | Why this study

02 Literature Review
Seminal works | Summary | Research Gap

03 Problem Statement
Business Problem | Analytics Solution

04 Objectives of this Study
Primary & Secondary Objectives | Expected Outcome

05 Project Methodology
Conceptual Framework | Research Design

06 Data Collection
Data Collection | Variables

07 Data Preparation
Pre-processing | Process \| Techniques

08 Data Analysis
Univariate | Bivariate | Hypothesis

09 Modeling
Machine Learning | Model Evaluation | Insights

10 Results and Insights
Applications | Demo

11 Conclusions and Suggestions
Insights | Next Step \| Future Scope

12 Annexure
References | Publications | Plagiarism Score

Introduction - Bangalore Traffic

- Traffic is one of the biggest problems that people in Bangalore are facing
- Bangalore is the world's most traffic-congested city(2020).
- The hottest topic of conversation in the city is 'Traffic'

1
Most traffic congested City – 2020

71%
More time than any other city - 2020

49%
Still stands at 10th position – 2021

An average person spends in traffic

Bangalore Traffic

According to a report by WRI

243hrs

An average man spend in traffic

65bn/year

Cost to city

35kmph

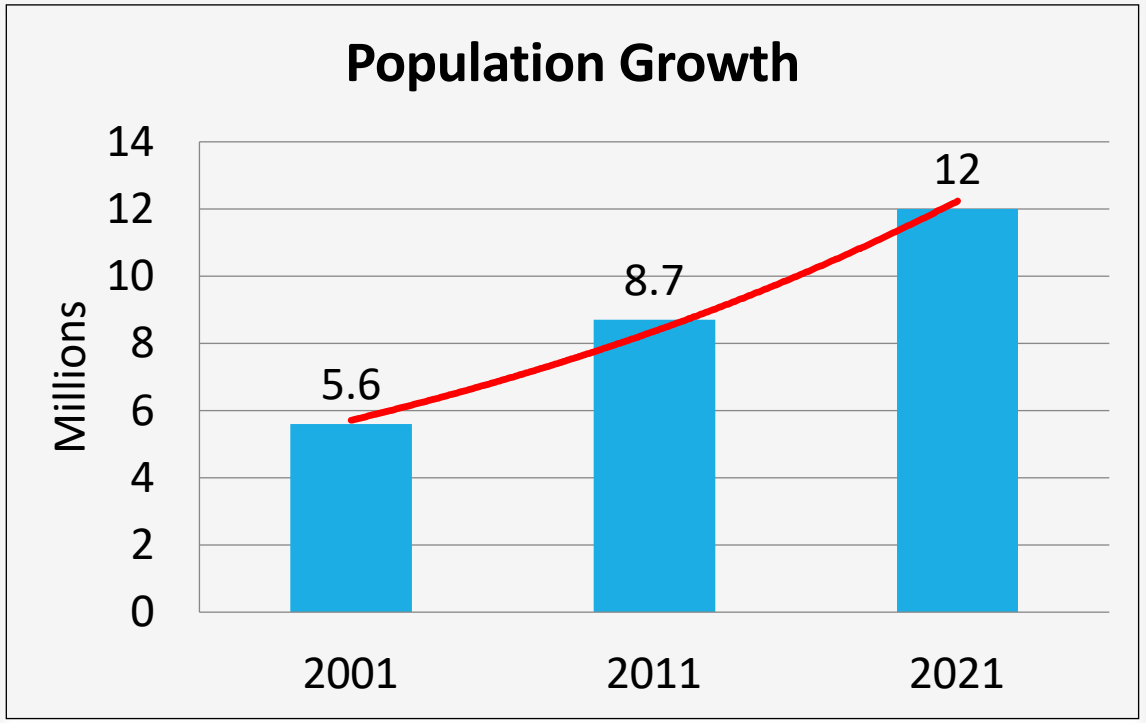
Speed of traffic - 2005

9.2kmph

Speed of traffic – 2014

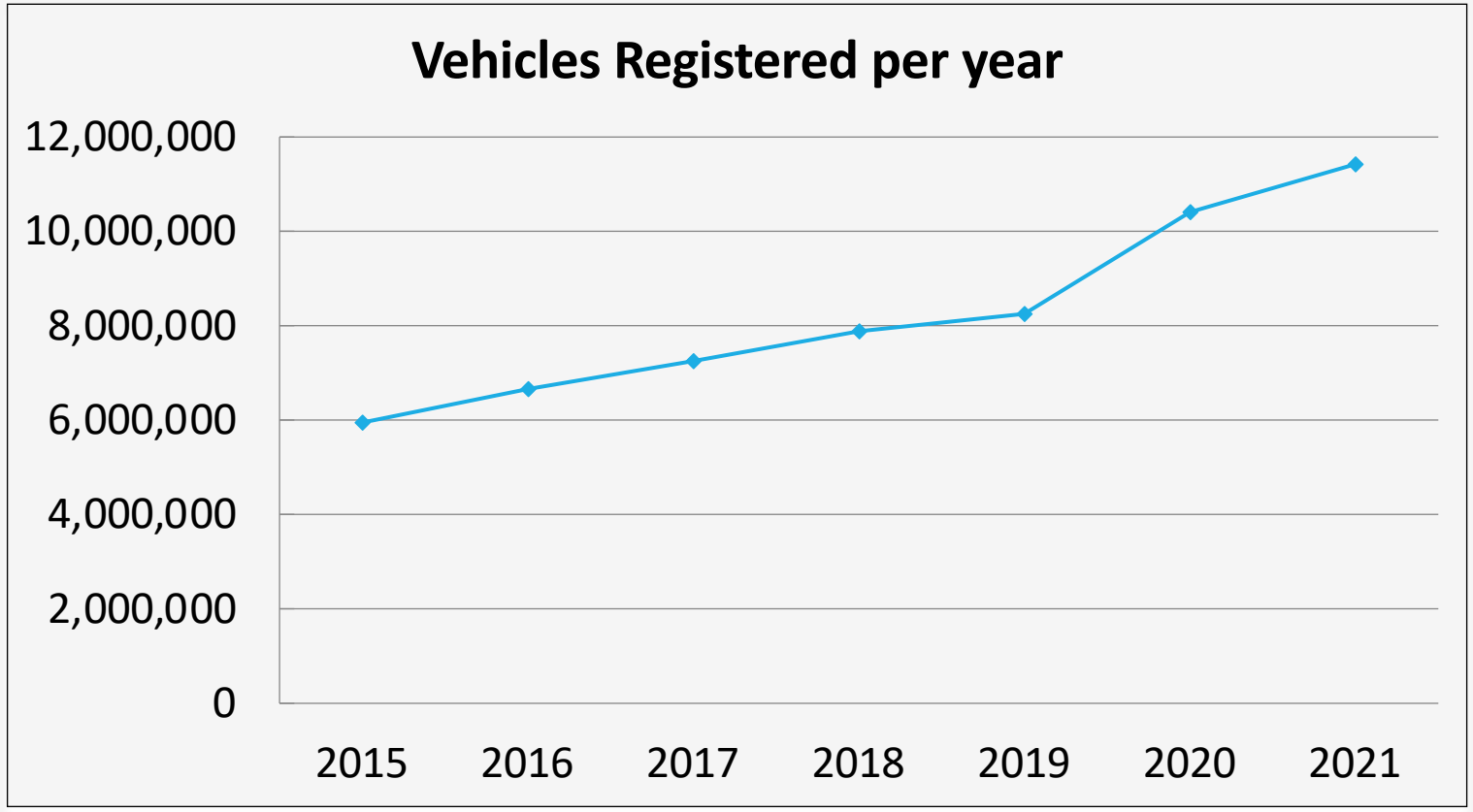
4.1kmph

Speed of traffic - 2019



Population Growth in Bangalore is increasing at an average of 46%

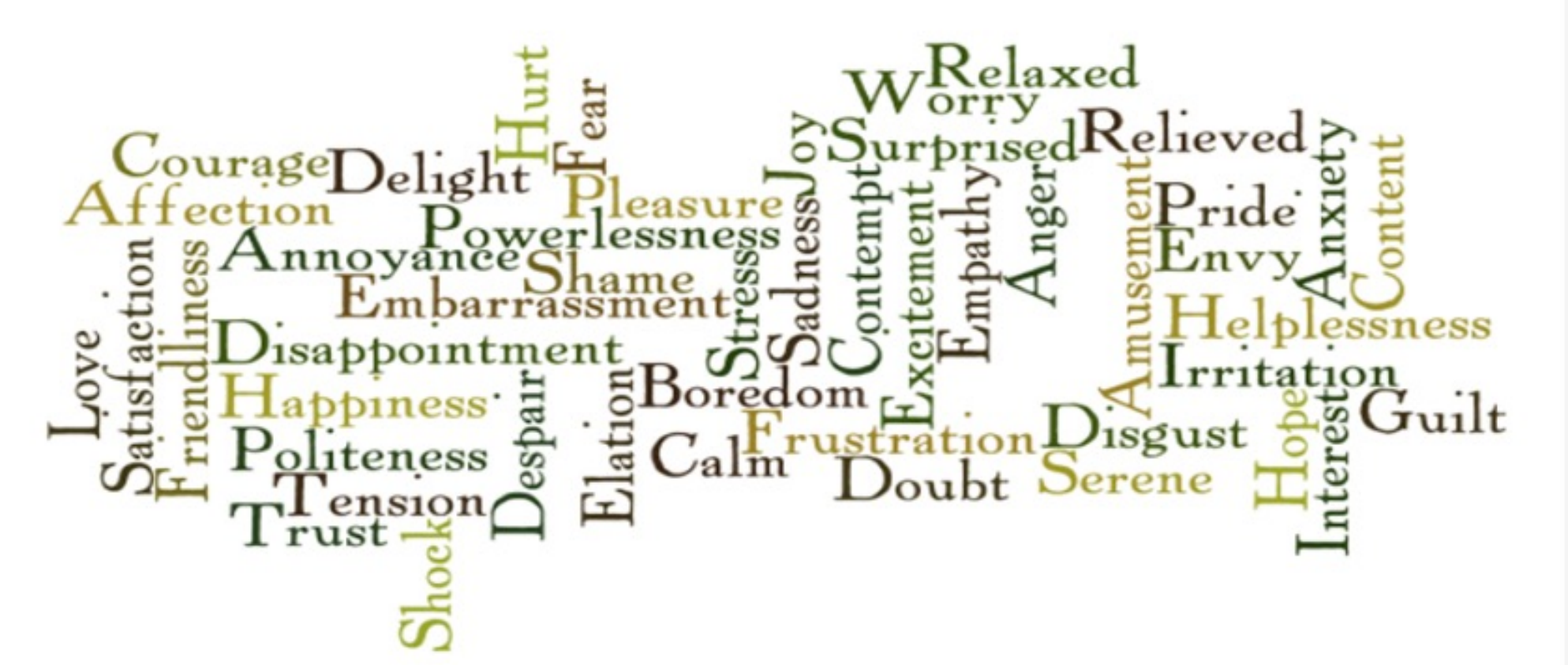
Bangalore Traffic



Source

- Data from mixed sources
- Embedded content
- Instantaneous coverage

Emotion


Relaxed
Worry
Courage
Delight
Hurt
Fear
Surprised
Relieved
Affection
Pleasure
Joy
Contempt
Pride
Powerlessness
Annoyance
Shame
Sadness
Excitement
Empathy
Anger
Amusement
Envy
Anxiety
Content
Satisfaction
Embarrassment
Stress
Helplessness
Love
Friendliness
Disappointment
Irritation
Happiness
Boredom
Elation
Despair
Frustration
Disgust
Guilt
Politeness
Calm
Hope
Tension
Doubt
Serene
Interest
Trust
Shock

Literature Review

Title	Author & Year	Source	Insights	Research gap
