FakeNewsNetClassifier

July 14, 2024

1 FakeNewsNetClassifier

This notebook builds and evaluates machine learning models for detecting fake news using the FakeNewsNet dataset, that contians two subsets, GossipCop and PolitiFact. I will be using three different classifiers: Logistic Regression, Support Vector Machine (SVM), and Naive Bayes. I will also balance the dataset using SMOTE and perform hyperparameter tuning to improve models' performance

1.1 Import necessary libraries

[1]:
```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score,␣
 ↪GridSearchCV
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, precision_score, recall_score,␣
 ↪f1_score, confusion_matrix, roc_auc_score, roc_curve
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.feature_extraction.text import TfidfVectorizer
from imblearn.over_sampling import SMOTE
import joblib
```

1.2 Load Preprocessed Data from Google Drive

[2]:
```
# Mount Google Drive
from google.colab import drive
drive.mount('/content/drive')

# Load the preprocessed datasets from Google Drive
gossipcop_combined = pd.read_csv('/content/drive/My Drive/
 ↪gossipcop_preprocessed.csv')
politifact_combined = pd.read_csv('/content/drive/My Drive/
 ↪politifact_preprocessed.csv')
```

```
# Quick check of the data loaded
print(gossipcop_combined.head())
print(politifact_combined.head())
```

```
Mounted at /content/drive
                                               title  label
0  lea michel hairstylist mix textur spray coconu…      1
1  thoma markl princ harri polit miss daughter me…      0
2  2019 sag award nomin see full list nomine varieti      1
3  see meghan markl royal coat arm symbol hide wi…      1
4    kyli jenner visit shaman life kyli season final      1
                                               title  label
0         world popular candi remov shelv octob 2017      0
1                             brows congression bill      1
2                          suprem court vacanc video      1
3                                    u import export      1
4  die 78 year old cia agent admit kill marilyn m…      0
```

1.3 Check loaded data

[3]:
```
# Check for NaN values in datasets
print("NaN values in GossipCop dataset:\n", gossipcop_combined.isnull().sum())
print("NaN values in PolitiFact dataset:\n", politifact_combined.isnull().sum())

# Ensure there are no NaN values
gossipcop_combined.dropna(inplace=True)
politifact_combined.dropna(inplace=True)

# Re-check for NaN values in datasets
print("NaN values in GossipCop dataset after dropna:\n", gossipcop_combined.
 ↪isnull().sum())
print("NaN values in PolitiFact dataset after dropna:\n", politifact_combined.
 ↪isnull().sum())
```

```
NaN values in GossipCop dataset:
 title    1
label    0
dtype: int64
NaN values in PolitiFact dataset:
 title    2
label    0
dtype: int64
NaN values in GossipCop dataset after dropna:
 title    0
label    0
dtype: int64
NaN values in PolitiFact dataset after dropna:
 title    0
```

```
label    0
dtype: int64
```

1.4 Split Data into Training and Testing Sets

I am splitting the dataset the same way as Shu et al. [7]: "We use 80% of data for training and 20% for testing."

[4]:
```
# Define features (X) and labels (y)
X_gossipcop = gossipcop_combined['title']
y_gossipcop = gossipcop_combined['label']

X_politifact = politifact_combined['title']
y_politifact = politifact_combined['label']

# Split the data into training and testing sets (80% train, 20% test)
X_train_gossipcop, X_test_gossipcop, y_train_gossipcop, y_test_gossipcop =␣
↪train_test_split(X_gossipcop, y_gossipcop, test_size=0.2, random_state=42)
X_train_politifact, X_test_politifact, y_train_politifact, y_test_politifact =␣
↪train_test_split(X_politifact, y_politifact, test_size=0.2, random_state=42)
```

1.5 Vectorize the Text Data Using TF-IDF

Shen et al [6] applied TF-IDF technique for feature extraction, so I will use this too. TF-IDF should be applied after splitting the data to avoid data leakage. The vectorizer is only fitted on the training data and then applied to the test data.

[5]:
```
# Initialize the TF-IDF Vectorizer
vectorizer = TfidfVectorizer(max_features=1000)

# Fit and transform the training data, transform the test data
X_train_gossipcop_tfidf = vectorizer.fit_transform(X_train_gossipcop)
X_test_gossipcop_tfidf = vectorizer.transform(X_test_gossipcop)

X_train_politifact_tfidf = vectorizer.fit_transform(X_train_politifact)
X_test_politifact_tfidf = vectorizer.transform(X_test_politifact)
```

1.6 Balance classes using SMOTE

SMOTE is a popular technique, that can be applied to balance an imbalanced dataset as descibed by Chawla et al.[14]. It generates synthetic examples for the minority class. The method combines minority class examples to create synthetic examples and balances the class distribution. It improves classifier performance on imbalanced datasets. I apply this technique to balance the class imbalance present in the original datasets.

[6]:
```
# Apply SMOTE to balance the classes in the training data
smote = SMOTE(random_state=42)
```

```
X_train_gossipcop_resampled, y_train_gossipcop_resampled = smote.
 ↪fit_resample(X_train_gossipcop_tfidf, y_train_gossipcop)
X_train_politifact_resampled, y_train_politifact_resampled = smote.
 ↪fit_resample(X_train_politifact_tfidf, y_train_politifact)

# Check class distribution after SMOTE
print("Class distribution in resampled GossipCop training set:")
print(pd.Series(y_train_gossipcop_resampled).value_counts())

print("Class distribution in resampled PolitiFact training set:")
print(pd.Series(y_train_politifact_resampled).value_counts())
```

```
Class distribution in resampled GossipCop training set:
label
0    13428
1    13428
Name: count, dtype: int64
Class distribution in resampled PolitiFact training set:
label
0    492
1    492
Name: count, dtype: int64
```

1.7 Train and Evaluate Classifiers

[7]:
```
# Train and evaluate the model, returning the model and predictions
def train_and_evaluate(model, X_train, X_test, y_train, y_test):
    # Train the model
    model.fit(X_train, y_train)
    # Predict on the test data
    y_pred = model.predict(X_test)
    # Calculate evaluation metrics
    accuracy = accuracy_score(y_test, y_pred)
    precision = precision_score(y_test, y_pred)
    recall = recall_score(y_test, y_pred)
    f1 = f1_score(y_test, y_pred)
    roc_auc = roc_auc_score(y_test, y_pred)
    # Print evaluation metrics
    print(f'Accuracy: {accuracy:.4f}')
    print(f'Precision: {precision:.4f}')
    print(f'Recall: {recall:.4f}')
    print(f'F1-score: {f1:.4f}')
    print(f'ROC AUC: {roc_auc:.4f}')
    # Confusion Matrix
    cm = confusion_matrix(y_test, y_pred)
    print('Confusion Matrix:')
    print(cm)
```

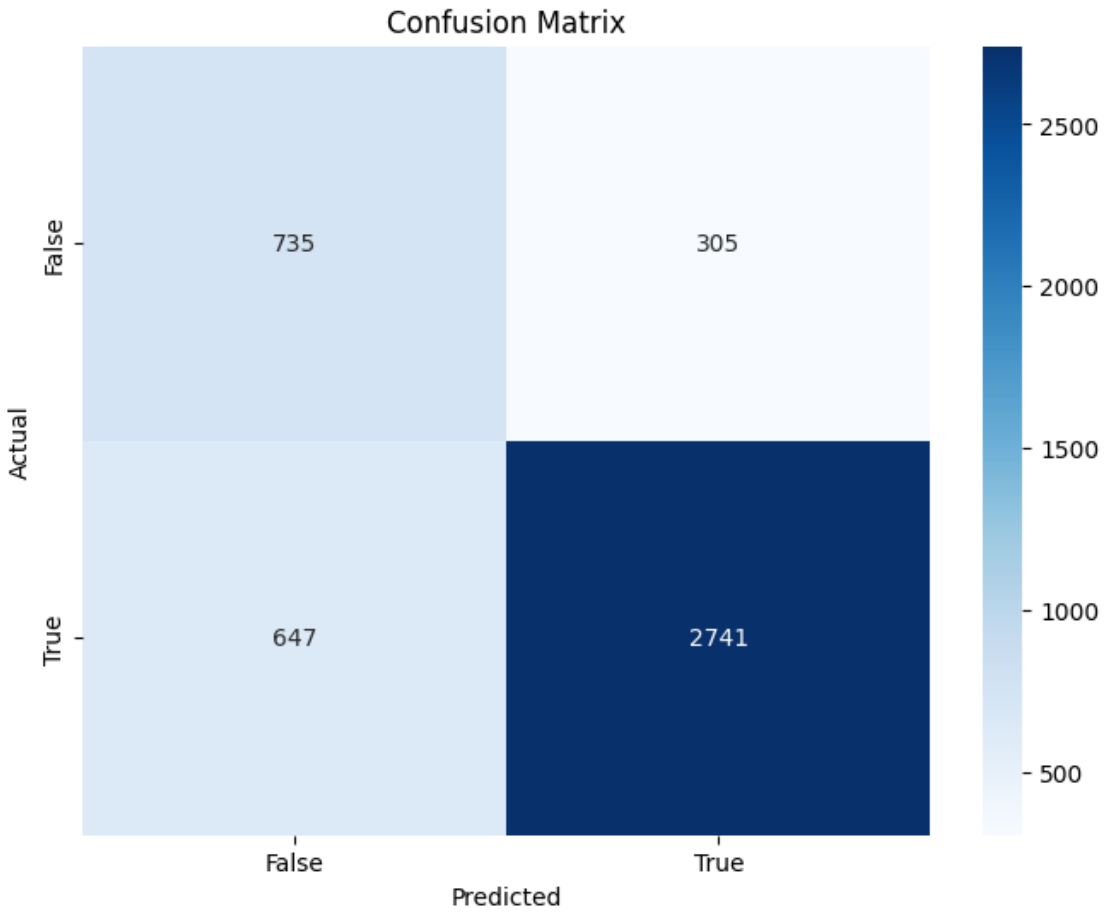
```
    # Plot Confusion Matrix
    cm = confusion_matrix(y_test, y_pred)
    plt.figure(figsize=(8, 6))
    sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=['False',␣
 ↪'True'], yticklabels=['False', 'True'])
    plt.xlabel('Predicted')
    plt.ylabel('Actual')
    plt.title('Confusion Matrix')
    plt.show()
    # Plot ROC Curve
    if hasattr(model, "predict_proba"):
        y_proba = model.predict_proba(X_test)[:, 1]
    else:  # Use decision_function if predict_proba is not available
        y_proba = model.decision_function(X_test)
        y_proba = (y_proba - y_proba.min()) / (y_proba.max() - y_proba.min())
    fpr, tpr, thresholds = roc_curve(y_test, y_proba)
    plt.figure(figsize=(8, 6))
    plt.plot(fpr, tpr, label='ROC Curve (area = %0.2f)' % roc_auc)
    plt.plot([0, 1], [0, 1], 'k--')
    plt.xlim([0.0, 1.0])
    plt.ylim([0.0, 1.05])
    plt.xlabel('False Positive Rate')
    plt.ylabel('True Positive Rate')
    plt.title('Receiver Operating Characteristic')
    plt.legend(loc="lower right")
    plt.show()
    # Return the trained model and predictions
    return model, y_pred
```

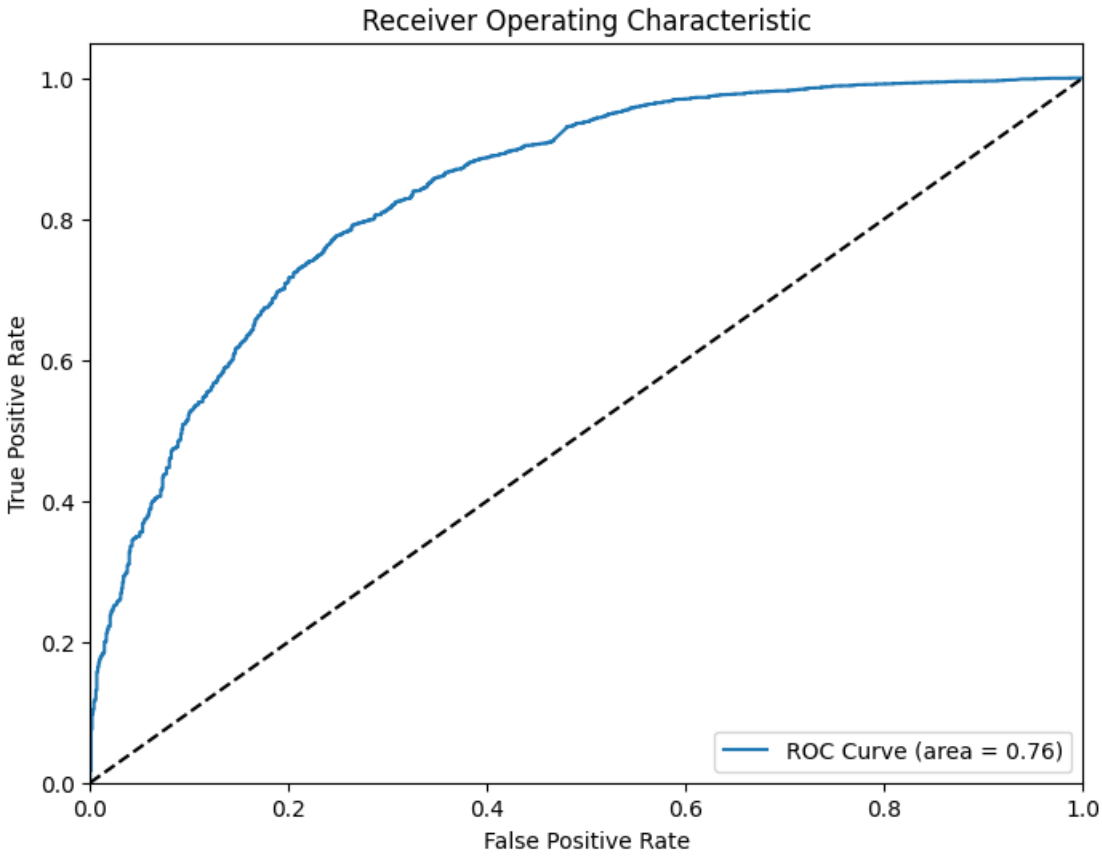
1.7.1 Model Training and Evaluation

[8]:
```
# Logistic Regression
print("Logistic Regression (GossipCop):")
logistic_model_gossipcop = LogisticRegression(max_iter=1000)
train_and_evaluate(logistic_model_gossipcop, X_train_gossipcop_resampled,␣
 ↪X_test_gossipcop_tfidf, y_train_gossipcop_resampled, y_test_gossipcop)

print("\nLogistic Regression (PolitiFact):")
logistic_model_politifact = LogisticRegression(max_iter=1000)
train_and_evaluate(logistic_model_politifact, X_train_politifact_resampled,␣
 ↪X_test_politifact_tfidf, y_train_politifact_resampled, y_test_politifact)

# Support Vector Machine (SVM)
print("\nSupport Vector Machine (GossipCop):")
svm_model_gossipcop = SVC(kernel='linear')
```

