

#### ML TextBooks:

- <https://github.com/josephmisiti/awesome-machine-learning>

#### Word Databases:

- Stop words
  - [https://code.google.com/p/pmtkdata/source/browse/trunk/20news\\_w1000/stopwords.txt?r=86](https://code.google.com/p/pmtkdata/source/browse/trunk/20news_w1000/stopwords.txt?r=86)
  - <http://www.ranks.nl/stopwords>
  - <http://xpo6.com/list-of-english-stop-words/>
  - <http://dev.mysql.com/doc/refman/5.1/en/fulltext-stopwords.html>
  - <https://code.google.com/p/stop-words/>
- Word Relationships
  - WordNet:
    - <http://wordnet.princeton.edu/wordnet/>
  - <http://www.phon.ucl.ac.uk/home/dick/tta/wf/wf.htm>

#### Algorithm Help:

- Naive Bayes:
  - <http://openclassroom.stanford.edu/MainFolder/DocumentPage.php?course=MachineLearning&doc=exercises/ex6/ex6.html>
- Forest for classification:
  - [http://research.microsoft.com/pubs/155552/decisionForests\\_MSR\\_TR\\_2011\\_114.pdf](http://research.microsoft.com/pubs/155552/decisionForests_MSR_TR_2011_114.pdf)
- Random Forests:
  - [http://www.whrc.org/education/indonesia/pdf/DecisionTrees\\_RandomForest\\_v2.pdf](http://www.whrc.org/education/indonesia/pdf/DecisionTrees_RandomForest_v2.pdf)
  - [http://nymetro.chapter.informs.org/prac\\_cor\\_pubs/RandomForest\\_SteinbergD.pdf](http://nymetro.chapter.informs.org/prac_cor_pubs/RandomForest_SteinbergD.pdf)
  - [https://www.stat.berkeley.edu/~breiman/RandomForests/cc\\_home.htm](https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm)
  - <https://www.stat.berkeley.edu/~breiman/RandomForests/>
  - <https://www.kaggle.com/wiki/RandomForests>
  - <http://citizennet.com/blog/2012/11/10/random-forests-ensembles-and-performance-metrics/>

#### Feature Choosing:

- <http://blog.christianperone.com/?p=1589>

#### Text Parsing:

- Matlab help:
  - <http://www.mathworks.com/help/matlab/string-parsing.html>
  - <http://www.mathworks.com/matlabcentral/answers/47477-parsing-formatted-text-file-quickly>
- <http://stackoverflow.com/questions/15750351/parse-text-file-in-matlab>
- <http://stackoverflow.com/questions/12352737/parsing-text-file-in-matlab?rq=1>

- <http://stackoverflow.com/questions/13666276/matlab-parsing-text-from-a-file-and-converting-to-a-mathematical-set>
- <http://www.mathworks.com/matlabcentral/answers/157855-how-to-extract-text-from-pdf-and-convert-to-excel>

#### Code For Text Analysis:

- [http://en.pudn.com/downloads251/sourcecode/graph/text\\_recognize/detail1164723\\_en.html](http://en.pudn.com/downloads251/sourcecode/graph/text_recognize/detail1164723_en.html)

#### Getting Text From PDFs:

- <http://www.wikihow.com/Convert-Images-and-PDF-Files-to-Editable-Text>
- <http://www.free-ocr.com/>
- <http://www.ocrconvert.com/>
- [http://www.downscripts.com/extract-text-from-a-pdf-document\\_matlab-script.html](http://www.downscripts.com/extract-text-from-a-pdf-document_matlab-script.html)

#### Matlab Libraries:

- Machine Learning:
  - <http://www.mathworks.com/matlabcentral/fileexchange/42744-machine-learning-with-matlab/content/Machine%20Learning/Classification/html/MachineLearning.html>
  - <http://www.mathworks.com/machine-learning/>
- Random Forests:
  - <https://code.google.com/p/randomforest-matlab/>
  - <https://github.com/karpathy/Random-Forest-Matlab>
  - <http://www.mathworks.com/help/stats/treebagger.html>