Assignment 2: Oulipian language modelling

JADEN THOMPSON
CART 498 C - Generative AI
January 30th, 2025

A short reflection (250–350 words) explaining how altering the x value impacted the output of your P+x version. Additionally, discuss how you would implement a P+7 technique in which all nouns are replaced with their seventh-highest probability alternatives.

Altering the X-value of the P+X version of the code produced a wide variety of different results depending on how large of an integer was given. Exponentially, the wording produced by the generative text and tokenizer became more incomprehensible as the context shifted away from the actual text itself to something unrecognizable. At lower X-value conditions the words produced by the generative text still remained somewhat relevant to the story, words like cliff, middle, could, separated, and sun. While the sentences have become more jumbled due to these new words being randomly inserted into the end of each line, it still remains within the boundary of contextual understanding. At the early double digits of the X-value condition sentences began to become fragmented and much more distorted from the actual context given from the story itself, prepositional words became much more common at the end of the sentences. Eventually, raising the probability to the maximum would often produce completely nonsensical words, grammatical errors, and out of place punctuation.

Implementing a P+7 technique in which all nouns are replaced with their seventh-highest probability alternatives would require an accurate way to properly identify the nouns within the given text, selectively split those words from the rest of the sentence, then generate and tokenize a new word to replace it, and finally truncate the sentence back together to properly display the newly generated story.