Model Performance Report

Overview

The following report presents the evaluation metrics for a binary classification model. The metrics used to evaluate the model include precision, recall, F1-score, and support for each class (0 and 1). Additionally, overall accuracy, macro average, and weighted average metrics are provided.

NB: Class 0 is for Healthy and Class 1 is for Unhealthy

Evaluation Metrics

1. **Precision:** Precision is the ratio of correctly predicted positive observations to the total predicted positives. It indicates how many of the predicted positives are actually positive.
   - Class 0: 0.73
   - Class 1: 0.90

2. **Recall:** Recall is the ratio of correctly predicted positive observations to the all observations in actual class. It measures the ability of the model to find all the relevant cases within a dataset.
   - Class 0: 0.93
   - Class 1: 0.64

3. **F1-Score:** The F1-score is the weighted average of Precision and Recall. It is particularly useful when the class distribution is imbalanced.
   - Class 0: 0.82
   - Class 1: 0.75

4. **Support:** Support is the number of actual occurrences of the class in the dataset.
   - Class 0: 288
   - Class 1: 279

5. **Accuracy:** Accuracy is the ratio of correctly predicted observations to the total observations.
   - Accuracy: 0.79

6. **Macro Average:** The macro average is the average of precision, recall, and F1-score computed for each class independently.
   - Macro Average Precision: 0.81
   - Macro Average Recall: 0.78
   - Macro Average F1-Score: 0.78

7. **Weighted Average:** The weighted average considers the support of each class while calculating the average of precision, recall, and F1-score.
   - Weighted Average Precision: 0.81

- Weighted Average Recall: 0.79
- Weighted Average F1-Score: 0.78

Detailed Analysis

Class 0 Performance:
- The precision for class 0 is 0.73, indicating that 73% of the instances predicted as class 0 are actually class 0.
- The recall for class 0 is 0.93, meaning the model successfully identifies 93% of the actual class 0 instances.
- The F1-score of 0.82 balances precision and recall, showing a strong performance in predicting class 0.

Class 1 Performance:
- The precision for class 1 is 0.90, suggesting a high level of accuracy in predicting class 1 instances when they are predicted.
- The recall for class 1 is 0.64, indicating that the model captures 64% of the actual class 1 instances. This is relatively lower compared to class 0, which suggests potential room for improvement in detecting all class 1 instances.
- The F1-score of 0.75 reflects the trade-off between precision and recall for class 1, showing that while the precision is high, the lower recall impacts the F1-score.

Overall Performance:
- The overall accuracy of the model is 0.79, meaning that 79% of the total predictions are correct.
- The macro average metrics (precision, recall, F1-score) provide a balanced view of the model performance across both classes, highlighting an overall decent performance.
- The weighted average metrics account for the class distribution, reaffirming the overall robustness of the model with similar values to the macro averages.

Recommendations

- **Improving Recall for Class 1:** Given that the recall for class 1 is relatively low, efforts should be made to improve the model's ability to correctly identify all class 1 instances. Techniques such as adjusting the decision threshold, using more balanced class weights, or exploring different algorithms might help enhance recall.

- **Evaluating Class Imbalance:** It's important to review the class distribution and possibly apply techniques like oversampling the minority class (class 1) or undersampling the majority class (class 0) to address any imbalance, which could improve the model's performance.

- **Model Tuning and Validation:** Conducting hyperparameter tuning and validating the model using cross-validation can provide insights into how to better optimise the model for both classes.

By focusing on these areas, the model's predictive performance can be enhanced, leading to better overall accuracy and more reliable predictions across both classes.