Dependencies and Versions Required to Run the Code:
To compile and run the provided classification pipeline, the following libraries and packages must be installed:

Python version:
- Python 3.7+

Required libraries:
- pandas
- Path
- numpy
- matplotlib
- re
- math
- scikit-learn
  - sklearn.feature_extraction.text.TfidfVectorizer
  - sklearn.model_selection.train_test_split, GridSearchCV
  - sklearn.metrics (accuracy, precision, recall, F1, ROC, AUC)
  - sklearn.ensemble.RandomForestClassifier, VotingClassifier
  - sklearn.linear_model.LogisticRegression
  - sklearn.naive_bayes.GaussianNB

- imbalanced-learn (for SMOTE)
  - imblearn.over_sampling.SMOTE

- nltk (for stopword removal)
  - Requires downloading stopwords:

```python
import nltk
nltk.download('stopwords')
```