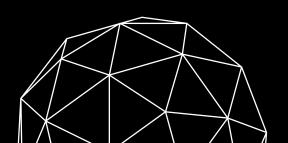
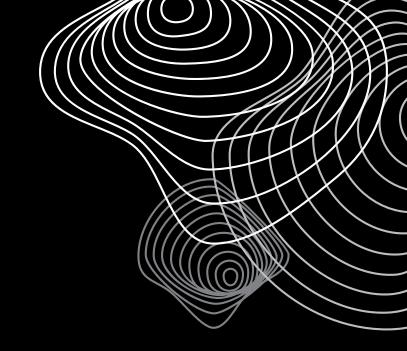
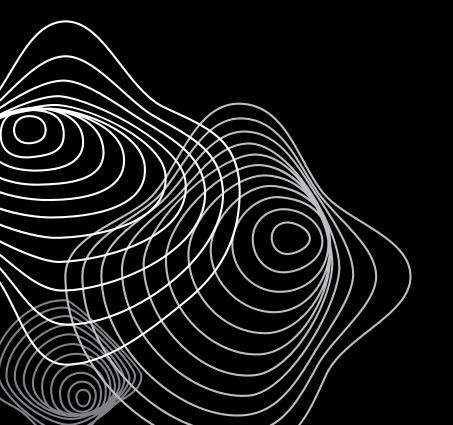


