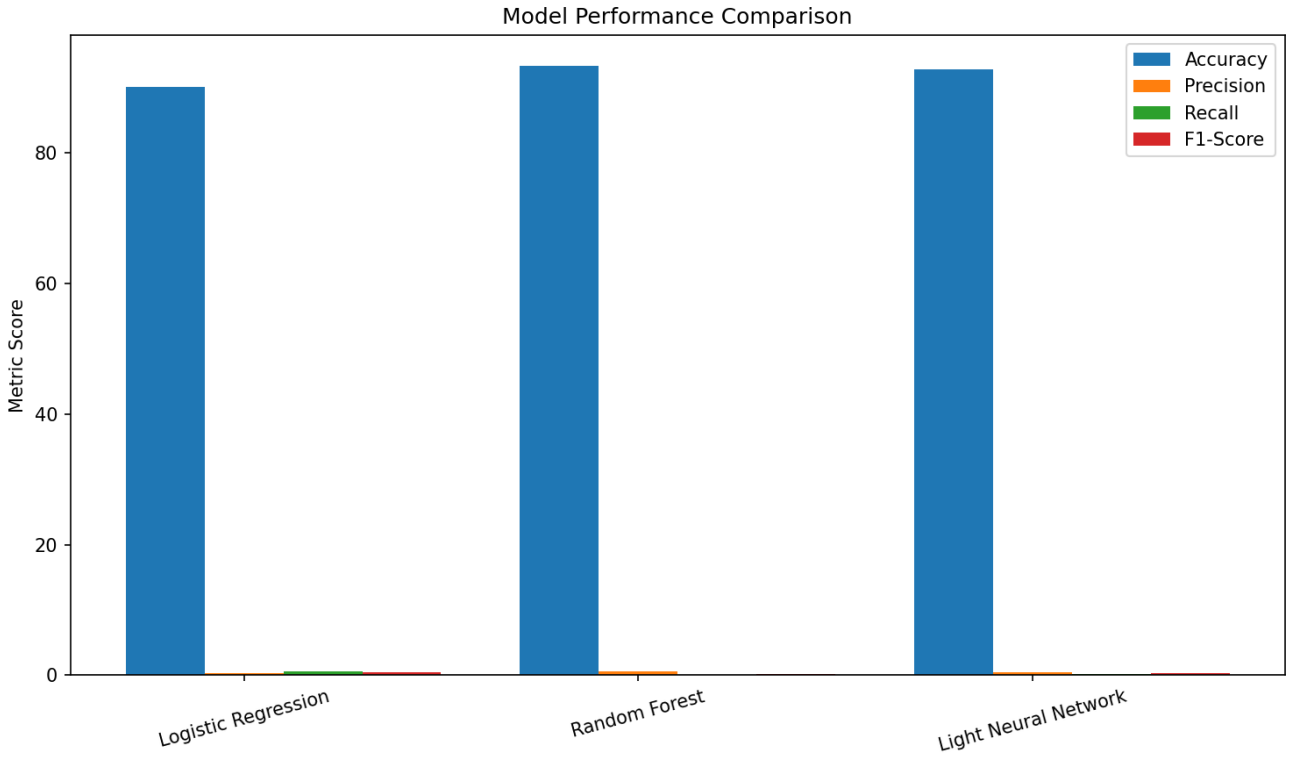
