1. Why did you choose the particular algorithm?
   The Random Forest Classifier should work well because it can handle outliers and multicollinearity. With lots of features, it is easy to overfit, but Random Forest can avoid that problem. Also, it's good at understanding complex relationships between the different features and the target.

2. What are the different tuning methods used for the algorithm?
   The hyperparameter tuning can be done by grid search or randomized search.
   I am fine-tuning the Random Forest Classifier by testing different settings for three main parameters: the number of trees (**n_estimators**), tree depth (**max_depth**), and minimum samples to split (min_samples_split). By experimenting with values like 100, 150, 200, and 250 for n_estimators, depths of 5, 7, 9, and 11 for max_depth, and sample sizes of 11, 13, 15, and 17 for **min_samples_split**, I aim to find the best combination for optimal classifier performance on our dataset.

3. Did you consider any other choice of algorithm? Why or why not?
   I also explored Support Vector Classifier (SVC) and decision trees due to their ability to handle multicollinearity and outliers. However, I encountered overfitting issues with these models. Consequently, I opted for Random Forest, which not only effectively addresses multicollinearity and outliers but also mitigates overfitting. As a result, Random Forest emerged as the good choice, delivering better performance on the dataset.

4. What is the accuracy?

Models	Accuracy
Random Forest Classifier	0.917930
Decision Tree Classifier	0.879247
Logistic Regression	0.601673
Support Vector Classifier	0.506012

5. What are the different types of metrics that can be used to evaluate the model?
   The different types of metrics that can be used to evaluate the model are F1-score (which is the harmonic mean of precision and recall) and balanced accuracy. These metrics will be better for evaluating the performance of the model since the dataset is imbalanced.

Models	Accuracy	Balanced Accuracy	F1 Score
Random Forest Classifier	0.917930	0.889242	0.917845
Decision Tree Classifier	0.879247	0.837020	0.879040
Logistic Regression	0.601673	0.553186	0.594968
Support Vector Classifier	0.506012	0.333333	0.340034