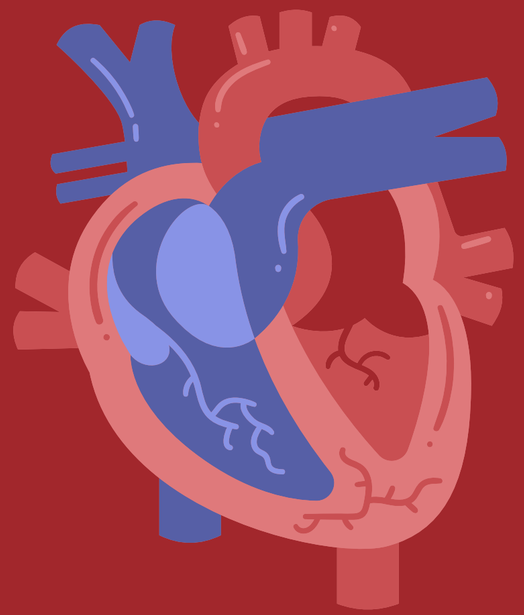
HEART DISEASE PREDICTION



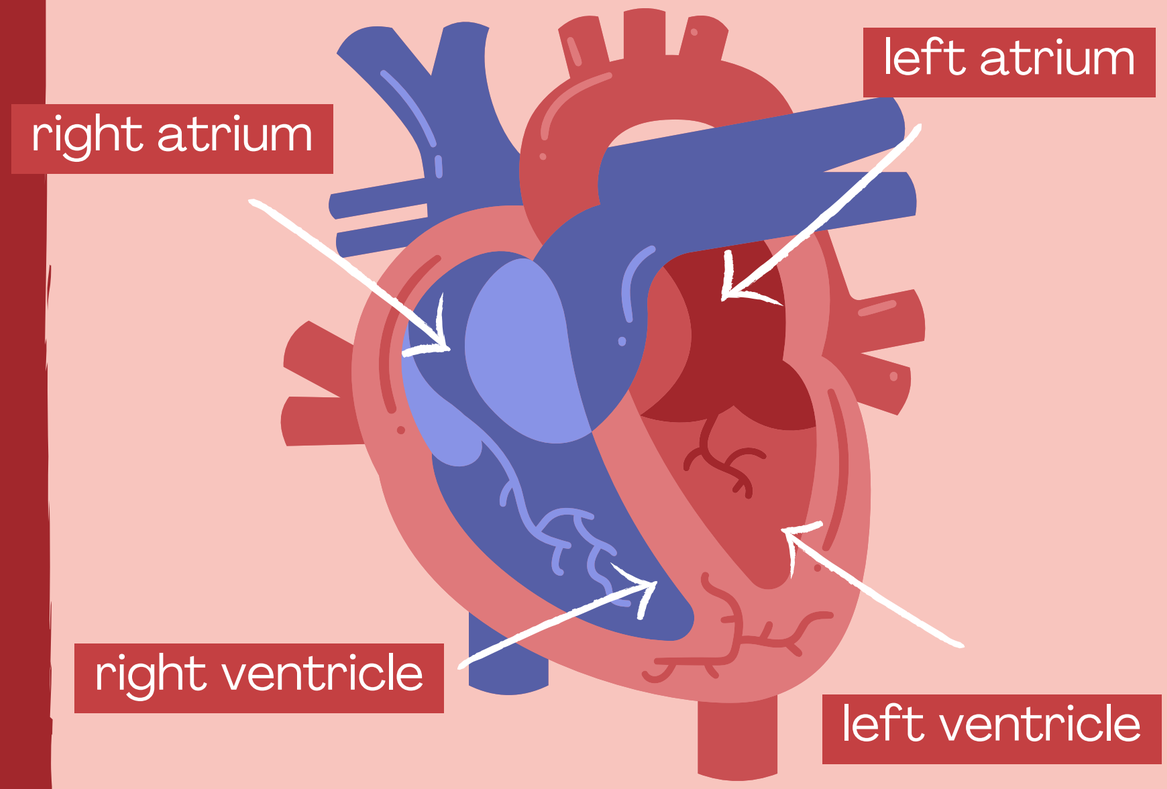
Sweety R
(RA2211026010083)

ABSTRACT

The primary objective of this project is to develop a binary classification model using a neural network to predict the presence or absence of heart disease based on various medical parameters. The project utilizes the popular Keras library and TensorFlow backend, implementing a deep learning approach. The dataset used is the heart.csv, which contains relevant health and medical data that serve as features for predicting heart disease. The model is trained and evaluated to analyze its performance, presenting an effective way to assist in early detection of heart disease.