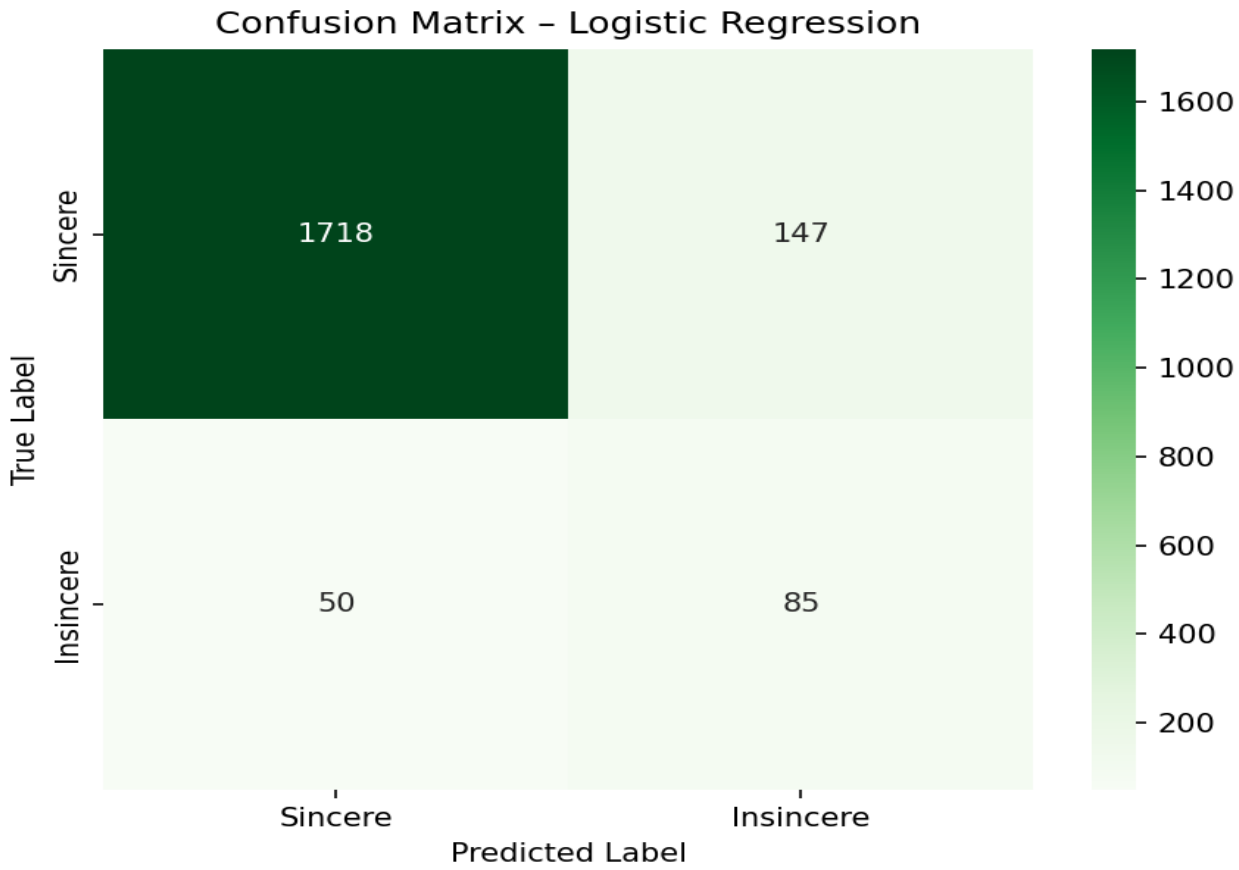
ML ASSIGNMENT OUTPUT SCREENSHOTS

