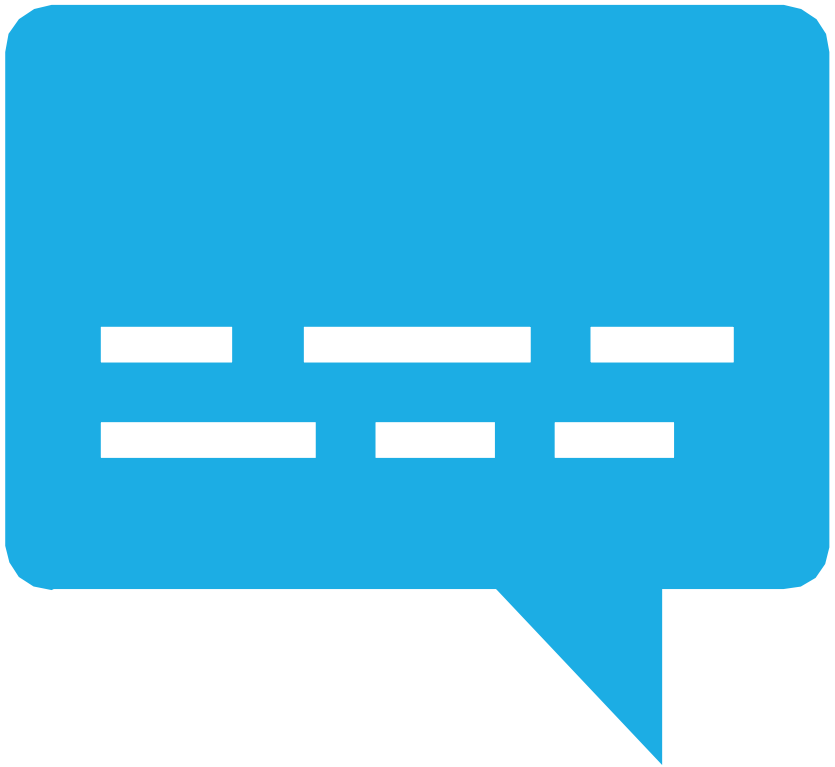
Emotion Detection in Tweet Replies

Faculty Advisor:
Mr. Rishabh Kaushal

Submitted by:
Manisha Keim
M.Tech Information Security

OUTLINE

- Problem Statement
- Literature Review
- Proposed Methodology
- Dataset Description
- Manual Annotation
- Binary & Multi-class Classifier
- Results
- Future Work



PROBLEM STATEMENT

- Online content has become a major issue in today's world due to an exponential increase in the use of internet by people of different cultures and educational background.
- Considering the time, a modern human spends on the Internet, one can expect that this online information, when combined, can reveal significant information.
- Research Question: How to automatically know the sentiment/emotion of tweet-replies in commenting platforms?

Examples of Tweet & Replies

RELATED WORK

[1] Ekman, P., Friesen, W. V., Ellsworth, P. (1972). Emotion in the Human Face. 1 ed. Perga-mon.

[2] Ribeiro, M. H.; Calais, P. H.; Santos, Y. A.; Almeida, V. A. and Meira Jr, W (2017). "LikeSheep Among Wolves": Characterizing Hateful Users on Twitter.

[3] B Mathew, N Kumar, P Goyal, A Mukherjee. (2018). Analyzing the hate and counter speech accounts on Twitter.

[4] D Chatzakou, N Kourtellis, J Blackburn. (2017). Detecting aggressors and bullies on Twitter.

[5] T. C. Li, J. Gharibshah, E. E. Papalexakis, and M. Faloutsos. (2018). TrollSpot: Detecting misbehavior in commenting platforms. ASONAM '17: Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 Pg.171-175.

RELATED WORK

[6] Hasan, M., Rundensteiner, E., and Agu, E. (2014). EMOTEX: Detecting Emotions in Twitter Messages. Cybersecurity Conference, Stanford University.

[7] Badugu, S. and Suhasini, M. (2017). Emotion Detection on Twitter Data using Knowledge Base Approach. International Journal of Computer Applications (0975 – 8887)Volume 162– No 10, March 2017.

[8] Ranganathan, J., Hedge, N., Irudayaraj A.S. and Tzacheva A.A. (2018). Automatic Detection of Emotions in Twitter Data - A Scalable Decision Tree Classification Method. ACMRevOpID '2018, Baltimore, Maryland, USA.

