Model Performance

Linear Regression Dataset 1

Linear Regression Model (Input is Decade)	Linear Regression Model Metrics
Schizophrenia Percent	Mean Absolute Error 0.029 Mean Squared Error 0.0017
Depression Percent	Mean Absolute Error 0.78 Mean Squared Error 0.93
Anxiety Percent	Mean Absolute Error 0.77 Mean Squared Error 1.17
Bipolar Percent	Mean Absolute Error 0.21 Mean Squared Error 0.065
Eating Disorder Percent	Mean Absolute Error 0.10 Mean Squared Error 0.015

Linear Regression Dataset 1

Linear Regression Model (Input is Decade)	Linear Regression Model Metrics
Schizophrenia DALYs	Mean Absolute Error 21.22 Mean Squared Error 846.91
Depression DALYs	Mean Absolute Error 149.022 Mean Squared Error 35241.67
Anxiety DALYs	Mean Absolute Error 78.41 Mean Squared Error 11602.19
Bipolar DALYs	Mean Absolute Error 47.36 Mean Squared Error 3147.43
Eating Disorder DALYs	Mean Absolute Error 22.92 Mean Squared Error 760.18

Linear Regression Dataset 1

Linear Regression Model (Input is Decade)	Linear Regression Model Metrics
Schizophrenia AVG DALYs	Mean Absolute Error 20.22 Mean Squared Error 735.48
Depression AVG DALYs	Mean Absolute Error 156.73 Mean Squared Error 38054.16
Anxiety AVG DALYs	Mean Absolute Error 76.03 Mean Squared Error 11154.11
Bipolar AVG DALYs	Mean Absolute Error 46.78 Mean Squared Error 3101.95
Eating Disorder AVG DALYs	Mean Absolute Error 21.78 Mean Squared Error 717.24

Linear Regression Dataset 1

Linear Regression Model (Input is Decade)	Linear Regression Model Metrics
Schizophrenia AVG Percent	Mean Absolute Error 0.030 Mean Squared Error 0.0017
Depression AVG Percent	Mean Absolute Error 0.76 Mean Squared Error 0.90
Anxiety AVG Percent	Mean Absolute Error 0.79 Mean Squared Error 1.22
Bipolar AVG Percent	Mean Absolute Error 0.21 Mean Squared Error 0.066
Eating Disorder AVG Percent	Mean Absolute Error 0.10 Mean Squared Error 0.015

SVM, Decision Tree, Random Forest Dataset 2

SVM	SVM Metrics
Gender as prediction	Accuracy: 0.49 Precision: 0.49 f1score: 0.65 recallscore: 0.98
Treatment as prediction	Accuracy: 0.57 Precision: 0.57 f1score: 0.73 recallscore: 1.0

Decision Tree	
Gender as prediction	Accuracy: 0.50 Precision: 0.50 f1score: 0.58 recallscore: 0.6975
Treatment as prediction	Accuracy: 0.75 Precision: 0.78 f1score: 0.78 recallscore: 0.77

Random Forest	
Gender as prediction	Accuracy: 0.60 Precision: 0.56 f1score: 0.69 recallscore: 0.90
Treatment as prediction	Accuracy: 0.76 Precision: 0.76 f1score: 0.80 recallscore: 0.84

It should be noted these differ from Milestone 2. I found an error when going through these models for Milestone 3, and updated them accordingly.

SVM, Decision Tree, Random Forest for Dataset 3

SVM	SVM Metrics
Depression as prediction	Accuracy: 0.78 Precision: 0.83 f1score: 0.52 recallscore: 0.38
Gender as prediction	Accuracy: 0.78 Precision: 0.78 f1score: 0.87 recallscore: 1.0

Decision Tree	Decision Tree Metrics
Depression as prediction	Accuracy: 0.90 Precision: 0.90 f1score: 0.83 recallscore: 0.76
Gender as prediction	Accuracy: 0.70 Precision: 0.77 f1score: 0.82 recallscore: 0.87

Random Forest	
Depression as prediction	Accuracy: 0.82 Precision: 0.75 f1score: 0.71 recallscore: 0.69
Gender as prediction	Accuracy: 0.65 Precision: 0.75 f1score: 0.79 recallscore: 0.84

As I also mentioned in dataset 2, there was a bug in my code that I noticed going through it for milestone 3. I updated the metrics to match the true result.

Interpretation

When interpreting the models, for dataset 1, when predicting percent and average percent with linear regression, the models performed well with low mean squared error and low mean absolute error. However, when predicting DALYs and average DALYs, the linear regression models performed horribly with mean squared error and mean absolute error in the thousands.

When interpreting the models, for dataset 2, for both prediction outputs, SVM performed at or worse than random guessing, so not a model I would use in the future. For decision tree, with gender as the output, it performed about that of random guessing, making it a useless model. However, with treatment as an output, it performed decently well, at about .75 across the board. For random forest, again, gender as an output performed just a little better than random guessing, making it a poor model. Random forest performed better with treatment as an output, with output scores of .75-.8, which indicates a decent model.

When interpreting the models, for dataset 3, for both outputs, depression and gender, SVM performed decently well, with scores .75 - .85 range. Decision Tree for depression as an output performed the best of any of the models, with scores around .9, the best of any model, which indicates this is a model worth using with useful insights and outputs. For gender as an output, it performed decently well, with scores around .78, indicating a decent model. For random forest, with depression as an output, it performed well, with scores around .8. For gender as an output, the model performed poorly, with scores around .65.

Across the board, it was a toss up on how well the models performed. This could be due to the large spread of data (dataset 1) the test and train set (making a smaller train-test dataset for dataset 2) or the small total size of data (dataset 3). Though it is important to note there was success with several of the models, so there were good tools that were developed amid the other failures. As far as potential biases, there was no immediately obvious bias for dataset 1, a perfect and balanced split for dataset 2, so no biases there, and there definitely was potential bias in dataset 3 for gender, as the data was predominantly female, but there was a unbiased split according to depression, likely why it performed the best.