

De-bugging:

<https://docs.microsoft.com/en-us/visualstudio/debugger/debugging-absolute-beginners?view=vs-2019&tabs=csharp>

Python Cheat Sheets:

<https://hakin9.org/python-cheat-sheet-for-hackers-and-developers/>

<https://towardsdatascience.com/collecting-data-science-cheat-sheets-d2cdff092855>

Python Help:

[https://www.geeksforgeeks.org/\\_init\\_-in-python/](https://www.geeksforgeeks.org/_init_-in-python/)

[https://www.tutorialspoint.com/python/file\\_close.htm](https://www.tutorialspoint.com/python/file_close.htm)

<https://stackabuse.com/read-a-file-line-by-line-in-python/>

<https://packaging.python.org/en/latest/tutorials/installing-packages/>

<https://linuxize.com/post/python-get-change-current-working-directory/>

<https://www.tutorialsteacher.com/python/python-read-write-file>

Web scraping/Text mining:

<https://towardsdatascience.com/a-short-practical-how-to-guide-to-scrape-data-from-a-website-using-python-888373227d4f>

<http://librarycarpentry.org/lc-tdm/aio/index.html>

TF-IDF:

<https://towardsdatascience.com/tf-idf-for-document-ranking-from-scratch-in-python-on-real-world-dataset-796d339a4089>

<https://towardsdatascience.com/how-to-make-word-clouds-in-python-that-dont-suck-86518cdcb61f>

Text cleaning/Preprocessing:

<https://www.projectpro.io/recipes/add-custom-stopwords-and-then-remove-them-from-text>

<https://towardsdatascience.com/pdf-preprocessing-with-python-19829752af9f>

<https://www.guru99.com/tokenize-words-sentences-nltk.html>

<https://towardsdatascience.com/nlp-learning-series-part-1-text-preprocessing-methods-for-deep-learning-20085601684b>

Word Embeddings:

<https://neptune.ai/blog/word-embeddings-guide>

<https://www.shanelynn.ie/word-embeddings-in-python-with-spacy-and-gensim/>

<https://www.kaggle.com/datasets/danielwillgeorge/glove6b100dtxt>

<https://radimrehurek.com/gensim/models/word2vec.html>

Topic Modeling and LDA:

<https://towardsdatascience.com/topic-modeling-and-latent-dirichlet-allocation-in-python-9bf156893c24>

<https://www.tidyttextmining.com/topicmodeling.html>

<https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0>

<https://towardsdatascience.com/hierarchical-topic-modeling-with-bigartm-library-6f2ff730689f>

Part of Speech:

<https://spacy.io/usage/linguistic-features>

Predictive Modeling:

<http://appliedpredictivemodeling.com/>

<https://neptune.ai/blog/the-ultimate-guide-to-evaluation-and-selection-of-models-in-machine-learning>

<https://towardsdatascience.com/data-normalization-in-machine-learning-395fdec69d02>

[https://www.dropbox.com/s/xa1oxdqoya7hwkm/predictionError\\_toshare.pdf?dl=0](https://www.dropbox.com/s/xa1oxdqoya7hwkm/predictionError_toshare.pdf?dl=0)

LSTM:

<https://rubikscore.net/2018/03/26/two-ways-to-implement-lstm-network-using-python-with-tensorflow-and-keras/>

<https://analyticsindiamag.com/how-to-implement-lstm-rnn-network-for-sentiment-analysis/>

Text Classification:

[https://www.tensorflow.org/text/tutorials/text\\_classification\\_rnn](https://www.tensorflow.org/text/tutorials/text_classification_rnn)

[https://www.tensorflow.org/text/tutorials/classify\\_text\\_with\\_bert](https://www.tensorflow.org/text/tutorials/classify_text_with_bert)

<https://towardsdatascience.com/machine-learning-nlp-text-classification-using-scikit-learn-python-and-nltk-c52b92a7c73a>

<https://jesusleal.io/2020/10/20/RoBERTA-Text-Classification/>

[https://pytorch.org/text/main/tutorials/sst2\\_classification\\_non\\_distributed.html](https://pytorch.org/text/main/tutorials/sst2_classification_non_distributed.html)

<https://towardsdatascience.com/fine-tuning-bert-and-roberta-for-high-accuracy-text-classification-in-pytorch-c9e63cf64646>

Sentiment analysis:

[https://www.nltk.org/\\_modules/nltk/sentiment/vader.html](https://www.nltk.org/_modules/nltk/sentiment/vader.html)

<https://towardsdatascience.com/a-beginners-guide-to-sentiment-analysis-in-python-95e354ea84f6>

Train, Validation, Test:

[https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.train\\_test\\_split.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html)

<https://towardsdatascience.com/how-to-split-data-into-three-sets-train-validation-and-test-and-why-e50d22d3e54c>

Cross-validation:

[https://scikit-learn.org/stable/modules/cross\\_validation.html](https://scikit-learn.org/stable/modules/cross_validation.html)

Regularization:

<https://towardsdatascience.com/ridge-and-lasso-regression-a-complete-guide-with-python-scikit-learn-e20e34bcbf0b>

<https://towardsdatascience.com/penalized-regression-with-classification-f60bc663e8dc>

Grid Search:

[https://scikit-learn.org/stable/modules/grid\\_search.html](https://scikit-learn.org/stable/modules/grid_search.html)

Model evaluation:

<https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc>

Platforms:

<https://medium.com/deep-learning-turkey/google-colab-free-gpu-tutorial-e113627b9f5d>

<https://neptune.ai/blog/google-colab-dealing-with-files>

GitHub files:

[https://github.com/skdreier/NIrelandNLP/tree/master/model\\_training\\_and\\_evaluation](https://github.com/skdreier/NIrelandNLP/tree/master/model_training_and_evaluation)