COMP30027 Reflection
Assignment 2: Book Rating Prediction

James La Fontaine, 1079860

During the completion this project, I feel like I gained some valuable insight into the process of solving a machine learning problem. This project has allowed me to gain some understanding of the thought process that must be applied when building machine learning models and the importance of being familiar with different models that can be used given a particular problem. I have also come to realise the importance of understanding why models may or may not be performing well and subsequently knowing how to address any performance issues. This project also allowed me to gain further familiarity with utilising Python for the coding aspect of building machine learning models.

Overall, I was satisfied with:

- My ability to take advantage of the machine learning libraries provided by Python such as sklearn.
- The level of testing performed during feature selection, fine tuning of hyperparameters, and comparison and evaluation of models.
- The choice of models and performance of final models produced considering the data.
- The quality of the report in terms of structure and polish.

Things that I think could be improved:

- Experimenting with more complex models and techniques such as ensemble classifiers and XGBoost.
- Further usage of figures and metrics (e.g. PCA, learning curves) to support model performance explanations.
- More rigorous testing and investigation of explanations posed in error analysis with figures that support whether there were any improvements in model performance after some of the proposed adjustments were made.
- Could likely have streamlined the feature selection, hyperparameter tuning and cross validation process to complete the assignment faster and reduce time spent waiting on model training.
- Further experimentation with different feature extractions from the original textual features.
- Further investigation of the importance of different features in relation to book rating (i.e. model interpretation).
- Further attempts to counteract the class imbalance and produce models which better predict the minority classes.
- Further usage of references to draw inspiration for how to approach this problem and potentially better explain the results observed