[9] Gaind, B. and Syal, V. and Padgalwar S. (2019). Emotion Detection and Analysis on Social Media. Global Journal of Engineering Science and Researches(ISSN 2348 - 8034, Pg. 78-89).

[10] Gaydhani, A., Doma, V., Kendre, S., and Bhagwat, L. (2018). Emotion Detection and Analysis on Social Media. Detecting hate speech and offensive language on twitter using machine learning: An n-gram and tfidf based approach.

[11] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Emotion Models

EMOTIONS	REFERENCES
Anger, Disgust, Fear, Happiness, Sadness, and Surprise	[6]
Nervous, Anxious, Tension, Afraid, and Fear	[7]
Anger, Fear, Anticipation, Trust, Surprise, Sadness, Joy, and Disgust	[8]
Surprise, Disgust, Happiness, Fear, Sadness, and Anger	[9]

Feature Analysis

FEATURE	REFERENCES
Offensive words/ Hostile language	[1], [2]
Number of statuses	[1], [4]
No. of followers	[1], [2], [4]
No. of favorites	[1], [2]
No. of URLs/tweet	[1], [4]
No. of hastags/tweet	[1], [2]
Suspended & deleted accounts	[1], [2]
No. of posts since account creation	[2], [4]

Evaluation Metrics

ALGORITHM	REFERENCES
Logistic Regression	[3], [10]
Naive Bayes	[4], [6]
Support Vector Machines	[3], [10]
Random Forest (RF)	[3], [4] , [5]
XGBoost(XGB)	[3]
CatBoost(CB)	[3]
Extra-Tree (ET)	[3]
Decision Tree	[4], [8]
TF-IDF	[10]

DATASETS

REFERENCES	DATASET
[2]	100386 users
[3]	558 hate tweets received a total of 1711 replies
[4]	Twitter Streaming API (a) 1M random tweets (b) hate-related 650K tweets based on 309 hashtags
[5]	7M comments
[6]	16,5300 tweets
[7]	10,48576 tweets

Proposed Methodology

- Twitter API
- Data Collection
- Manual Annotation
- Pre-processing
- Emotion classification

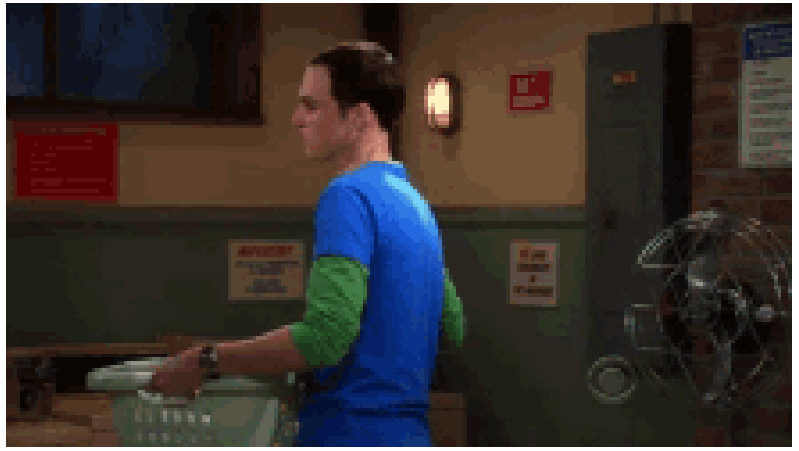
Data Collection

- We searched verified profiles for the target political leaders and manually collected 15 unique tweets for each of the leader.
- we collected tweet replies from Twitter API, web-scraping using Selenium, BeautifulSoup. We have focused on tweets related to Indian Politics. We collect all tweet-replies commented at tweets published by politicians like Arvind Kjeriwal & Amit Shah from September 2019 till June 2020.

