

Deepfake Video Classifiers – The Comparison of Image and Text Features in Random Forest Model

Yunting Chiu

M.S. in Data Science, Department of Mathematics & Statistics at American University

IMAGE + TEXT FEATURES

Hyperparameters Tuning the Random Forest Model

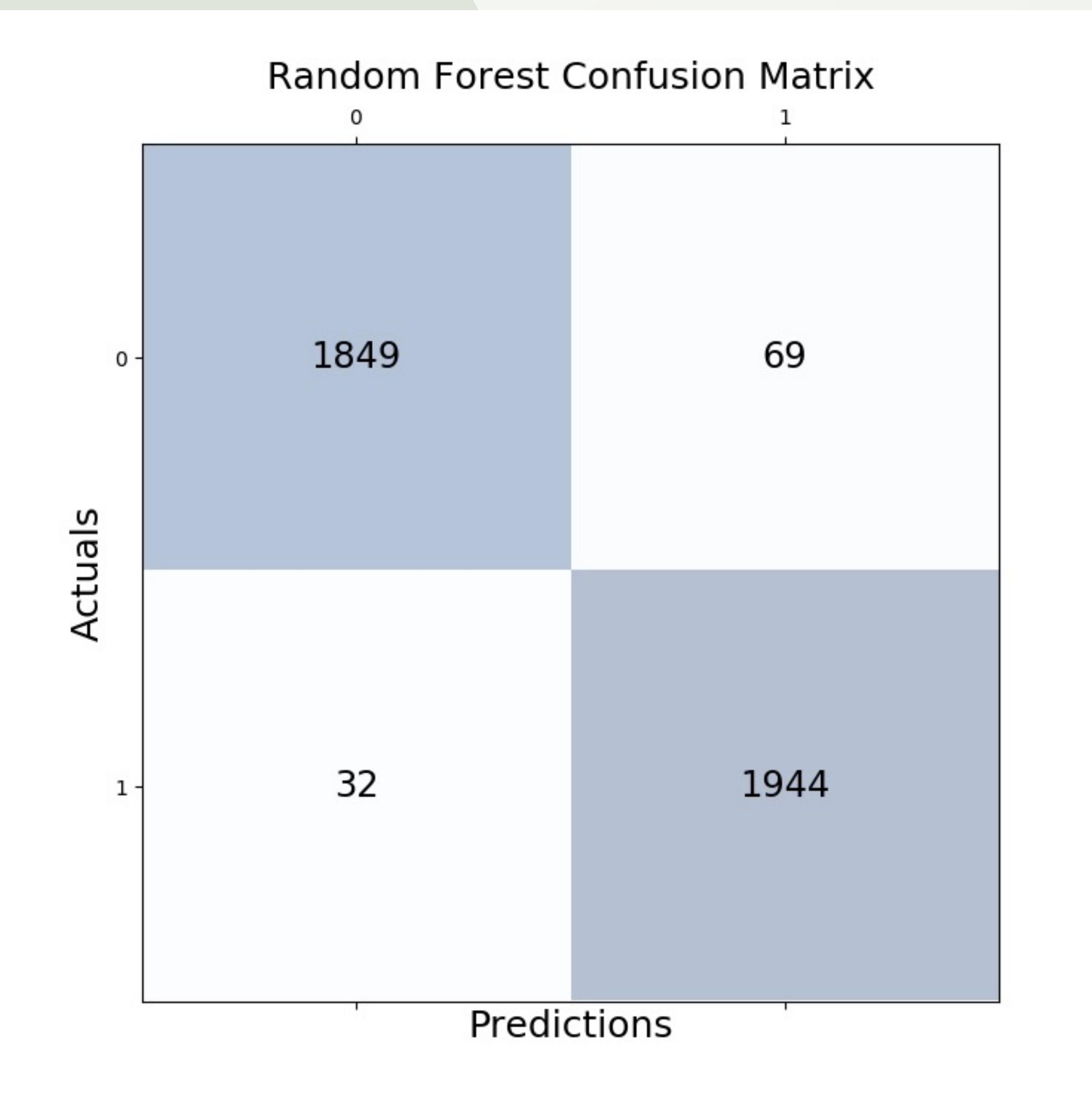
I leverage grid search in training set, the selected parameters are:

- grid['n_estimators'] = [100, **500**]
- grid['max_depth'] = [5, 10, 50, **100**]

Model Results

Testing Accuracy: 97.41%

Confusion Matrix



Classification Report

```
-----  
Confusion Matrix-----  
[[1849  69]  
 [ 32 1944]]  
-----  
Accuracy Score-----  
0.9740626605033385  
-----  
Classification Report-----  
 precision    recall   f1-score   support  
  
 fake          0.98      0.96      0.97     1918  
 real          0.97      0.98      0.97     1976  
  
 accuracy       0.97      0.97      0.97     3894  
 macro avg     0.97      0.97      0.97     3894  
 weighted avg  0.97      0.97      0.97     3894
```

