LINK TO STANFORD NLP COURSE:
https://web.stanford.edu/class/cs224n/
-- This will provide you with almost everything you need to know, but is *very* dense and math-heavy, requires a knowledge of calculus. However, it's a good jumping off point, and if there's something you don't understand, you can google "easy introduction to _____" and normally find a blog post that will give you some intuition before diving back in.

Jurafsky and Martin book on Speech and Language Processing:
https://web.stanford.edu/~jurafsky/slp3/

N-gram language modeling:
https://web.stanford.edu/~jurafsky/slp3/3.pdf

Padding for NLP:
https://github.com/blester125/A2D-NLP-Talk-Feb-27-2020

RNNs for NLP:
http://karpathy.github.io/2015/05/21/rnn-effectiveness/

CNNs for NLP:
http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp

Intro to Contextual Embeddings (BERT, ELMo):
http://jalammar.github.io/illustrated-bert/

t-SNE visualization: https://lvdmaaten.github.io/tsne/

Information Entropy:
https://machinelearningmastery.com/what-is-information-entropy/

Named Entity Recognition:
https://www.kdnuggets.com/2018/12/introduction-named-entity-recognition.html

Part of Speech tagging:
https://www.cs.umd.edu/~nau/cmsc421/part-of-speech-tagging.pdf

Dependency parsing:
https://web.stanford.edu/~jurafsky/slp3/15.pdf

Sentence embeddings:
https://medium.com/huggingface/universal-word-sentence-embeddings-ce48ddc8fc3a

Optimizers:

https://towardsdatascience.com/types-of-optimization-algorithms-used-in-neural-networks-and-ways-to-optimize-gradient-95ae5d39529f

Dropout Regularization:
https://machinelearningmastery.com/dropout-for-regularizing-deep-neural-networks/

Transformers:
http://jalammar.github.io/illustrated-transformer/
https://www.machinecurve.com/index.php/2020/12/28/introduction-to-transformers-in-machine-learning

Sequence-to-Sequence Models:
https://nlp.stanford.edu/~johnhew/public/14-seq2seq.pdf
https://machinelearningmastery.com/encoder-decoder-recurrent-neural-network-models-neural-machine-translation/

Attention:
https://nlp.seas.harvard.edu/2018/04/03/attention.html
http://mlexplained.com/2017/12/29/attention-is-all-you-need-explained/

ELMo:
https://medium.com/saarthi-ai/elmo-for-contextual-word-embedding-for-text-classification-24c9693b0045