- The dataset consists of:
  (a) 30 tweets,
  (b) 3200 tweet replies

EMOTIONS OF USERS & GROUND TRUTH

SAD



ANGRY

DISGUST

Manual Annotation of Tweet-Replies

We aim to build two types of classifiers:

Binary classifier (Label 1)

Multi-class classifier (Label 2)

Label 1: To classify a tweet reply as happy or unhappy:

0: Happy

1: Unhappy

Label 2: If the value of Label 1 is 'unhappy', then we define 3 emotional roles:

1: Sad

2: Angry

3: Disgust

Tweet-Reply	Label
Wastage of funds. Build statues and Temples instead.	Unhappy
U people r doing less n speak much louder n other hand @narendramodi @TajinderBagga he work hard to unite India n u people donate like its really scary situation again let u ask to think about ur strategy n make sure India first n All ur other agenda comes second.	Unhappy

EXAMPLES WHERE LABEL 1 IS MARKED 1(UNHAPPY)

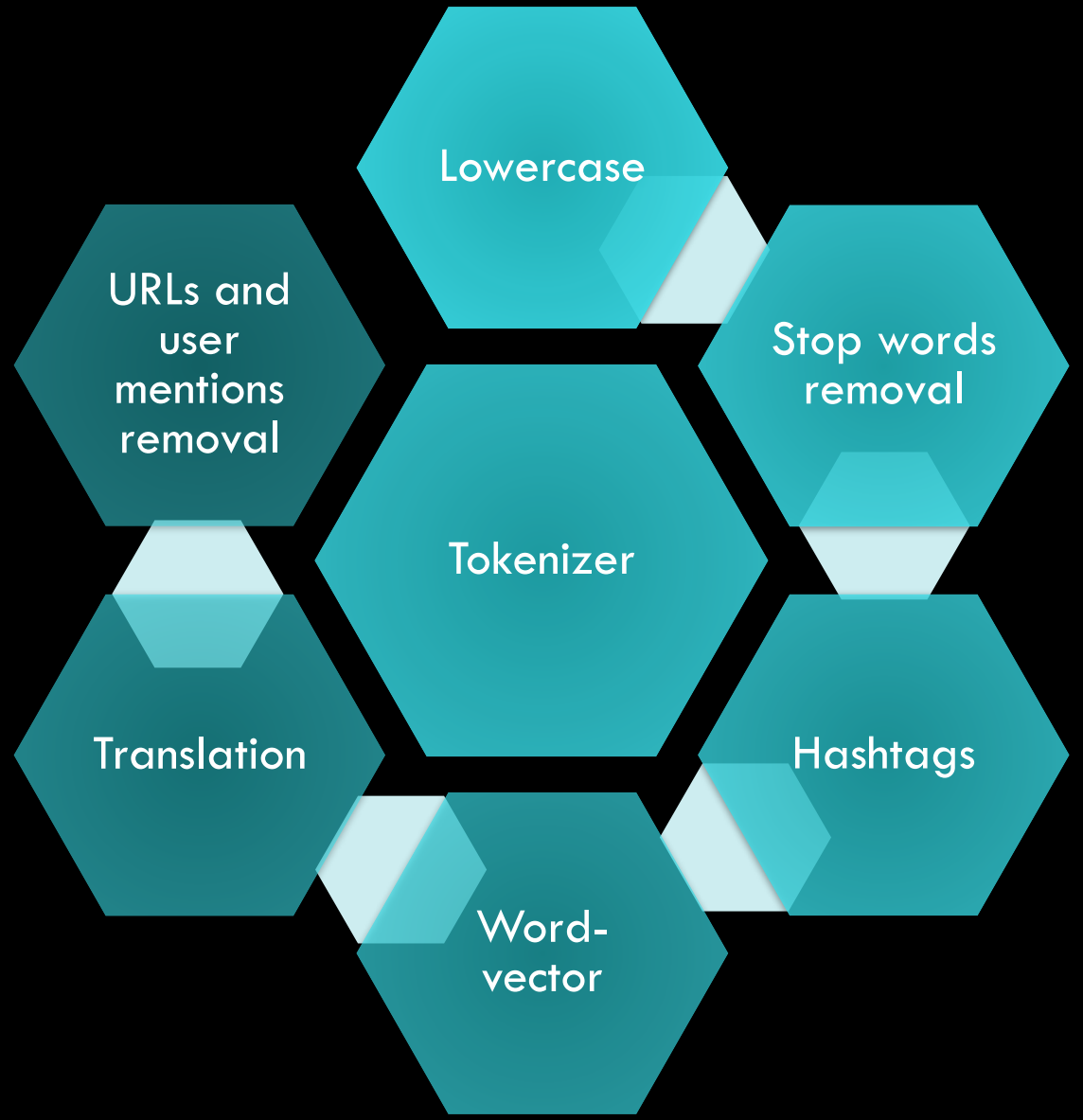
SOME EXAMPLES WHERE LABEL 1 IS MARKED 0(HAPPY)