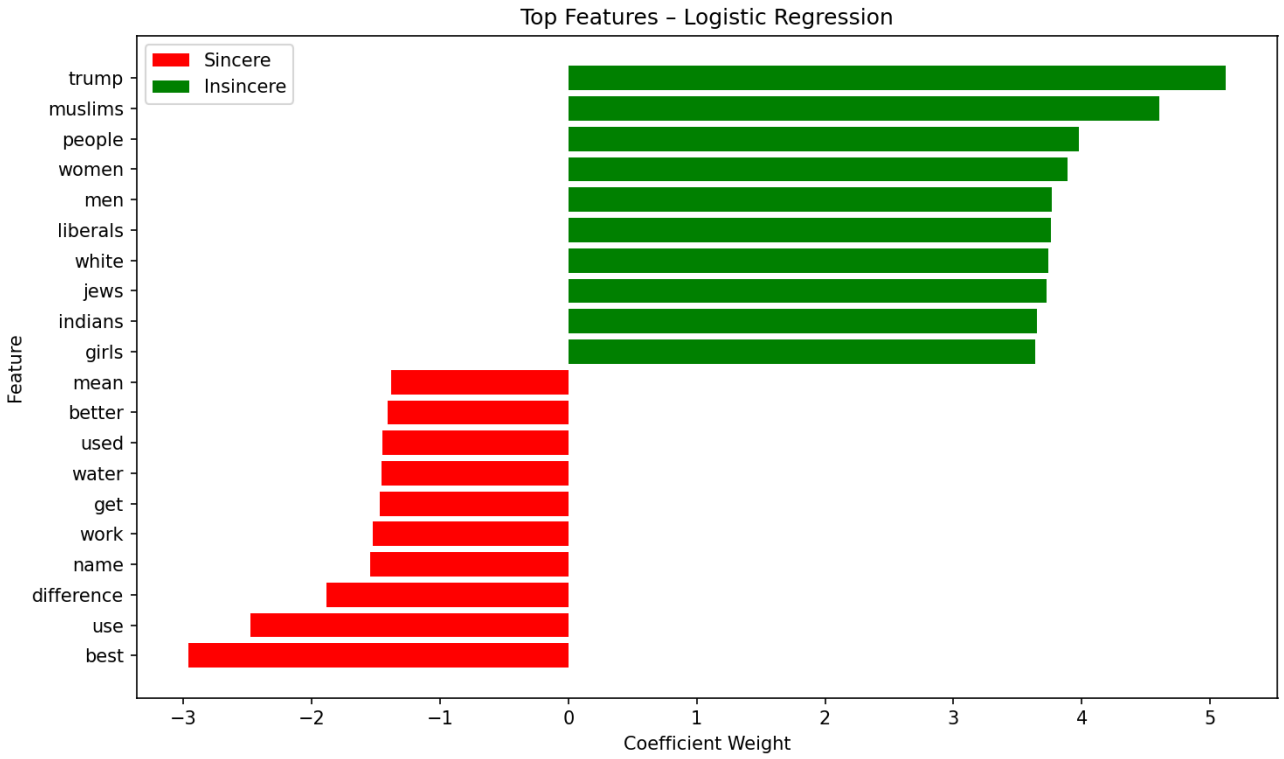
Team ID : 16