- 1. N CHARAN SAI RA2311003011343
- 2. AKASH SINGH RA2311026010638

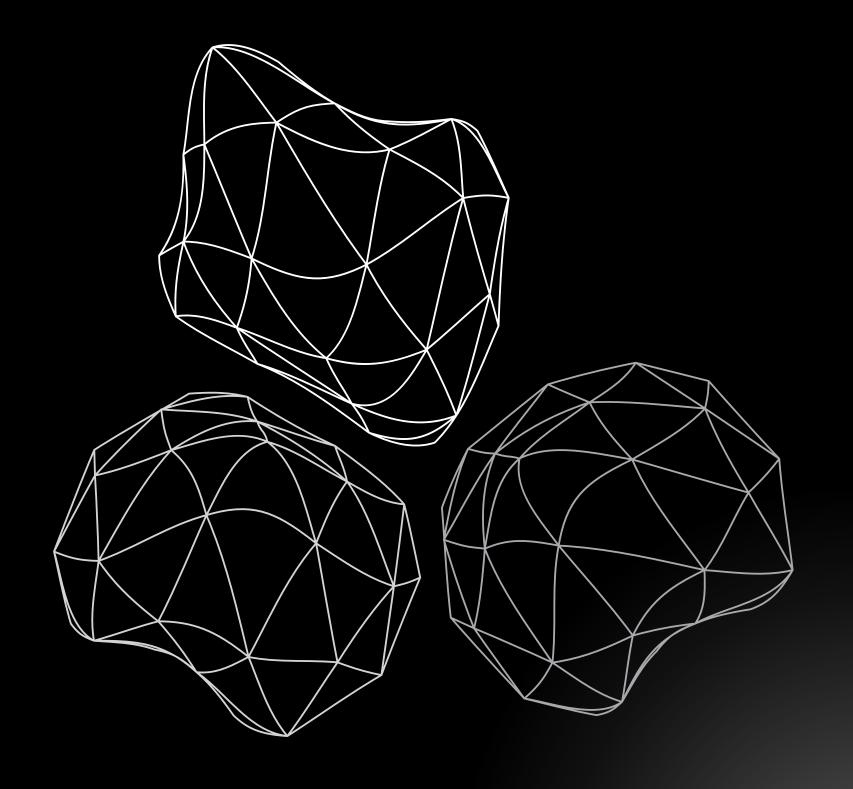


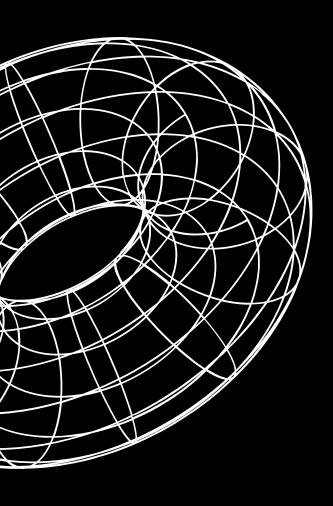




PROBLEM STATEMENT

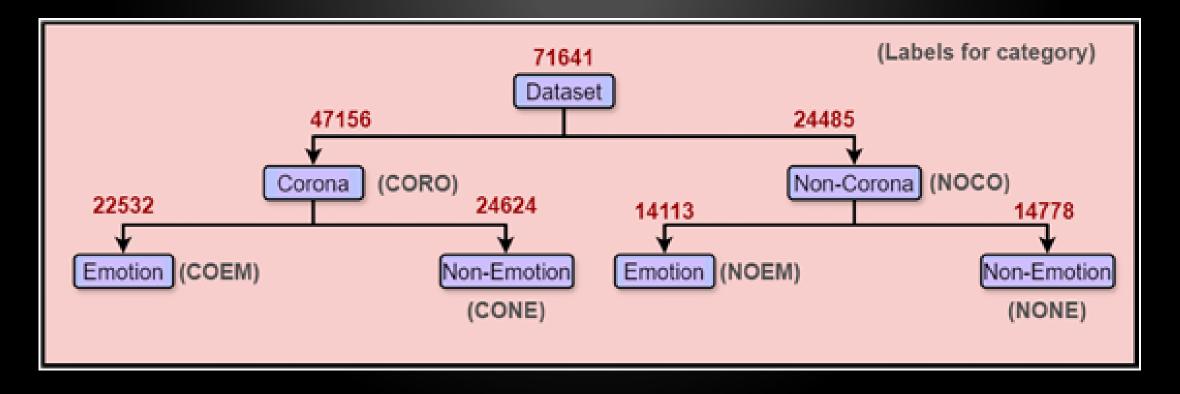
Multi-Level Classification of Emotional Well-being Tweets Based on Content and Sentiment.

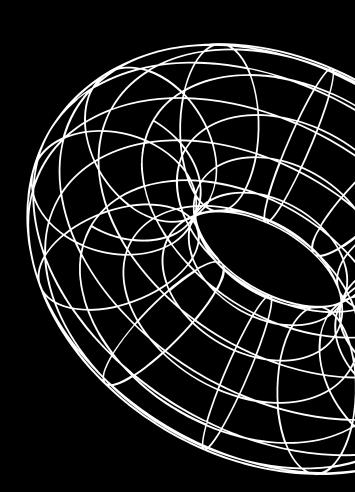




MISSION

Distinguishing between COVID-19-related and non-COVID-19 tweets while further classifying them based on emotions is essential for understanding public sentiment and misinformation trends.

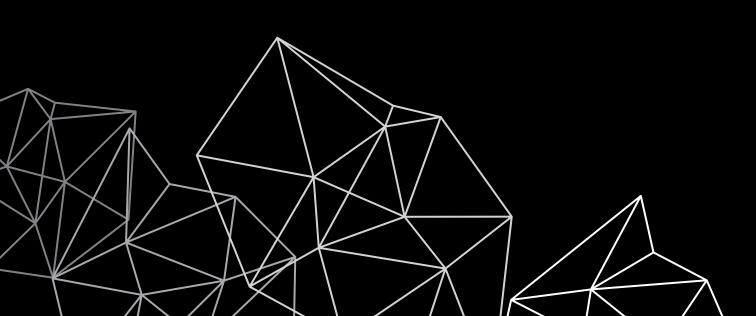


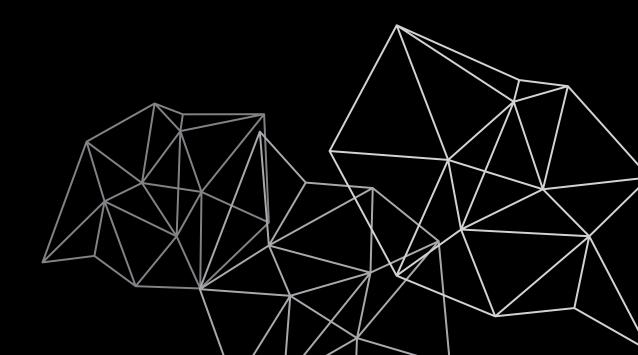


STEP-BY-STEP PROCESS FOR MULTI-LEVEL CLASSIFICATION

This process involves:

- 1.Loading and Understanding the Dataset
- 2. Data Preprocessing (Cleaning & Tokenization)
- 3. Feature Engineering (TF-IDF/Word Embeddings)
- 4. Building Classification Models (First-Level & Second-Level)
- 5. Evaluating Model Performance
- 6. Making Predictions on the basis of Dataset