IMAGE FEATURES ONLY

Hyperparameters Tuning the Random Forest Model

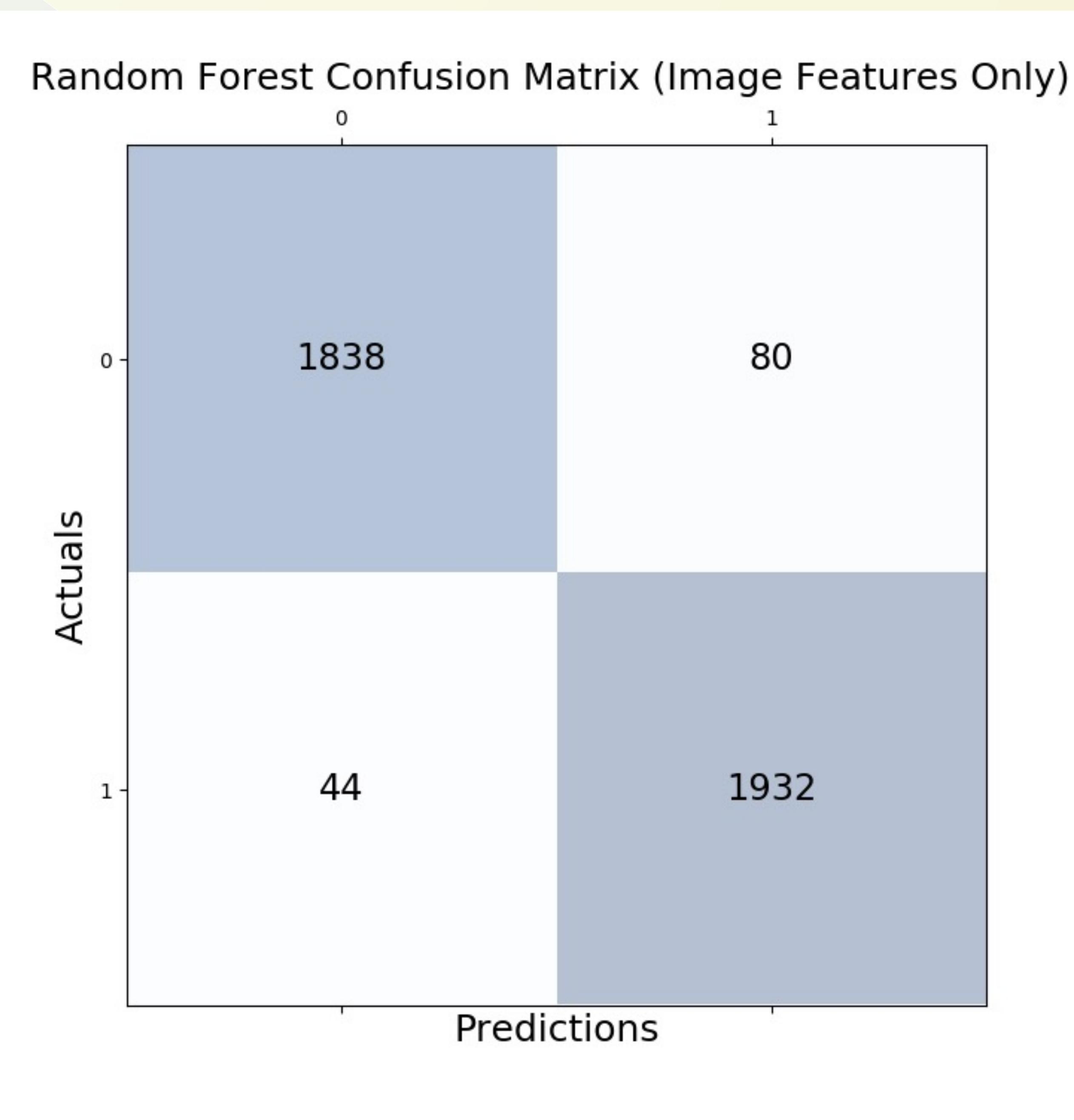
I leverage grid search in training set, the selected parameters are:

- grid['n_estimators'] = [100, **500**]
- grid['max_depth'] = [5, 10, 50, **100**]

Model Results

Testing Accuracy: 96.82%

Confusion Matrix



Classification Report

```
-----  
Confusion Matrix-----  
[[1838  80]  
 [ 44 1932]]  
-----  
Accuracy Score-----  
0.9681561376476631  
-----  
Classification Report-----  
 precision    recall   f1-score   support  
  
 fake          0.98      0.96      0.97     1918  
 real          0.96      0.98      0.97     1976  
  
 accuracy       0.97      0.97      0.97     3894  
 macro avg     0.97      0.97      0.97     3894  
 weighted avg  0.97      0.97      0.97     3894
```

