# 2.2 Build Models

For this section we considered accuracy to be our primary criterion, interpretability as our second criterion, and robustness as our third criterion. Therefore, we used the following 3 models: XGBoost, K-Nearest Neighbours, and Random Forests. We chose two ensemble methods: XGBoost and Random Forests, because they are robust to outliers and non-linear data, and they produce results with high classification accuracy. As for KNN, we chose it because we felt the results were very simple to interpret and it is robust regarding search space since classes do not have to be linearly separable like in SVM.

# 2.3 Evaluation

As we considered accuracy to be our primary criterion when choosing the models, we used prediction accuracy as the main metric to evaluate our models, especially due to the problem statement for this project being able to accurately predict outcomes of COVID patients. For XGBoost, K-Nearest Neighbours, and Random Forests, the accuracy of our model’s prediction values compared to the actual test data’s outcomes were 87.89%, 88.71%, and 89.19% respectively.