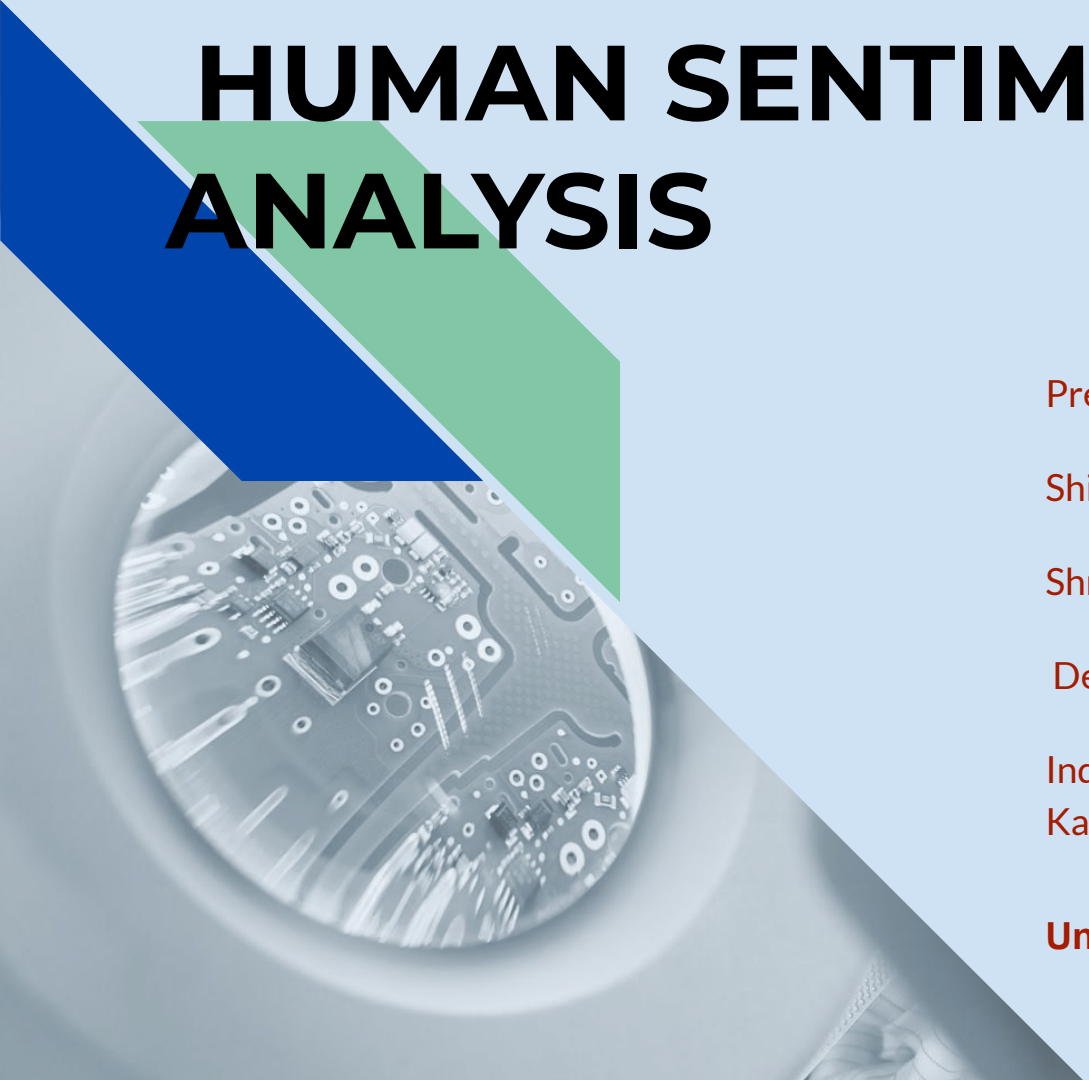
HUMAN SENTIMENT ANALYSIS



Presented By :

