# Machine Learning Spam Detection Report

Automatically generated evaluation report.
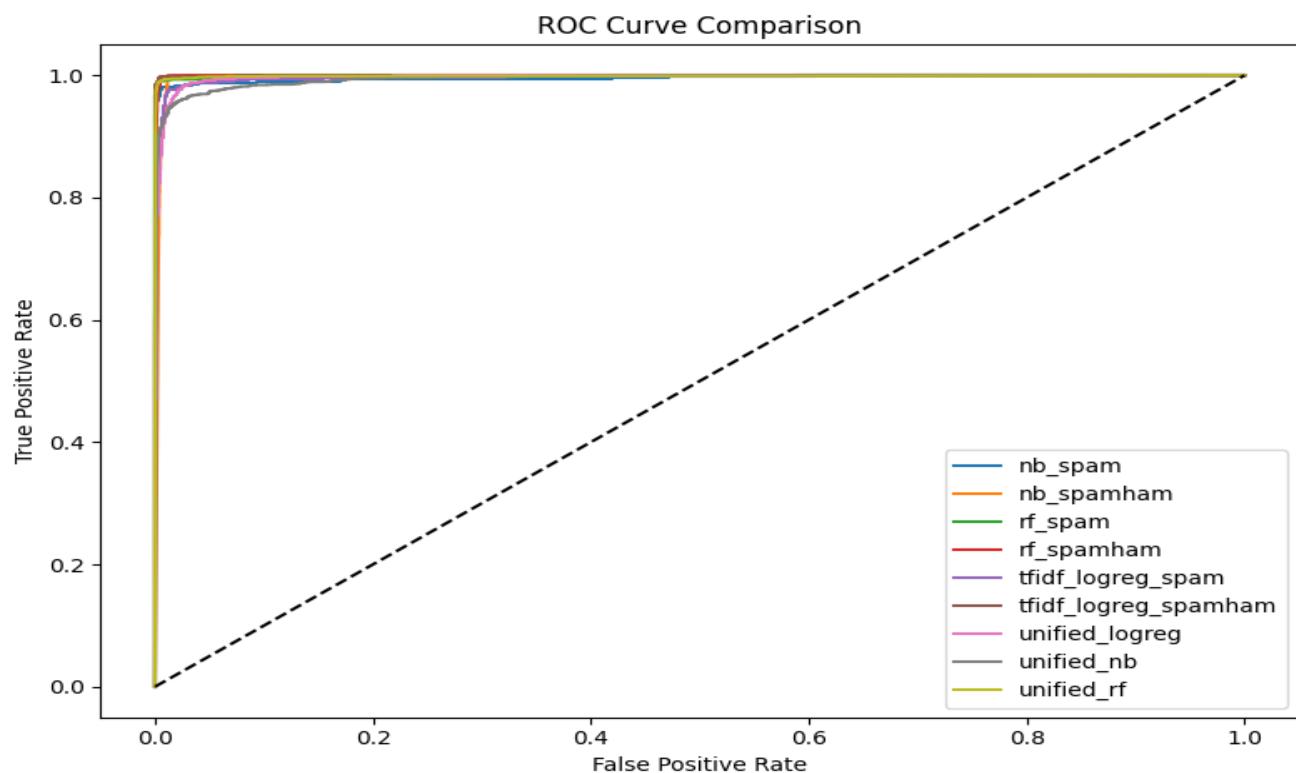
Model Performance Summary

```
model dataset accuracy auc
tfidf_logreg_spam data/spam.csv 0.961953 0.996062
tfidf_logreg_spamham data/spam_ham_dataset.csv 0.995359 0.999683
rf_spam data/spam.csv 0.993719 0.998332
rf_spamham data/spam_ham_dataset.csv 0.994005 0.999554
nb_spam data/spam.csv 0.990309 0.995761
nb_spamham data/spam_ham_dataset.csv 0.983369 0.997154
unified_logreg data/unified_spam_dataset.csv 0.951398 0.996604
unified_rf data/unified_spam_dataset.csv 0.994863 0.998784
unified_nb data/unified_spam_dataset.csv 0.965820 0.994207
```