Tweet-Reply	Label
Sir Indeed.. Ladies in Delhi are really enjoying and overwhelmed.	Happy
Wish each state & central govt. in our country someday throws away their political ideologies & EGOs for the sake of #INDIA's future and implement similar educational structure... We didn't get a #smart_educational_system.. Hope the next gen gets it... #India_will_shine_again...	Happy
Modi hain to mumkin hai.	Happy

Tweet-Reply	Label
Wastage of funds. Build statues and Temples instead.	Sad
I am a modi supporter but really like your work. Just keep distance from congress. I will support you.	Sad
What for. They did nothing. They are making fool . They are liars. No college was opened by them. They are outsiders.	Angry
Bribing voters will not help	Disgust
In unauthorized colonies largely illegal constructions carried out by property mafia. Govt of Delhi do not want to touch this issue as govt land is sold by builder mafia and big vote bank politics loss	Disgust

SOME EXAMPLES OF LABEL 2 ANNOTATION

PREPROCESSING



Word Vector Representation

Bag of Words

- BoW is an unigram model where we consider individual words into account and give each word a unique number.
- CountVectorizer is a part of sklearn package, which converts collection of documents to tokens.

	1 This	2 movie	3 is	4 very	5 scary	6 and	7 long	8 not	9 slow	10 spooky	11 good	Length of the review(in words)
Review 1	1	1	1	1	1	1	1	0	0	0	0	7
Review 2	1	1	2	0	0	1	1	0	1	0	0	8
Review 3	1	1	1	0	0	0	1	0	0	1	1	6

TF-IDF

- This is another method which is based on the frequency method but it is different to the bag-of-words approach in the sense that it takes into account, not just the occurrence of a word in a single document (or tweet) but in the entire corpus.
- TF-IDF works by penalizing the common words by assigning them lower weights while giving importance to words which are rare in the entire corpus but appear in good numbers in few documents.

$$\text{tfidf}_{i,j} = \text{tf}_{i,j} \times \log\left(\frac{N}{\text{df}_i}\right)$$

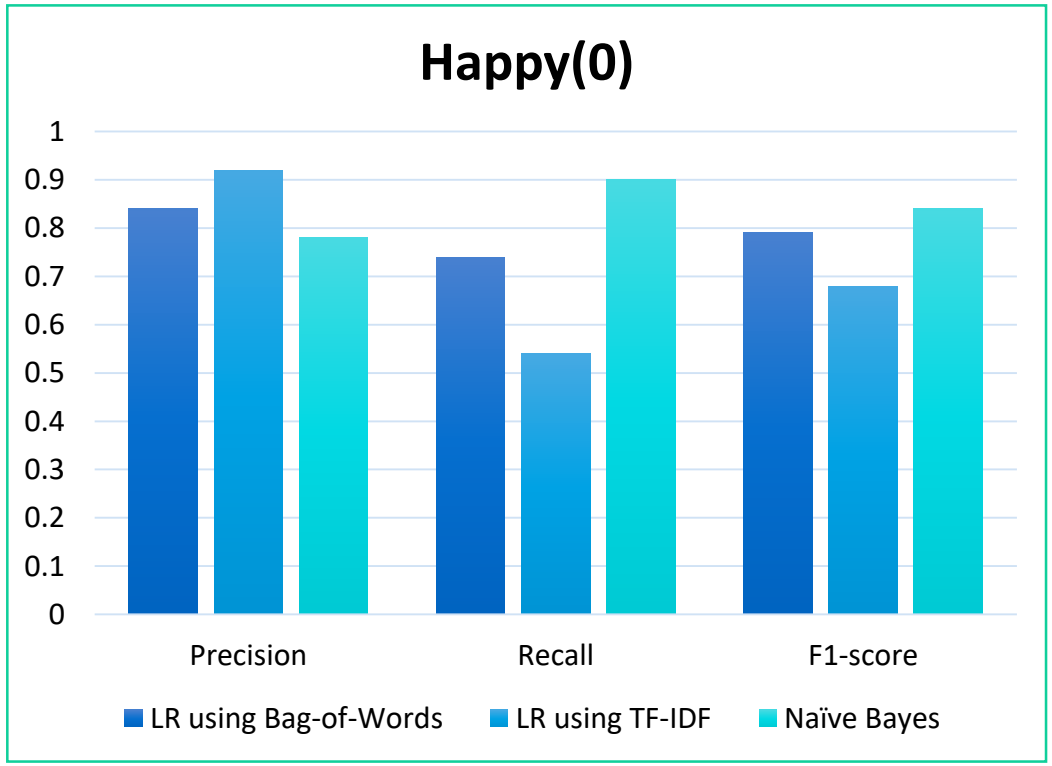
$\text{tf}_{i,j}$ = total number of occurences of i in j

