Lesson 27 (NaiveBayes20Newsgroups) Train a Naive Bayes classifier to predict the newsgroup of a news article.

- Use the `20news-bydate_py3.pkz` file containing news articles from 20 different news groups.
- Use tf-idf vectorization to extract features (attributes) from the news articles.
- Complete an explicit grid search of the range of n-grams to use in tf-idf vectorization by filling out the table below. (You may need to use Gauss.)
- Use grid search cross-validation to optimize the pseudo-count hyper-parameter (alpha) of the Naive Bayes classifier and compute the validation error.
- Use the optimal value of alpha and all of the data to compute the training error.
- Note: The number of attributes equals the size of the vocabulary.
- Note: Sklearn's MultinomialNB classifier normally works with counts, i.e. attribute values are suppose to be integers, but it can work with fractional counts. tf-idf can be interpreted by MultinomialNB as a fractional count.

Naive Bayes Classifier Using TF-IDF Vectorization:

n-grams	# attributes	alpha	validation error	training error	baseline error
(1,1)					
(1,2)					
(1,3)					