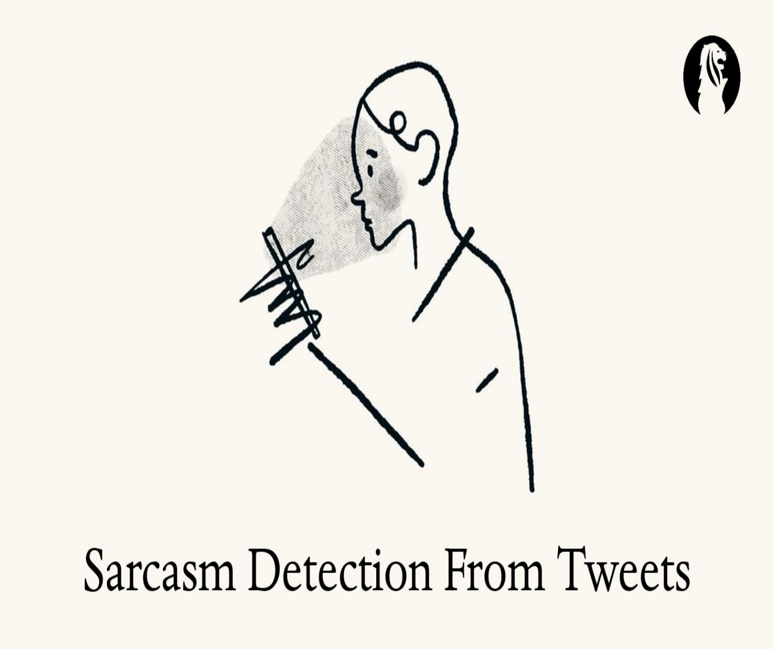
Sarcasm Detection in Tweets to Reduce Misinformation

By – Group 5

- Divya Khatarkar
- Sanskriti Shukla
- Jiya Goyal


Sarcasm Detection From Tweets

Business Problem:

Sarcasm is a common linguistic tool on social media, often used to convey humor or criticism. However, its subtle nature makes it easily misinterpreted, especially by automated systems. When sarcastic tweets spread false information, it becomes difficult to distinguish between genuine news and fabricated content. This leads to rapidly disseminating misinformation, impacting public opinion and decision-making.

Proposed Solution:

Developing accurate sarcasm detection technology can help mitigate the spread of misinformation due to tweets. This technology can be integrated into social media platforms to flag potentially sarcastic content, allowing users to interpret information correctly.

Dataset Description:

The dataset contain a collection of tweets with correspond to a labels indicating whether the tweets are sarcastic and non sarcastic.

Data Format:

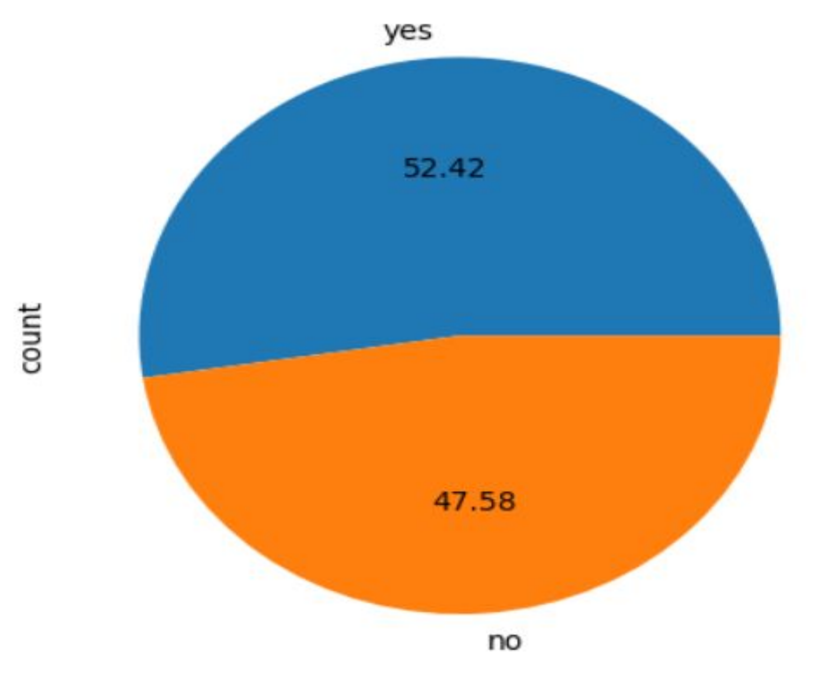
- Tweet: The text content of the tweet.
- Sarcasm: A label (yes or no) indicating Sarcasm (yes) or non-sarcasm(no).