# Accuracy Comparison Chart

ROC Curve Comparison
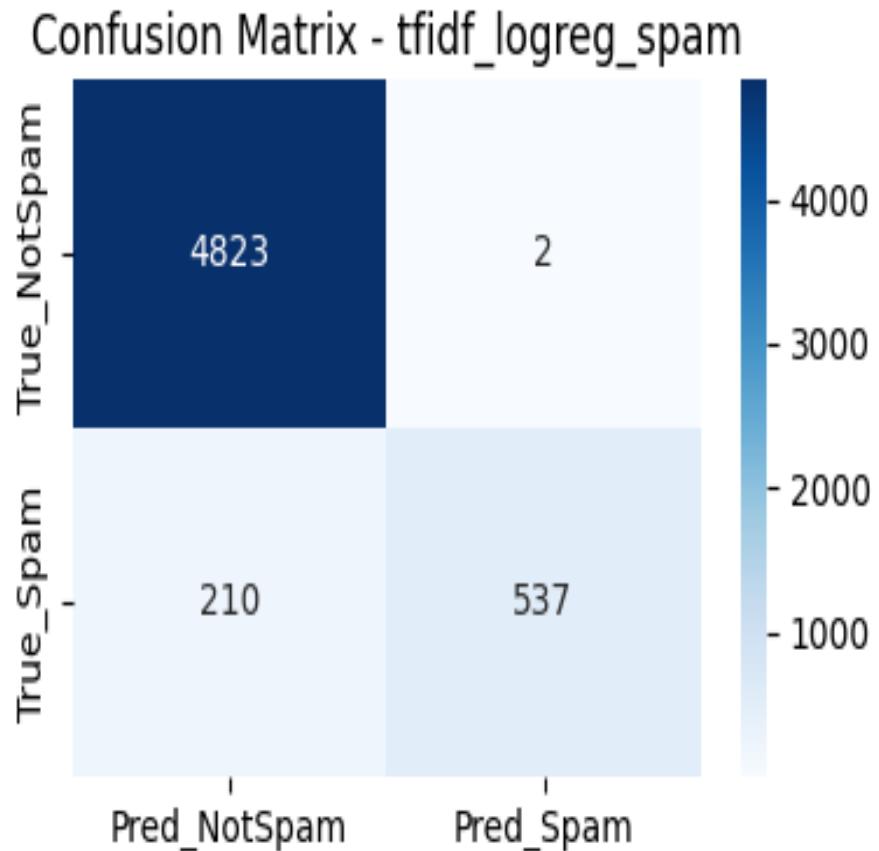
ROC Curve Comparison
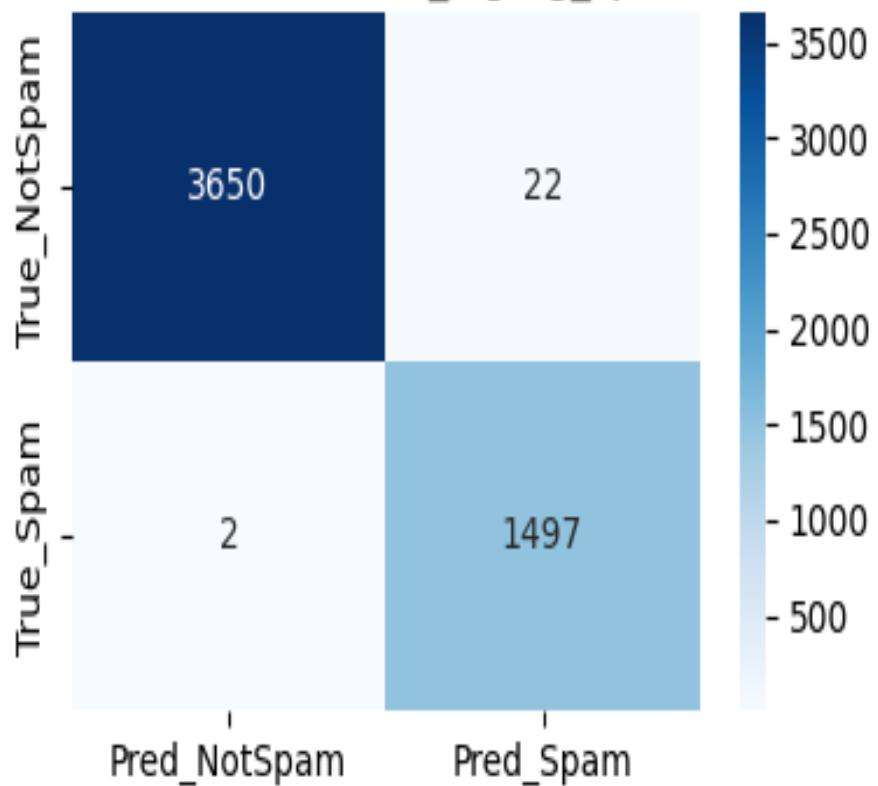
Confusion Matrices (Per Model)

