Project Proposal

Objective: To build and train a neural network with dictionary data to create a homophone generation tool that is able to deconstruct an unknown word into proper IPA, and then generate unique homophones rel to the input value.

The English IPA Lookup Tool is a standard for finding the phonetic pronunciation of many common words. This tool searches a dictionary to find the proper phonetic pronunciation. When it comes to more uncommon words, the tool is often unable to find the correct phonetic pronunciation.

Names are often derived from other languages, and can be anglicized into an English pronunciation. This can cause many possible name spellings to be used. This can complicate things when looking at the real-world application of Facebook’s search engine tool. Say in a scenario two people meet. They both know the phonetic pronunciation of the name they search for, but are unsure of the spelling. A homophone generation tool could suggest alternative name spellings and improving search results when many homophones can exist in a search.

Research will be done to determine the type of neural network structure that is to be used to give the best results when suggesting homophones. The type of network will need to be able to take a word, generate the IPA based on the learned dataset, and then generate another similar word with a different spelling. Using bi-grams will allow the dataset to contain relationships between two characters and the resulting phonetic characters.

**Links**

Just out of interest as to how Facebook search is conducted:

<https://www.facebook.com/notes/facebook-engineering/under-the-hood-indexing-and-ranking-in-graph-search/10151361720763920>

Definition of Phoneme:

<https://en.wikipedia.org/wiki/Phoneme>

Maybe not relevant? Framewise phoneme classification with bidirectional LSTM and other neural network architectures

<https://www.sciencedirect.com/science/article/pii/S0893608005001206#sec6>

Types of Networks structures.

<https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464>

Variational Auto Encoder network:

http://kvfrans.com/variational-autoencoders-explained/