Emotional Analysis using Twitter Data during Pandemic Situation: COVID-19	Amrita Mathur, Purnima Kubde, Sonali Vaidya 2020	Fifth International Conference on Communication and Electronics Systems (ICCES 2020)	Research identifies the sentiments and emotions on twitter data.	Research is done to analyse the emotions from Twitter data on Covid-19 impact. Used AFINN, NRC lexicons
Speech Sentiment Analysis for Citizen's Engagement in Smart Cities' Events	Christine Janel Sora 2022	7th International Conference on Smart and Sustainable Technologies, 2022	presents a comprehensive deep learning approach for sentiment and emotion interpretation in speech	Used Deep learning techniques to identify sentiments and emotions from videos
Incident Detection From Social Media Targeting Indian Traffic Scenario Using Transfer Learning	Priyambada Ambastha, Maunendra Sankar Desarkar 2020	IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC) 2020	Results indicated how the tweets are classified as Traffic incidents and Non-Traffic incidents.	Identify the incidents occurred based on tweets using Transfer learning

Literature Review

Title	Author & Year	Source	Insights	Research gap
Prediction of Traffic Density for Congestion Analysis under Indian Traffic Conditions	Ameena Padiath, Lelitha Vanajakshi, Shankar C. Subramanian, 2009	12th International IEEE Conference on Intelligent Transportation Systems, 2009	An Artificial Neural Network (ANN) approach of identifying the traffic congestion from video	Identifies traffic congestion from recorded video
Study and Analysis of Emotion Classification on Textual Data	Huzaib Avez Sayyed, Dr. Shounak Rushikesh Sugave 2021	6th International Conference on Communication and Electronics Systems 2021	This research examines the methods of machine and deep learning in textual emotion classification. Based on the study of different papers, it concludes Deep learning techniques provide promising results	Not able to provide which ML/DL model can be used.
Traffic Forecasting with Deep Learning	Shounak Kundu, Maunendra Sankar Desarkar 2020	IEEE Region 10 Symposium 2020	Time series analysis on traffic using LSTM.	Publicly available U.K traffic data. Not real time.

