

## **APPENDIX 1 – FIVE MACHINE LEARNING ALGORITHMS USED TO TRAIN DATASET**

Leave-One-out Cross Validation with Random Forest:

We issued a command to run random sampling without replacement such that each observation in the old data was only selected once in the new data. Next, we created a loop to run the random forest model 255 times.

Leave-One-out Cross Validation with Logistic Regression:

We create a loop to run the logistic regression model 255 times.

Leave-One-out Cross Validation with K Nearest Neighbour Classification:

We tested several values of k (number of nearest neighbours) and obtained k=3, which yielded the highest level of sensitivity.

Leave-One-out Cross Validation with K Nearest Neighbour Regression:

We tested several values of k and arrived at k=3, which yielded the highest level of sensitivity.

Leave-One-out Cross Validation with Adaptive KNN:

We used a package developed at the Institute of Health Professions of Massachusetts General Hospital (Boston, MA, and USA).