DA Assignment 5

IIT2022004-Vatsal Bhuva

**Summary of Results and Analysis**

### **1. Performance Metrics**

The analysis yielded the following performance measures:

* **Accuracy**: 0.8646
* **Precision**: 0.5161
* **Recall**: 0.5000
* **F1 Score**: 0.5079

**Confusion Matrix**:  
 [[182 15]

[ 16 16]]

### **2. Interpretation of Metrics**

* **Accuracy (86.46%)** indicates that the model correctly classified most instances.
* **Precision (51.61%)** suggests that when the model predicted a positive class, it was correct about 51.61% of the time.
* **Recall (50.00%)** shows that the model correctly identified half of the actual positive cases.
* **F1 Score (50.79%)** represents the harmonic mean of precision and recall, indicating a moderate balance between false positives and false negatives.
* The confusion matrix highlights that the model made **15 false positives and 16 false negatives**, suggesting room for improvement in distinguishing between classes.

### **3. Potential Areas for Improvement**

* **Feature Engineering**: Enhancing feature selection and representation might improve classification performance.
* **Hyperparameter Tuning**: Optimizing the model's parameters can enhance precision and recall.
* **Handling Class Imbalance**: If class imbalance exists, techniques such as oversampling, undersampling, or adjusting class weights should be considered.
* **Alternative Models**: Exploring different algorithms like ensemble methods or deep learning could improve overall predictive accuracy.

### **4. Conclusion**

The model performs well in terms of overall accuracy but struggles with precision and recall, suggesting potential misclassification of positive instances. Future improvements should focus on refining features, adjusting model hyperparameters, and addressing class imbalances to achieve better results.