Problem Statement

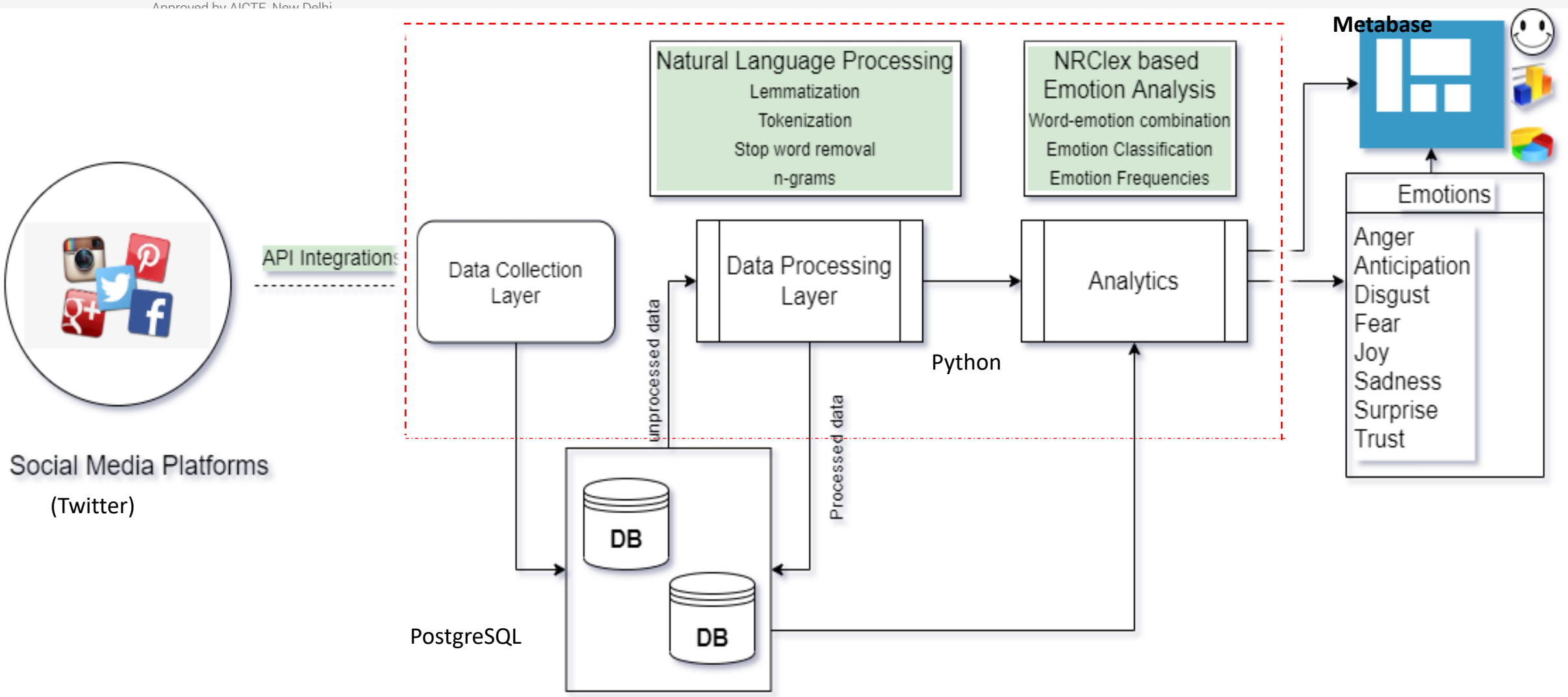
Business Problem | Analytics Solution

- Represent the Voice of Citizen on Bangalore Traffic.
- Identify the Emotions from the Tweets.

Project Objectives

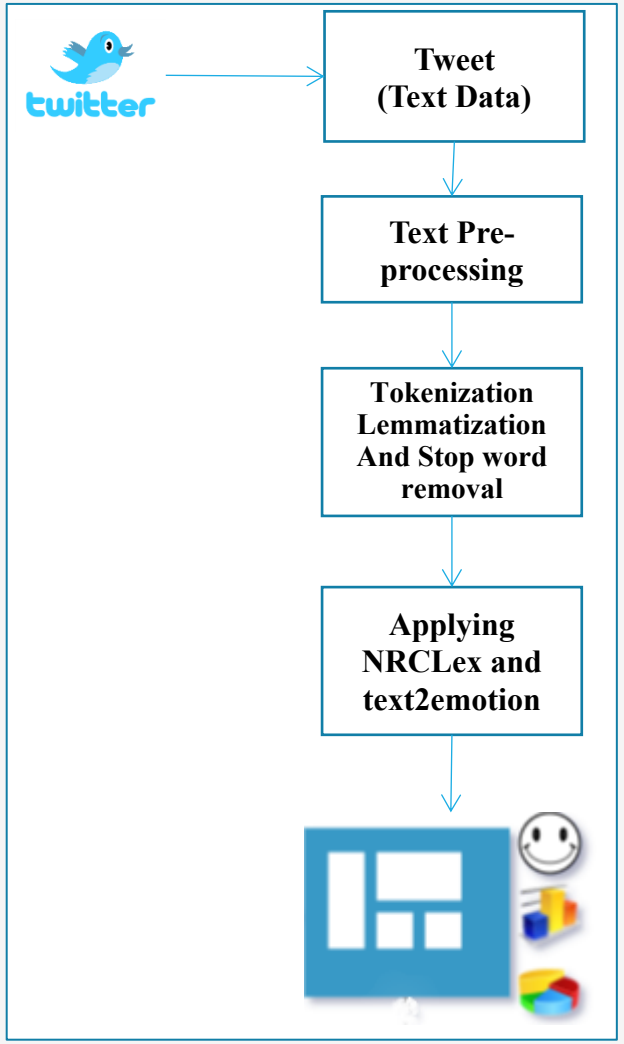
- Create a corpus of Twitter data using Twitter API
- Utilize NLP techniques to pre-process data
- Identify the emotions from tweets
- Create a dashboard identifying different emotions from the Tweet which can help officials to identify the most congested areas.