tfidf_logreg_spam

Confusion Matrix - tfidf_logreg_spam

|	Pred_NotSpam	Pred_Spam
True_NotSpam	4823	2
True_Spam	210	537
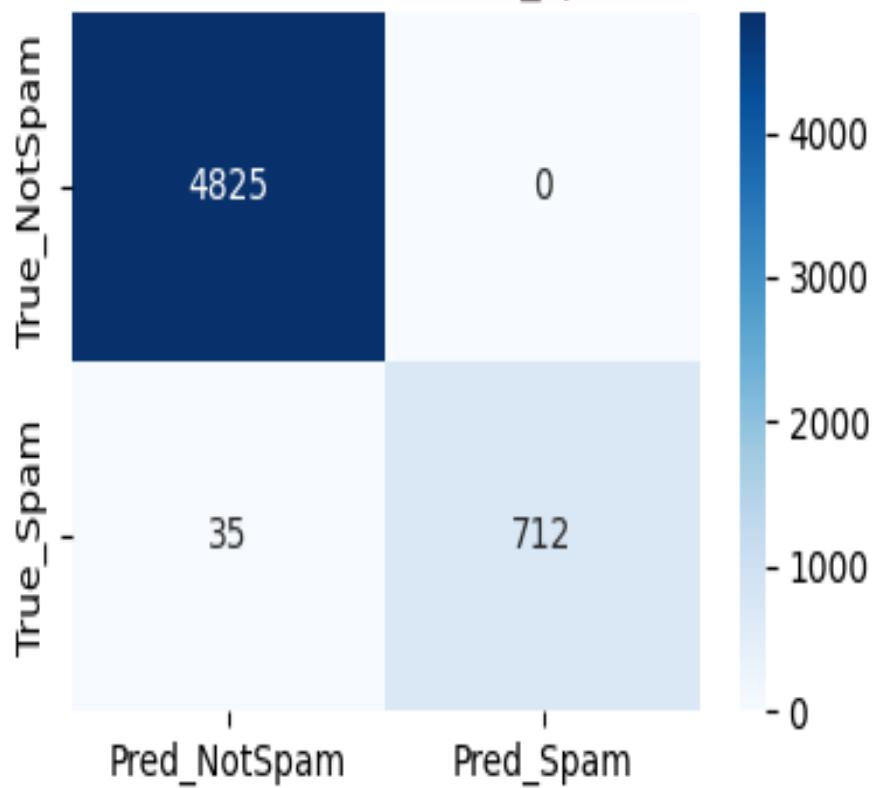
tfidf_logreg_spamham