Tweet	Sarcasm
Fantastic service yet again from EE. 1st you u...	yes
Not sure if that was or. I will take it! face_...	yes
Barely 9 am and already shaking with rage.	yes
I guess that proves it then. Black folks have ...	yes
Does this tweet need a tag	yes
...	...
just home from uni and im knackered have no en...	no
as we have been approved by god to be entruste...	no
i love these boys so much and id be so gratefu...	yes
if youre in ny wed love to have you speak on w...	no
i learned so much chatting w u today thank you...	no

Data Visualization:

The dataset used for this sarcasm detection model comprises 6,930 Twitter comments.

There are 3633 sarcastic tweets and 3297 non-sarcastic tweets in the dataset.

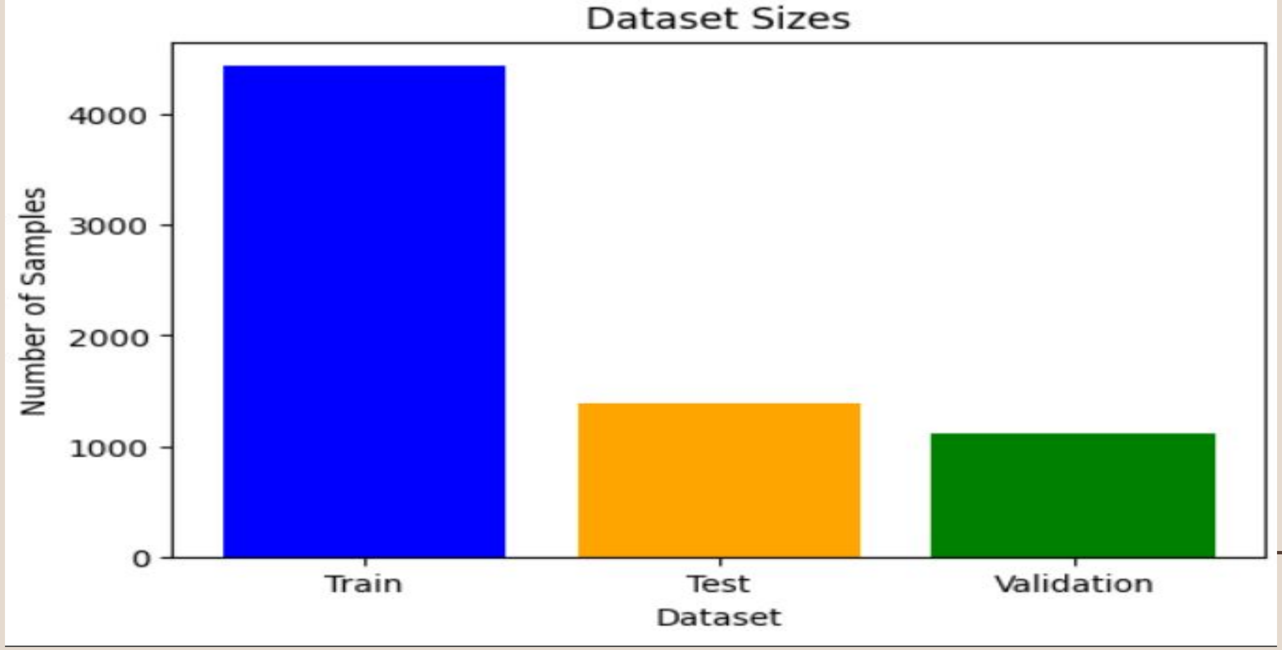


Train, Test and Validation data size:

Train size: 4435

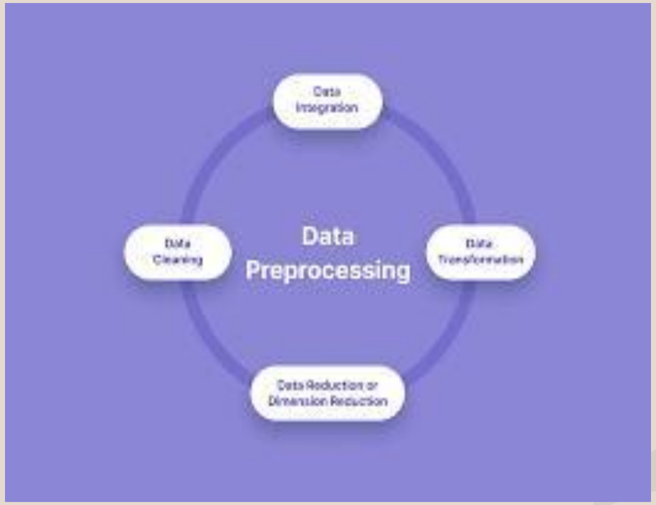
Test size: 1386

Validation size: 1109



Data Preprocessing:

- Drop unwanted column.
- Replace emoji with its text description.
- Replace abbreviations with their expansions.
- Remove special characters like #, $, %, *, @ etc.
- Check for duplicates and null values.
- Removed URL’s.
- Numeric labels are replace with textual labels.
- Merged two dataset.



Tokenization and Embedding Techniques:

Word Tokenization: Useful for most NLP tasks where words are treated as basic units, such as text classification or sentiment analysis. Word tokenization breaks text into individual words or terms.