TEXT FEATURES ONLY

Hyperparameters Tuning the Random Forest Model

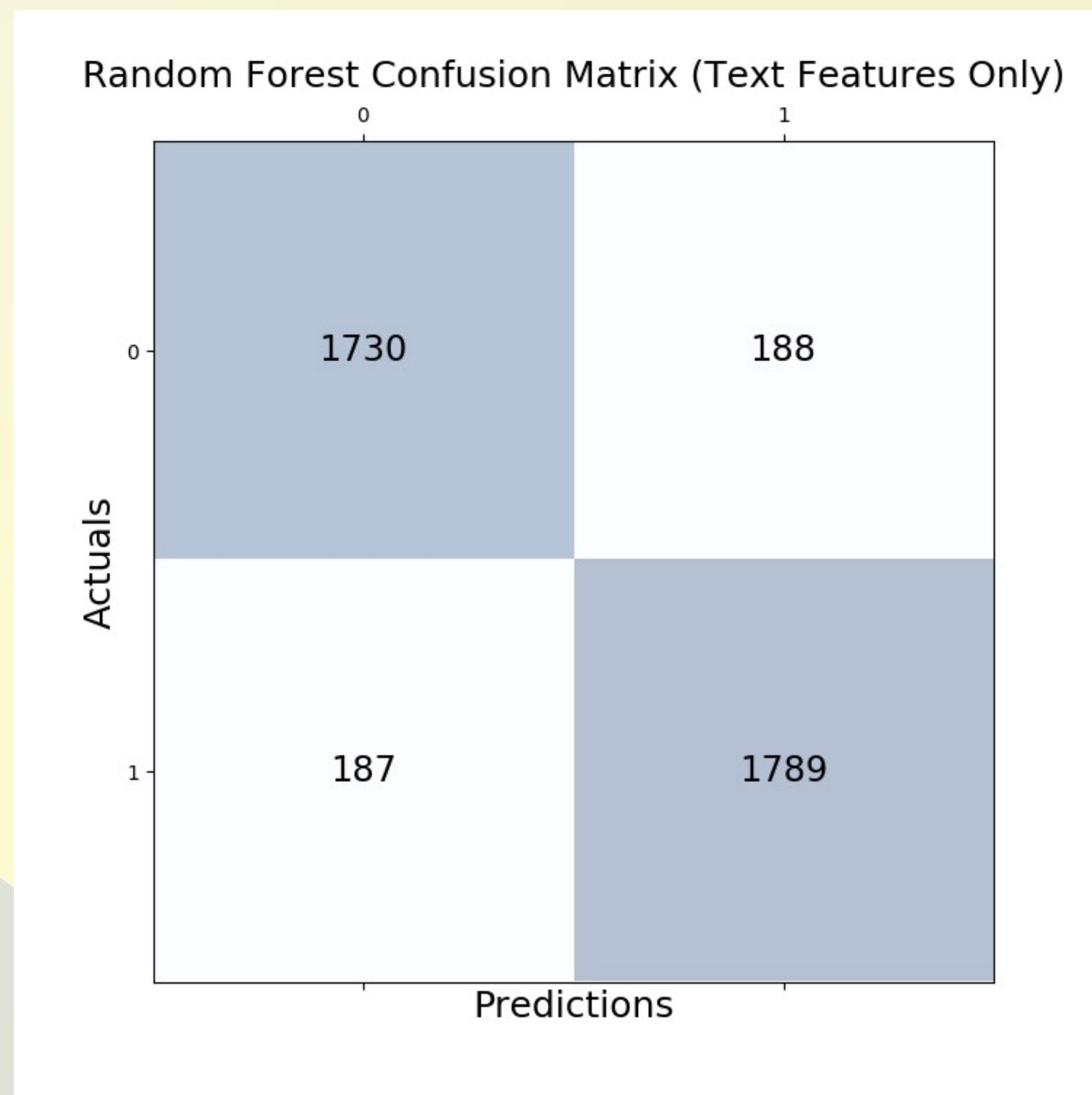
I leverage grid search in training set, the selected parameters are:

- grid['n_estimators'] = [100, **500**]
- grid['max_depth'] = [5, 10, 50, **100**]

Model Results

Testing Accuracy: 90.37%

Confusion Matrix



Classification Report

```
-----  
Confusion Matrix-----  
[[1730 188]  
 [ 187 1789]]  
-----  
Accuracy Score-----  
0.9036979969183359  
-----  
Classification Report-----  
 precision    recall   f1-score   support  
  
 fake          0.90      0.90      0.90     1918  
 real          0.90      0.91      0.91     1976  
  
 accuracy       0.90      0.90      0.90     3894  
 macro avg     0.90      0.90      0.90     3894  
 weighted avg  0.90      0.90      0.90     3894
```