Confusion Matrix - tfidf_logreg_spamham
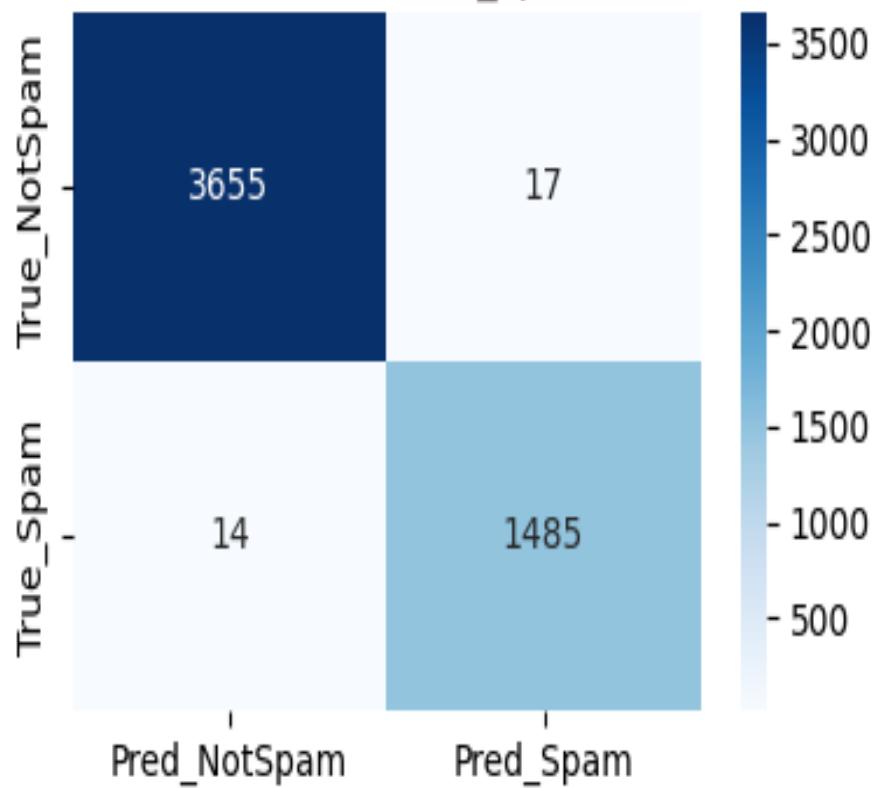
rf_spam