Shikha Sinha (0000190) and

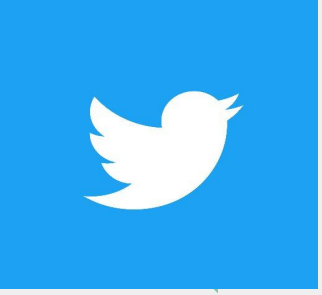
Shruti Sharma (0000192)

Department of CSE

Indian Institute of Information Technology, Kalyani

Under Guidance Of : Dr. Sanjoy Pratihar

Table Of Contents



Overview

- Twitter is a micro-blogging site where people post their opinion and express their views/sentiments on current topics.
- Many polls are done on it to extract people's opinion and analyse sentiment on a particular topic
- The goal is to collect such data which pertains to our interest, extract useful information from it, and then summarise the overall sentiment.

Problem Statement

01 The problem is to classify the polarity of a feature/dataset in a tweet. Also to detect the level of sarcasm present in dataset.

02 To implement an algorithm to automatically classify the text as positive, negative or neutral and to determine the interest of mass towards a particular topic of interest.

03 Graphical representation of sentiment analysis.

Project Objective

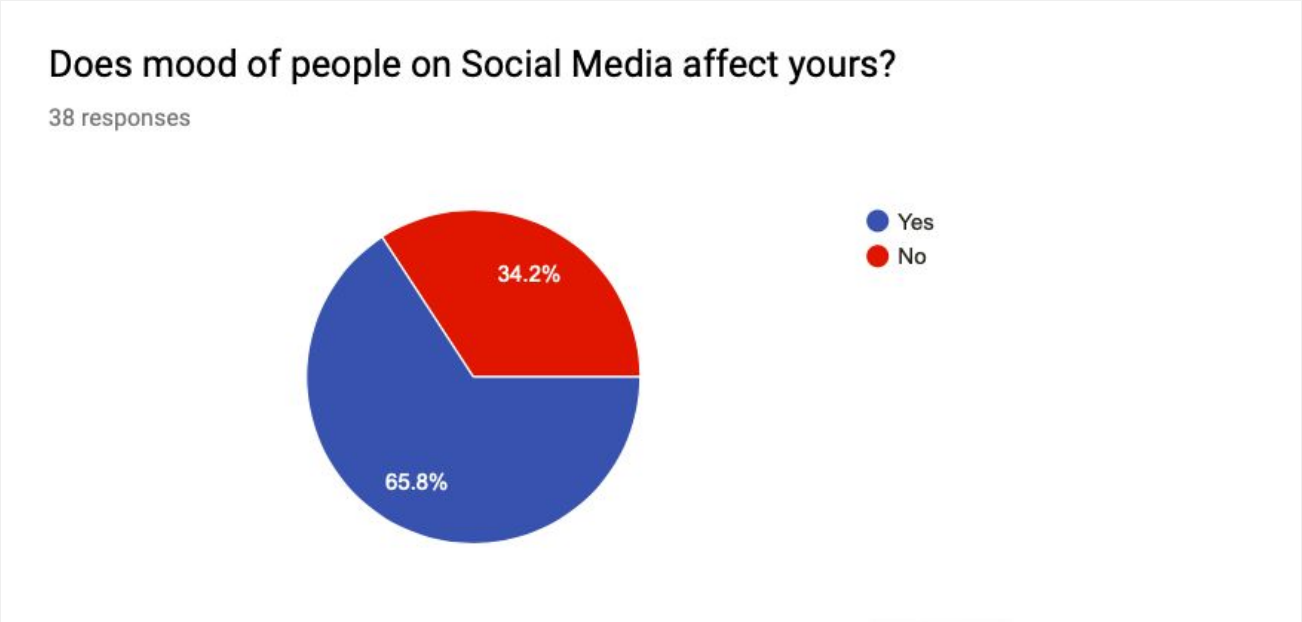
- Sentiments are subjective, and not facts. The amount of data generated on Internet these days is humongous.
- Sarcasm is a subtle linguistic trait, where usually the author states the opposite of what they mean.
- Sentiment analysis is in demand for its efficiency - which will take lots of time when done manually.