Project Methodology



Project Methodology

- Data is collected from Twitter using Twitter API
- Pre-process the data
- Tokenization
- Lemmatization
- Stop word removal
- Identify emotion using lexicon based approach and ML based approach
- Create a Dashboard



Data collection

```
consumer_key= 'XXXX'
consumer_secret= 'XXXX'

access_token= 'XXXX'
access_token_secret= 'XXXX'

bearer_token='XXXX'
```

Define the keys for API

```
auth = tw.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tw.API(auth, wait_on_rate_limit=True)
```

Authenticate to access the API

```
search_term = "XXXXXXXXXXXXXXX -filter:retweets"
```

Provide the Search Keyword

```
tweets = tw.Cursor(api.search,
                   q=search_term,
                   lang="en",
                   since='YYYY-MM-DD').items(n)
```

Execute the API cursor provided via Twitter API Handler

Data collection

Tweet data

	user	tweet	location	time
0	CodeKrafter	Getting #stuck in #traffic on every single #we...	India	2022-07-10 12:11:02
1	EarthDate2109	@kritikatwtss You still will find a place to s...	Bengaluru, India	2022-07-10 11:48:34
2	VBS_seeker	@jointcptraffic 🙏 Please do visit #SilkBoard ...		2022-07-08 15:19:19
3	VBS_seeker	@jointcptraffic Hope the new plan will deconge...		2022-07-07 17:08:15
4	Badass_Superdad	@JMawaali @whosthatmiss This traffic looks ver...	Silkboard Flyover	2022-07-07 16:15:06
5	sajhm13	Bad drains proving last straw for #SilkBoard j...	Bengaluru, India	2022-07-07 14:50:07
6	Clemenza2020	@vilakudy @minicnair Flat 50% discount it seem...		2022-07-07 03:08:49

	tweet_id	text	retweet_count	created_at
0	1546104710595477506	Getting #stuck in #traffic on every single #we...	0	Sun Jul 10 12:11:02 +0000 2022
1	1545427316494729216	@jointcptraffic 🙏 Please do visit #SilkBoard ...	0	Fri Jul 08 15:19:19 +0000 2022
2	1545092340749963264	@jointcptraffic Hope the new plan will deconge...	0	Thu Jul 07 17:08:15 +0000 2022
3	1545057580074307589	Bad drains proving last straw for #SilkBoard j...	3	Thu Jul 07 14:50:07 +0000 2022
4	1543608727936180224	#SilkBoard traffic 😑	0	Sun Jul 03 14:52:54 +0000 2022

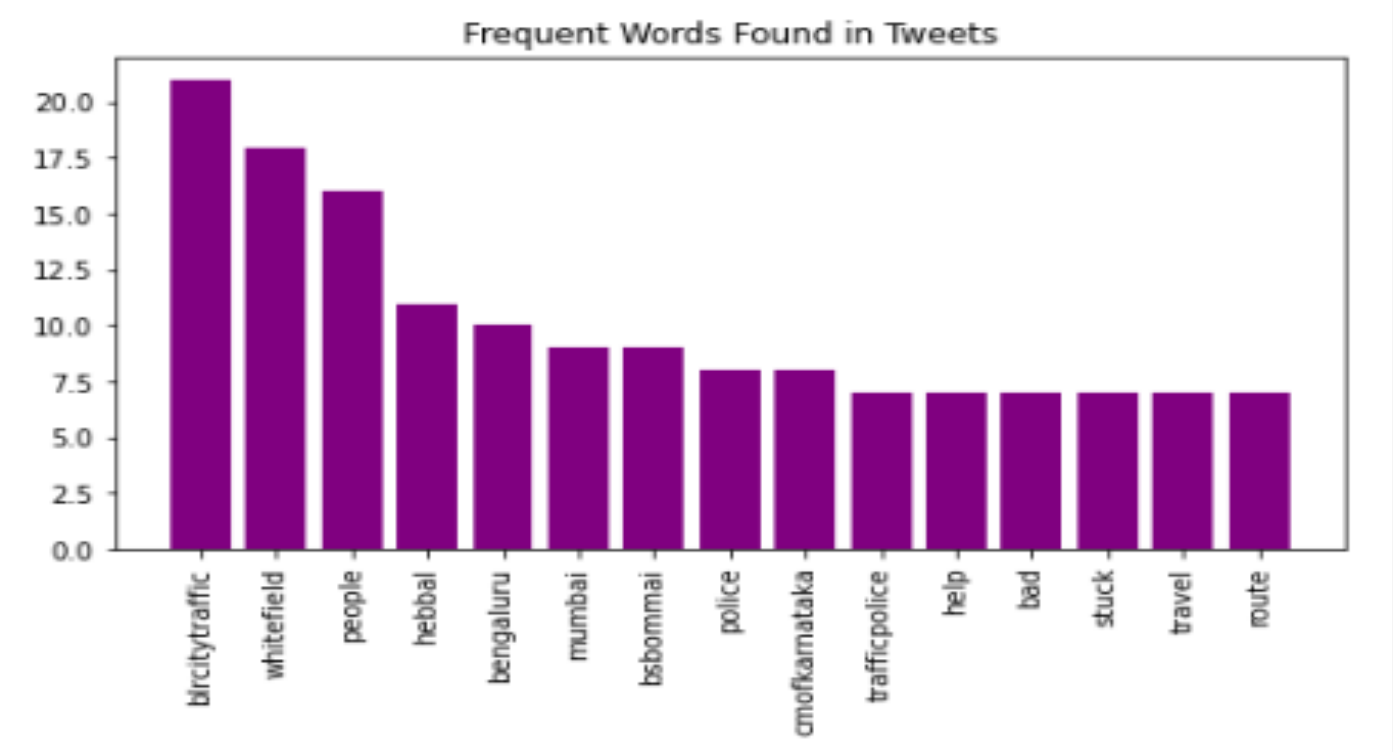
```
['@CMofKarnataka @BSBommai The only way to solve traffic congestion and providing way to silk board need parallel roa… https://
t.co/Mr3HUT1MeZ',
 "@ChristinMP_TOI Sir sorry it's not like that:-)\nPl check how this 1 of d world class flyover… https://t.co/C8Z4vITnTH"]
```