Confusion Matrix - rf_spam

rf_spamham

Confusion Matrix - rf_spamham

	Pred_NotSpam	Pred_Spam
True_NotSpam	3655	17
True_Spam	14	1485
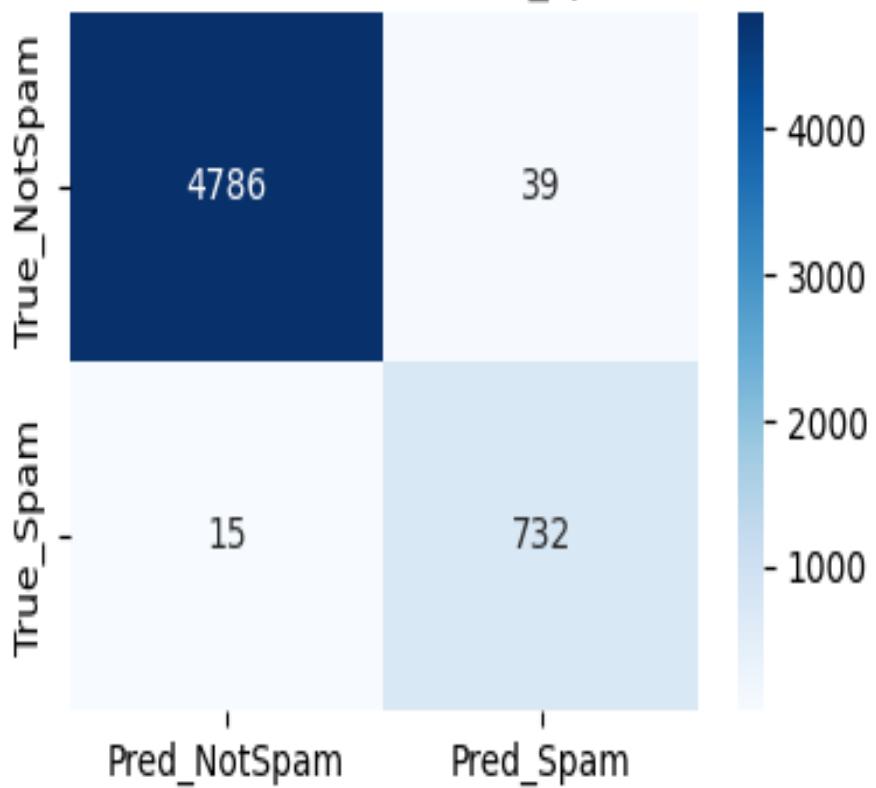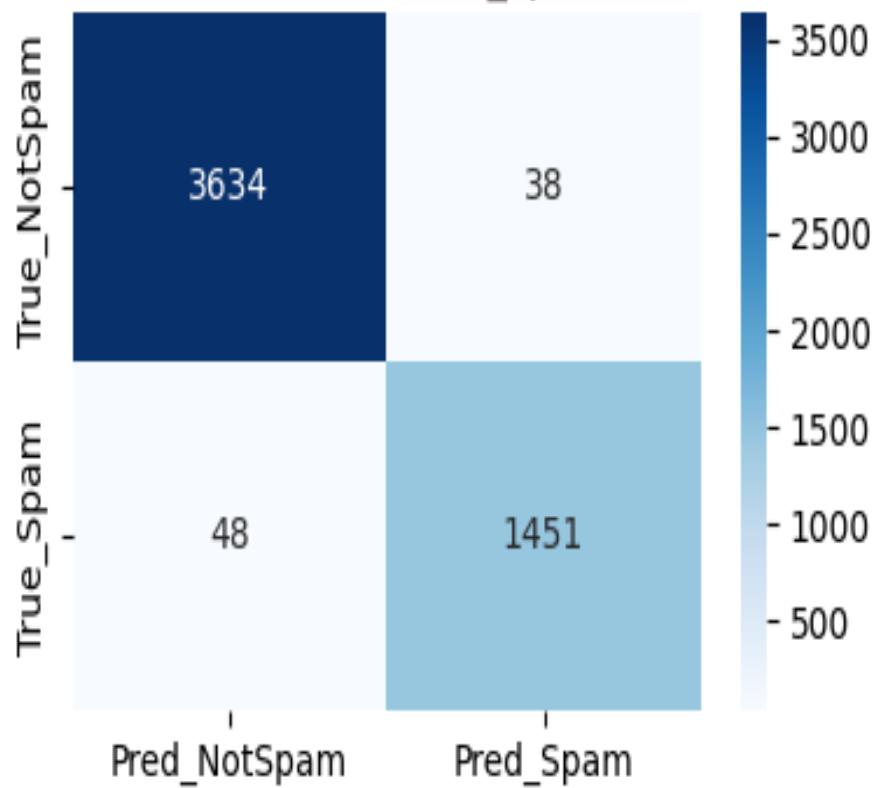
nb_spam