Target Audience

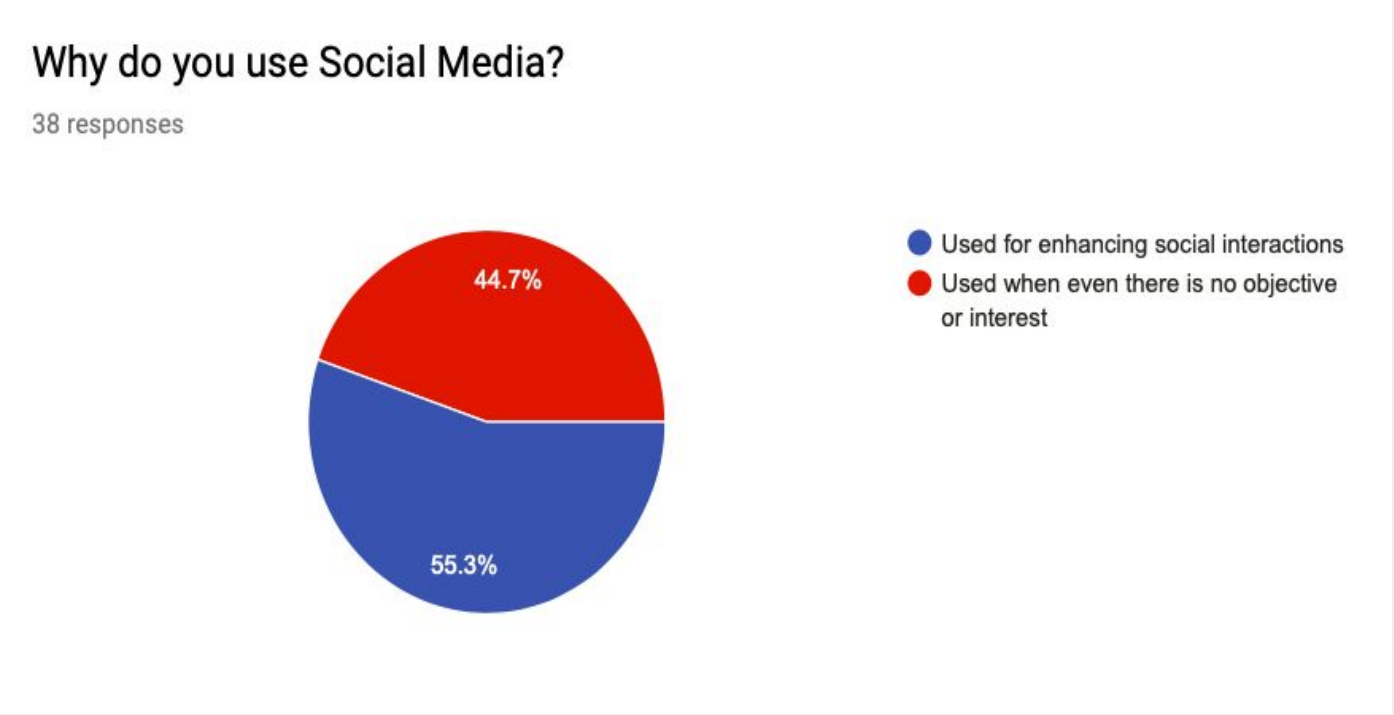
The questions Sentiment Analysis might ask are :

- Is this product positive or negative?
- Is this movie of my taste?
- Based on a sample of tweets, how are people responding to the campaign/ person/ event etc.
- Is this customer email satisfied or not?