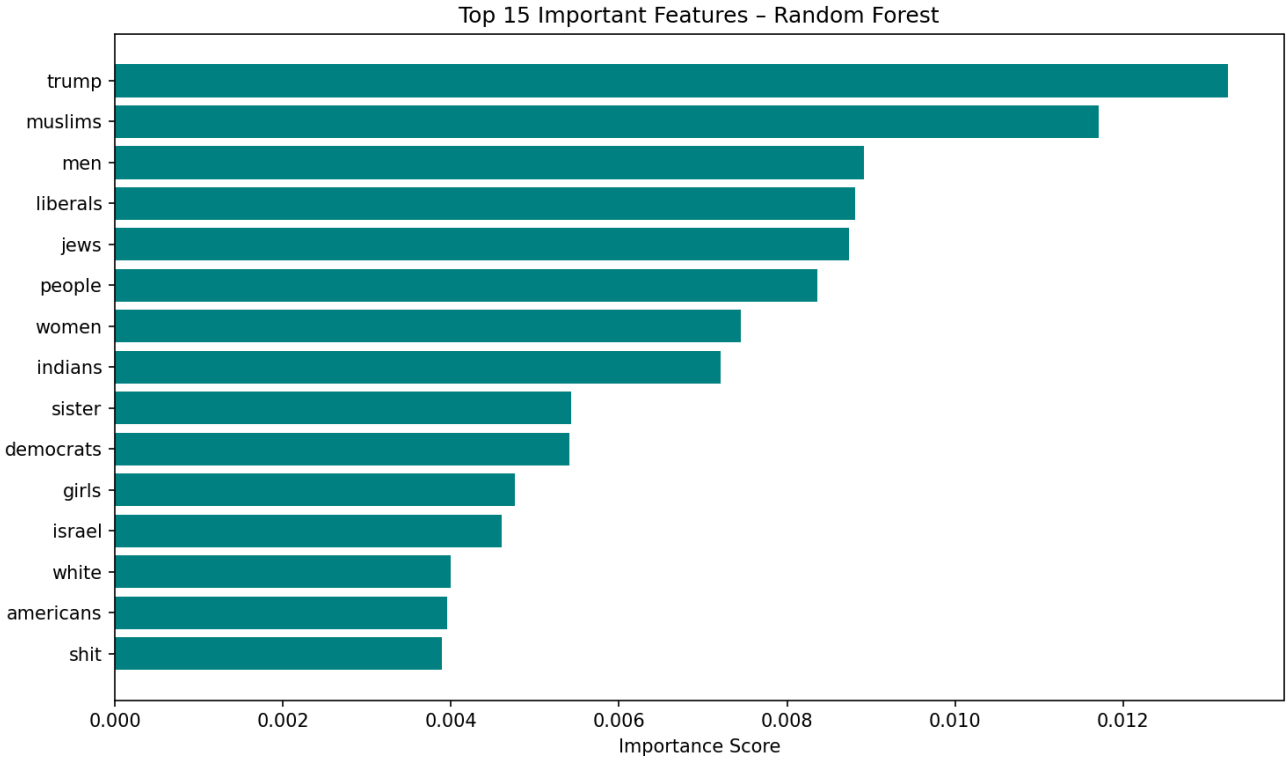
1) Run – "python main.py"