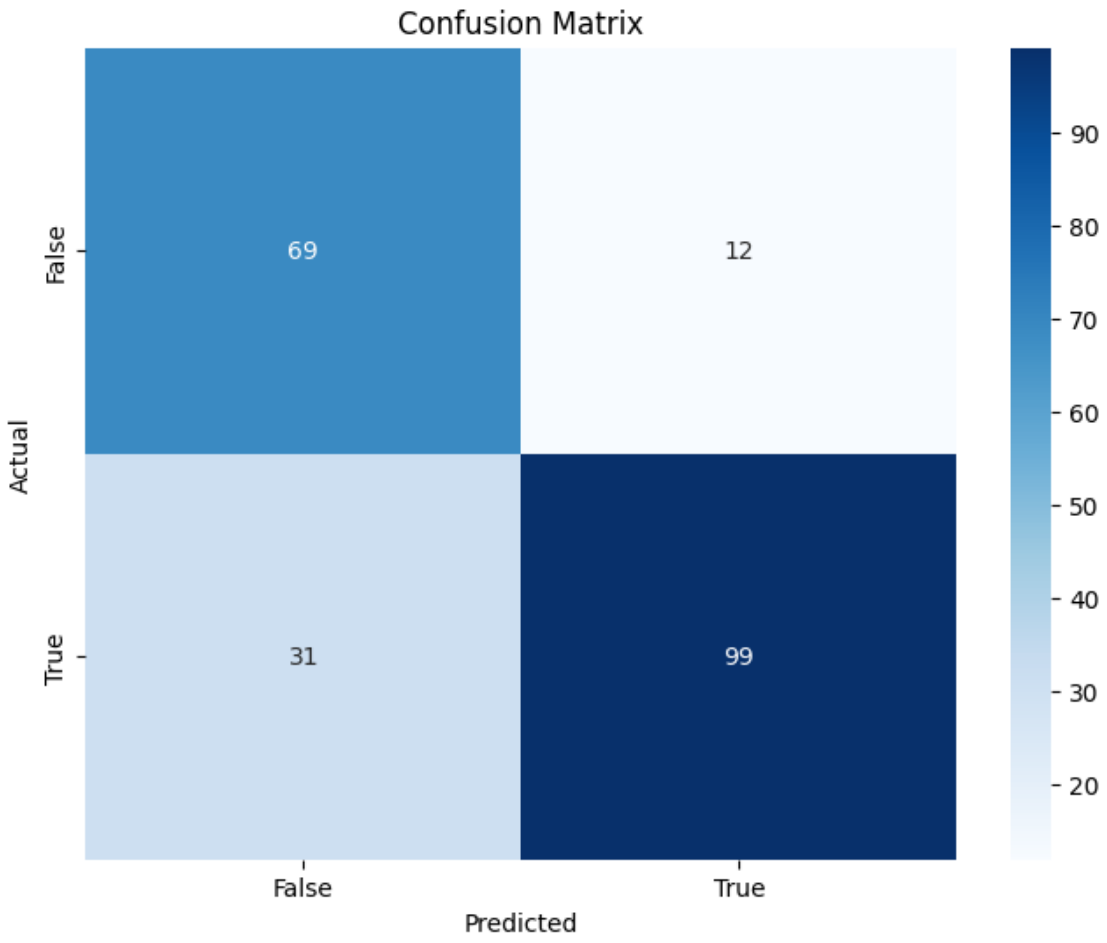
```
train_and_evaluate(svm_model_gossipcop, X_train_gossipcop_resampled,␣
 ↪X_test_gossipcop_tfidf, y_train_gossipcop_resampled, y_test_gossipcop)

print("\nSupport Vector Machine (PolitiFact):")
svm_model_politifact = SVC(kernel='linear')
train_and_evaluate(svm_model_politifact, X_train_politifact_resampled,␣
 ↪X_test_politifact_tfidf, y_train_politifact_resampled, y_test_politifact)

# Naive Bayes
print("\nNaive Bayes (GossipCop):")
nb_model_gossipcop = MultinomialNB()
train_and_evaluate(nb_model_gossipcop, X_train_gossipcop_resampled,␣
 ↪X_test_gossipcop_tfidf, y_train_gossipcop_resampled, y_test_gossipcop)

print("\nNaive Bayes (PolitiFact):")
nb_model_politifact = MultinomialNB()
train_and_evaluate(nb_model_politifact, X_train_politifact_resampled,␣
 ↪X_test_politifact_tfidf, y_train_politifact_resampled, y_test_politifact)
```

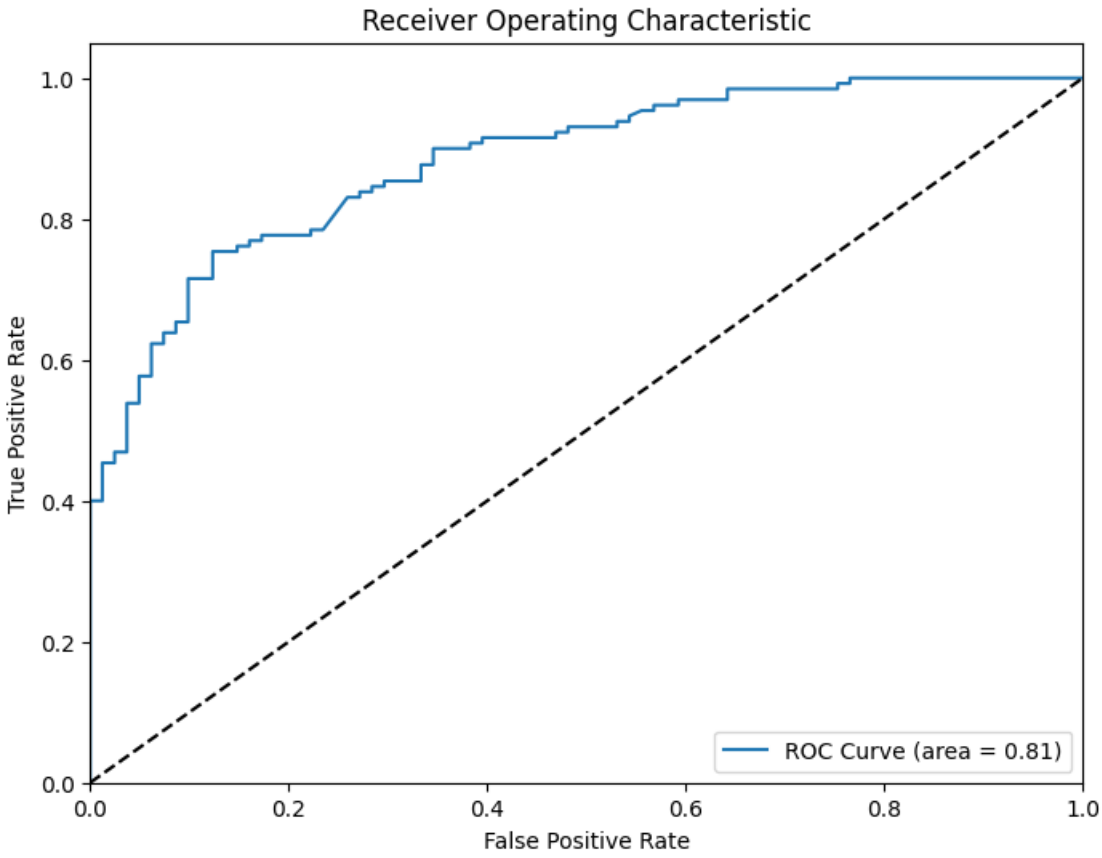
```
Logistic Regression (GossipCop):
Accuracy: 0.7850
Precision: 0.8999
Recall: 0.8090
F1-score: 0.8520
ROC AUC: 0.7579
Confusion Matrix:
[[ 735  305]
 [ 647 2741]]
```


Confusion Matrix
735
305
647
2741
False
True
Actual
False
True
Predicted
2500
2000
1500
1000
500

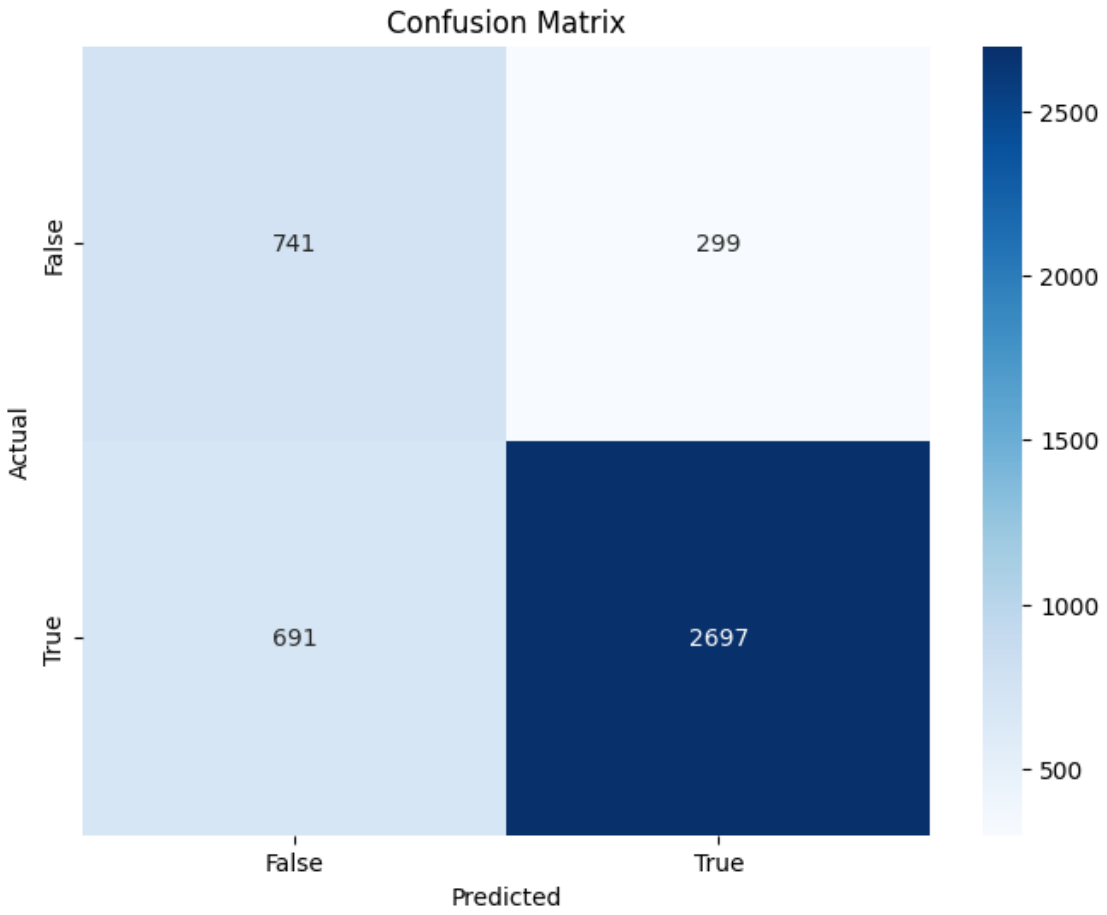


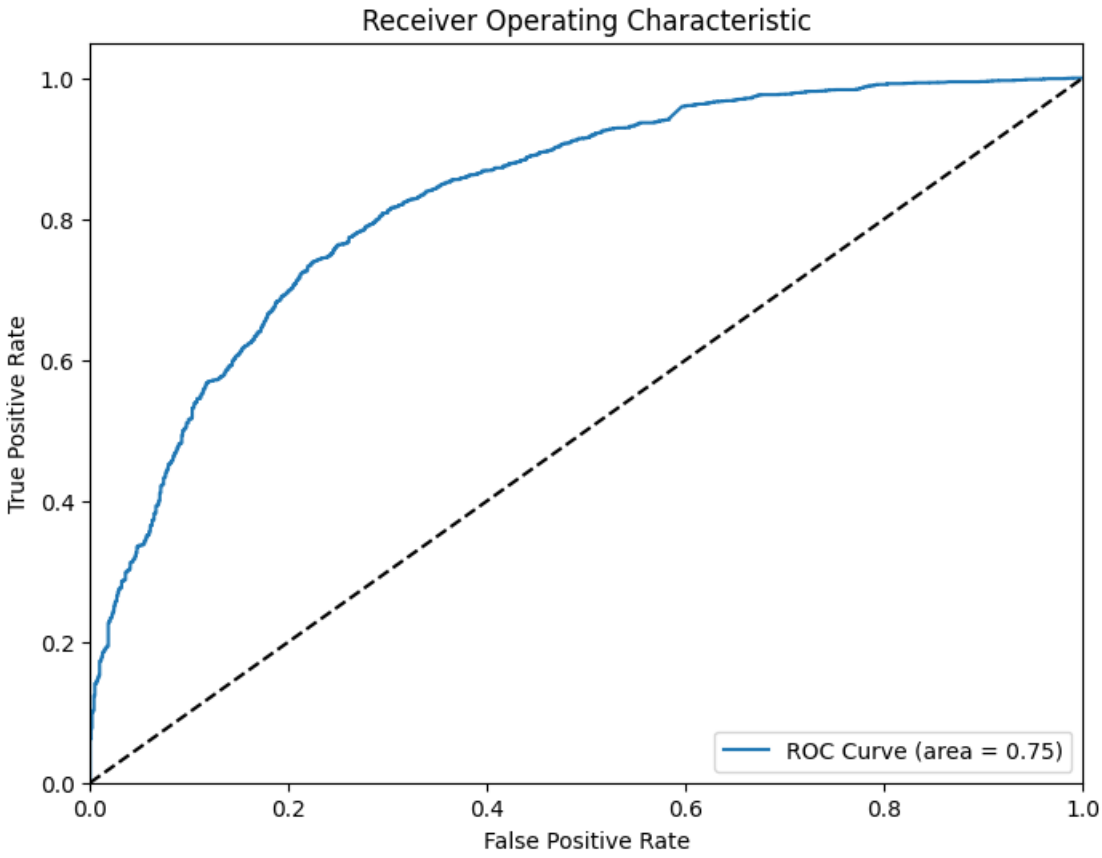

```
Logistic Regression (PolitiFact):
Accuracy: 0.7962
Precision: 0.8919
Recall: 0.7615
F1-score: 0.8216
ROC AUC: 0.8067
Confusion Matrix:
[[69 12]
 [31 99]]
```


Confusion Matrix
69
12
31
99
False
True
Actual
False
True
Predicted
90
80
70
60
50
40
30
20

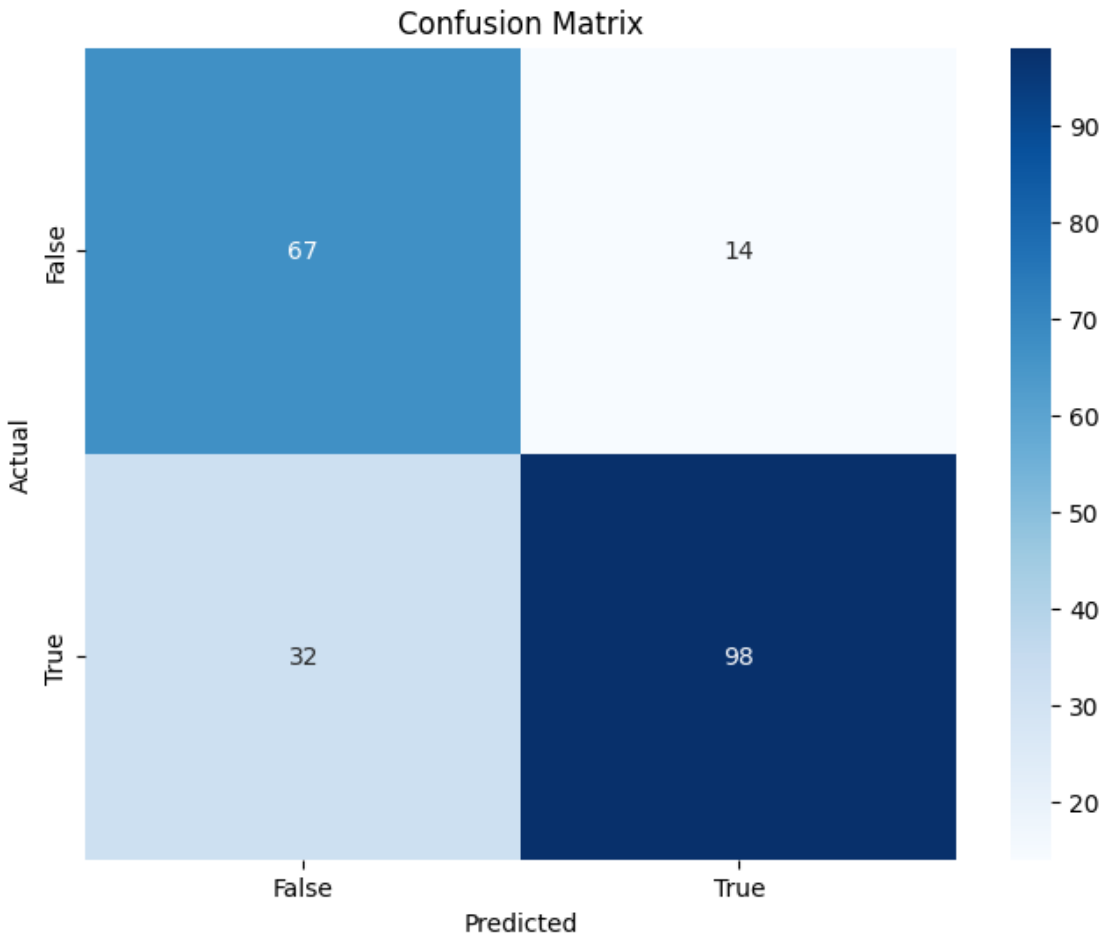


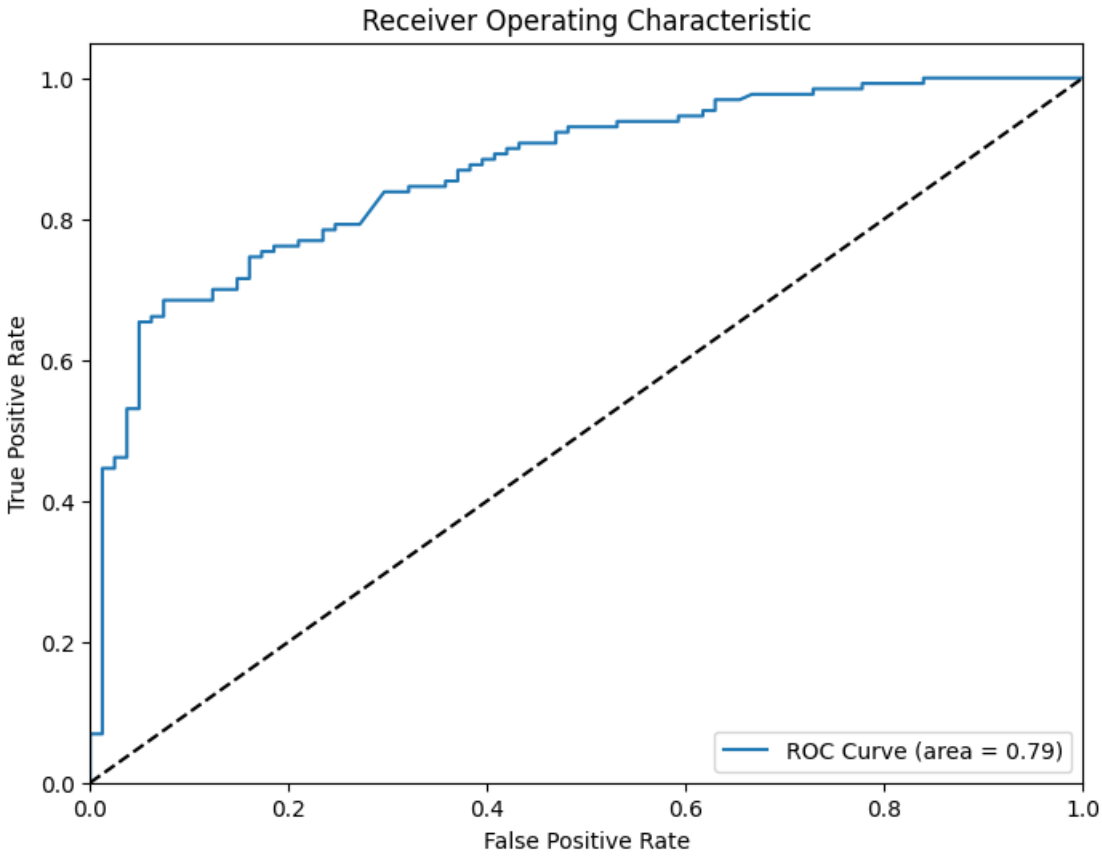

```
Support Vector Machine (GossipCop):
Accuracy: 0.7764
Precision: 0.9002
Recall: 0.7960
F1-score: 0.8449
ROC AUC: 0.7543
Confusion Matrix:
[[ 741  299]
 [ 691 2697]]
```