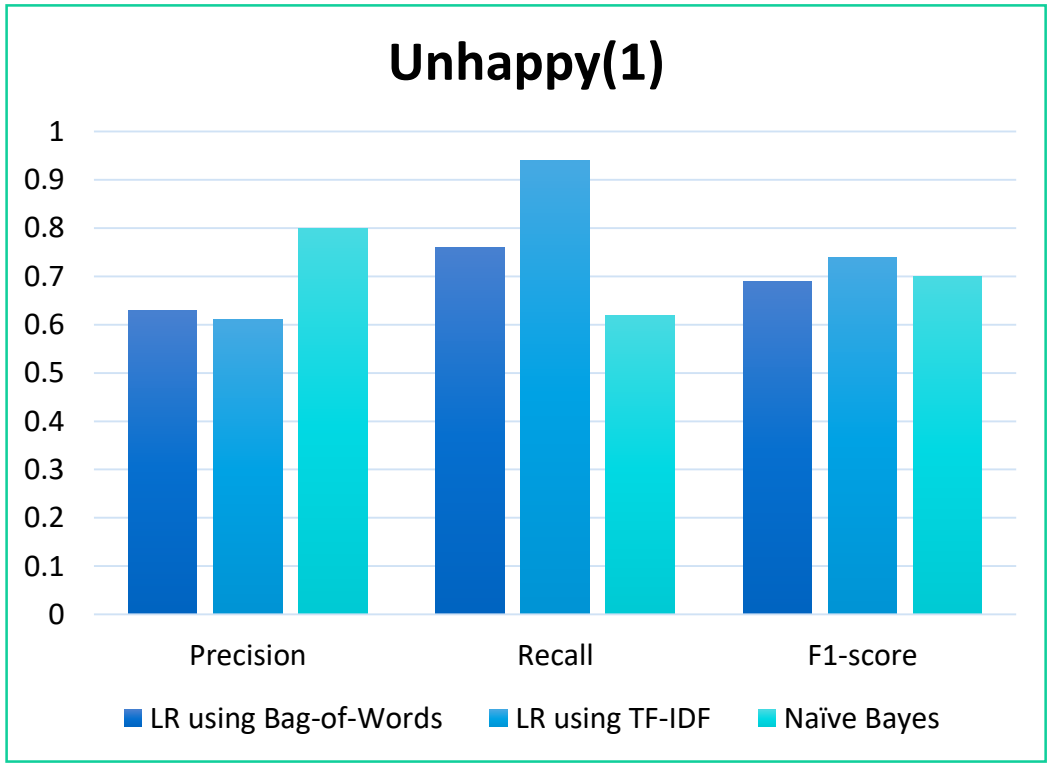
df_i = total number of documents (speeches) containing i

N = total number of documents (speeches)

Binary Classifier

Classifiers	Label	precision	recall	f1-score
LR using Bag-of-Words	Happy(0)	0.84	0.74	0.79
	Unhappy(1)	0.63	0.76	0.69
LR using TF-IDF	Happy(0)	0.92	0.54	0.68
	Unhappy(1)	0.61	0.94	0.74
Naive Bayes	Happy(0)	0.78	0.90	0.84
	Unhappy(1)	0.80	0.62	0.70





Binary Emotion Classifier

Logistic Regression – BoW

Accuracy: 75%

Logistic Regression – TF-IDF

Accuracy: 71%

Naïve Bayes

Accuracy: 79%

Multi-class Classifier

'Angry':

Most correlated unigrams:

- elections
- jhoot

Most correlated bigrams:

- ground reality
- sewer line

'Disgust':