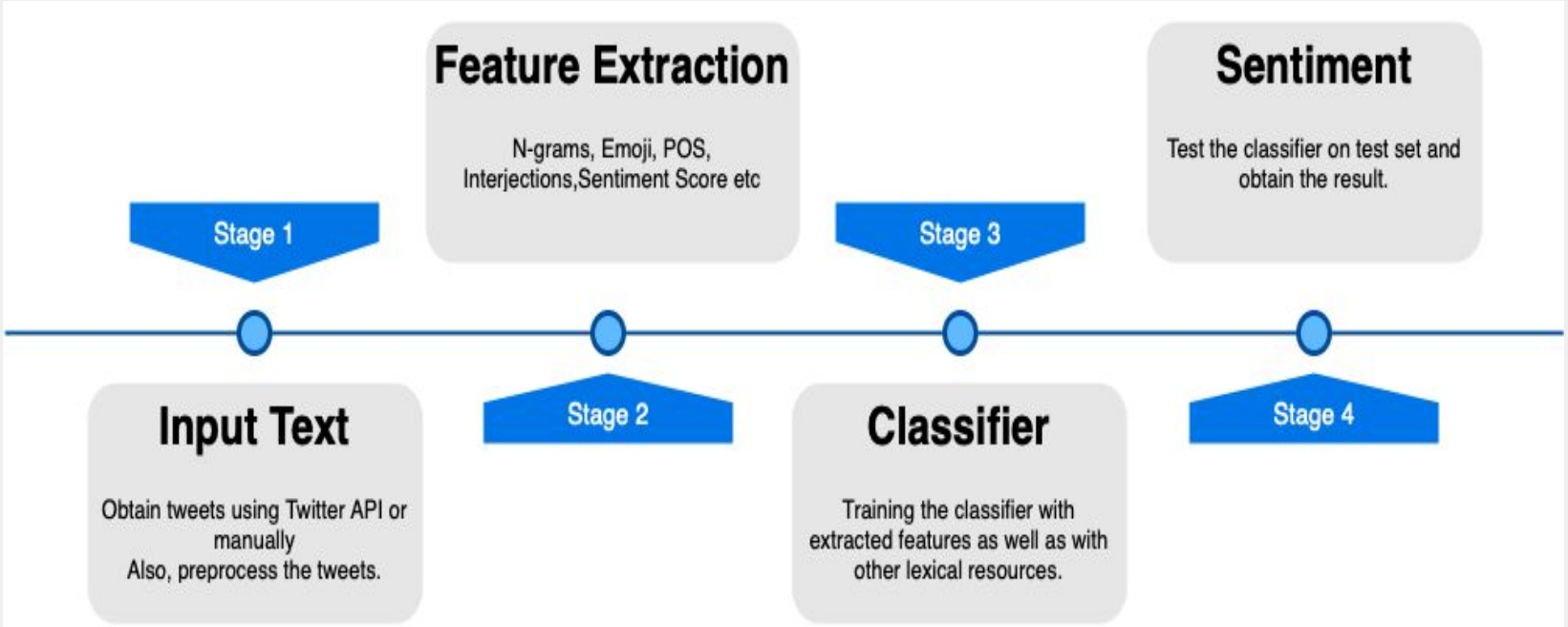
Survey Work


Does mood of people on Social Media affect yours?
38 responses
34.2%
65.8%
Yes
No

Survey Work (Cont.)



How to proceed?



Techniques and Tools

Techniques

- Text mining
- Topic modelling
- Sentiment analysis
- Social network analysis

Tools

- Twitter API
- R and its packages
- Python
- NLTK library
- WEKA

Methodology

1. We limit our study to English tweets as more resources are available for the processing of text in English language.
2. In our project, we integrate emojis as a feature to detect sarcasm or at least a change in polarity.
3. User mentions and URLs are removed from the tweet as they are not indicative of the original nature of the tweet.

Methodology

Feature Engineering:

In this project we use 21 special features along with usual unigrams and bigrams for classification.