Confusion Matrix
741
299
691
2697
False
True
Actual
False
True
Predicted
2500
2000
1500
1000
500

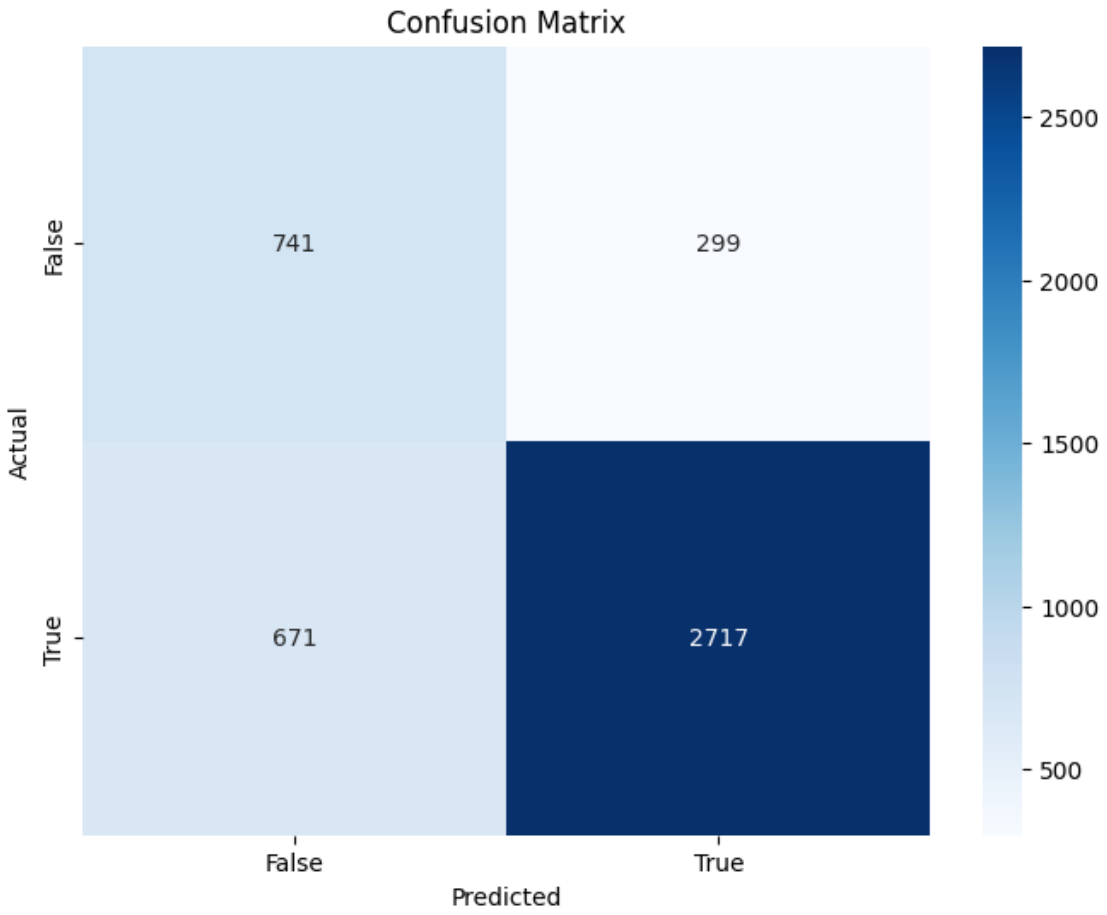


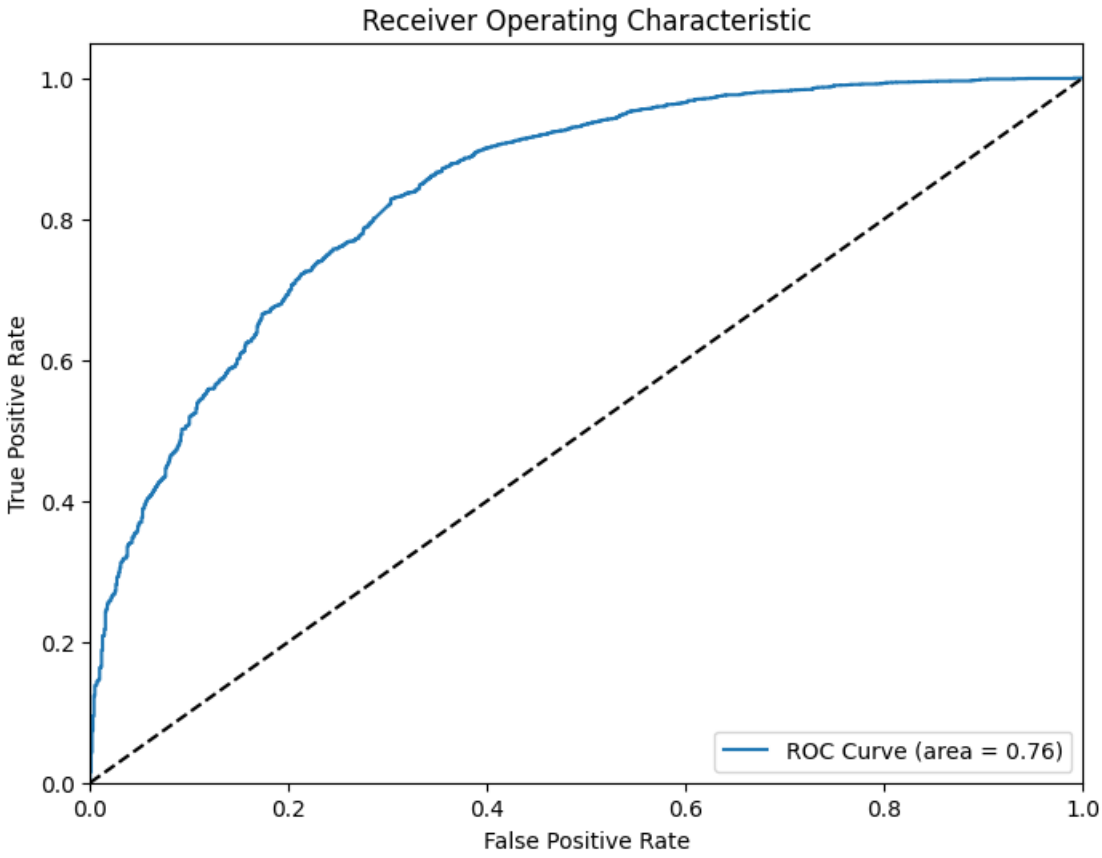

```
Support Vector Machine (PolitiFact):
Accuracy: 0.7820
Precision: 0.8750
Recall: 0.7538
F1-score: 0.8099
ROC AUC: 0.7905
Confusion Matrix:
[[67 14]
 [32 98]]
```


Confusion Matrix
67
14
32
98
False
True
Actual
False
True
Predicted
90
80
70
60
50
40
30
20

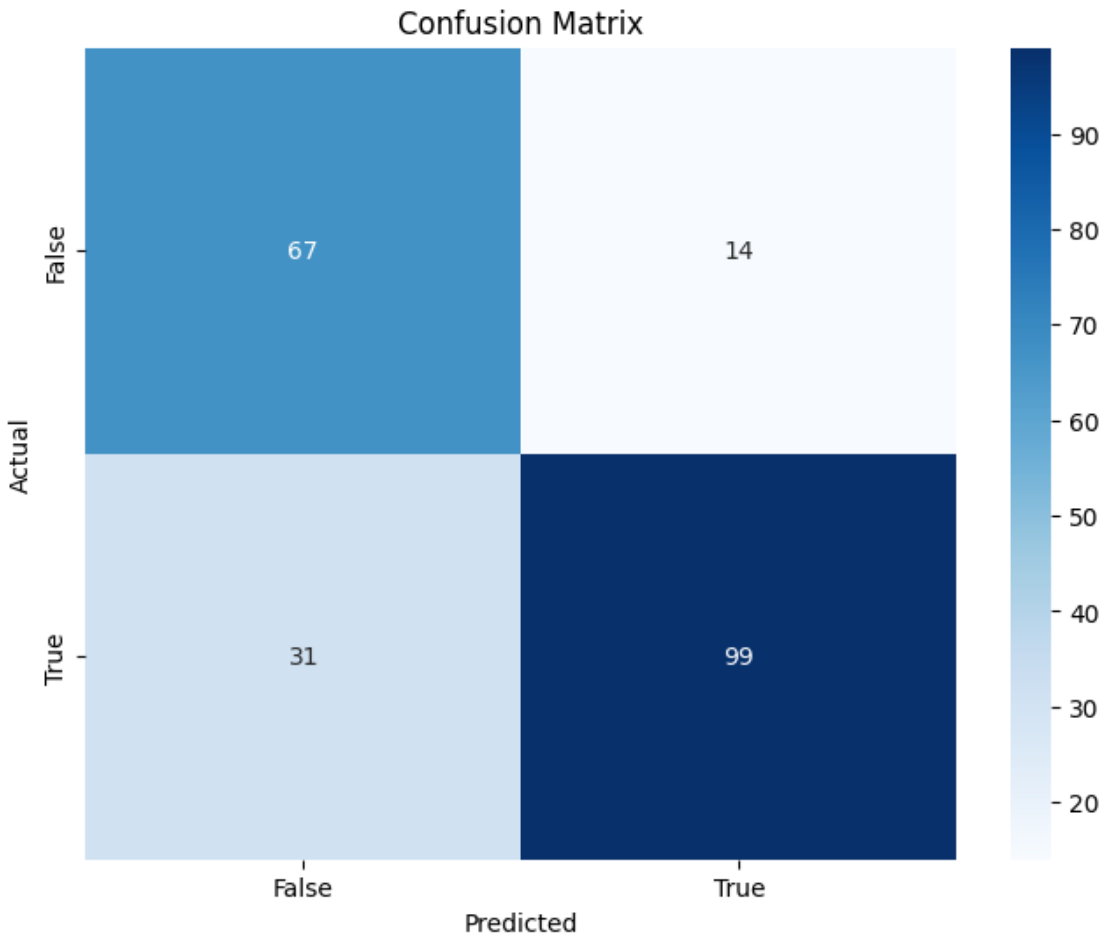


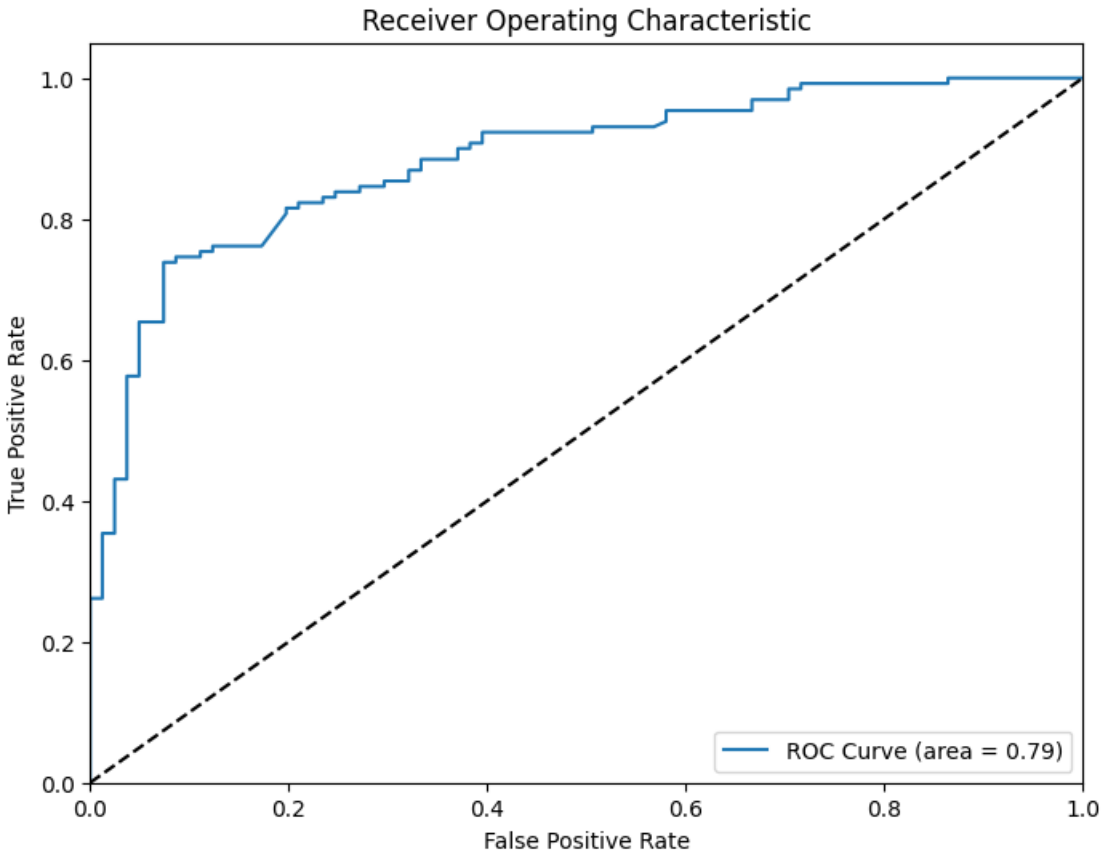

```
Naive Bayes (GossipCop):
Accuracy: 0.7809
Precision: 0.9009
Recall: 0.8019
F1-score: 0.8485
ROC AUC: 0.7572
Confusion Matrix:
[[ 741  299]
 [ 671 2717]]
```


Confusion Matrix
741
299
671
2717
False
True
Actual
False
True
Predicted
2500
2000
1500
1000
500




```
Naive Bayes (PolitiFact):
Accuracy: 0.7867
Precision: 0.8761
Recall: 0.7615
F1-score: 0.8148
ROC AUC: 0.7943
Confusion Matrix:
[[67 14]
 [31 99]]
```


Confusion Matrix
67
14
31
99
False
True
Actual
False
True
Predicted
90
80
70
60
50
40
30
20




```
[8]: (MultinomialNB(),
      array([1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1,
             0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0,
             0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1,
             1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0,
             1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1,
             1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1,
             0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0,
             1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1,
             0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1,
             0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0]))
```

1.7.2 Comparison of my results with baseline Shu et al. [7]

GossipCop Dataset

Support Vector Machine (SVM)

```
My Results:
    Accuracy: 0.7764
    Precision: 0.9002
```

```
    Recall: 0.7960
    F1-score: 0.8449
Baseline Results:
    Accuracy: 0.497
    Precision: 0.511
    Recall: 0.713
    F1-score: 0.595
```

Logistic Regression

```
My Results:
    Accuracy: 0.7850
    Precision: 0.8999
    Recall: 0.8090
    F1-score: 0.8520
Baseline Results:
    Accuracy: 0.648
    Precision: 0.675
    Recall: 0.619
    F1-score: 0.646
```

Naive Bayes

```
My Results:
    Accuracy: 0.7809
    Precision: 0.9009
    Recall: 0.8019
    F1-score: 0.8485
Baseline Results:
    Accuracy: 0.624
    Precision: 0.631
    Recall: 0.669
    F1-score: 0.649
```

PolitiFact Dataset

Support Vector Machine (SVM)

```
My Results:
    Accuracy: 0.7820
    Precision: 0.8750
    Recall: 0.7538
    F1-score: 0.8099
Baseline Results:
    Accuracy: 0.580
    Precision: 0.611
    Recall: 0.717
    F1-score: 0.659
```

Logistic Regression

```
My Results:
```

```
    Accuracy: 0.7962
    Precision: 0.8919
    Recall: 0.7615
    F1-score: 0.8216
Baseline Results:
    Accuracy: 0.642
    Precision: 0.757
    Recall: 0.543
    F1-score: 0.633
```

Naive Bayes

```
My Results:
    Accuracy: 0.7867
    Precision: 0.8761
    Recall: 0.7615
    F1-score: 0.8148
Baseline Results:
    Accuracy: 0.617
    Precision: 0.674
    Recall: 0.630
    F1-score: 0.651
```

My models outperform the baseline results significantly across all metrics (accuracy, precision, recall, F1-score) for both datasets (GossipCop and PolitiFact). After applying SMOTE to balance the datasets, the models show a more balanced performance across precision and recall, which results in higher F1-scores.

