**1. Why did you choose this particular algorithm?**

* The Random Forest Classifier was chosen due to its ability to handle large, high-dimensional datasets effectively. It is robust, performs well on structured (tabular) data, and its ensemble approach reduces overfitting, which can be particularly beneficial with larger datasets.
* Additionally, it is widely used for both classification and regression tasks, making it an ideal fit for this classification problem.

**2. What tuning methods were used for the algorithm?**

* A grid search was conducted to optimize several key hyperparameters, including:
  + n\_estimators: Controls the number of trees, which affects model robustness and reduces variance.
  + criterion: Helps select the quality measure used to split nodes.
  + class\_weight: Adjusts weights for classes to help handle any class imbalance, improving the model's sensitivity to minority classes.
* These hyperparameters were chosen to strike a balance between accuracy, model complexity, and handling class imbalance.

**3. Did you consider other algorithms? Why or why not?**

* No, other algorithms were not considered for this task. The Random Forest Classifier is widely regarded for its effectiveness in both classification and regression tasks and is especially advantageous for structured data. This, along with its ability to handle large datasets, made it a preferred choice without the need for extensive comparisons.

**4. What was the model’s accuracy?**

* The model achieved an accuracy score of **0.91**, calculated using the accuracy\_score metric. This metric provided a straightforward measure of the proportion of correct predictions, making it suitable for initial model evaluation.
* accuracy\_score = TP + TN / TP + TN + FP + FN

**5. What metrics were used to evaluate the model, and why?**

* Alongside accuracy, **precision** and **recall** were also used to gain a more detailed view of the model's performance:
  + **Precision**: Measures the accuracy of positive predictions, indicating the proportion of true positives out of all predicted positives.
  + **Recall**: Reflects the model’s ability to capture actual positive instances, showing the proportion of true positives out of all actual positives.