Confusion Matrix - nb_spam

nb_spamham

Confusion Matrix - nb_spamham

	Pred_NotSpam	Pred_Spam
True_NotSpam	3634	38
True_Spam	48	1451
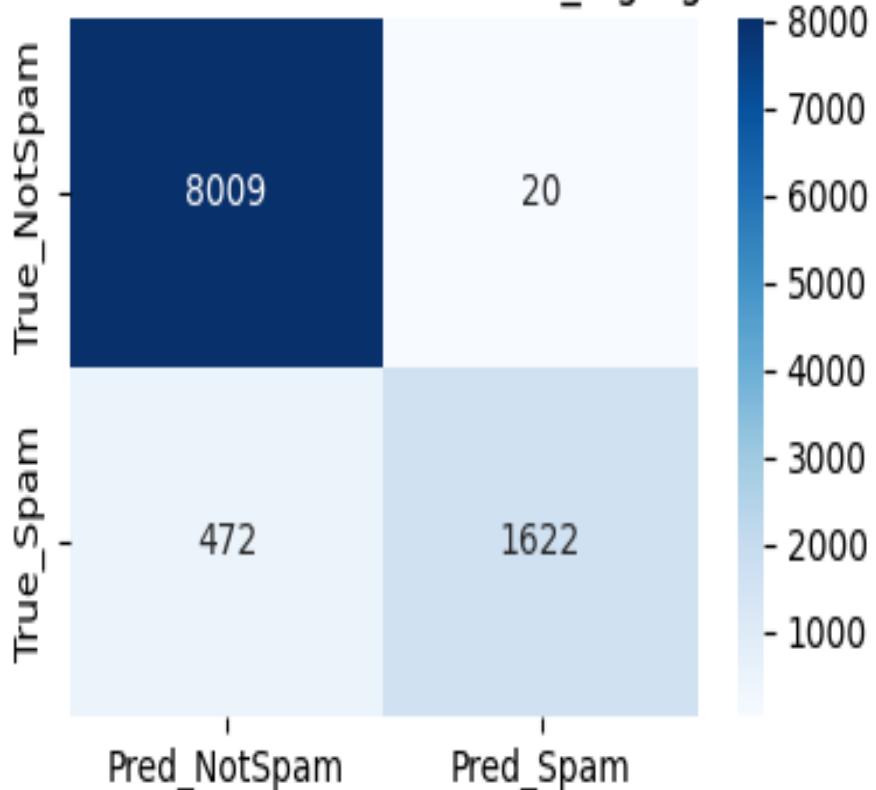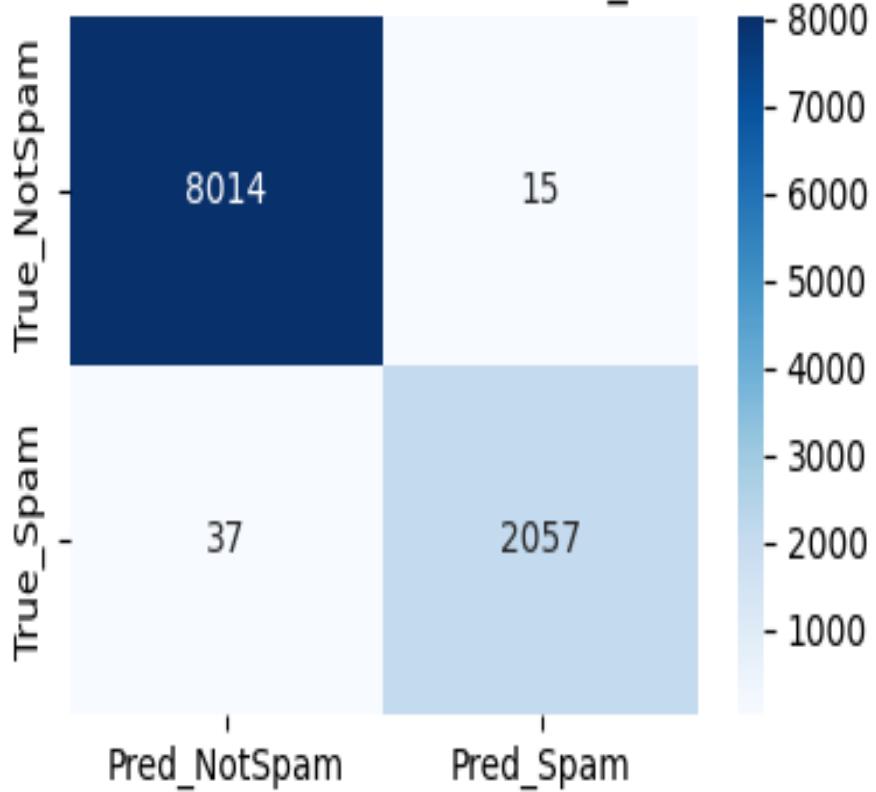
unified_logreg

Confusion Matrix - unified_logreg

	Pred_NotSpam	Pred_Spam
True_NotSpam	8009	20
True_Spam	472	1622

unified_rf

Confusion Matrix - unified_rf

unified_nb

Confusion Matrix - unified_nb