1.8 Implement 5-Fold Cross-Validation

I will randomly split the datasets into five parts and conduct 5-fold cross-validation to obtain robust results, as Bian et al.[5] did. Cross-validation is a statistical method used to estimate the performance of machine learning models. As Browne et al.[15] explain: "In its simplest form, the leaving one out at a time method, this involves partitioning a sample of size N into a calibration sample of size N-1 and a validation sample of size 1 and repeating the process N times. An average of the N cross-validation index values is then used." This method involves splitting the data into a number of subsets (folds), training the model on some subsets while testing it on the remaining subset, and repeating this process several times. The performance metrics are then averaged over all iterations to provide a more robust evaluation.

According to Powers et al.[16], "The F1-score, which is the harmonic mean of precision and recall, is particularly recommended for imbalanced datasets because it provides a balance between false positives and false negatives, thus giving a more comprehensive measure of a model's performance." For this reason, I have selected the F1 score as the scoring parameter in 5-Fold Cross-Validation, because the datasets are imbalanced.

1.8.1 Cross-Validation Function

[9]:
```
# Function to perform cross-validation
def cross_validate_model(model, X, y, cv=5, scoring='f1'):
    # Perform cross-validation
    scores = cross_val_score(model, X, y, cv=cv, scoring=scoring)
    print(f'Cross-Validation Scores: {scores}')
    print(f'Average Cross-Validation Score: {np.mean(scores):.4f}')
```

1.8.2 Cross-Validation Scores

[10]:
```
# Logistic Regression
print("Logistic Regression (Cross-Validation) (GossipCop):")
logistic_model_gossipcop = LogisticRegression(max_iter=1000)
cross_validate_model(logistic_model_gossipcop, X_train_gossipcop_resampled,␣
 ↪y_train_gossipcop_resampled)

print("\nLogistic Regression (Cross-Validation) (PolitiFact):")
logistic_model_politifact = LogisticRegression(max_iter=1000)
cross_validate_model(logistic_model_politifact, X_train_politifact_resampled,␣
 ↪y_train_politifact_resampled)

# Support Vector Machine (SVM)
print("\nSupport Vector Machine (Cross-Validation) (GossipCop):")
svm_model_gossipcop = SVC(kernel='linear')
cross_validate_model(svm_model_gossipcop, X_train_gossipcop_resampled,␣
 ↪y_train_gossipcop_resampled)

print("\nSupport Vector Machine (Cross-Validation) (PolitiFact):")
svm_model_politifact = SVC(kernel='linear')
cross_validate_model(svm_model_politifact, X_train_politifact_resampled,␣
 ↪y_train_politifact_resampled)

# Naive Bayes
print("\nNaive Bayes (Cross-Validation) (GossipCop):")
nb_model_gossipcop = MultinomialNB()
cross_validate_model(nb_model_gossipcop, X_train_gossipcop_resampled,␣
 ↪y_train_gossipcop_resampled)

print("\nNaive Bayes (Cross-Validation) (PolitiFact):")
nb_model_politifact = MultinomialNB()
cross_validate_model(nb_model_politifact, X_train_politifact_resampled,␣
 ↪y_train_politifact_resampled)
```

```
Logistic Regression (Cross-Validation) (GossipCop):
Cross-Validation Scores: [0.78099694 0.8046788  0.83555041 0.82404748
0.82903981]
```

```
Average Cross-Validation Score: 0.8149

Logistic Regression (Cross-Validation) (PolitiFact):
Cross-Validation Scores: [0.84153005 0.8172043  0.84848485 0.83243243
0.85714286]
Average Cross-Validation Score: 0.8394

Support Vector Machine (Cross-Validation) (GossipCop):
Cross-Validation Scores: [0.77716995 0.80158282 0.83418669 0.82606989
0.82816229]
Average Cross-Validation Score: 0.8134

Support Vector Machine (Cross-Validation) (PolitiFact):
Cross-Validation Scores: [0.84324324 0.77419355 0.84102564 0.79569892
0.87292818]
Average Cross-Validation Score: 0.8254

Naive Bayes (Cross-Validation) (GossipCop):
Cross-Validation Scores: [0.77412321 0.79001628 0.79508493 0.77641007
0.77808832]
Average Cross-Validation Score: 0.7827

Naive Bayes (Cross-Validation) (PolitiFact):
Cross-Validation Scores: [0.85263158 0.83673469 0.86734694 0.84656085
0.87150838]
Average Cross-Validation Score: 0.8550
```

The results are consistent across the folds, which is a good sign that the models are generalizing well and not just memorizing the training data. The average performance metrics (F1-score) are very close to the fold-specific metrics. This consistency suggests that the models performed similarly on unseen data, that is another good sign.

1.9 Hyperparameter Tuning

My use of grid search and hyperparameter tuning was inspired by the methodology described by Chong et al. [17].

1.9.1 Hyperparameter Tuning Function

[11]:
```python
# Hyperparameter tuning using Grid search - computationally intense, will run␣
↪for 1 h
# Define Parameter Grids
logistic_params = {
    'C': [0.1, 1, 10, 100],
    'solver': ['liblinear', 'saga']
}

svm_params = {
```

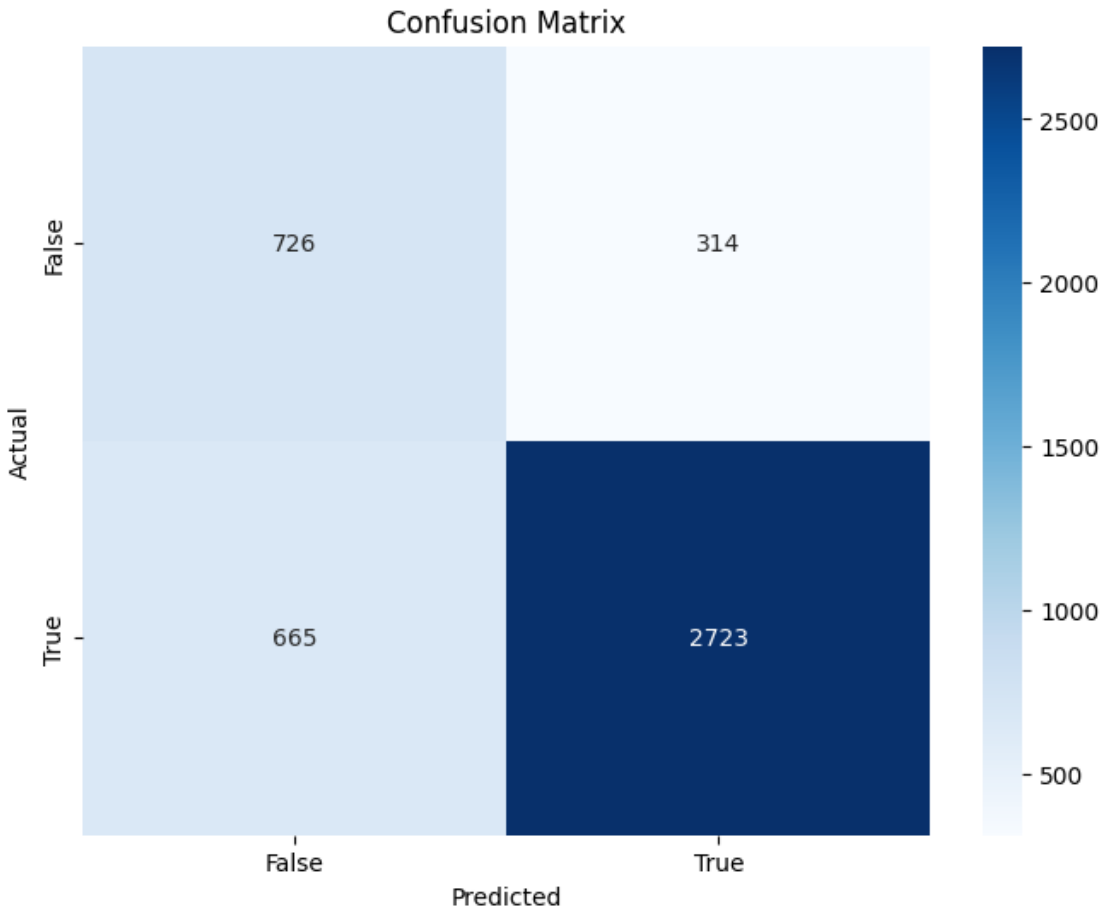
```
    'C': [0.1, 1, 10, 100],
    'kernel': ['linear', 'rbf']
}

nb_params = {
    'alpha': [0.1, 0.5, 1, 5, 10]
}

# Function to perform hyperparameter tuning and evaluation with 5-fold␣
 ↪cross-validation
def tune_and_evaluate(model, param_grid, X_train, X_test, y_train, y_test,␣
 ↪scoring='f1'):
    grid_search = GridSearchCV(model, param_grid, cv=5, scoring=scoring)
    grid_search.fit(X_train, y_train)
    best_model = grid_search.best_estimator_
    print(f'Best Parameters: {grid_search.best_params_}')
    # Evaluate the best model using cross-validation scores
    cv_results = grid_search.cv_results_
    mean_cv_score = grid_search.best_score_
    print(f'Best cross-validation score: {mean_cv_score:.4f}')
    # Evaluate the best model on the test set
    return train_and_evaluate(best_model, X_train, X_test, y_train, y_test)
```

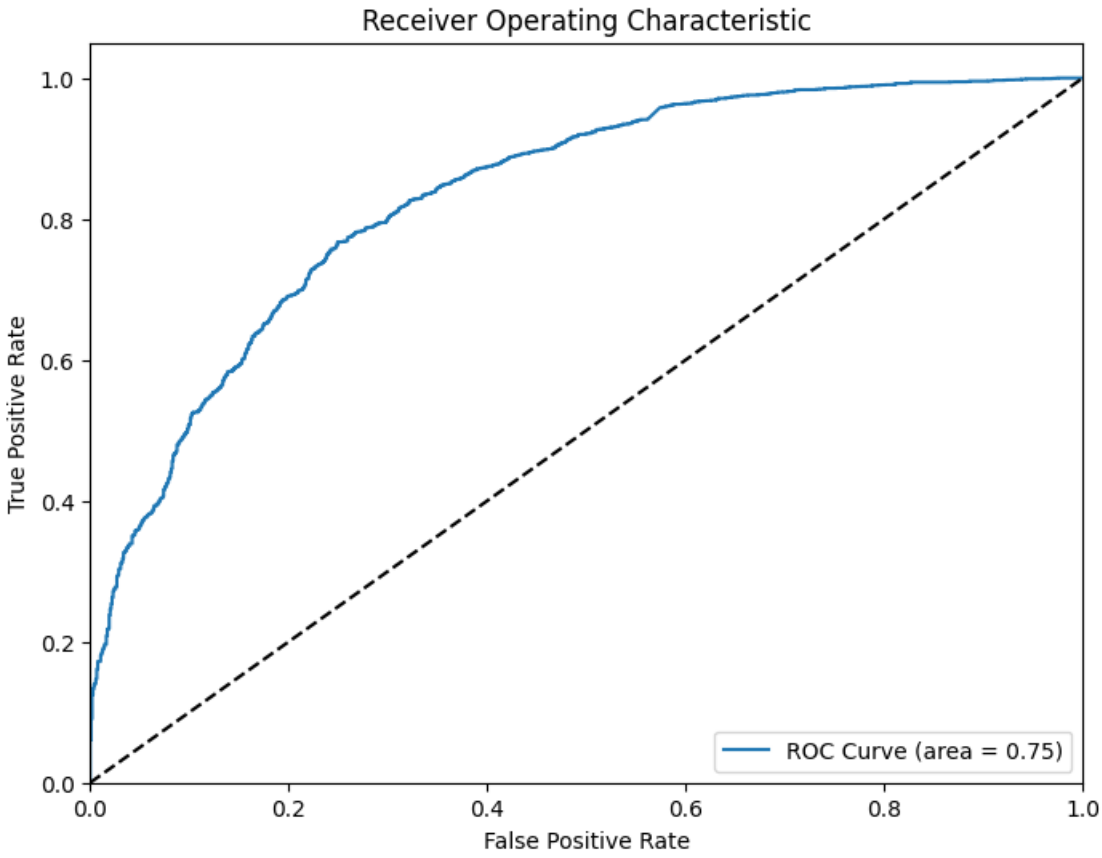
1.9.2 Hyperparameter Tuning Results

[12]:
```
# Logistic Regression with Hyperparameter Tuning
print("Logistic Regression with Hyperparameter Tuning (GossipCop):")
tune_and_evaluate(LogisticRegression(max_iter=1000, random_state=42),␣
 ↪logistic_params,
                  X_train_gossipcop_resampled, X_test_gossipcop_tfidf,␣
 ↪y_train_gossipcop_resampled, y_test_gossipcop)

print("\nLogistic Regression with Hyperparameter Tuning (PolitiFact):")
tune_and_evaluate(LogisticRegression(max_iter=1000, random_state=42),␣
 ↪logistic_params,
                  X_train_politifact_resampled, X_test_politifact_tfidf,␣
 ↪y_train_politifact_resampled, y_test_politifact)

# Support Vector Machine (SVM) with Hyperparameter Tuning
print("\nSupport Vector Machine with Hyperparameter Tuning (GossipCop):")
tune_and_evaluate(SVC(random_state=42), svm_params,
                  X_train_gossipcop_resampled, X_test_gossipcop_tfidf,␣
 ↪y_train_gossipcop_resampled, y_test_gossipcop)

print("\nSupport Vector Machine with Hyperparameter Tuning (PolitiFact):")
tune_and_evaluate(SVC(random_state=42), svm_params,
```

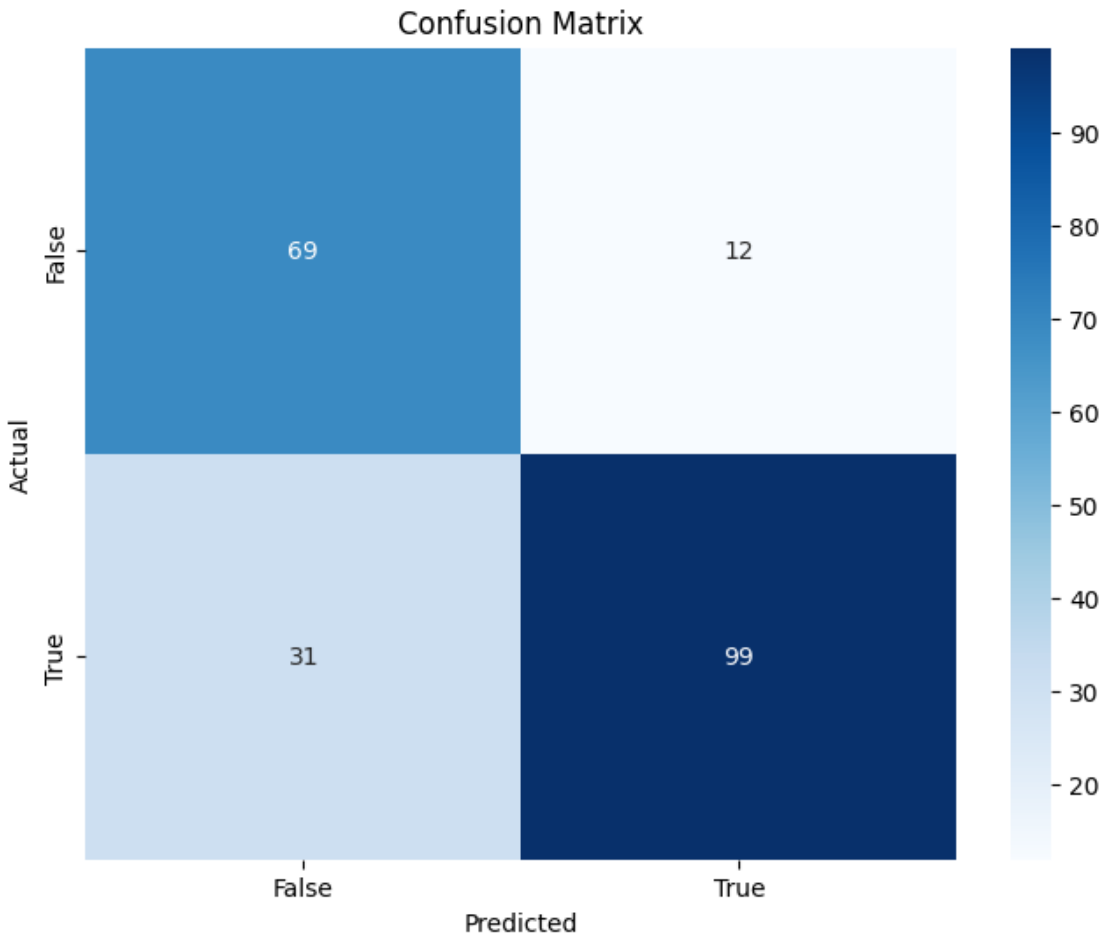
```
                  X_train_politifact_resampled, X_test_politifact_tfidf,␣
 ↪y_train_politifact_resampled, y_test_politifact)

# Naive Bayes with Hyperparameter Tuning
print("\nNaive Bayes with Hyperparameter Tuning (GossipCop):")
tune_and_evaluate(MultinomialNB(), nb_params,
                  X_train_gossipcop_resampled, X_test_gossipcop_tfidf,␣
 ↪y_train_gossipcop_resampled, y_test_gossipcop)

print("\nNaive Bayes with Hyperparameter Tuning (PolitiFact):")
tune_and_evaluate(MultinomialNB(), nb_params,
                  X_train_politifact_resampled, X_test_politifact_tfidf,␣
 ↪y_train_politifact_resampled, y_test_politifact)
```

