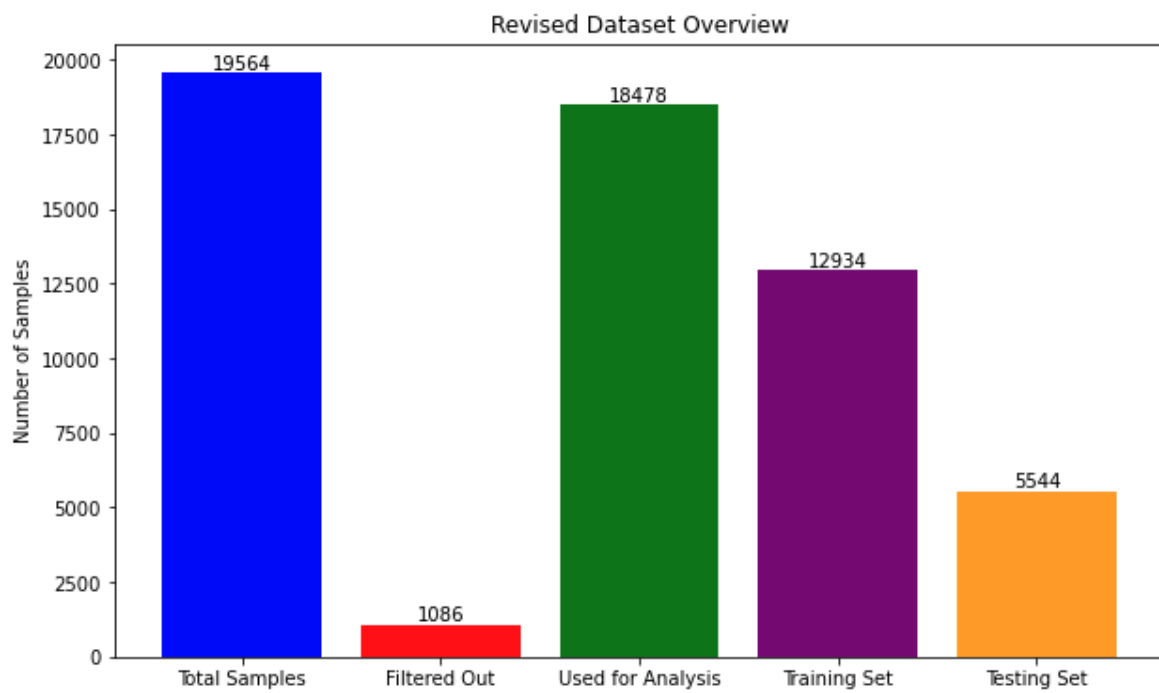Main Results

The primary objective that I was hoping to accomplish was to predict the employment status of post-secondary employment status using factors such as part-time or full-time status, GPA, volunteer activities and disability status. Using supervised machine learning and logistic regression I got the results of Accuracy being 52.11%, Precision being 51.01%, Recall being 52.11% and F1 score being 44.17%. Given these results it seems that the model is pretty much as good as randomly guessing the status of a person.