PROBLEM STATEMENT

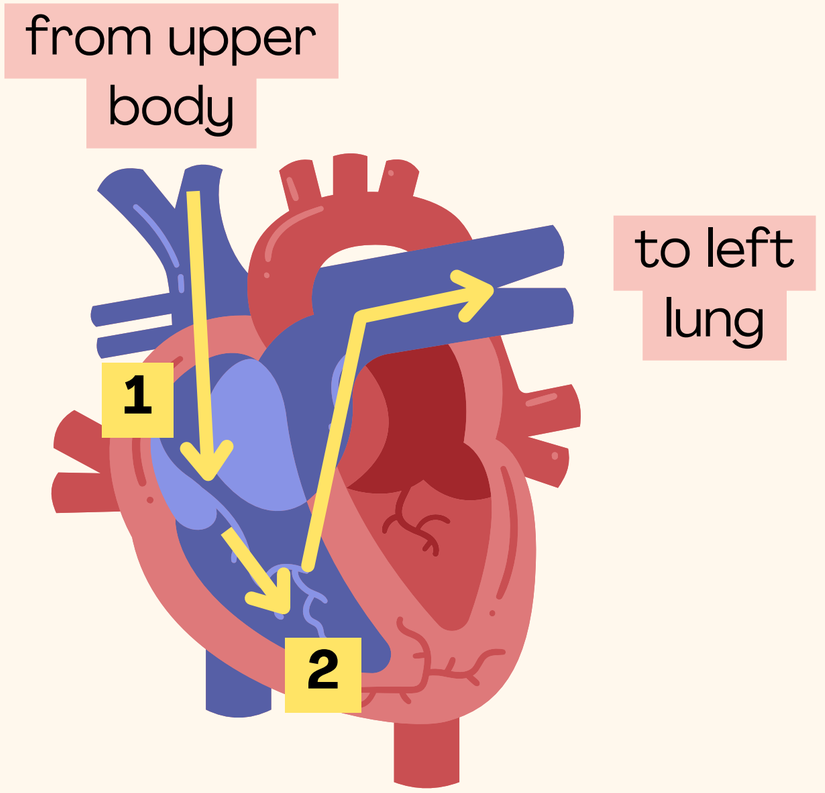
Heart disease is one of the leading causes of mortality worldwide. Early prediction and diagnosis are crucial in preventing severe health complications and improving patient outcomes. This project aims to create a neural network model to predict the likelihood of a patient having heart disease based on clinical data. The model should achieve high accuracy to be clinically relevant and aid healthcare professionals in making informed decisions.