FIRST LEVEL CLASSIFICATION

.9942

.99

.99

ACCURACY

MACRO AVERAGE WEIGHTED AVERAGE

SECOND LEVEL CLASSIFICATION

.9482

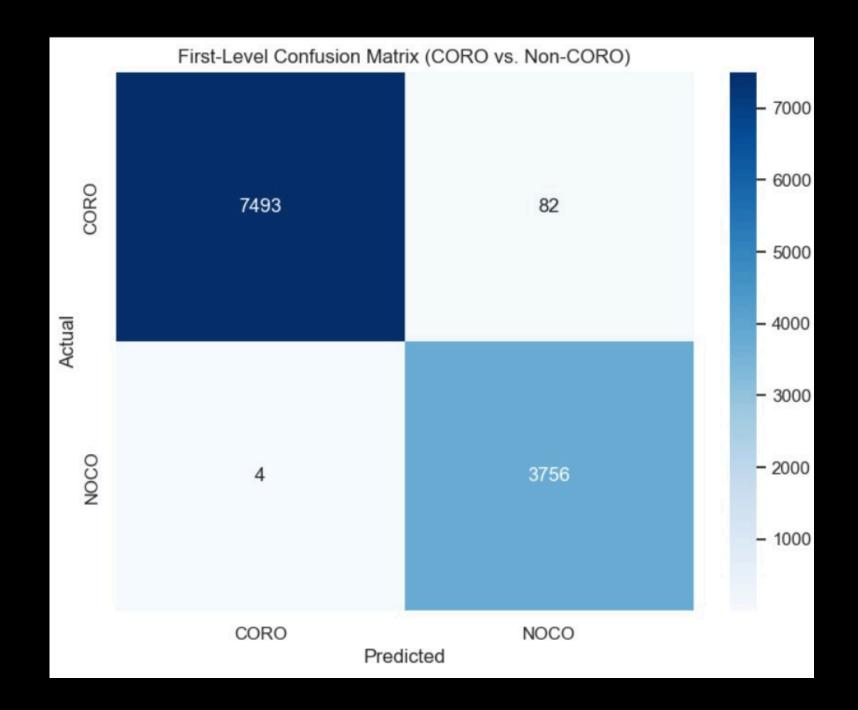
.94

.95

ACCURACY

MACRO AVERAGE WEIGHTED AVERAGE

```
First-Level Classification (CORO vs. Non-CORO):
Best Parameters: {'C': 10, 'penalty': 'l2', 'solver': 'liblinear'}
Test Accuracy: 0.9924
                                        0.99
                                                 11335
    accuracy
                                        0.99
                   0.99
                             0.99
                                                 11335
   macro avg
weighted avg
                   0.99
                             0.99
                                        0.99
                                                 11335
```



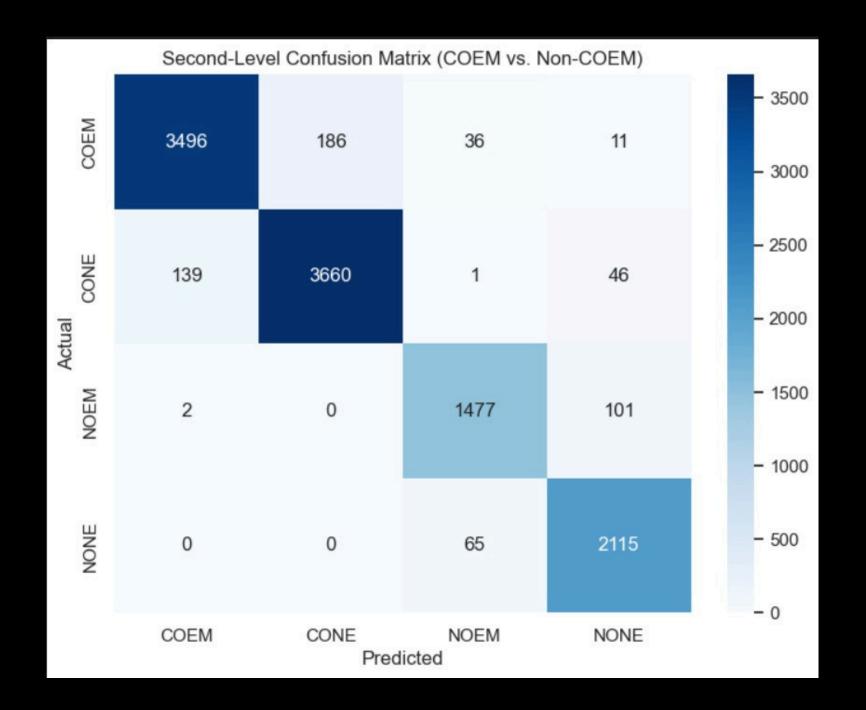
Second-Level Classification (COEM vs. Non-COEM):

Best Parameters: {'C': 1, 'penalty': 'l1', 'solver': 'liblinear'}

Test Accuracy: 0.9482

Classification Report:

	precision	recall	f1-score	support	
COEM	0.96	0.94	0.95	3729	
CONE	0.95	0.95	0.95	3846	
NOEM	0.94	0.93	0.94	1580	
NONE	0.93	0.97	0.95	2180	
accuracy			0.95	11335	
macro avg	0.94	0.95	0.95	11335	
weighted avg	0.95	0.95	0.95	11335	





We have used different libraries of python:-

- 1.Pandas
- 2.numpy
- 3.re
- 4.matplotlib
- 5.seaborn
- 6.scikit learn

We have used Logical Regression for the categorical outcomes and TF-IDF

OUTPUT

```
Welcome to the Tweet Classification CLI!

Type 'exit' to quit.

First-Level Prediction (CORO/Non-CORO): CORO

Second-Level Prediction (COEM/Non-COEM): COEM

Exiting the program.
```