TF-IDF Vectorization (Term Frequency-Inverse Document Frequency): Represent documents as vectors based on word importance, where high frequency terms across documents are weighted lower.

```
['Fantastic', 'service', 'yet', 'again', 'from...
['Not', 'sure', 'if', 'that', 'was', 'or', '.'...
['Barely', '9', 'am', 'and', 'already', 'shaki...
['I', 'guess', 'that', 'proves', 'it', 'then',...
    ['Does', 'this', 'tweet', 'need', 'a', 'tag']
                                                ...
['just', 'home', 'from', 'uni', 'and', 'im', '...
['as', 'we', 'have', 'been', 'approved', 'by',...
['i', 'love', 'these', 'boys', 'so', 'much', '...
['if', 'youre', 'in', 'ny', 'wed', 'love', 'to...
['i', 'learned', 'so', 'much', 'chatting', 'w'...
```

Modeling

Different model used and their corresponding Accuracy:-

ML Model:

1. Gradient Boosting Machine - 79.56%
2. Naïve Bayes Model- 78.22%
3. Logistic Regression - 77.89%
4. Support Vector Machine- 77.55%
5. Decision tree - 76.72%
6. K - Nearest Neighbor - 43.22%

DL Model:

7. CNN - 90.84%
8. GRU - 89.83%
9. Bidirectional LSTM - 89.25%
10. BERT - 67.31%
11. RNN - 53.25%

Evaluation metrics of Best Model

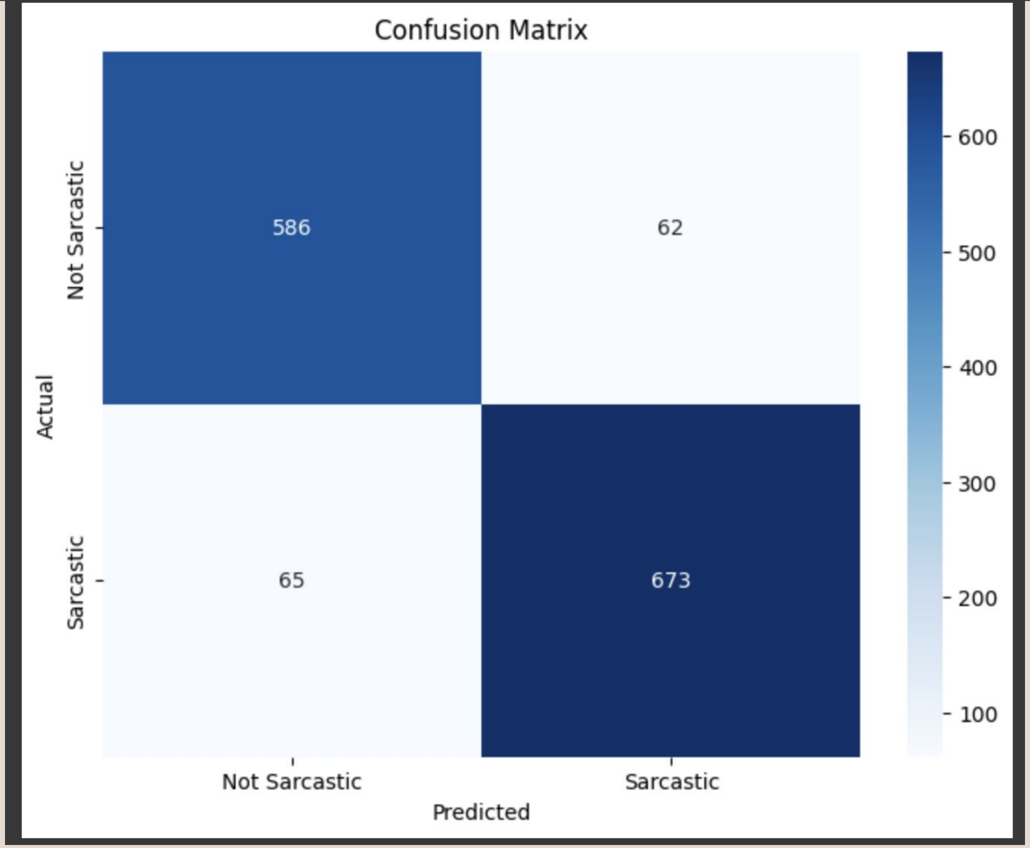
Best Model-

CNN(Convolutional Neural Network)

Reason – CNNs excel at capturing local patterns in data. For sarcasm, they can identify key phrases or word combinations indicative of sarcastic intent without relying on predefined features.

Classification Report:

	precision	recall	f1-score	support
0	0.90	0.90	0.90	648
1	0.92	0.91	0.91	738
accuracy			0.91	1386
macro avg	0.91	0.91	0.91	1386
weighted avg	0.91	0.91	0.91	1386



Thank you