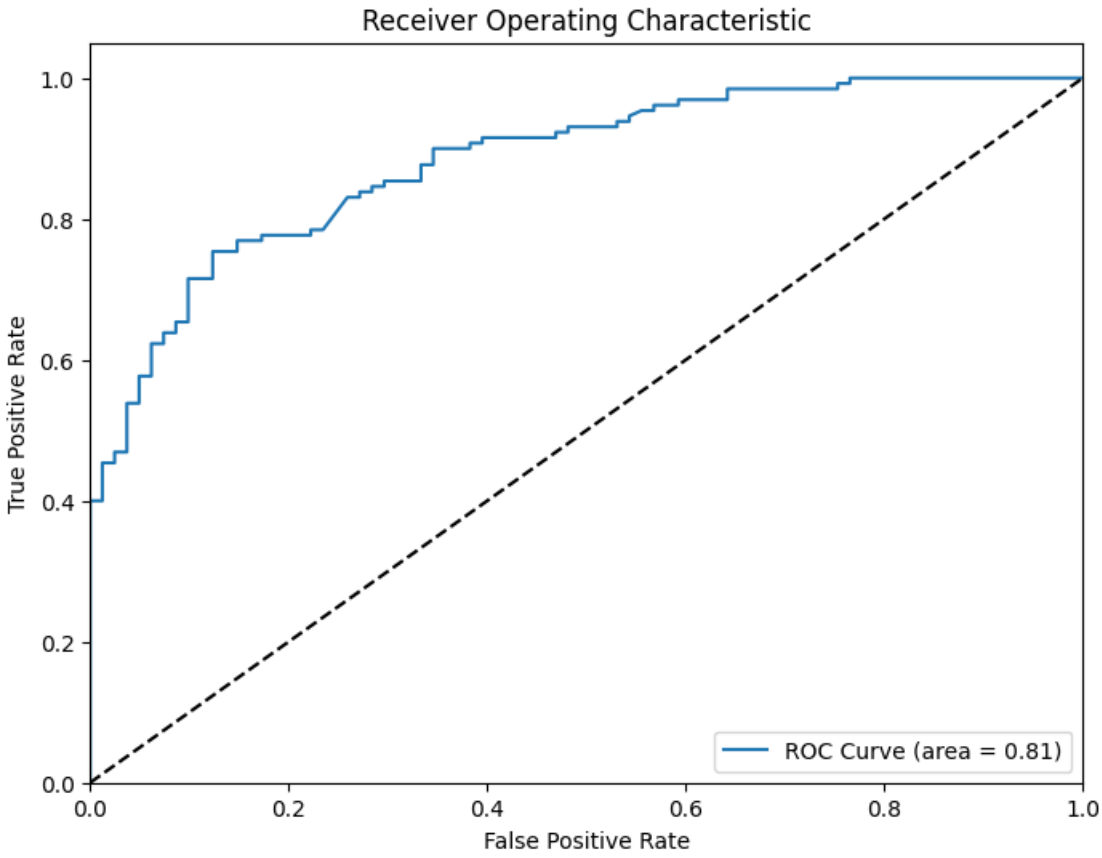
```
Logistic Regression with Hyperparameter Tuning (GossipCop):
Best Parameters: {'C': 10, 'solver': 'liblinear'}
Best cross-validation score: 0.8158
Accuracy: 0.7789
Precision: 0.8966
Recall: 0.8037
F1-score: 0.8476
ROC AUC: 0.7509
Confusion Matrix:
[[ 726  314]
 [ 665 2723]]
```


Confusion Matrix
726
314
665
2723
False
True
Actual
False
True
Predicted
2500
2000
1500
1000
500

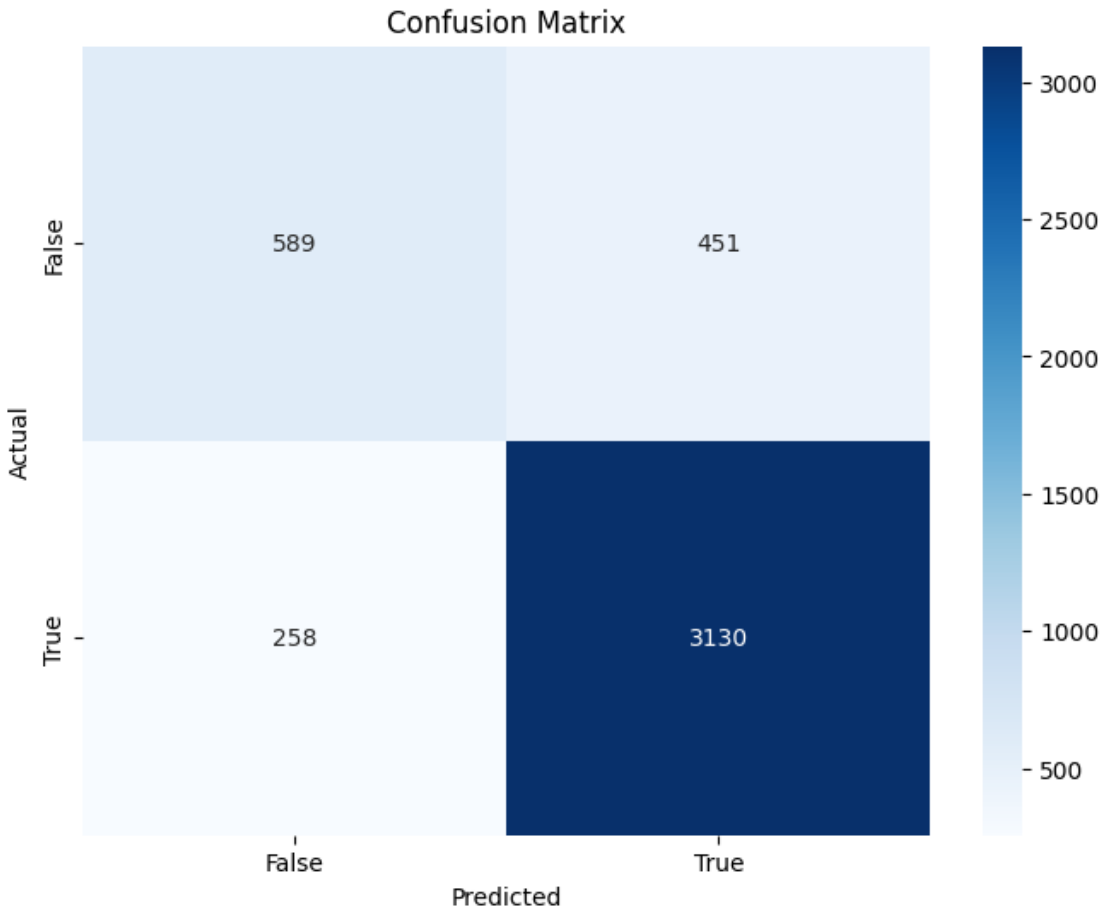


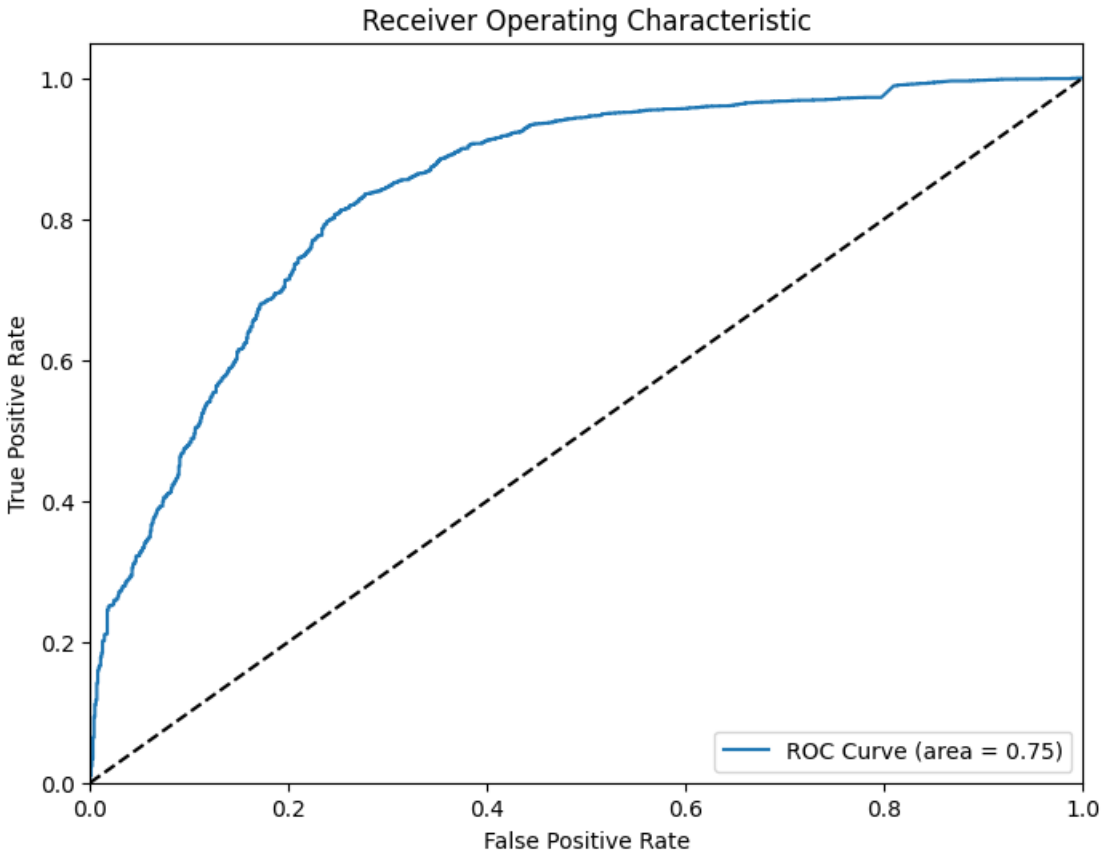

```
Logistic Regression with Hyperparameter Tuning (PolitiFact):
Best Parameters: {'C': 1, 'solver': 'liblinear'}
Best cross-validation score: 0.8394
Accuracy: 0.7962
Precision: 0.8919
Recall: 0.7615
F1-score: 0.8216
ROC AUC: 0.8067
Confusion Matrix:
[[69 12]
 [31 99]]
```


Confusion Matrix
69
12
31
99
False
True
Actual
False
True
Predicted
90
80
70
60
50
40
30
20

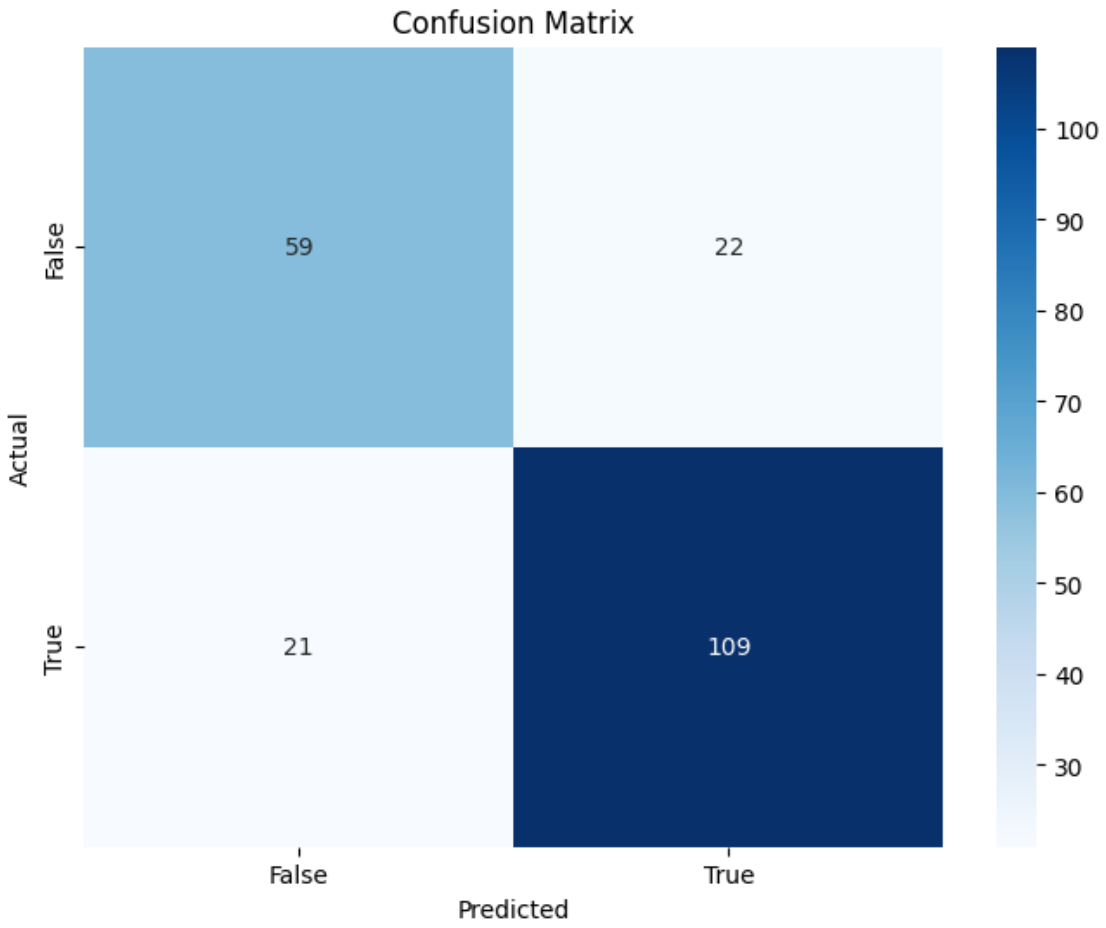


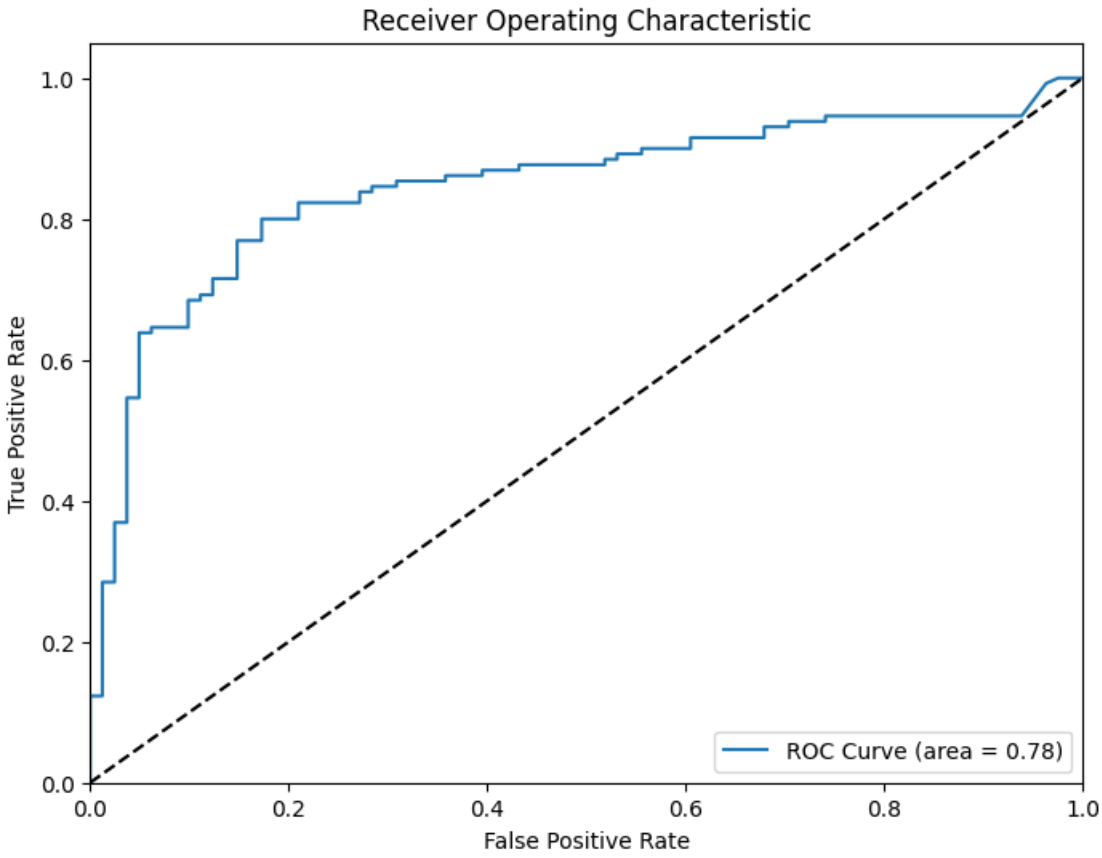

```
Support Vector Machine with Hyperparameter Tuning (GossipCop):
Best Parameters: {'C': 10, 'kernel': 'rbf'}
Best cross-validation score: 0.9207
Accuracy: 0.8399
Precision: 0.8741
Recall: 0.9238
F1-score: 0.8983
ROC AUC: 0.7451
Confusion Matrix:
[[ 589  451]
 [ 258 3130]]
```