Data Pre-processing

- Pass the Tweet as input
- Convert to lower case
- Tokenize
- Lemmatize
- Remove stop words
- Compare the token with Word-emotion Lexicon
- Calculate below
  - Number of Tokens,
  - Number of emotion lexicons in each tweet,
  - Lexicon Ratio
  - Lexicon Dynamics for each tweet

Analysis

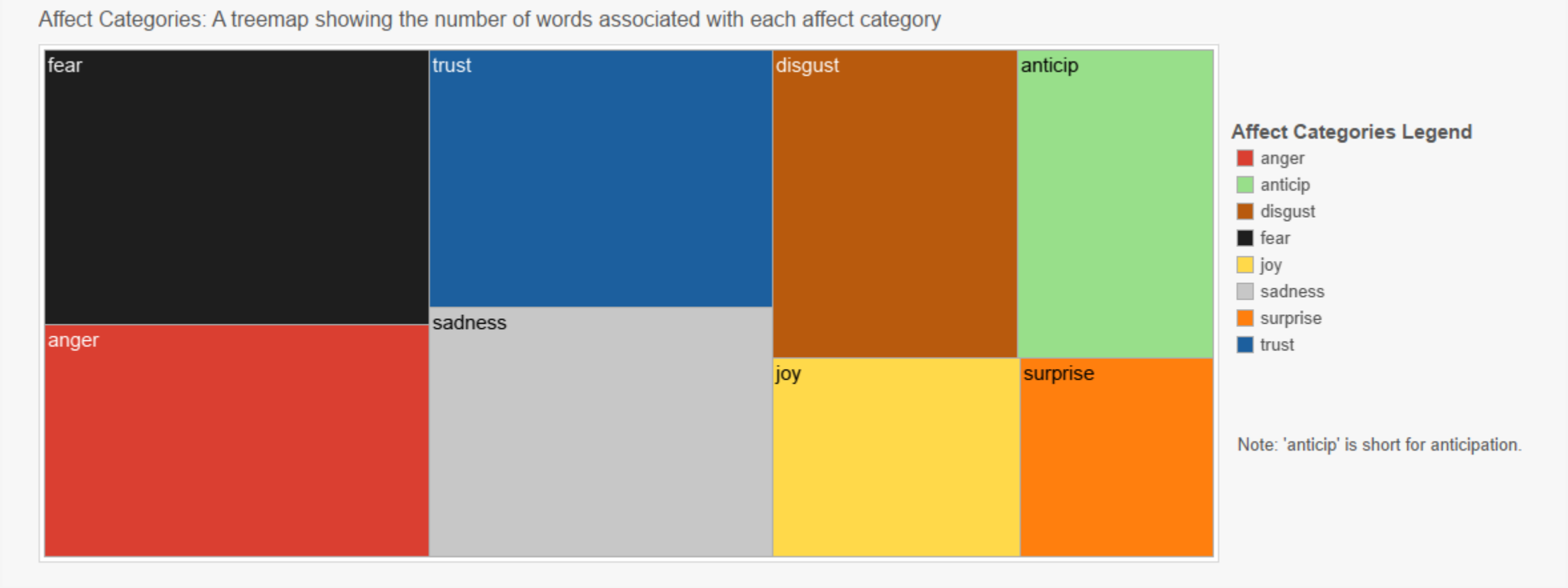
Whitefield and Hebbal localities are more congested areas than Silk board



Word Cloud


narendramodi
near
bad
govt
whitefield
blockbine
today
infrastructure
trafficpolice
cpblr
travel
stuck
help
life
people
problem
check
cities
airport
hebbal
live
bbmpcomm
start
pathetic
cmofkarnataka
mumbai
police
metro
bmtc
silkboard
highway
yes
blrcitytraffic
lot
bsbommai
delhi
good
bengaluru
getting
karnataka
morning
blrcitypolice
route
driving
parking

Modelling


Affect Categories: A treemap showing the number of words associated with each affect category
fear
trust
disgust
anticip
anger
sadness
joy
surprise
Affect Categories Legend
anger
anticip
disgust
fear
joy
sadness
surprise
trust
Note: 'anticip' is short for anticipation.

Modelling

Tweet_lc character varying	emotion character varyir	numTokens bigint	numLexTokens bigint	lexRatio double precision	avgLexVal double precision
nikateen slvppy bitch i live in whitefield you ain t got nothing on traffic	fear	14	2	0.142857143	0.5
opencity in citizenmatters reapbenefit wriindia wricitiesindia while koramangala is famous for...	trust	14	2	0.142857143	0.5
bengaluru traffic police implements changes to reduce snarls near hebbal flyover	fear	11	2	0.181818182	0.5
reminds me of silkboard 2018 a hilarious short film about bangalore traffic	anger	11	4	0.363636364	0
getting stuck in traffic on every single weekends in front of pheonixmall we are having our on ...	anger	19	4	0.210526316	0
kritikatwtss you still will find a place to stay but the greatest hazard is the traffic snarl pre covi...	anger	21	7	0.333333333	0.142857143
pmoindia narendramodi you made big announcement of fixing bangalore traffic but your cmo...	anticipation	16	4	0.25	0.25
traffic exiting from nh44 from airport towards hebbal is a mess some officer s brilliant idea of ...	anticipation	19	6	0.315789474	0.333333333
blrcitytraffic dgpkarnataka cpblr jointcptraffic blrcitypolice dcptrnorthbcp dcptreastbcp dcptr...	anticipation	9	1	0.111111111	0
blrcitytraffic hebbal flyover frm airport i d advise to open one entry for vehicle which are comi...	anticipation	19	6	0.315789474	0.333333333
check out the new traffic plan to ease hebbal gridlock	disgust	10	4	0.4	0
bmtc buses used this cheat route b w the esteem mall amp hebbal to avoid some traffic cmof...	disgust	20	7	0.35	0.142857143
hebbaltrafficps the bus stop near esteem mall kempapur cross should be shifted to a wider pl...	joy	19	6	0.315789474	0.166666667
blrcitytraffic cpblr blrcitypolice sir in your video watch the left side traffic on the ring road how ...	joy	20	7	0.35	0
bbmpcomm bbmpadmn cpblr cmofkarnataka bsbommai pl find enclosed a few tips to impro...	joy	17	3	0.176470588	0.333333333