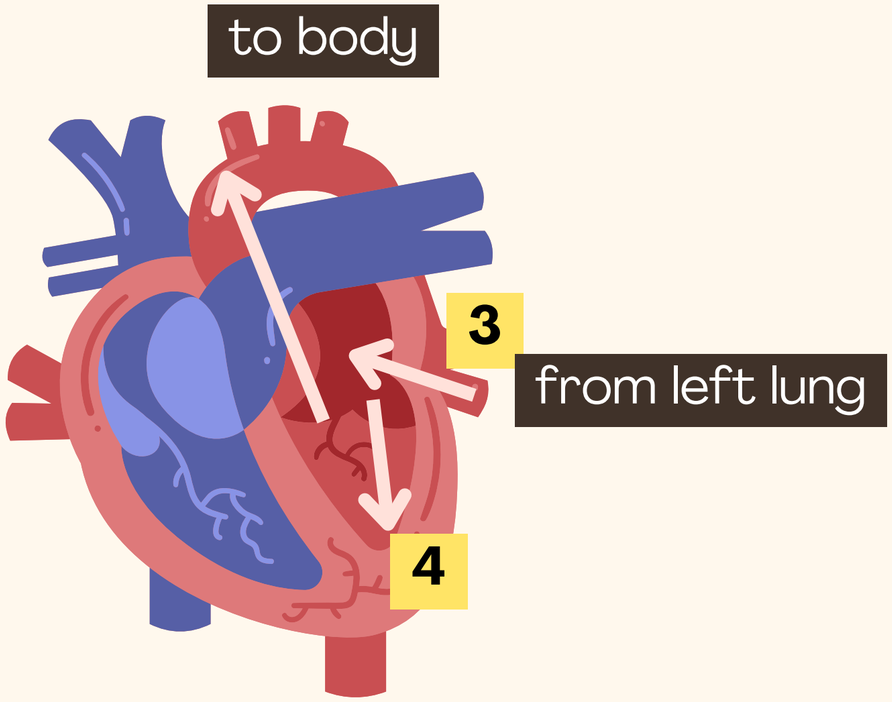
OBJECTIVES

1. To preprocess and analyze heart disease data to ensure suitability for model training.

2. To design and implement a neural network model using Keras for binary classification.



OBJECTIVES



3 To evaluate model performance in terms of accuracy, precision, and confusion matrix.

4 To discuss the model's potential impact in clinical settings for heart disease prediction.

Paper title	Techniques Used	Publication	Inference
1. Heart Disease Prediction Using Artificial Neural Networks	ANN with backpropagation	IEEE Xplore	Demonstrates high accuracy in heart disease prediction using ANN; backpropagation algorithm is crucial for optimizing the model's predictive performance.
2. A Robust Heart Disease Prediction System Using Hybrid Deep NNs	Hybrid Deep Neural Networks (DNN)	IEEE Xplore	Combines DNNs with preprocessing techniques to handle imbalanced datasets, improving accuracy and robustness in heart disease prediction.

3. Early Prediction of Heart Disease with Data Analysis and Stochastic Gradient Boosting	ANN and gradient boosting, SMOTE for class balancing	Springer	Highlights the effectiveness of SMOTE in addressing class imbalance, combined with gradient boosting for reliable heart disease predictions.
4. Heart Disease Prediction using CNN and ANN Ensemble	Convolutional Neural Networks (CNN) + ANN Ensemble	arXiv	Integrates CNNs and ANN models for feature extraction and classification, achieving improved prediction accuracy over standalone ANN models.

5. Comparative Study on Heart Disease Prediction Models	ANN, Decision Trees, Random Forest, and Ensemble Methods	Journal of Engineering Science	Compares multiple algorithms, finding ANN with ensemble methods offers better accuracy and insights for handling different data distributions.
6. Application of Neural Networks in Medical Diagnostics	ANN with ReLU activation and dropout regularization	Springer	Uses dropout to avoid overfitting in ANN for heart disease prediction, suggesting practical improvements in predictive power for clinical applications.

MODULES EXPLANATION

Data Preprocessing

The dataset is cleaned and standardized, with an ID column added for easy tracking. Features are scaled using StandardScaler to improve model performance.

Training

The model is trained using binary cross-entropy as the loss function and the Adam optimizer for faster convergence.

Model Construction

A neural network is built with the Keras Sequential model. The model includes two dense hidden layers, each using Leaky ReLU for non-linearity, and an output layer with sigmoid activation to output probabilities.

Evaluation

After training, the model predicts on the test data, and results are compared with actual values to calculate accuracy, confusion matrix, and other performance metrics.

ALGORITHM EXPLANATION (MATHEMATICAL MODEL OF ALGORITHM)