Confusion Matrix
589
451
258
3130
False
True
Actual
False
True
Predicted
3000
2500
2000
1500
1000
500

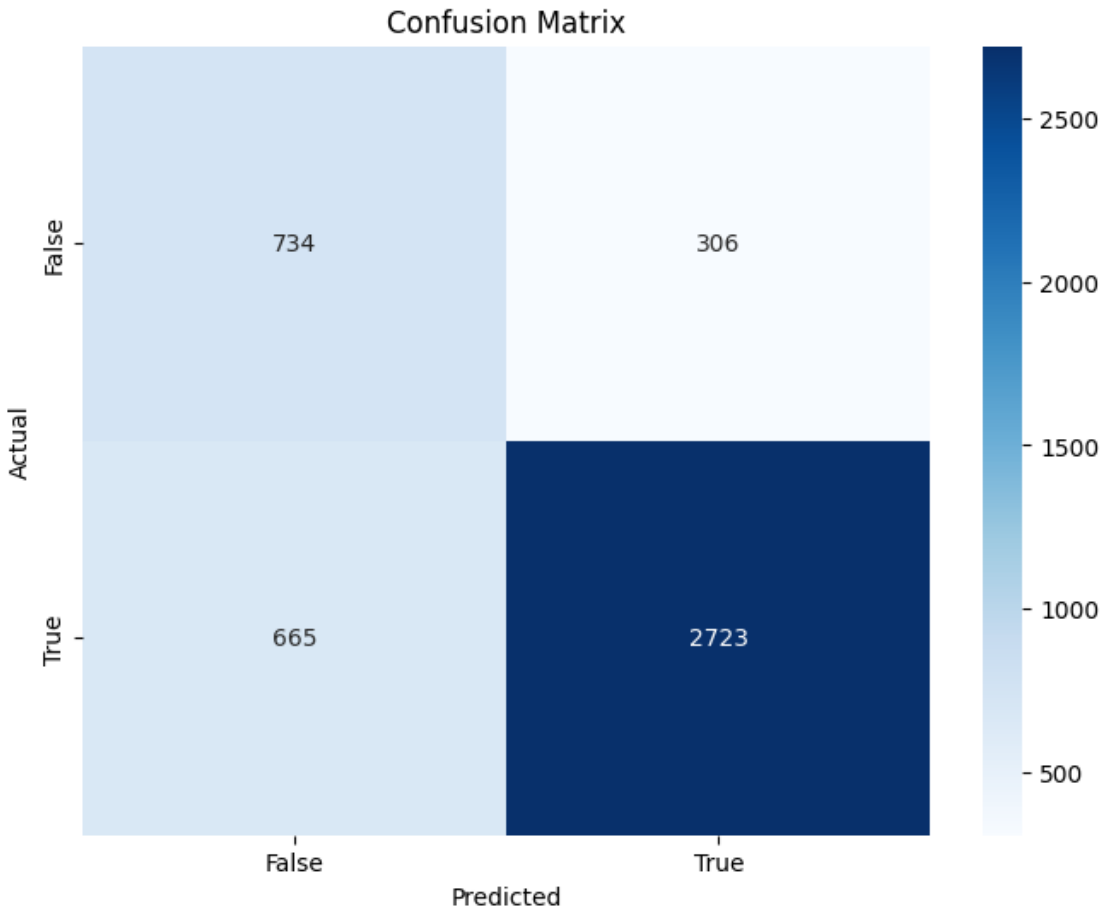


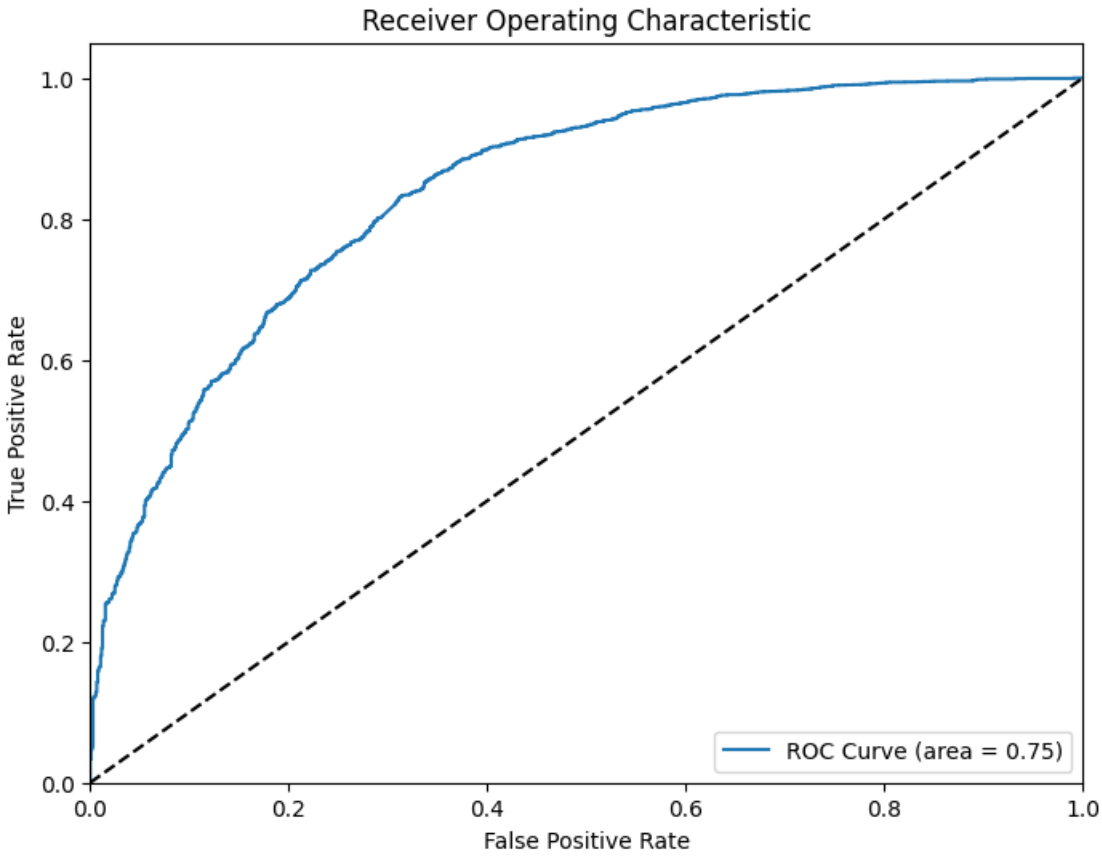

```
Support Vector Machine with Hyperparameter Tuning (PolitiFact):
Best Parameters: {'C': 1, 'kernel': 'rbf'}
Best cross-validation score: 0.8650
Accuracy: 0.7962
Precision: 0.8321
Recall: 0.8385
F1-score: 0.8352
ROC AUC: 0.7834
Confusion Matrix:
[[ 59  22]
 [ 21 109]]
```


Confusion Matrix
59
22
21
109
False
True
Actual
False
True
Predicted
100
90
80
70
60
50
40
30

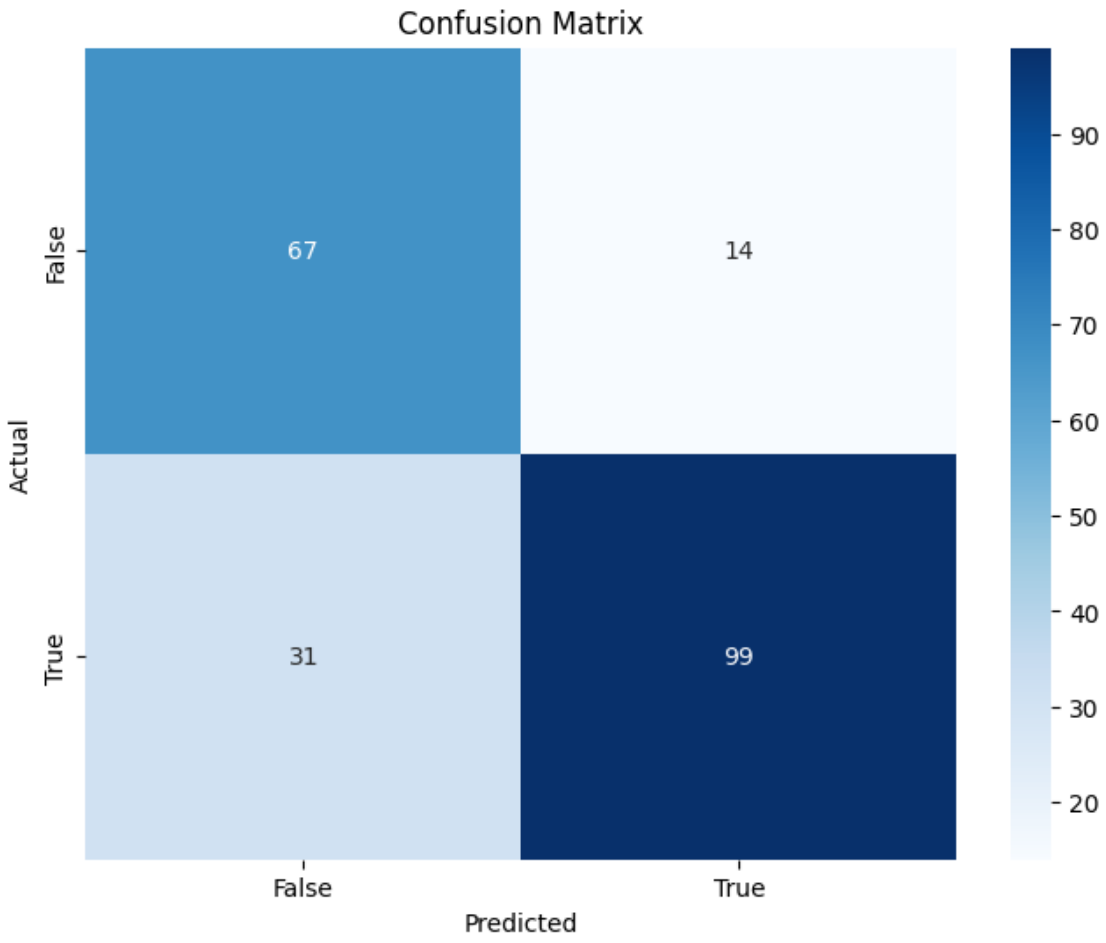


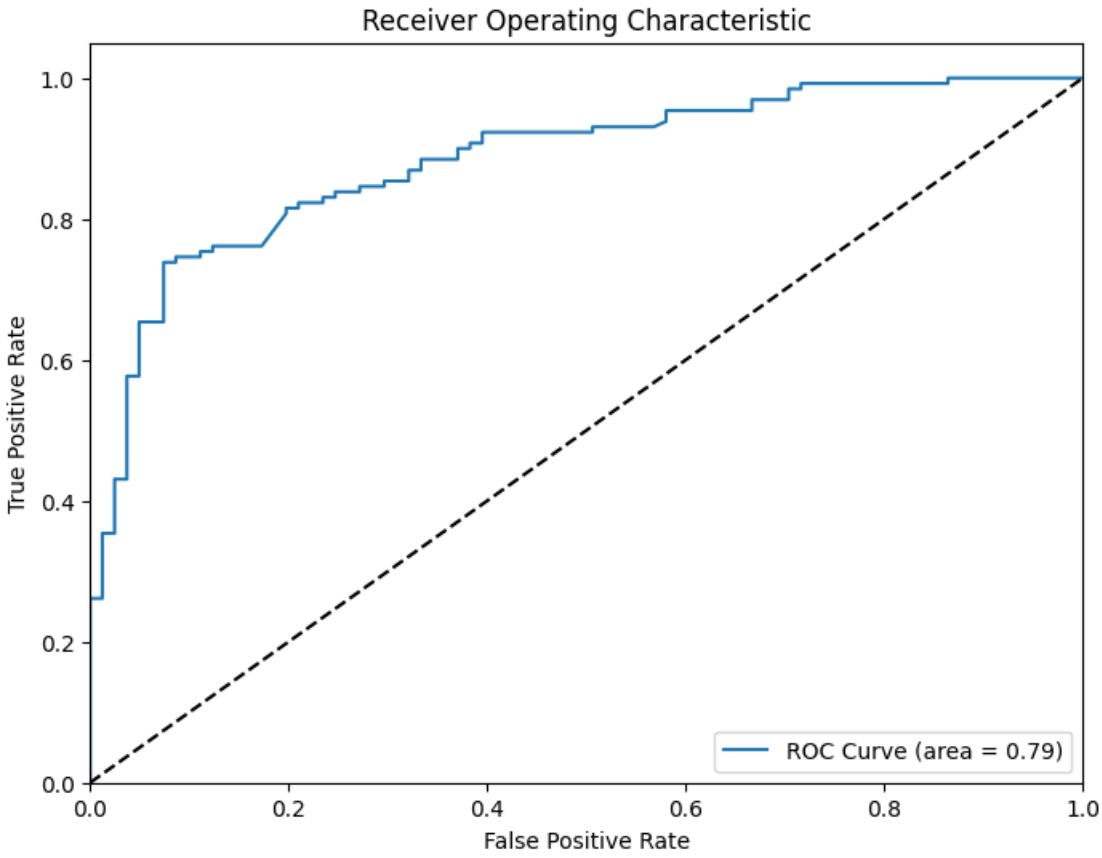

```
Naive Bayes with Hyperparameter Tuning (GossipCop):
Best Parameters: {'alpha': 0.1}
Best cross-validation score: 0.7837
Accuracy: 0.7807
Precision: 0.8990
Recall: 0.8037
F1-score: 0.8487
ROC AUC: 0.7547
Confusion Matrix:
[[ 734  306]
 [ 665 2723]]
```


Confusion Matrix
734
306
665
2723
False
True
Actual
False
True
Predicted
2500
2000
1500
1000
500




```
Naive Bayes with Hyperparameter Tuning (PolitiFact):
Best Parameters: {'alpha': 1}
Best cross-validation score: 0.8550
Accuracy: 0.7867
Precision: 0.8761
Recall: 0.7615
F1-score: 0.8148
ROC AUC: 0.7943
Confusion Matrix:
[[67 14]
 [31 99]]
```


Confusion Matrix
67
14
31
99
False
True
Actual
False
True
Predicted
90
80
70
60
50
40
30
20



[12]:
```
(MultinomialNB(alpha=1),
 array([1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1,
        0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0,
        0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1,
        1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0,
        1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1,
        1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1,
        0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0,
        1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1,
        0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1,
        0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0]))
```

Logistic Regression

GossipCop:

```
Original:
    Accuracy: 0.7850
    F1-score: 0.8520
    ROC AUC: 0.7579
Fine-Tuned:
```

```
    Accuracy: 0.7789 (slightly worse)
    F1-score: 0.8476 (slightly worse)
    ROC AUC: 0.7509 (worse)
```

PolitiFact:

```
Both original and fine-tuned models have identical metrics.
```

Support Vector Machine

GossipCop:

```
Original:
    Accuracy: 0.7764
    F1-score: 0.8449
    ROC AUC: 0.7543
Fine-Tuned:
    Accuracy: 0.8399 (improved)
    F1-score: 0.8983 (improved)
    ROC AUC: 0.7451 (slightly worse)
```

PolitiFact:

```
Original:
    Accuracy: 0.7820
    F1-score: 0.8099
    ROC AUC: 0.7905
Fine-Tuned:
    Accuracy: 0.7962 (improved)
    F1-score: 0.8352 (improved)
    ROC AUC: 0.7834 (slightly worse)
```

Naive Bayes

GossipCop:

```
Original:
    Accuracy: 0.7809
    F1-score: 0.8485
    ROC AUC: 0.7572
Fine-Tuned:
    Accuracy: 0.7807 (about the same)
    F1-score: 0.8487 (about the same)
    ROC AUC: 0.7547 (slightly worse)
```

PolitiFact:

```
Both original and fine-tuned models have identical metrics.
```

Hyperparameter tuning has led to some improvements, especially in the case of SVM on the GossipCop dataset. However, for Logistic Regression and Naive Bayes, the changes are minimal.