Most correlated unigrams:

- corruption
- muft

Most correlated bigrams:

- ground reality
- Clean water

'Sad':

Most correlated unigrams:

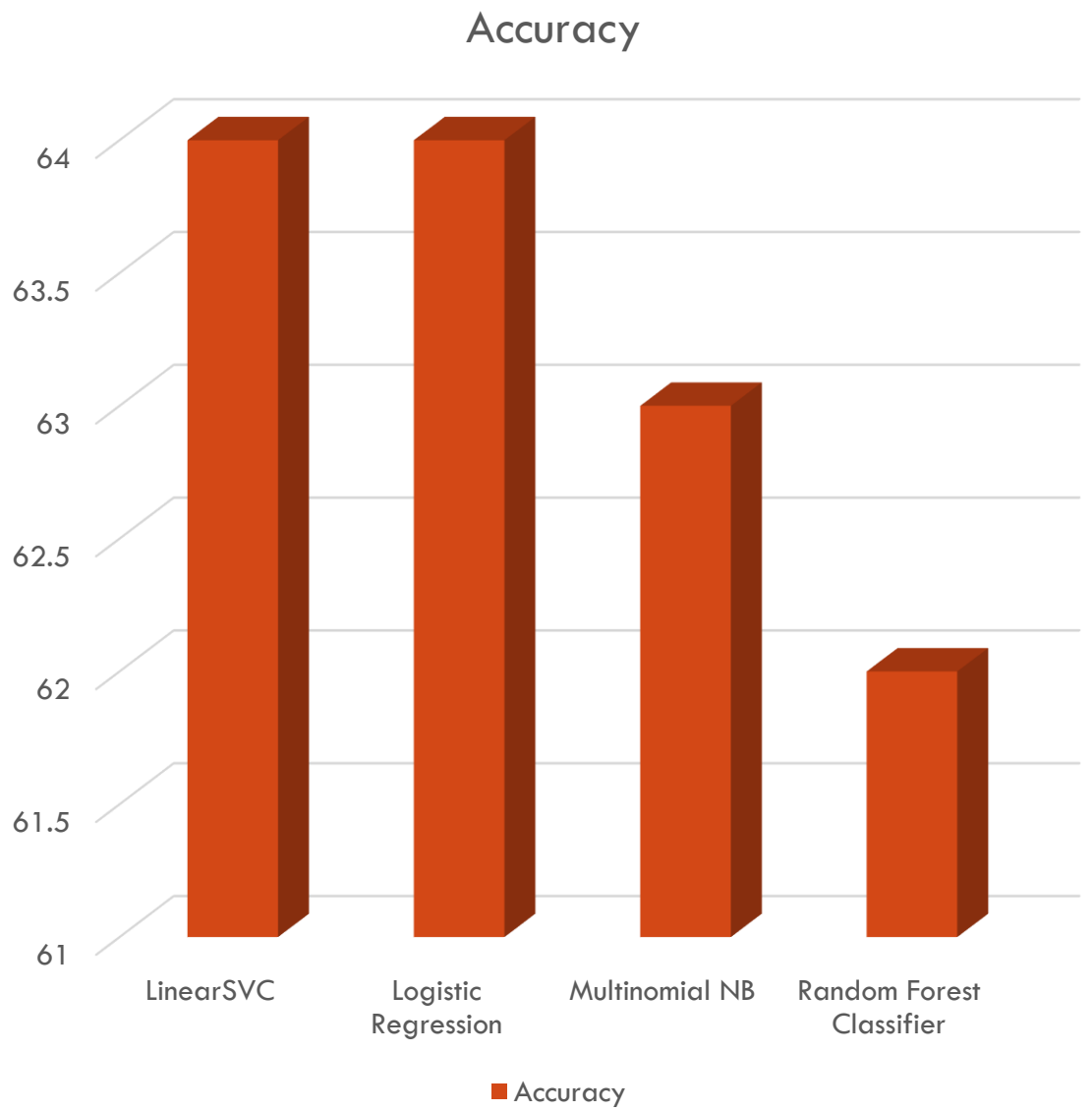
- respect
- help

Most correlated bigrams:

- vote nahi
- kisi vote

MULTI-CLASS CLASSIFIER-RESULTS

Model Name	Accuracy
Linear SVC	0.64
Logistic Regression	0.64
Multinomial NB	0.63
Random Forest Classifier	0.62



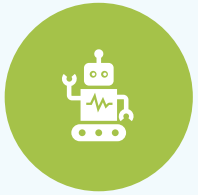
Future Work

We may increase the size of our Hinglish dataset.