Hidden Layer Computation

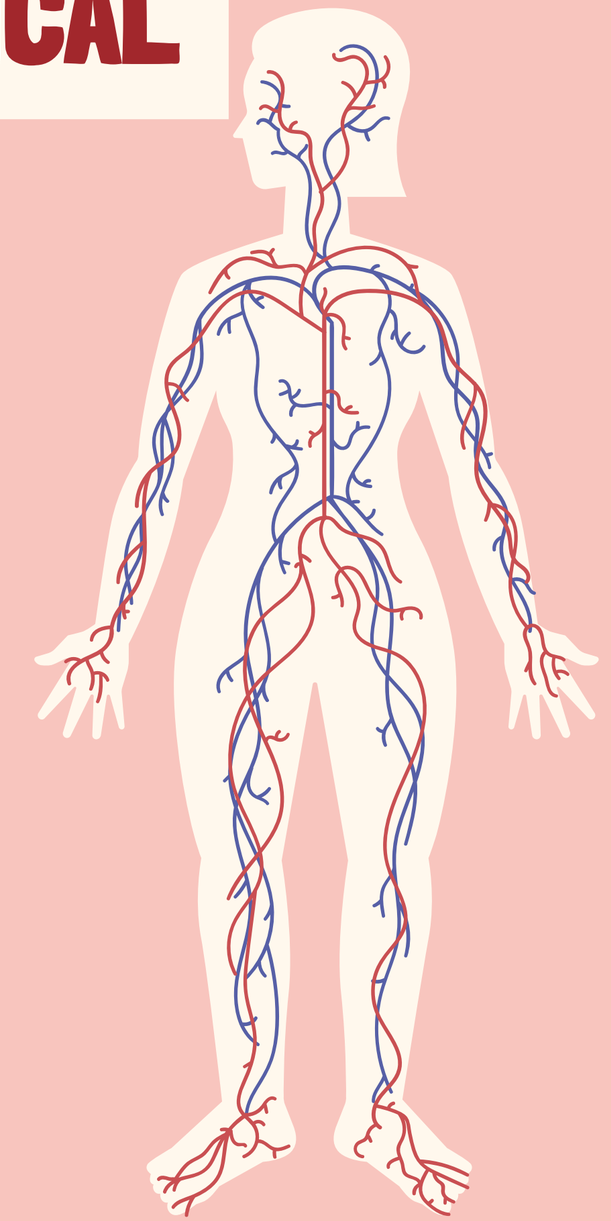
Hidden layers:

$$\text{Leaky ReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{if } x \leq 0 \end{cases}$$

Output layer

Output layer:

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

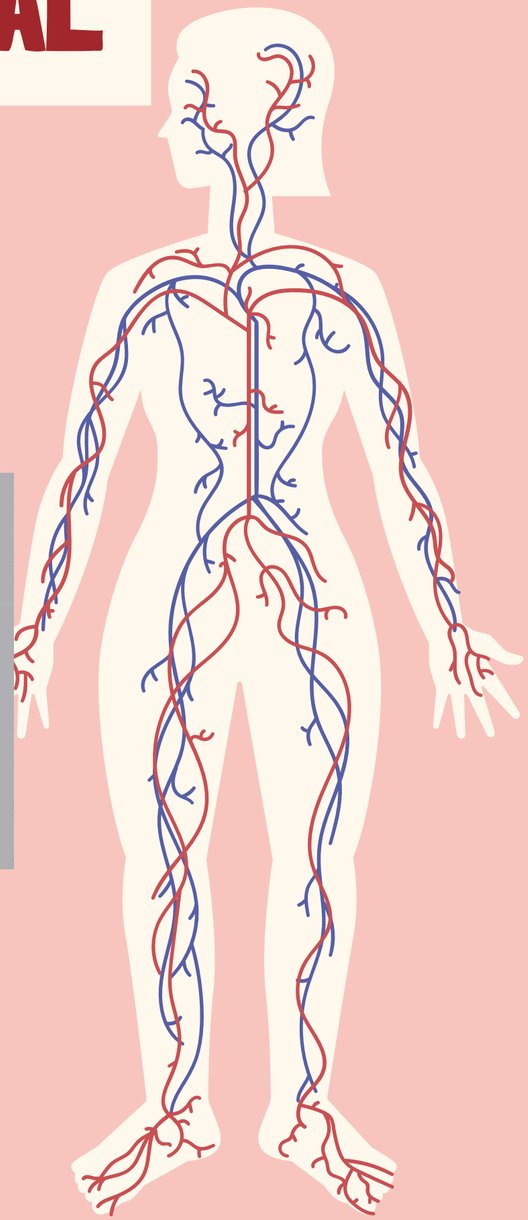


ALGORITHM EXPLANATION (MATHEMATICAL MODEL OF ALGORITHM)