1.10 Save the Best Model to Google Drive

[13]:
```
# Selecting svm_model_gossipcop as the best model after hyperparameter tuning
best_model = SVC(C=10, kernel='rbf', random_state=42)
best_model.fit(X_train_gossipcop_resampled, y_train_gossipcop_resampled)

# Save the model to a file
model_filename = '/content/drive/My Drive/best_svm_model_gossipcop.pkl'
joblib.dump(best_model, model_filename)

# Save the TF-IDF vectorizer to a file
vectorizer_filename = '/content/drive/My Drive/tfidf_vectorizer_gossipcop.pkl'
joblib.dump(vectorizer, vectorizer_filename)

print("Model and vectorizer saved to Google Drive.")
```

```
Model and vectorizer saved to Google Drive.
```

1.11 Conclusion

My models significantly outperform the baseline results across all metrics for both datasets (GossipCop and PolitiFact). Balancing the datasets with SMOTE and tuning hyperparameters further improved the models' performance.

Based on the evaluation metrics (accuracy, precision, recall, F1-score, and ROC AUC) for each classifier, the Support Vector Machine (SVM) with hyperparameter tuning on the GossipCop dataset seems to perform the best overall.

1.12 References

[5] BIAN, T., XIAO, X., XU, T., ZHAO, P., HUANG, W., RONG, Y., and HUANG, J. 2020. Rumor Detection on Social Media with Bi-Directional Graph Convolutional Networks. arXiv preprint arXiv:2001.06362. https://doi.org/10.48550/arXiv.2001.06362

[6] SHEN, Y., LIU, Q., GUO, N., YUAN, J., and YANG, Y. 2023. Fake News Detection on Social Networks: A Survey. Applied Sciences, 13, 11877. https://doi.org/10.3390/app132111877

[7] SHU, K., MAHUDESWARAN, D., WANG, S., LEE, D., and LIU, H. 2018. FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. https://doi.org/10.48550/arXiv.1809.01286

[14] CHAWLA, N.V., BOWYER, K.W., HALL, L.O., and KEGELMEYER, W.P. 2002. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321-357. https://doi.org/10.48550/arXiv.1106.1813

[15] BROWNE, M.W. 2000. Cross-Validation Methods. Journal of Mathematical Psychology, 44, 108–132. https://doi.org/10.1006/jmps.1999.1279

[16] POWERS, D.M.W. 2011. Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. International Journal of Machine Learning Technology, 2(1), pp. 37-63. https://doi.org/10.48550/arXiv.2010.16061

[17] CHONG, K. and SHAH, N. 2022. Comparison of Naive Bayes and SVM Classification in Grid-Search Hyperparameter Tuned and Non-Hyperparameter Tuned Healthcare Stock Market Sentiment Analysis. International Journal of Advanced Computer Science and Applications, 13(12), pp. 90-94. DOI: 10.14569/IJACSA.2022.0131213

