**Recommended Books for Year Levels**

Prepared by

Carlo Carbonilla

Date

25/09/2021

**Executive Summary**

* We’ll be going over the process of the data exploration, model training, evaluation, and results of data from CommonLit Readability.
* This will help aid teachers to signify the readability on texts and which year levels to recommend them for.

**Introduction**

* We’ll be using data analysis and machine learning to determine the readability score of certain texts, and based on that data, we’ll recommend books for certain year levels.
* Some limitations of this analysis are:
* The recommended books come from the top 100 books on project Gutenberg, and so books not in there weren’t included in this analysis

**Analysis**

* We trained an Multi Layer Perceptron Regressor Model on the CommonLit data provided by Kaggle which gives an excerpt and a readability score which is the result of a Bradley-Terry analysis of more than 111,000 pairwise comparisons between excerpts.
* We then got the top 100 books on project Gutenberg, extracted 10 paragraphs from them, and used our existing model to give them a readability score.
* Given that the teachers that helped with the CommonLit data teach from Grades 3-12, we sorted our books by score and evenly split them by 10 to signify the year levels they belong to.

**Conclusion**

* Based on the scores outputted by my machine learning models, here are some recommended books per year level
* Year 9 – Oliver Twist
* Year 10 – The Odyssey
* Year 11 – Moby Dick
* Year 12 – Les Misérables
* Year 13 – Beyond Good and Evil