Model Training and Evaluation Report

1. Introduction

This report summarizes the training process, performance metrics, and evaluation of an Artificial Neural Network (ANN) model developed to classify patients as either **No Disease** or **Heart Disease** based on clinical data. The dataset consists of 54 samples used for testing, with the model trained over 100 epochs.

2. Training Summary

The model was trained for 100 epochs, and the loss steadily decreased, indicating effective learning.
The rapid and consistent drop in loss values over epochs demonstrates the model's successful convergence and improved ability to minimize prediction errors.

3. Confusion Matrix

	Predicted No Disease	**Predicted Heart Disease**
Actual No Disease	24 (True Negative)	6 (False Positive)
Actual Heart Disease	2 (False Negative)	22 (True Positive)

Interpretation:

- **True Negatives (TN):** 24 patients correctly classified as healthy.
- **False Positives (FP):** 6 patients without disease incorrectly labeled as having heart disease.
- **False Negatives (FN):** 2 patients with heart disease missed by the model (classified as healthy).
- **True Positives (TP):** 22 patients with heart disease correctly identified.

Clinical significance:
Minimizing false negatives is crucial in medical diagnosis to avoid missing true cases of heart disease. The model achieves only 2 false negatives, demonstrating strong sensitivity.

4. Classification Metrics

Class	Precision	Recall	F1-Score	Support
No Disease	0.92	0.80	0.86	30
Heart Disease	0.79	0.92	0.85	24

- **Precision:**
  - *No Disease:* 92% of predicted healthy cases were actually healthy.
  - *Heart Disease:* 79% of predicted heart disease cases were true positives.

- **Recall (Sensitivity):**
  - *No Disease:* 80% of actual healthy patients were correctly identified.

- ➢ *Heart Disease:* 92% of actual heart disease patients were correctly detected.

❖ **F1-Score:** The harmonic mean of precision and recall, showing balance:
- ➢ *No Disease:* 0.86
- ➢ *Heart Disease:* 0.85

Support: Number of actual samples in each class (30 healthy, 24 diseased).

5. Overall Model Performance

- The accuracy of 85% reflects that 85 out of 100 predictions (or equivalent) were correct.
- Macro averages treat both classes equally, showing balanced performance.
- Weighted averages account for class imbalance, confirming robust results across classes.

6. Detailed Analysis

❖ **Strengths:**
- ➢ High recall (0.92) for heart disease means the model effectively identifies most diseased patients, critical to reduce missed diagnoses.
- ➢ The low false negative count (2 cases) further supports this strength.
- ➢ Training loss dropped sharply and stabilized at a very low value (0.0206), showing efficient training.

❖ **Weaknesses:**
- ➢ False positives (6 cases) indicate the model sometimes misclassifies healthy patients as diseased. While this is less dangerous than false negatives in medical contexts, it could lead to unnecessary further tests or anxiety.

- ➢ Precision for heart disease (0.79) is slightly lower, meaning about 21% of predicted diseased cases were false alarms.

❖ **Possible improvements:**

- ➢ Tune the classification threshold to reduce false positives without increasing false negatives too much.
- ➢ Collect more data to improve model generalization and reduce overfitting risks.
- ➢ Explore ensemble models or additional feature engineering to enhance precision.

7. Conclusion

This Artificial Neural Network model demonstrates strong ability to detect heart disease with a good balance of sensitivity and specificity. The low final loss and balanced classification metrics support its suitability for preliminary diagnostic support systems.

8. Next Steps

- ➔ **Validation:** Test on a larger, external dataset to confirm generalization.
- ➔ **Deployment:** Consider integration into clinical decision support software with human oversight.
- ➔ **Monitoring:** Track model performance over time to detect data drift or performance degradation.