[ ]:
```
# Install LaTeX packages necessary for converting notebooks to PDF
!apt-get update
!apt-get install -y texlive-xetex texlive-fonts-recommended␣
  ↪texlive-plain-generic texlive-latex-extra pandoc

# Convert the notebook to PDF
!jupyter nbconvert --to pdf "/content/drive/My Drive/Colab Notebooks/
  ↪FakeNewsNetClassifier.ipynb"
```

```
Get:1 https://cloud.r-project.org/bin/linux/ubuntu jammy-cran40/ InRelease
[3,626 B]
Hit:2 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86_64
InRelease
Get:3 http://security.ubuntu.com/ubuntu jammy-security InRelease [129 kB]
Hit:4 http://archive.ubuntu.com/ubuntu jammy InRelease
Ign:5 https://r2u.stat.illinois.edu/ubuntu jammy InRelease
Get:6 http://archive.ubuntu.com/ubuntu jammy-updates InRelease [128 kB]
Get:7 https://r2u.stat.illinois.edu/ubuntu jammy Release [5,713 B]
Hit:8 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Get:9 https://r2u.stat.illinois.edu/ubuntu jammy Release.gpg [793 B]
Get:10 http://security.ubuntu.com/ubuntu jammy-security/main amd64 Packages
[1,998 kB]
Get:11 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy
InRelease [24.3 kB]
Hit:12 http://archive.ubuntu.com/ubuntu jammy-backports InRelease
Hit:13 https://ppa.launchpadcontent.net/ubuntugis/ppa/ubuntu jammy InRelease
Get:14 https://r2u.stat.illinois.edu/ubuntu jammy/main all Packages [8,181 kB]
Get:15 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 Packages
[1,410 kB]
Get:16 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy/main
amd64 Packages [48.1 kB]
Get:17 http://security.ubuntu.com/ubuntu jammy-security/universe amd64 Packages
[1,127 kB]
Get:18 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 Packages [2,263
kB]
Get:19 https://r2u.stat.illinois.edu/ubuntu jammy/main amd64 Packages [2,544 kB]
Fetched 17.9 MB in 4s (4,259 kB/s)
Reading package lists… Done
W: Skipping acquire of configured file 'main/source/Sources' as repository
'https://r2u.stat.illinois.edu/ubuntu jammy InRelease' does not seem to provide
it (sources.list entry misspelt?)
Reading package lists… Done
Building dependency tree… Done
```

```
Reading state information… Done
The following additional packages will be installed:
  dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono fonts-
texgyre
  fonts-urw-base35 libapache-pom-java libcmark-gfm-extensions0.29.0.gfm.3
libcmark-gfm0.29.0.gfm.3
  libcommons-logging-java libcommons-parent-java libfontbox-java libfontenc1
libgs9 libgs9-common
  libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsynctex2
  libteckit0 libtexlua53 libtexluajit2 libwoff1 libzzip-0-13 lmodern pandoc-data
poppler-data
  preview-latex-style rake ruby ruby-net-telnet ruby-rubygems ruby-webrick ruby-
xmlrpc ruby3.0
  rubygems-integration t1utils teckit tex-common tex-gyre texlive-base texlive-
binaries
  texlive-latex-base texlive-latex-recommended texlive-pictures tipa xfonts-
encodings xfonts-utils
Suggested packages:
  fonts-noto fonts-freefont-otf | fonts-freefont-ttf libavalon-framework-java
  libcommons-logging-java-doc libexcalibur-logkit-java liblog4j1.2-java texlive-
luatex
  pandoc-citeproc context wkhtmltopdf librsvg2-bin groff ghc nodejs php python
libjs-mathjax
  libjs-katex citation-style-language-styles poppler-utils ghostscript fonts-
japanese-mincho
  | fonts-ipafont-mincho fonts-japanese-gothic | fonts-ipafont-gothic fonts-
arphic-ukai
  fonts-arphic-uming fonts-nanum ri ruby-dev bundler debhelper gv | postscript-
viewer perl-tk xpdf
  | pdf-viewer xzdec texlive-fonts-recommended-doc texlive-latex-base-doc
python3-pygments
  icc-profiles libfile-which-perl libspreadsheet-parseexcel-perl texlive-latex-
extra-doc
  texlive-latex-recommended-doc texlive-pstricks dot2tex prerex texlive-
pictures-doc vprerex
  default-jre-headless tipa-doc
The following NEW packages will be installed:
  dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono fonts-
texgyre
  fonts-urw-base35 libapache-pom-java libcmark-gfm-extensions0.29.0.gfm.3
libcmark-gfm0.29.0.gfm.3
  libcommons-logging-java libcommons-parent-java libfontbox-java libfontenc1
libgs9 libgs9-common
  libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsynctex2
  libteckit0 libtexlua53 libtexluajit2 libwoff1 libzzip-0-13 lmodern pandoc
pandoc-data
```

```
  poppler-data preview-latex-style rake ruby ruby-net-telnet ruby-rubygems ruby-
webrick ruby-xmlrpc
  ruby3.0 rubygems-integration t1utils teckit tex-common tex-gyre texlive-base
texlive-binaries
  texlive-fonts-recommended texlive-latex-base texlive-latex-extra texlive-
latex-recommended
  texlive-pictures texlive-plain-generic texlive-xetex tipa xfonts-encodings
xfonts-utils
0 upgraded, 58 newly installed, 0 to remove and 45 not upgraded.
Need to get 202 MB of archives.
After this operation, 728 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu jammy/main amd64 fonts-droid-fallback all
1:6.0.1r16-1.1build1 [1,805 kB]
Get:2 http://archive.ubuntu.com/ubuntu jammy/main amd64 fonts-lato all 2.0-2.1
[2,696 kB]
Get:3 http://archive.ubuntu.com/ubuntu jammy/main amd64 poppler-data all
0.4.11-1 [2,171 kB]
Get:4 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tex-common all 6.17
[33.7 kB]
Get:5 http://archive.ubuntu.com/ubuntu jammy/main amd64 fonts-urw-base35 all
20200910-1 [6,367 kB]
Get:6 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libgs9-common
all 9.55.0~dfsg1-0ubuntu5.7 [752 kB]
Get:7 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libidn12 amd64
1.38-4ubuntu1 [60.0 kB]
Get:8 http://archive.ubuntu.com/ubuntu jammy/main amd64 libijs-0.35 amd64
0.35-15build2 [16.5 kB]
Get:9 http://archive.ubuntu.com/ubuntu jammy/main amd64 libjbig2dec0 amd64
0.19-3build2 [64.7 kB]
Get:10 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libgs9 amd64
9.55.0~dfsg1-0ubuntu5.7 [5,028 kB]
Get:11 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libkpathsea6
amd64 2021.20210626.59705-1ubuntu0.2 [60.4 kB]
Get:12 http://archive.ubuntu.com/ubuntu jammy/main amd64 libwoff1 amd64
1.0.2-1build4 [45.2 kB]
Get:13 http://archive.ubuntu.com/ubuntu jammy/universe amd64 dvisvgm amd64
2.13.1-1 [1,221 kB]
Get:14 http://archive.ubuntu.com/ubuntu jammy/universe amd64 fonts-lmodern all
2.004.5-6.1 [4,532 kB]
Get:15 http://archive.ubuntu.com/ubuntu jammy/main amd64 fonts-noto-mono all
20201225-1build1 [397 kB]
Get:16 http://archive.ubuntu.com/ubuntu jammy/universe amd64 fonts-texgyre all
20180621-3.1 [10.2 MB]
Get:17 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libapache-pom-java
all 18-1 [4,720 B]
Get:18 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libcmark-
gfm0.29.0.gfm.3 amd64 0.29.0.gfm.3-3 [115 kB]
Get:19 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libcmark-gfm-
```

```
extensions0.29.0.gfm.3 amd64 0.29.0.gfm.3-3 [25.1 kB]
Get:20 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libcommons-parent-
java all 43-1 [10.8 kB]
Get:21 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libcommons-logging-
java all 1.2-2 [60.3 kB]
Get:22 http://archive.ubuntu.com/ubuntu jammy/main amd64 libfontenc1 amd64
1:1.1.4-1build3 [14.7 kB]
Get:23 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libptexenc1
amd64 2021.20210626.59705-1ubuntu0.2 [39.1 kB]
Get:24 http://archive.ubuntu.com/ubuntu jammy/main amd64 rubygems-integration
all 1.18 [5,336 B]
Get:25 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 ruby3.0 amd64
3.0.2-7ubuntu2.7 [50.1 kB]
Get:26 http://archive.ubuntu.com/ubuntu jammy/main amd64 ruby-rubygems all
3.3.5-2 [228 kB]
Get:27 http://archive.ubuntu.com/ubuntu jammy/main amd64 ruby amd64 1:3.0~exp1
[5,100 B]
Get:28 http://archive.ubuntu.com/ubuntu jammy/main amd64 rake all 13.0.6-2 [61.7
kB]
Get:29 http://archive.ubuntu.com/ubuntu jammy/main amd64 ruby-net-telnet all
0.1.1-2 [12.6 kB]
Get:30 http://archive.ubuntu.com/ubuntu jammy/universe amd64 ruby-webrick all
1.7.0-3 [51.8 kB]
Get:31 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 ruby-xmlrpc all
0.3.2-1ubuntu0.1 [24.9 kB]
Get:32 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libruby3.0
amd64 3.0.2-7ubuntu2.7 [5,113 kB]
Get:33 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libsynctex2
amd64 2021.20210626.59705-1ubuntu0.2 [55.6 kB]
Get:34 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libteckit0 amd64
2.5.11+ds1-1 [421 kB]
Get:35 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libtexlua53
amd64 2021.20210626.59705-1ubuntu0.2 [120 kB]
Get:36 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libtexluajit2
amd64 2021.20210626.59705-1ubuntu0.2 [267 kB]
Get:37 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libzzip-0-13 amd64
0.13.72+dfsg.1-1.1 [27.0 kB]
Get:38 http://archive.ubuntu.com/ubuntu jammy/main amd64 xfonts-encodings all
1:1.0.5-0ubuntu2 [578 kB]
Get:39 http://archive.ubuntu.com/ubuntu jammy/main amd64 xfonts-utils amd64
1:7.7+6build2 [94.6 kB]
Get:40 http://archive.ubuntu.com/ubuntu jammy/universe amd64 lmodern all
2.004.5-6.1 [9,471 kB]
Get:41 http://archive.ubuntu.com/ubuntu jammy/universe amd64 pandoc-data all
2.9.2.1-3ubuntu2 [81.8 kB]
Get:42 http://archive.ubuntu.com/ubuntu jammy/universe amd64 pandoc amd64
2.9.2.1-3ubuntu2 [20.3 MB]
Get:43 http://archive.ubuntu.com/ubuntu jammy/universe amd64 preview-latex-style
```

```
all 12.2-1ubuntu1 [185 kB]
Get:44 http://archive.ubuntu.com/ubuntu jammy/main amd64 t1utils amd64
1.41-4build2 [61.3 kB]
Get:45 http://archive.ubuntu.com/ubuntu jammy/universe amd64 teckit amd64
2.5.11+ds1-1 [699 kB]
Get:46 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tex-gyre all
20180621-3.1 [6,209 kB]
Get:47 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 texlive-
binaries amd64 2021.20210626.59705-1ubuntu0.2 [9,860 kB]
Get:48 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-base all
2021.20220204-1 [21.0 MB]
Get:49 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-fonts-
recommended all 2021.20220204-1 [4,972 kB]
Get:50 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-base
all 2021.20220204-1 [1,128 kB]
Get:51 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libfontbox-java all
1:1.8.16-2 [207 kB]
Get:52 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libpdfbox-java all
1:1.8.16-2 [5,199 kB]
Get:53 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-
recommended all 2021.20220204-1 [14.4 MB]
Get:54 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-pictures
all 2021.20220204-1 [8,720 kB]
Get:55 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-extra
all 2021.20220204-1 [13.9 MB]
Get:56 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-plain-
generic all 2021.20220204-1 [27.5 MB]
Get:57 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tipa all 2:1.3-21
[2,967 kB]
Get:58 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-xetex all
2021.20220204-1 [12.4 MB]
Fetched 202 MB in 17s (11.9 MB/s)
Extracting templates from packages: 100%
Preconfiguring packages …
Selecting previously unselected package fonts-droid-fallback.
(Reading database … 123576 files and directories currently installed.)
Preparing to unpack …/00-fonts-droid-fallback_1%3a6.0.1r16-1.1build1_all.deb
…
Unpacking fonts-droid-fallback (1:6.0.1r16-1.1build1) …
Selecting previously unselected package fonts-lato.
Preparing to unpack …/01-fonts-lato_2.0-2.1_all.deb …
Unpacking fonts-lato (2.0-2.1) …
Selecting previously unselected package poppler-data.
Preparing to unpack …/02-poppler-data_0.4.11-1_all.deb …
Unpacking poppler-data (0.4.11-1) …
Selecting previously unselected package tex-common.
Preparing to unpack …/03-tex-common_6.17_all.deb …
Unpacking tex-common (6.17) …
```

```
Selecting previously unselected package fonts-urw-base35.
Preparing to unpack …/04-fonts-urw-base35_20200910-1_all.deb …
Unpacking fonts-urw-base35 (20200910-1) …
Selecting previously unselected package libgs9-common.
Preparing to unpack …/05-libgs9-common_9.55.0~dfsg1-0ubuntu5.7_all.deb …
Unpacking libgs9-common (9.55.0~dfsg1-0ubuntu5.7) …
Selecting previously unselected package libidn12:amd64.
Preparing to unpack …/06-libidn12_1.38-4ubuntu1_amd64.deb …
Unpacking libidn12:amd64 (1.38-4ubuntu1) …
Selecting previously unselected package libijs-0.35:amd64.
Preparing to unpack …/07-libijs-0.35_0.35-15build2_amd64.deb …
Unpacking libijs-0.35:amd64 (0.35-15build2) …
Selecting previously unselected package libjbig2dec0:amd64.
Preparing to unpack …/08-libjbig2dec0_0.19-3build2_amd64.deb …
Unpacking libjbig2dec0:amd64 (0.19-3build2) …
Selecting previously unselected package libgs9:amd64.
Preparing to unpack …/09-libgs9_9.55.0~dfsg1-0ubuntu5.7_amd64.deb …
Unpacking libgs9:amd64 (9.55.0~dfsg1-0ubuntu5.7) …
Selecting previously unselected package libkpathsea6:amd64.
Preparing to unpack …/10-libkpathsea6_2021.20210626.59705-1ubuntu0.2_amd64.deb
…
Unpacking libkpathsea6:amd64 (2021.20210626.59705-1ubuntu0.2) …
Selecting previously unselected package libwoff1:amd64.
Preparing to unpack …/11-libwoff1_1.0.2-1build4_amd64.deb …
Unpacking libwoff1:amd64 (1.0.2-1build4) …
Selecting previously unselected package dvisvgm.
Preparing to unpack …/12-dvisvgm_2.13.1-1_amd64.deb …
Unpacking dvisvgm (2.13.1-1) …
Selecting previously unselected package fonts-lmodern.
Preparing to unpack …/13-fonts-lmodern_2.004.5-6.1_all.deb …
Unpacking fonts-lmodern (2.004.5-6.1) …
Selecting previously unselected package fonts-noto-mono.
Preparing to unpack …/14-fonts-noto-mono_20201225-1build1_all.deb …
Unpacking fonts-noto-mono (20201225-1build1) …
Selecting previously unselected package fonts-texgyre.
Preparing to unpack …/15-fonts-texgyre_20180621-3.1_all.deb …
Unpacking fonts-texgyre (20180621-3.1) …
Selecting previously unselected package libapache-pom-java.
Preparing to unpack …/16-libapache-pom-java_18-1_all.deb …
Unpacking libapache-pom-java (18-1) …
Selecting previously unselected package libcmark-gfm0.29.0.gfm.3:amd64.
Preparing to unpack …/17-libcmark-gfm0.29.0.gfm.3_0.29.0.gfm.3-3_amd64.deb …
Unpacking libcmark-gfm0.29.0.gfm.3:amd64 (0.29.0.gfm.3-3) …
Selecting previously unselected package libcmark-gfm-
extensions0.29.0.gfm.3:amd64.
Preparing to unpack …/18-libcmark-gfm-
extensions0.29.0.gfm.3_0.29.0.gfm.3-3_amd64.deb …
Unpacking libcmark-gfm-extensions0.29.0.gfm.3:amd64 (0.29.0.gfm.3-3) …
```

```
Selecting previously unselected package libcommons-parent-java.
Preparing to unpack …/19-libcommons-parent-java_43-1_all.deb …
Unpacking libcommons-parent-java (43-1) …
Selecting previously unselected package libcommons-logging-java.
Preparing to unpack …/20-libcommons-logging-java_1.2-2_all.deb …
Unpacking libcommons-logging-java (1.2-2) …
Selecting previously unselected package libfontenc1:amd64.
Preparing to unpack …/21-libfontenc1_1%3a1.1.4-1build3_amd64.deb …
Unpacking libfontenc1:amd64 (1:1.1.4-1build3) …
Selecting previously unselected package libptexenc1:amd64.
Preparing to unpack …/22-libptexenc1_2021.20210626.59705-1ubuntu0.2_amd64.deb
…
Unpacking libptexenc1:amd64 (2021.20210626.59705-1ubuntu0.2) …
Selecting previously unselected package rubygems-integration.
Preparing to unpack …/23-rubygems-integration_1.18_all.deb …
Unpacking rubygems-integration (1.18) …
Selecting previously unselected package ruby3.0.
Preparing to unpack …/24-ruby3.0_3.0.2-7ubuntu2.7_amd64.deb …
Unpacking ruby3.0 (3.0.2-7ubuntu2.7) …
Selecting previously unselected package ruby-rubygems.
Preparing to unpack …/25-ruby-rubygems_3.3.5-2_all.deb …
Unpacking ruby-rubygems (3.3.5-2) …
Selecting previously unselected package ruby.
Preparing to unpack …/26-ruby_1%3a3.0~exp1_amd64.deb …
Unpacking ruby (1:3.0~exp1) …
Selecting previously unselected package rake.
Preparing to unpack …/27-rake_13.0.6-2_all.deb …
Unpacking rake (13.0.6-2) …
Selecting previously unselected package ruby-net-telnet.
Preparing to unpack …/28-ruby-net-telnet_0.1.1-2_all.deb …
Unpacking ruby-net-telnet (0.1.1-2) …
Selecting previously unselected package ruby-webrick.
Preparing to unpack …/29-ruby-webrick_1.7.0-3_all.deb …
Unpacking ruby-webrick (1.7.0-3) …
Selecting previously unselected package ruby-xmlrpc.
Preparing to unpack …/30-ruby-xmlrpc_0.3.2-1ubuntu0.1_all.deb …
Unpacking ruby-xmlrpc (0.3.2-1ubuntu0.1) …
Selecting previously unselected package libruby3.0:amd64.
Preparing to unpack …/31-libruby3.0_3.0.2-7ubuntu2.7_amd64.deb …
Unpacking libruby3.0:amd64 (3.0.2-7ubuntu2.7) …
Selecting previously unselected package libsynctex2:amd64.
Preparing to unpack …/32-libsynctex2_2021.20210626.59705-1ubuntu0.2_amd64.deb
…
Unpacking libsynctex2:amd64 (2021.20210626.59705-1ubuntu0.2) …
Selecting previously unselected package libteckit0:amd64.
Preparing to unpack …/33-libteckit0_2.5.11+ds1-1_amd64.deb …
Unpacking libteckit0:amd64 (2.5.11+ds1-1) …
Selecting previously unselected package libtexlua53:amd64.
```

```
Preparing to unpack …/34-libtexlua53_2021.20210626.59705-1ubuntu0.2_amd64.deb
…
Unpacking libtexlua53:amd64 (2021.20210626.59705-1ubuntu0.2) …
Selecting previously unselected package libtexluajit2:amd64.
Preparing to unpack
…/35-libtexluajit2_2021.20210626.59705-1ubuntu0.2_amd64.deb …
Unpacking libtexluajit2:amd64 (2021.20210626.59705-1ubuntu0.2) …
Selecting previously unselected package libzzip-0-13:amd64.
Preparing to unpack …/36-libzzip-0-13_0.13.72+dfsg.1-1.1_amd64.deb …
Unpacking libzzip-0-13:amd64 (0.13.72+dfsg.1-1.1) …
Selecting previously unselected package xfonts-encodings.
Preparing to unpack …/37-xfonts-encodings_1%3a1.0.5-0ubuntu2_all.deb …
Unpacking xfonts-encodings (1:1.0.5-0ubuntu2) …
Selecting previously unselected package xfonts-utils.
Preparing to unpack …/38-xfonts-utils_1%3a7.7+6build2_amd64.deb …
Unpacking xfonts-utils (1:7.7+6build2) …
Selecting previously unselected package lmodern.
Preparing to unpack …/39-lmodern_2.004.5-6.1_all.deb …
Unpacking lmodern (2.004.5-6.1) …
Selecting previously unselected package pandoc-data.
Preparing to unpack …/40-pandoc-data_2.9.2.1-3ubuntu2_all.deb …
Unpacking pandoc-data (2.9.2.1-3ubuntu2) …
Selecting previously unselected package pandoc.
Preparing to unpack …/41-pandoc_2.9.2.1-3ubuntu2_amd64.deb …
Unpacking pandoc (2.9.2.1-3ubuntu2) …
Selecting previously unselected package preview-latex-style.
Preparing to unpack …/42-preview-latex-style_12.2-1ubuntu1_all.deb …
Unpacking preview-latex-style (12.2-1ubuntu1) …
Selecting previously unselected package t1utils.
Preparing to unpack …/43-t1utils_1.41-4build2_amd64.deb …
Unpacking t1utils (1.41-4build2) …
Selecting previously unselected package teckit.
Preparing to unpack …/44-teckit_2.5.11+ds1-1_amd64.deb …
Unpacking teckit (2.5.11+ds1-1) …
Selecting previously unselected package tex-gyre.
Preparing to unpack …/45-tex-gyre_20180621-3.1_all.deb …
Unpacking tex-gyre (20180621-3.1) …
Selecting previously unselected package texlive-binaries.
Preparing to unpack …/46-texlive-
binaries_2021.20210626.59705-1ubuntu0.2_amd64.deb …
Unpacking texlive-binaries (2021.20210626.59705-1ubuntu0.2) …
Selecting previously unselected package texlive-base.
Preparing to unpack …/47-texlive-base_2021.20220204-1_all.deb …
Unpacking texlive-base (2021.20220204-1) …
Selecting previously unselected package texlive-fonts-recommended.
Preparing to unpack …/48-texlive-fonts-recommended_2021.20220204-1_all.deb …
Unpacking texlive-fonts-recommended (2021.20220204-1) …
Selecting previously unselected package texlive-latex-base.
```

```
Preparing to unpack …/49-texlive-latex-base_2021.20220204-1_all.deb …
Unpacking texlive-latex-base (2021.20220204-1) …
Selecting previously unselected package libfontbox-java.
Preparing to unpack …/50-libfontbox-java_1%3a1.8.16-2_all.deb …
Unpacking libfontbox-java (1:1.8.16-2) …
Selecting previously unselected package libpdfbox-java.
Preparing to unpack …/51-libpdfbox-java_1%3a1.8.16-2_all.deb …
Unpacking libpdfbox-java (1:1.8.16-2) …
Selecting previously unselected package texlive-latex-recommended.
Preparing to unpack …/52-texlive-latex-recommended_2021.20220204-1_all.deb …
Unpacking texlive-latex-recommended (2021.20220204-1) …
Selecting previously unselected package texlive-pictures.
Preparing to unpack …/53-texlive-pictures_2021.20220204-1_all.deb …
Unpacking texlive-pictures (2021.20220204-1) …
Selecting previously unselected package texlive-latex-extra.
Preparing to unpack …/54-texlive-latex-extra_2021.20220204-1_all.deb …
Unpacking texlive-latex-extra (2021.20220204-1) …
Selecting previously unselected package texlive-plain-generic.
Preparing to unpack …/55-texlive-plain-generic_2021.20220204-1_all.deb …
Unpacking texlive-plain-generic (2021.20220204-1) …
Selecting previously unselected package tipa.
Preparing to unpack …/56-tipa_2%3a1.3-21_all.deb …
Unpacking tipa (2:1.3-21) …
Selecting previously unselected package texlive-xetex.
Preparing to unpack …/57-texlive-xetex_2021.20220204-1_all.deb …
Unpacking texlive-xetex (2021.20220204-1) …
Setting up fonts-lato (2.0-2.1) …
Setting up fonts-noto-mono (20201225-1build1) …
Setting up libwoff1:amd64 (1.0.2-1build4) …
Setting up libtexlua53:amd64 (2021.20210626.59705-1ubuntu0.2) …
Setting up libijs-0.35:amd64 (0.35-15build2) …
Setting up libtexluajit2:amd64 (2021.20210626.59705-1ubuntu0.2) …
Setting up libfontbox-java (1:1.8.16-2) …
Setting up rubygems-integration (1.18) …
Setting up libzzip-0-13:amd64 (0.13.72+dfsg.1-1.1) …
Setting up fonts-urw-base35 (20200910-1) …
Setting up poppler-data (0.4.11-1) …
Setting up tex-common (6.17) …
update-language: texlive-base not installed and configured, doing nothing!
Setting up libfontenc1:amd64 (1:1.1.4-1build3) …
Setting up libjbig2dec0:amd64 (0.19-3build2) …
Setting up libteckit0:amd64 (2.5.11+ds1-1) …
Setting up libapache-pom-java (18-1) …
Setting up ruby-net-telnet (0.1.1-2) …
Setting up xfonts-encodings (1:1.0.5-0ubuntu2) …
Setting up t1utils (1.41-4build2) …
Setting up libidn12:amd64 (1.38-4ubuntu1) …
Setting up fonts-texgyre (20180621-3.1) …
```

```
Setting up libkpathsea6:amd64 (2021.20210626.59705-1ubuntu0.2) …
Setting up ruby-webrick (1.7.0-3) …
Setting up libcmark-gfm0.29.0.gfm.3:amd64 (0.29.0.gfm.3-3) …
Setting up fonts-lmodern (2.004.5-6.1) …
Setting up libcmark-gfm-extensions0.29.0.gfm.3:amd64 (0.29.0.gfm.3-3) …
Setting up fonts-droid-fallback (1:6.0.1r16-1.1build1) …
Setting up pandoc-data (2.9.2.1-3ubuntu2) …
Setting up ruby-xmlrpc (0.3.2-1ubuntu0.1) …
Setting up libsynctex2:amd64 (2021.20210626.59705-1ubuntu0.2) …
Setting up libgs9-common (9.55.0~dfsg1-0ubuntu5.7) …
Setting up teckit (2.5.11+ds1-1) …
Setting up libpdfbox-java (1:1.8.16-2) …
Setting up libgs9:amd64 (9.55.0~dfsg1-0ubuntu5.7) …
Setting up preview-latex-style (12.2-1ubuntu1) …
Setting up libcommons-parent-java (43-1) …
Setting up dvisvgm (2.13.1-1) …
Setting up libcommons-logging-java (1.2-2) …
Setting up xfonts-utils (1:7.7+6build2) …
Setting up libptexenc1:amd64 (2021.20210626.59705-1ubuntu0.2) …
Setting up pandoc (2.9.2.1-3ubuntu2) …
Setting up texlive-binaries (2021.20210626.59705-1ubuntu0.2) …
update-alternatives: using /usr/bin/xdvi-xaw to provide /usr/bin/xdvi.bin
(xdvi.bin) in auto mode
update-alternatives: using /usr/bin/bibtex.original to provide /usr/bin/bibtex
(bibtex) in auto mode
Setting up lmodern (2.004.5-6.1) …
Setting up texlive-base (2021.20220204-1) …
/usr/bin/ucfr
/usr/bin/ucfr
/usr/bin/ucfr
/usr/bin/ucfr
mktexlsr: Updating /var/lib/texmf/ls-R-TEXLIVEDIST…
mktexlsr: Updating /var/lib/texmf/ls-R-TEXMFMAIN…
mktexlsr: Updating /var/lib/texmf/ls-R…
mktexlsr: Done.
tl-paper: setting paper size for dvips to a4:
/var/lib/texmf/dvips/config/config-paper.ps
tl-paper: setting paper size for dvipdfmx to a4:
/var/lib/texmf/dvipdfmx/dvipdfmx-paper.cfg
tl-paper: setting paper size for xdvi to a4: /var/lib/texmf/xdvi/XDvi-paper
tl-paper: setting paper size for pdftex to a4: /var/lib/texmf/tex/generic/tex-
ini-files/pdftexconfig.tex
Setting up tex-gyre (20180621-3.1) …
Setting up texlive-plain-generic (2021.20220204-1) …
Setting up texlive-latex-base (2021.20220204-1) …
Setting up texlive-latex-recommended (2021.20220204-1) …
Setting up texlive-pictures (2021.20220204-1) …
Setting up texlive-fonts-recommended (2021.20220204-1) …
```

```
Setting up tipa (2:1.3-21) …
Setting up texlive-latex-extra (2021.20220204-1) …
Setting up texlive-xetex (2021.20220204-1) …
Setting up rake (13.0.6-2) …
Setting up libruby3.0:amd64 (3.0.2-7ubuntu2.7) …
Setting up ruby3.0 (3.0.2-7ubuntu2.7) …
Setting up ruby (1:3.0~exp1) …
Setting up ruby-rubygems (3.3.5-2) …
Processing triggers for man-db (2.10.2-1) …
Processing triggers for fontconfig (2.13.1-4.2ubuntu5) …
Processing triggers for libc-bin (2.35-0ubuntu3.4) …
/sbin/ldconfig.real: /usr/local/lib/libtbbmalloc_proxy.so.2 is not a symbolic
link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_5.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbb.so.12 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_0.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbmalloc.so.2 is not a symbolic link

Processing triggers for tex-common (6.17) …
Running updmap-sys. This may take some time… done.
Running mktexlsr /var/lib/texmf … done.
Building format(s) --all.
        This may take some time…
```