Connor Burbridge

2021-11-22 – November 22nd, 2021

Results and Discussion:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Method | True Neg | False Neg | True Pos | False Pos | Misclass. | Precision | Recall | F1 Score |
| LDA | 19414 | 2959 | 25 | 32 | 0.13334 | 0.4705 | 0.0085 | 0.0166 |
| QDA | 17475 | 2496 | 488 | 1971 | 0.19915 | 0.1919 | 0.1593 | 0.1740 |
| Elastic Net | 19435 | 2975 | 9 | 11 | 0.13312 | 0.4444 | 0.0021 | 0.0041 |
| Random Forest | 19446 | 2984 | 0 | 0 | 0.13303 | 0 | 0 | NA |
| XGBoost | 19269 | 140 | 101 | 2920 | 0.13642 | 0.0334 | 0.4190 | 0.0618 |

Table X: Descriptive metrics of models run on the test data partition.

Results indicate, that while our models show a reasonable level of accuracy through the misclassification rate metric, the other metrics tell a very different story on overall performance. The LDA model shows one of the lowest misclassification rates, but misclassifies false negatives very frequently while detecting very few true positives. Precision is the highest of all models but recall is quite poor for this model. QDA does the best job of identifying true positive cases but misclassifies more cases than any of the other models, resulting in low precision and recall scores. While this model does predict more true positives, a higher number of misclassifications can result in significant downstream costs associated with asthma and healthcare. The Elastic Net model performs the second best out of all models when it comes to misclassification rate, but predicts very few true positive cases and misclassifies many false negatives. Random forest, definitively the most accurate of the models does not classify a single true positive or false positive, making the model virtually useless. Classifying every data point as negative is a poor way to create an accurate model.

Overall, these models performed adequately well when looking at the performance metrics, however, it would be difficult to select one as the best. Three of the five models show relatively decent performance in precision, while the other two show better performance in recall. When considering asthma status, the difference between using recall and precision as deciding metrics is very important. A model with higher precision means fewer false positives, which in the case of asthma means fewer people prescribed inhalers when unnecessary, saving costs on inhalers for the patients. A model optimized for recall might end up sending true asthmatic patients home without an inhaler, causing suffering and further problems down the line. F1 scores also tell a similar story with QDA being the highest performing model in this metric. Similar to recall and precision, F1 scores and accuracy have trade-offs between them as well.