User features like hashtags, emoticons, followers, tweet reply length, user mentions, username, date & time of tweet reply, can be considered to extract some meaning out of them.

Users usually post images, videos & GIFs as part of tweet-reply which can also be a part of corpus.

Language ambiguity.

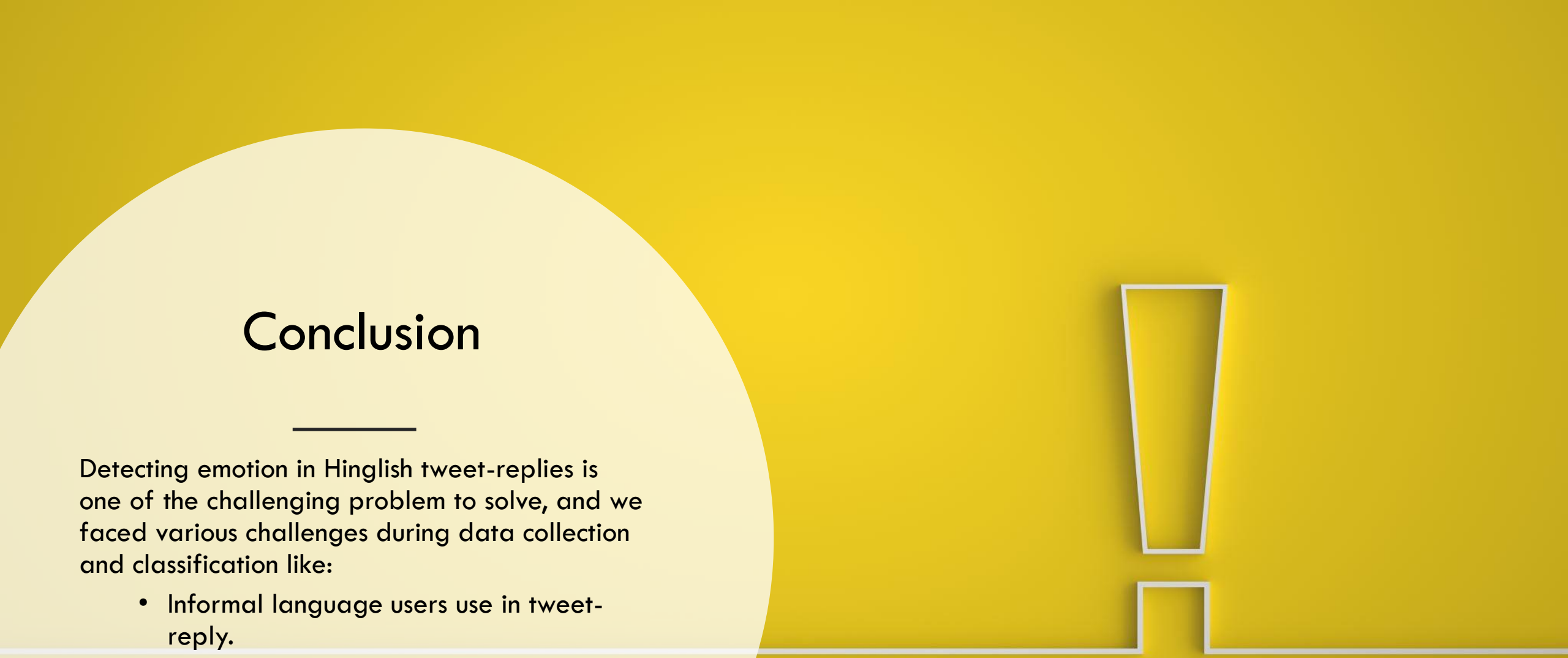


In the future, we can use Transfer Learning/Deep Learning approach to automatically detect emotion.

Conclusion

Detecting emotion in Hinglish tweet-replies is one of the challenging problem to solve, and we faced various challenges during data collection and classification like:

- Informal language users use in tweet-reply.
- Tweet-reply which don't have any emotion.
- Language ambiguity




Thank You
DIGITAL
INFORMATION
WORLD