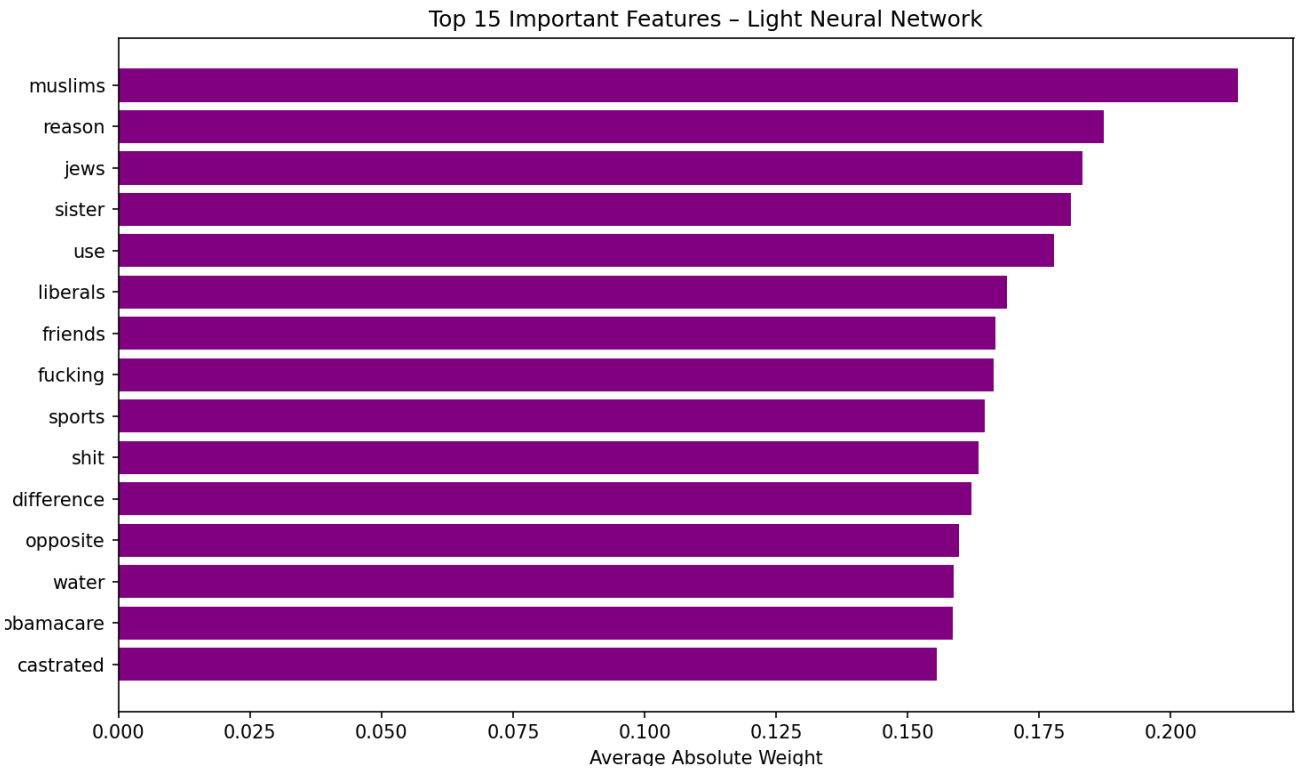
```
PS C:\Users\Chetan\Downloads\ML-PROJECT> python main.py
Loading dataset from train.csv...
Dataset Loaded: 10000 samples
Cleaning text data...
Applying TF-IDF vectorization...

Training Logistic Regression...
Logistic Regression done → Accuracy: 90.15%, Time: 0.17s

Training Random Forest...
Random Forest done → Accuracy: 93.35%, Time: 9.34s

Training Light Neural Network...
C:\Users\Chetan\AppData\Local\Programs\Python\Python312\Lib\site-packages\sklearn\neural_network\_multilayer_perceptron.py
:781: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (20) reached and the optimization hasn't converged yet.
  warnings.warn(
Light Neural Network done → Accuracy: 92.8%, Time: 33.04s

Model Performance Summary:

| Model               | Accuracy | Precision | Recall | F1 | FP | FN | Train Time (s) |
|:--------------------|:---------|:----------|:-------|:---|:---|:---|:---------------|
| Logistic Regression | 90.15%   | 0.37      | 0.63   | 0.46 | 147 | 50 | 0.17          |
| Random Forest       | 93.35%   | 0.56      | 0.07   | 0.12 | 7   | 126 | 9.34          |
| Light Neural Network | 92.8%   | 0.44      | 0.25   | 0.32 | 43  | 101 | 33.04         |

Best Performing Model: Logistic Regression with F1-Score = 0.46
Detailed Classification Report:

              precision    recall  f1-score   support

     Sincere       0.97      0.92      0.95      1865
   Insincere       0.37      0.63      0.46       135

    accuracy                           0.90      2000
   macro avg       0.67      0.78      0.70      2000
weighted avg       0.93      0.90      0.91      2000


Model and vectorizer saved successfully → Logistic Regression

Training, Evaluation, and Visualization Completed Successfully!
PS C:\Users\Chetan\Downloads\ML-PROJECT>
```


Model Performance Comparison
Accuracy
Precision
Recall
F1-Score
Metric Score
0
20
40
60
80
Logistic Regression
Random Forest
Light Neural Network


Confusion Matrix – Logistic Regression
1718
147
50
85
Sincere
Insincere
True Label
Sincere
Insincere
Predicted Label
1600
1400
1200
1000
800
600
400
200


Top Features – Logistic Regression
Sincere
Insincere
trump
muslims
people
women
men
liberals
white
jews
indians
girls
mean
better
used
water
get
work
name
difference
use
best
Feature
−3
−2
−1
0
1
2
3
4
5
Coefficient Weight