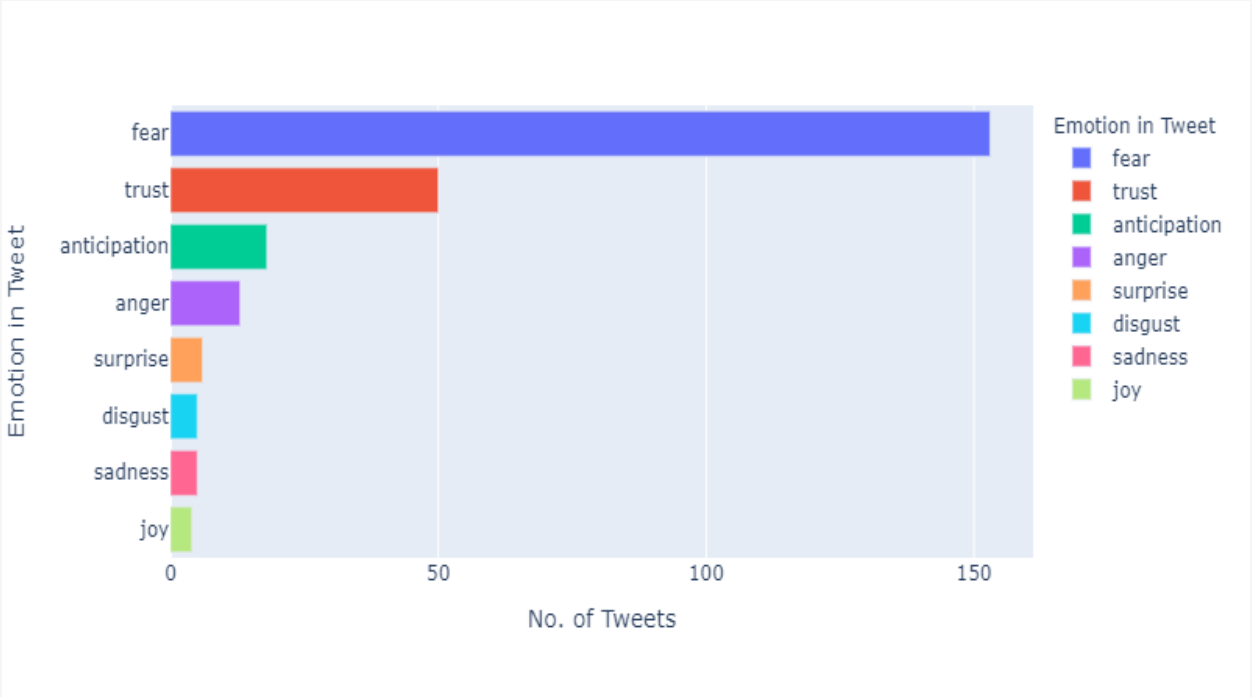
Emotion : emotion name
numTokens : number of tokens in the tweet.
numLexTokens : number of tokens in the tweet that are present in the emotion-word lexicon
lexRatio : numLexTokens/numTokens.
avgLexVal : emotion score of the tweet.

Modelling

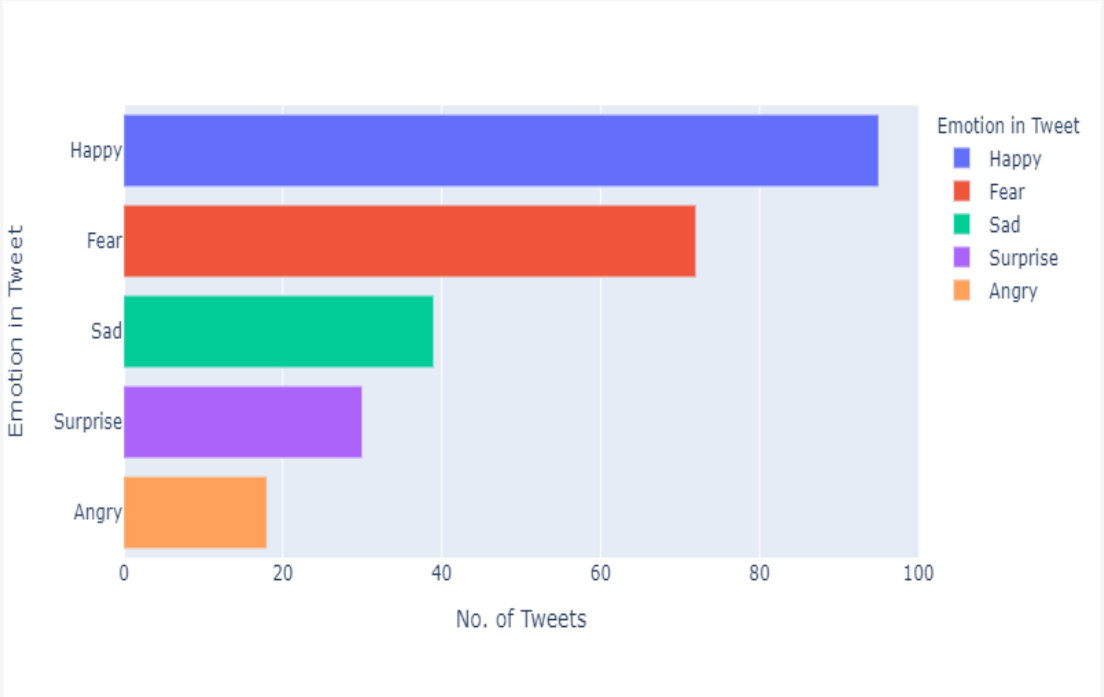
How each word of a Tweet is affected with NRC lexicon

```
{'happen': ['anticipation'],
 'risk': ['anticipation', 'fear'],
 'hilarious': ['joy', 'surprise'],
 'hazard': ['fear'],
 'snarl': ['anger', 'disgust'],
 'plan': ['anticipation'],
 'bad': ['anger', 'disgust', 'fear', 'sadness'],
 'law': ['trust'],
 'finally': ['anticipation', 'disgust', 'joy', 'surprise', 'trust'],
 'board': ['anticipation'],
 'government': ['fear'],
 'share': ['anticipation', 'joy', 'trust'],
 'hospital': ['fear', 'sadness', 'trust'],
 'time': ['anticipation'],
 'love': ['joy'],
 'perceive': ['trust'],
 'don': ['trust'],
 'offend': ['anger', 'disgust'],
 'trip': ['surprise'],
 'complain': ['anger', 'sadness'],
 'improvement': ['joy', 'trust'],
 'provide': ['trust'],
 'chaos': ['anger', 'fear', 'sadness'],
 'treat': ['anger',
  'anticipation',
  'disgust',
  'fear',
  'joy',
  'sadness',
  'surprise',
  'trust'],
 'urgent': ['anticipation', 'fear', 'surprise'],
 'horrible': ['anger', 'disgust', 'fear'],
 'deal': ['anticipation', 'joy', 'surprise', 'trust'],
```

