Naive Bayes text classification

Task

In text classification, the goal is to find the best class for the document.
Multinomial Naive Bayes or multinomial NB model, a probabilistic learning method is a supervised learning method.

Please, use the pdf file for more detailed instructions. For worked example look at 44 slide.
Use Naive Bayes model with TF/IDF algorithm to solve text classification problem. Be free to choose any task for applying this algorithm.
The datasets can be taken from broad set of links e.g., kaggle, etc.
For testing purposes you can use email spam filter dataset from here.

Validation

This section designed for testing / validating your model.

```
# Train data
data = [["Chinese Beijing Chinese","0"],
        ["Chinese Chinese Shanghai","0"],
        ["Chinese Macao","0"],
        ["Tokyo Japan Chinese","1"]]

# Create instance of NaiveBayes class
nb = NaiveBayes()

# Train our model
# Tips: inside fit method it would be nice to split input data into train / test (80/20) sets and return model' accuracy, e.g.:
Accuracy = nb.fit(data)  # return accuracy

# Try to predict class of text
nb.predict(["Chinese Chinese Chinese Tokyo Japan"])

# Must return[ ('Chinese Chinese Chinese Tokyo Japan', '0')]
# pobability {'1': 0.00013548070246744226, '0': 0.00030121377997263036}
# or log     {'1': -7.906681345001262, '0': -7.10769031284391}
```

Report

This task, as well as all tasks in this course, must include a report - document with main key points of what you do in scope of this task. The document' structure described in the previous task "Face recognition with ORL"