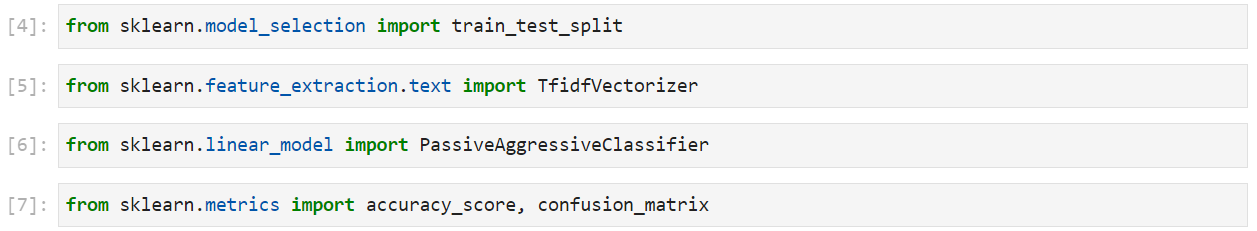
### Steps for detecting fake news with Python

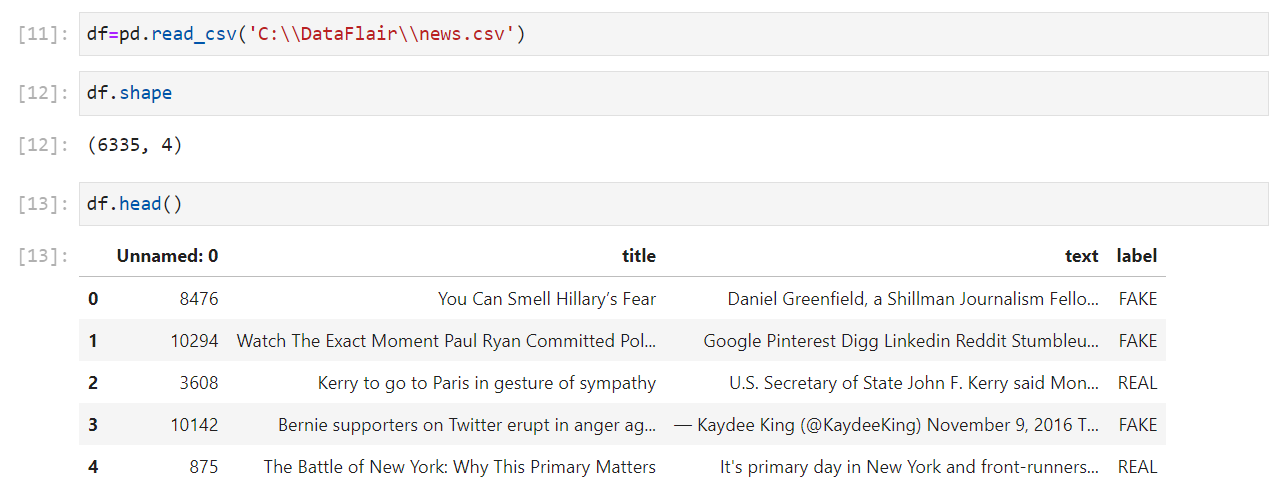
### Make necessary imports



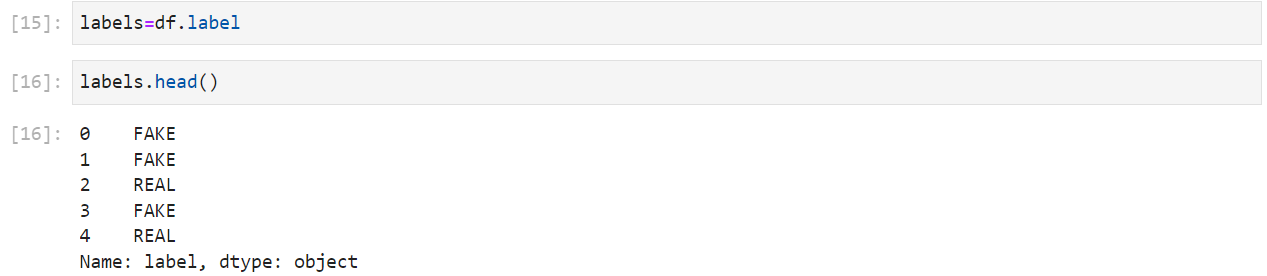




1. Read the input data into a DataFrame, and get the shape of the data and the first 5 records.



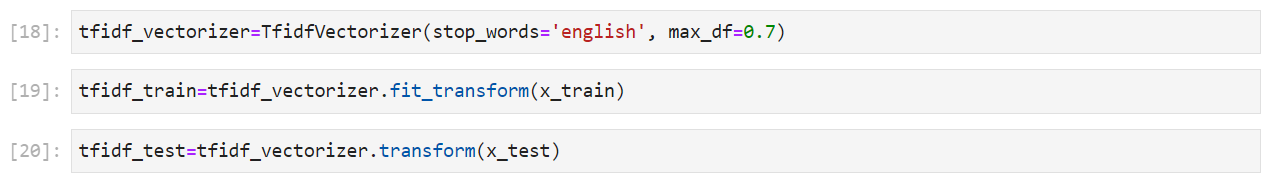
1. Get the labels from DataFrame



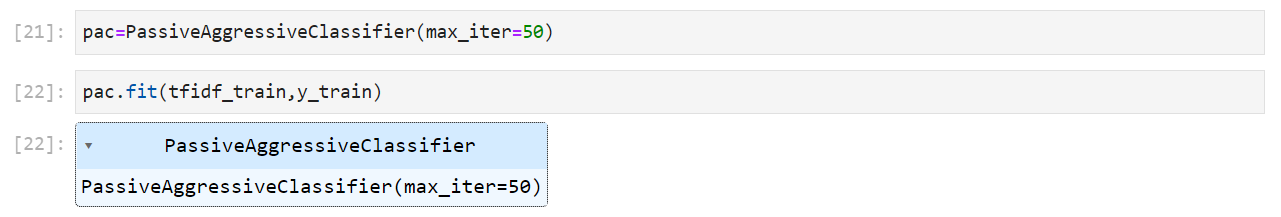
1. Split the data into training and testing sets



1. Initialize a TfidfVectorizer with stop words from the English language and a maximum document frequency of 0.7 (terms with a higher document frequency will be discarded). And a TfidfVectorizer turns a collection of raw documents into a matrix of TF-IDF features.

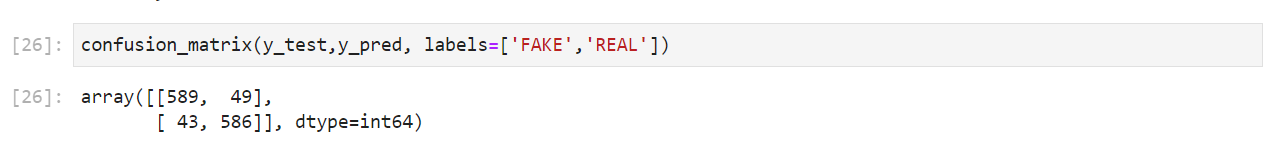


1. Initialize a PassiveAggressiveClassifier and fit this on tfidf\_train and y\_train.





1. We got an accuracy of 92.74% with this model. Finally, print out a confusion matrix to gain insight into the number of false and true negatives and positives.



So with this model, we have 589 true positives, 586 true negatives, 43 false positives, and 49 false negatives.