Modelling

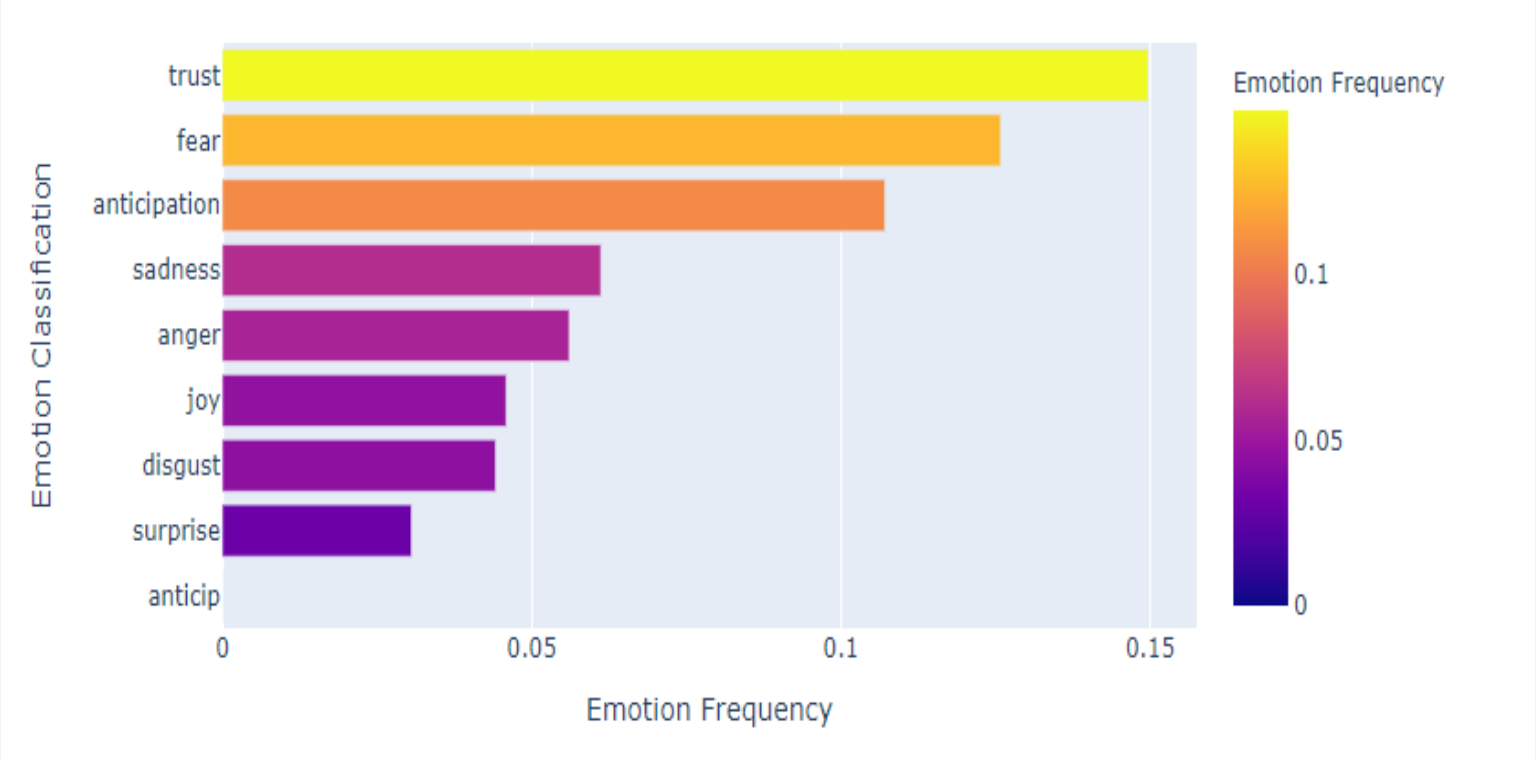


Emotion Classification using NRCLex



Emotion Classification using text2emotion

Modelling



Emotion Frequency of tokens

Comparison

Total Emotions on Tweets		
Emotion	**NRCLex**	**text2emotion**
Fear	**153**	**72**
Trust	**50**	
Joy	**4**	**95**
Sad	**5**	**39**
Anticipation	**18**	
Surprise	**6**	**30**
Angry	**13**	**18**
Disgust	**5**	

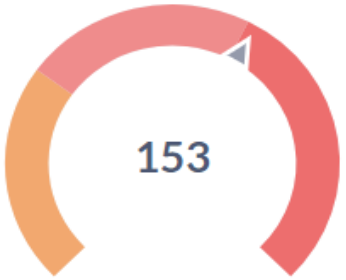
Dashboard

Voice of Citizen on Bangalore Traffic

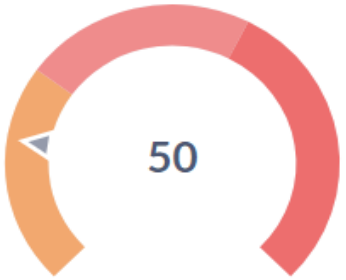
Edited 2 minutes ago by you

254
Total No. of Tweets

No. of Tweets with Fear as Em...



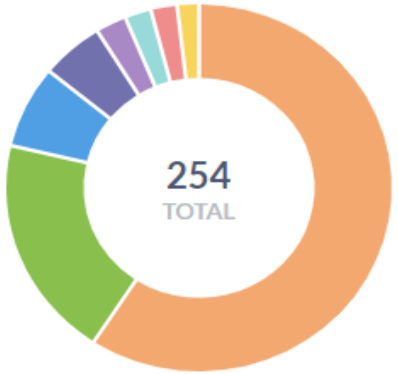
No. of tweets with Trust as Em...



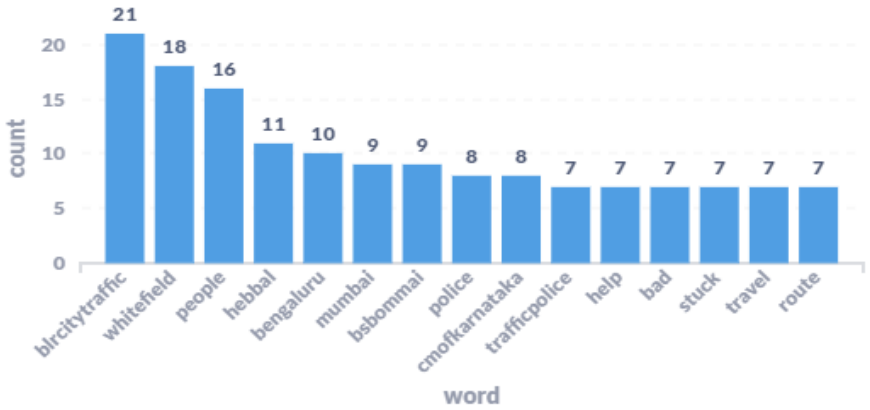
No. of Tweets with Anticipate ...

