### **Report: Predictive Model for Medical Condition Based on Urine Analysis Data**

#### **Objective**

The objective was to build a predictive model to accurately identify a medical condition based on patient

**Key Findings**

1. **Random Forest Model:**

* **Accuracy:** Achieved **0.6696**.
* **Performance Metrics:**
* **Precision:** Class 0 (Healthy): 0.69, Class 1 (Non-Healthy): 0.67
* **Recall:** Class 0 (Healthy): 0.65, Class 1 (Non-Healthy): 0.71
* **F1-Score:** Class 0 (Healthy): 0.67, Class 1 (Non-Healthy): 0.69

The Random Forest model showed balanced performance but struggled with false positives and false negatives

1. **Gradient Boosting Model:**

* **Accuracy:** Archived **0.70**, compared to the Random Forest model’s accuracy.
* **Precision:** Class 0 (Healthy): 0.75, Class 1 (Non-Healthy): 0.67
* **Recall:** Class 0 (Healthy): 0.61, Class 1 (Non-Healthy): 0.80
* **F1-Score:** Class 0 (Healthy): 0.67, Class 1 (Non-Healthy): 0.73

The Gradient Boosting model improved the ability to identify non-healthy individuals, with a better recall for positive cases and a higher overall F1-Score.

**Conclusion**

The Gradient Boosting model demonstrated better performance compared to the initial Random Forest Classifier. It achieved higher accuracy, improved recall for identifying non-healthy individuals, and a higher F1-Score overall.