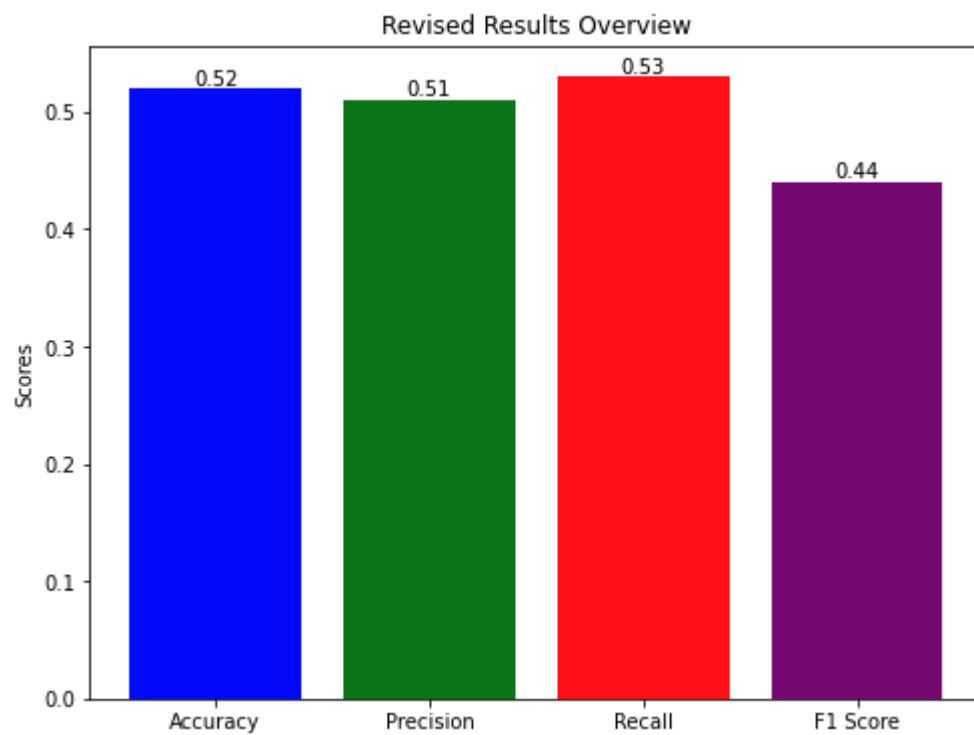Cross Validation and Hyperparameter study

I conducted a 5 fold cross to validate the stability and reliability of the model. By splitting the data into 5 subsets and using each one for testing while training the remaining. This process was repeated 5 times to ensure that every part of the dataset was tested on. I also conducted a hyperparameter study by varying the parameters of the model. The changes did not show any improvement in the model's performance.
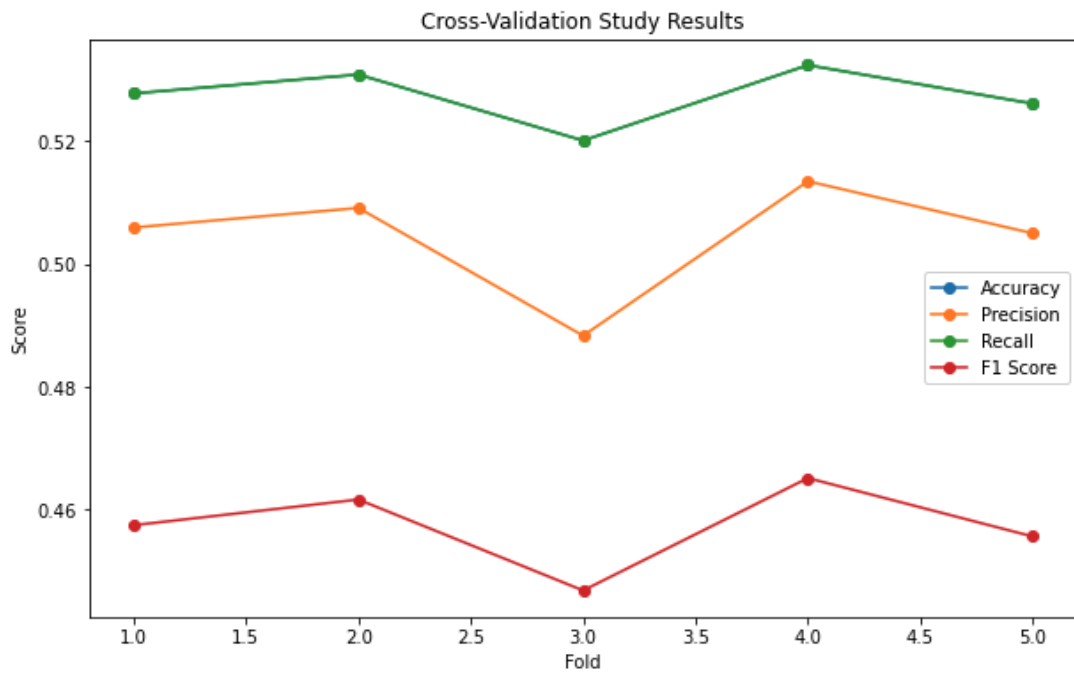
Dataset Figure:

Revised Dataset Overview

I filtered out any column that had not stated as an answer.

Revised Result Overview:

Revised Results Overview

Additional Figures:

Cross-Validation Study Results

Accuracy and recall are both giving the same values and are on the same line.