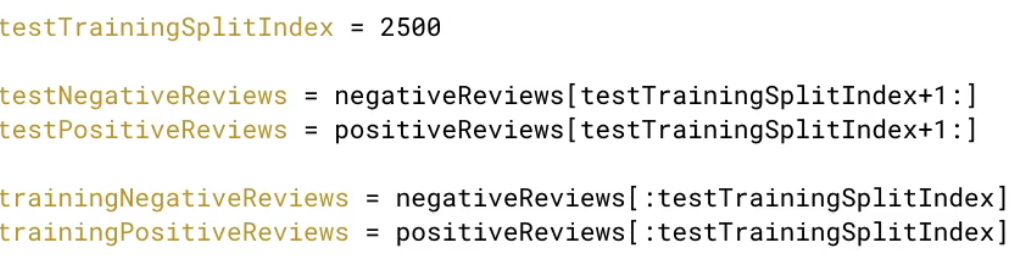
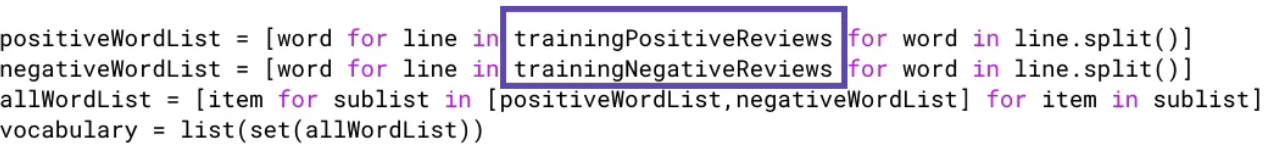
Implementing Sentiment Analysis with an ML Based Approach

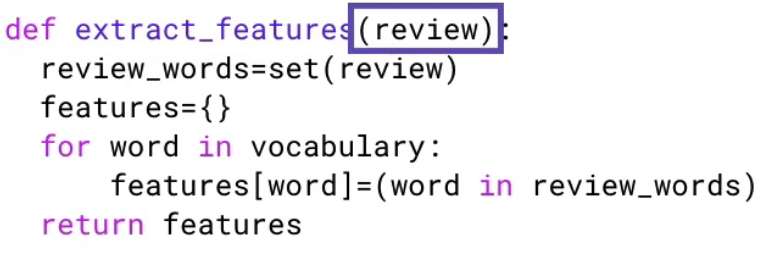
\*\*Have to start with a set of known positive and negative comments\*\*

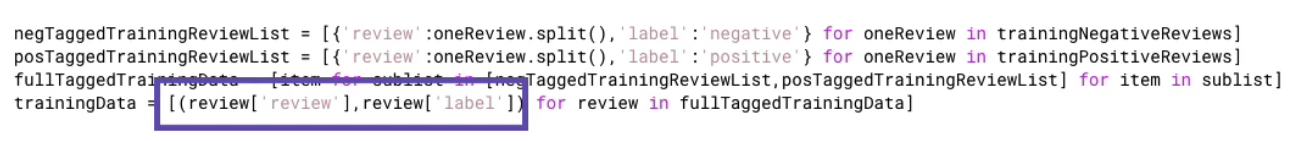
1. Split corpus into test and training data
   1. Split positive and negative reviews into test and training sections



1. Define vocabulary – use training data only
   1. Each word in training data occurs exactly once



1. Extract features – create word tuples
   1. Express each review as a tuple of 1,0 elements
   2. First element is review, second element is the label



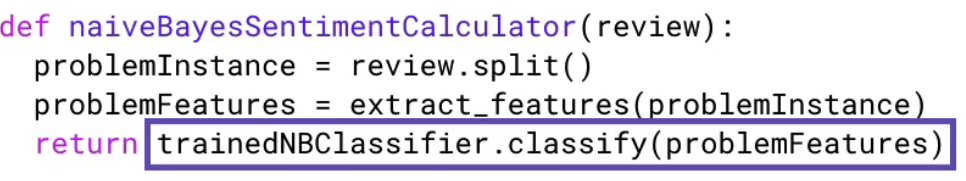
1. Train classifier – done by nltk
   1. Feature Extraction – input = training data and function, output = correct feature vector form



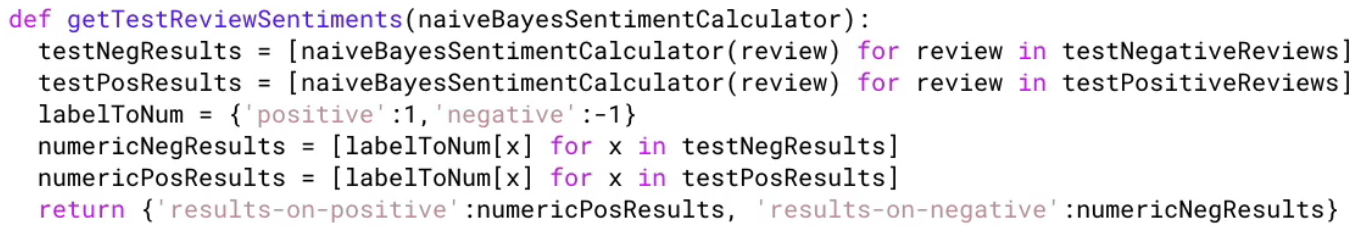


Output is a ready to use classifier

1. Classify test data – done by nltk
   1. Use only test data
   2. Input a feature vector, output is an assigned label



c.) Invoke using a test harness



1. Measure accuracy on test data