Top 15 Important Features – Random Forest
trump
muslims
men
liberals
jews
people
women
indians
sister
democrats
girls
israel
white
americans
shit
0.000
0.002
0.004
0.006
0.008
0.010
0.012
Importance Score

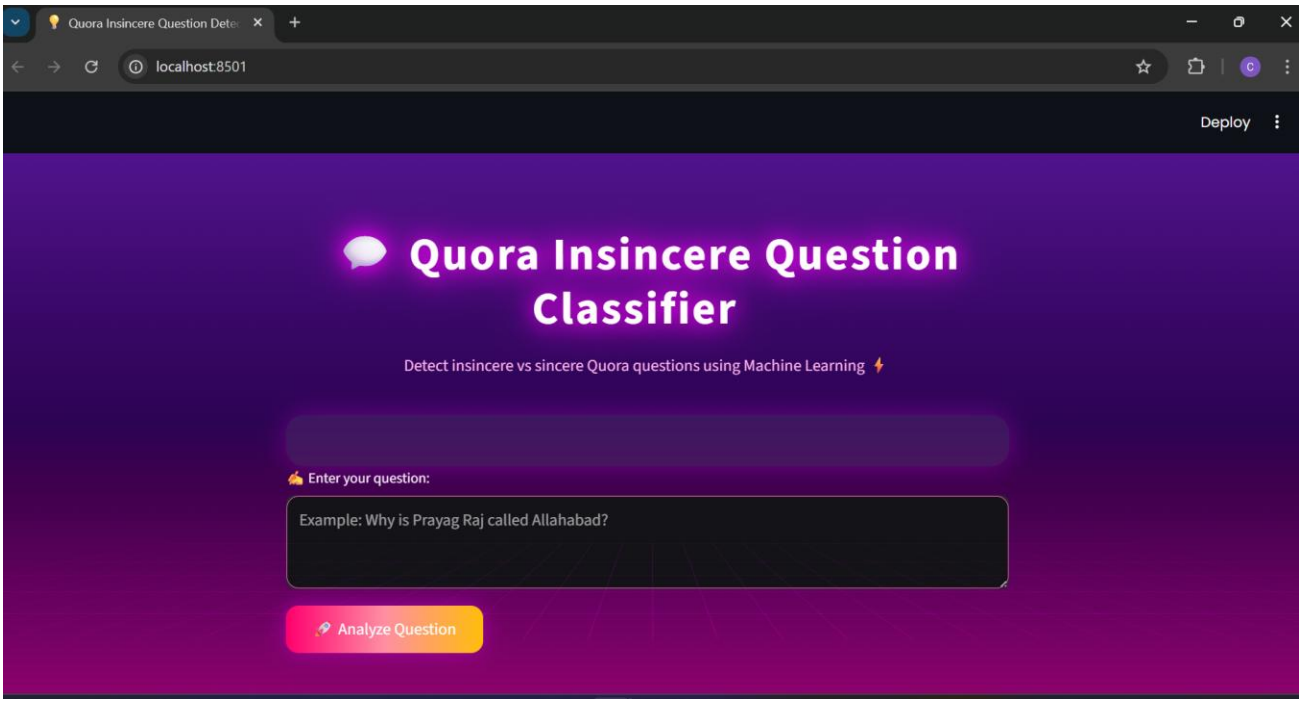


2) We made UI for our Project :

If we run command –

"streamlit run app.py"

```
PS C:\Users\Chetan\Downloads\ML-PROJECT> streamlit run app.py

  You can now view your Streamlit app in your browser.

  Local URL: http://localhost:8501
  Network URL: http://172.16.127.150:8501
```




```
PS C:\Users\Chetan\Downloads\ML-PROJECT> streamlit run app.py

  You can now view your Streamlit app in your browser.

  Local URL: http://localhost:8501
  Network URL: http://172.16.127.150:8501

  Stopping...
PS C:\Users\Chetan\Downloads\ML-PROJECT>
```

Thank You