**Requirements:**

Must create CSVs of data that list an author and associated simple stylometry variables across all text. This would be done programmatically as in working with the text as strings to parse it into different stylometric variables. For example, this would include the function word use and percentage use associated with an author and a group of text. Visualizations must be created from these stylometric variable data. For example, visualize the function word percentage relationship between two words and count for some authors (see below) Success measurement: visualization of data.

Must create CSVs of data that list an author and associated complex stylometry variables across all text. This would be done programmatically as in working with the text as strings to parse it into different stylometric variables. For example, this would include the N-grams, bigrams, trigrams use and percentage use associated with an author and a group of text. Visualizations must be created from these complex stylometric variable data. For example, plot the top frequency Bi-words and how much they are used for some authors Success measurement: visualization of data.

Must fit created stylometric text data to a machine learning model for The Book of Mormon as a whole and Joseph Smith’s other writings. This should include hyperparameter optimizations and trial and error with different train and test sizes for the data. Also include an explanation of metrics of performance for the model. Success measurement: Explanation of metrics and/or visualizations of model performance.

Must fit created stylometric text data to a XGBoost and Random Forest machine learning model for the inner voices and authors of The Book of Mormon. This should include hyperparameter optimizations and trial and error with different train and test sizes for the data. Also include an explanation of metrics of performance for the model. Success measurement: Explanation of metrics and/or visualizations of model performance.

Must fit created stylometric text data to a Neural Net machine learning model for the inner voices and authors of The Book of Mormon. This should include hyperparameter optimizations and trial and error with different train and test sizes for the data. Also include an explanation of metrics of performance for the model. Success measurement: Explanation of metrics and/or visualizations of model performance.

Must have a cohesive markdown file to present findings in and comprehensively covers my project: Success measurement: Markdown HTML file that covers my findings.

**Stretch Requirements:**

Should obtain different text that are similar in nature or around same time period of The Book of Mormon (Bible, View of The Hebrews, The Late War, Spauling manuscript, Writings of Joseph Smith and other that were important figures in early Mormonism) and extract stylometric variables from them, and prepare them for a machine learning model

Should fit text data to a Neural net machine learning model for The Book of Mormon and other text that are similar in nature or around same time period (Bible, View of The Hebrews, The Late War, Spauling manuscript, Writings of Joseph Smith and other that were important figures in early Mormonism)

Should fit the different voices of authorship in The Book of Mormon to a Recurrent Neural Net machine learning model for The Book of Mormon to attempt to produce writing in the voices of each individual author. Then use already trained models to predict authorship on that writing.

Should fit text data to a Neural net machine learning model of specific voices in The Book of Mormon (Jesus Christ, incarnate and pre-incarnate, Isaiah) to the same voice in the Bible, and other scriptures.