Classification algorithms used:

1. Naive Bayes
2. Logistic Regression
3. Support Vector Machines
4. Random Forest
5. Neural Networks
6. Decision Trees

Features

The features are :

User mention

Exclamation

Question mark

Ellipsis

Emoji tweet polarity flip

UpperCase

Verbs

Emoji Sentiment

Passive aggressive count

Repeat Letters

Sentiment Score

Positive word count

Negative word count

Polarity flip

Nouns

Positive/Negative Intensifier(Adjective)

Hashtag polarity

Dataset

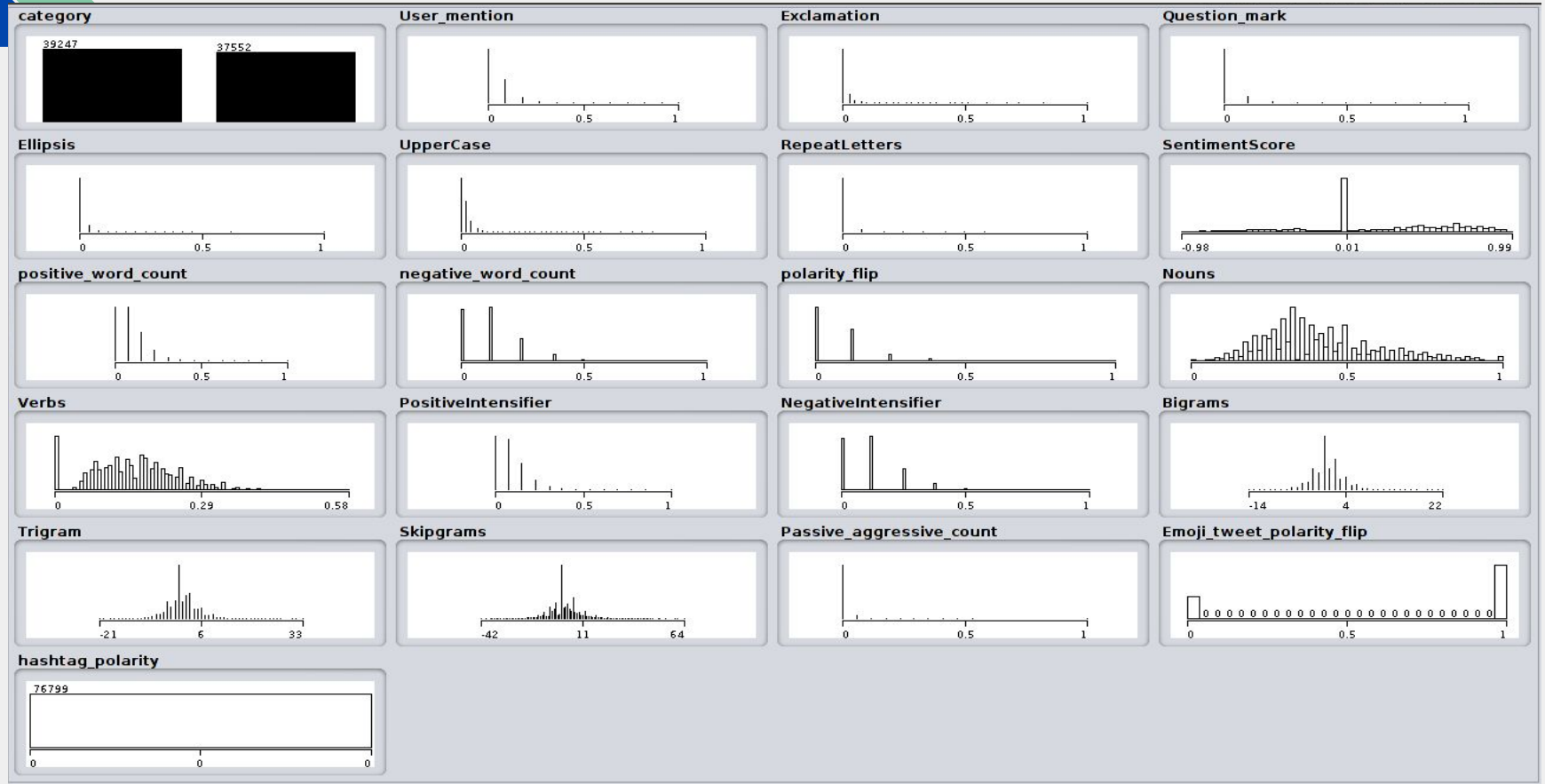
For evaluation purposes, we created a data set of manually annotated tweets. Even for people, it is not always easy to identify sarcasm in tweets because sarcasm often depends on conversational context that spans more than a single tweet.