Emotions based on Tweets

fear	60.24%
trust	19.69%
anticipation	7.09%
anger	5.12%
surprise	2.36%
disgust	1.97%
sadness	1.97%
joy	1.57%

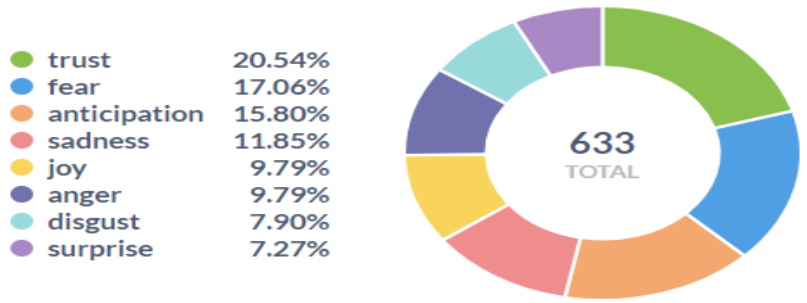


Most used Words

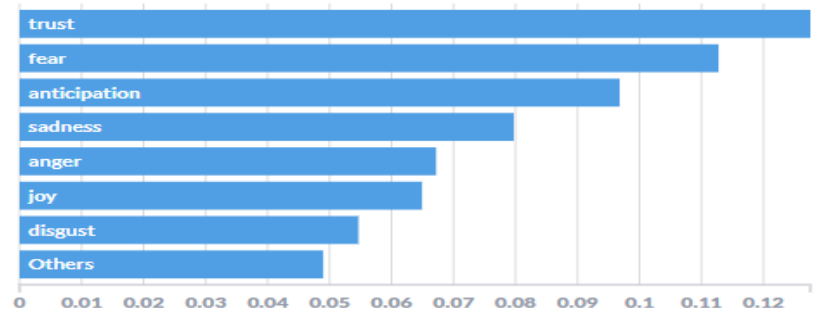


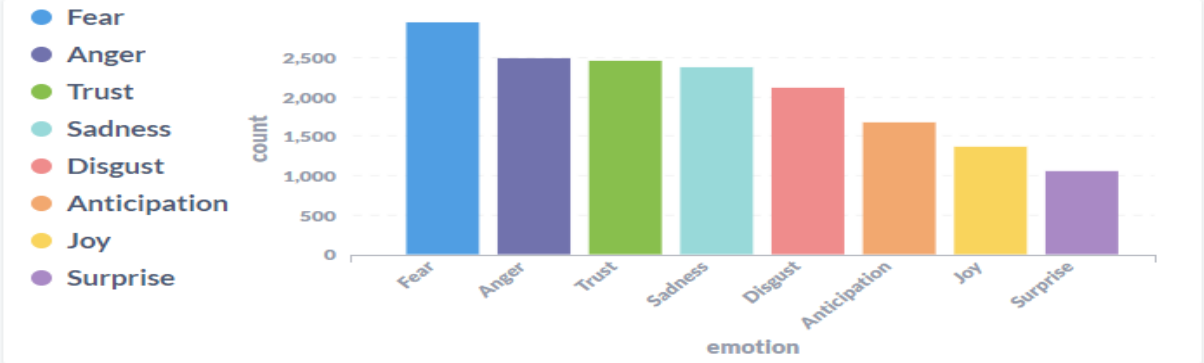
Dashboard

Emotion Classification based on Tokens in tweet



Frequency of occurrence of emotions





NRC Word-Emotion Association

word	anger	anticipation	disgust	fear	joy	sadness	surprise	trust
aback	0	0	0	0	0	0	0	0
abacus	0	0	0	0	0	0	0	1

Conclusion and Future Work

Proposed solutions | Scope for future work

- Most of the tweets carry an emotion of Fear followed by Trust
- Concludes that the traffic congestion is very high in Bangalore which needs attention
- Most of the words in the NRC lexicon are associated with Fear and Anger. Whereas, the word-emotion associations in text2emotion are mostly towards Happy.
- The Dashboard can be utilized by government officials and traffic police to identify the most traffic congested areas and take necessary action in diverting the traffic
- Add real time integration of API

References

Bibliography | Webliography

- *Bangalore has the worst traffic in the world*. (2020, January 29). Retrieved from Live mint: https://www.livemint.com/news/india/bengaluru-has-the-worst-traffic-in-the-world-report-11580278531554.html
- *Bangalore Traffic Report | TomTom traffic Index*. (2019, December 10). Retrieved from TomTom: https://www.tomtom.com/traffic-index/bengaluru-traffic/
- *BBC News*. (2016, December 7). Retrieved from Why is Bangalore stuck in traffic jams?: https://www.bbc.com/news/world-asia-india-38155635
- Dev, A. (2021, September 04). *35% increase in private vehicles on Bengaluru roads: Officials*. Retrieved from Hindustan Times: https://www.hindustantimes.com/india-news/35-increase-in-private-vehicles-on-bengaluru-roads-officials-101630695800757.html
- Ema Kusen ˇ 1, G. C. (2017). Identifying Emotions in Social Media:Comparison of Word-emotion Lexicons. *5th International Conference on Future Internet of Things and Cloud Workshops*, 6.

References

Bibliography | Webliography

- *Felipe Bravo-Marquez, S. M. (2016). Determining Word–Emotion Associations from Tweets by Multi-Label Classification. 2016 IEEE/WIC/ACM International Conference on Web Intelligence, 4.*
- *Mihalcea, S. a. (2004). WordNet-Affect: an Affective Extension of WordNe. ITC-irst, Istituto per la Ricerca Scientifica e Tecnologica, 4.*
- *Mohammad, S. M. (2013). NRC-Canada: Building the State-of-the-Art in Sentiment Analysis of Tweets. National Research Council Canada, 7.*
- *Mulukutla, A. M. (2015). When it Comes to Reducing Car Congestion, India's Cities Can Learn from its Businesses. Bangalore: World Resources Institute.*

Annexure

Additional Information | Plagiarism score

Similarity Index – 14%
Internet Sources – 14%
Publications – 3%
Student Papers – 7%

REVA UNIVERSITY

Bengaluru, India

Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Thank you!