We focus on identifying sarcasm that is self-contained in one tweet and does not depend on prior conversational context.

To ensure that our evaluation data had a healthy mix of both sarcastic and non-sarcastic tweets, we collected over 75000 tweets.

```
39211,0,"Never give up. Yes i Skipped 5 days of trainings! And i thought i will give up, but no!! Today I continue my
challenge and today - DAY 17 ✅"
39212,0,I just got full access to Code School for 48 hours! Get your Hall Pass now and go on a learning-spree for fre
39213,0,"bener kan tebakan gw, pasti ost, but still can't wait for it RT allkpop: Hwanhee to come back with 'Pretty
Man' OST
"
39214,0,@Cinsoft I know. It's very hard to work with those kind of people. :-(
39215,0,Vote for a Castle episode: (Still) vs. Beauty and the Beast (Man or Beast?)
39216,0,"@kndmolfese Hola! Look at my polish version of ""Podemos""  What do you think about it? :)"
39217,1,"b""So you text me, ask me a stupid question, don't bother asking how I am, then never text back?.... you're
great friend. #Sarcasm"""
39218,1,"b""can't wait for the second sociology test this week, next lesson #sarcasm #fuckedit"""
39219,1,"b'nothing better than taking a nap after work, to wake up at midnight, then to take a shower only to go to
sleep after that again #sarcasm'"
```

Feature Graph



Results

Classifier Model	Accuracy (10-fold cross-validation)
• Random Forest	94.54%
• Logistic Regression	91.55%
• SVM	91.18%
• Naive Bayes	81.07%
• Decision Tree	91.91%

Confusion Matrix for Random Forest (Acc = 94.5%)

```
=== Confusion Matrix ===

     a     b   <-- classified as
 36393  2854 |     a = 0
  1343 36209 |     b = 1
```

Applications And Future Goals

- Review-related analysis
- Developing ‘hate mail filters’ analogous to ‘spam mail filters’
- Question-answering (Opinion-oriented questions may involve different treatment)
- Tracking time spent on social media
- This can be further extended to use PYTHON for more analysis of big data.
- Using Deep-Learning based approaches for analysis
- Incorporating cultural/social aspects for sarcasm analysis

References

- https://begriffs.com/posts/2015-02-25-text-mining-in-r.html
- https://knightlab.northwestern.edu/2014/03/15/a-beginners-guide-to-collecting-twitter-https://joparga3.github.io/Udemy_text_analysis/#example-selecting-all-text-from-page---not-selecting-contents-images-etc
- Twitter Sentiment Analysis System - Shaunak Joshi and Deepali Deshpande , International Journal of Computer Applications (0975 – 8887) Volume 180 – No.47, June 2018
- P. Basile, V. Basile, M. Nissim, N. Novielli, V. Patti.“Sentiment Analysis of Microblogging Data”. To appearin Encyclopedia of Social Network Analysis and Mining, Springer. In press.data-and-a-bit-of-web-scraping/
- https://github.com/Mandar-Kulkarni/Sarcasm-Detection-in-Tweets
- Automatic Sarcasm Detection: A Survey - Aditya Joshi, Pushpak Bhattacharyya, Mark J Carman

Special Thanks to our Mentor - Dr. Sanjoy Pratihar