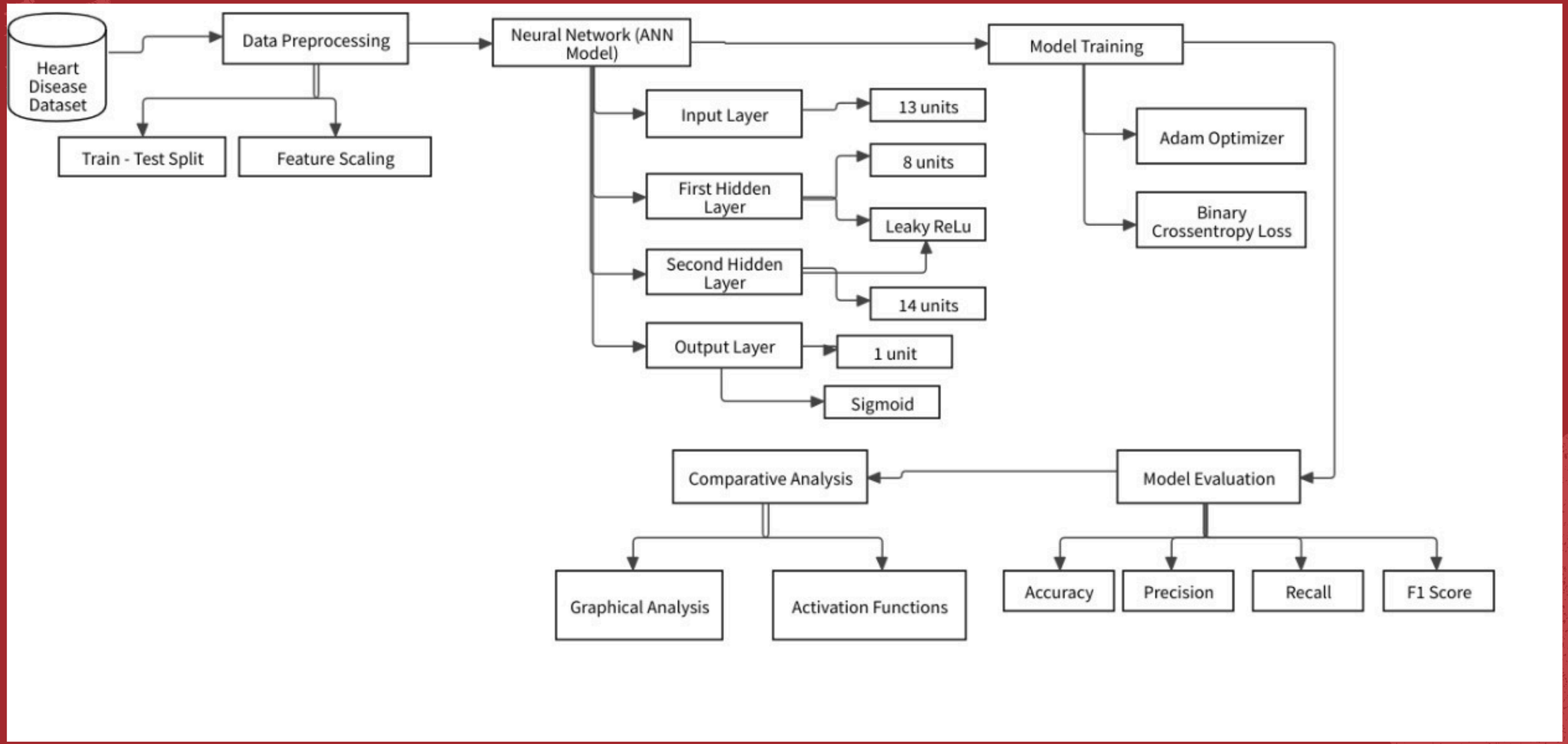
Loss function

Loss Function (Binary Cross-entropy)

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot \log(p_i) + (1-y_i) \cdot \log(1-p_i)\right]$$



ARCHITECTURE DIAGRAM



RESULT AND DISCUSSION

Results

Accuracy

The model's accuracy on the test dataset is calculated as shown in the final output, demonstrating the overall performance.

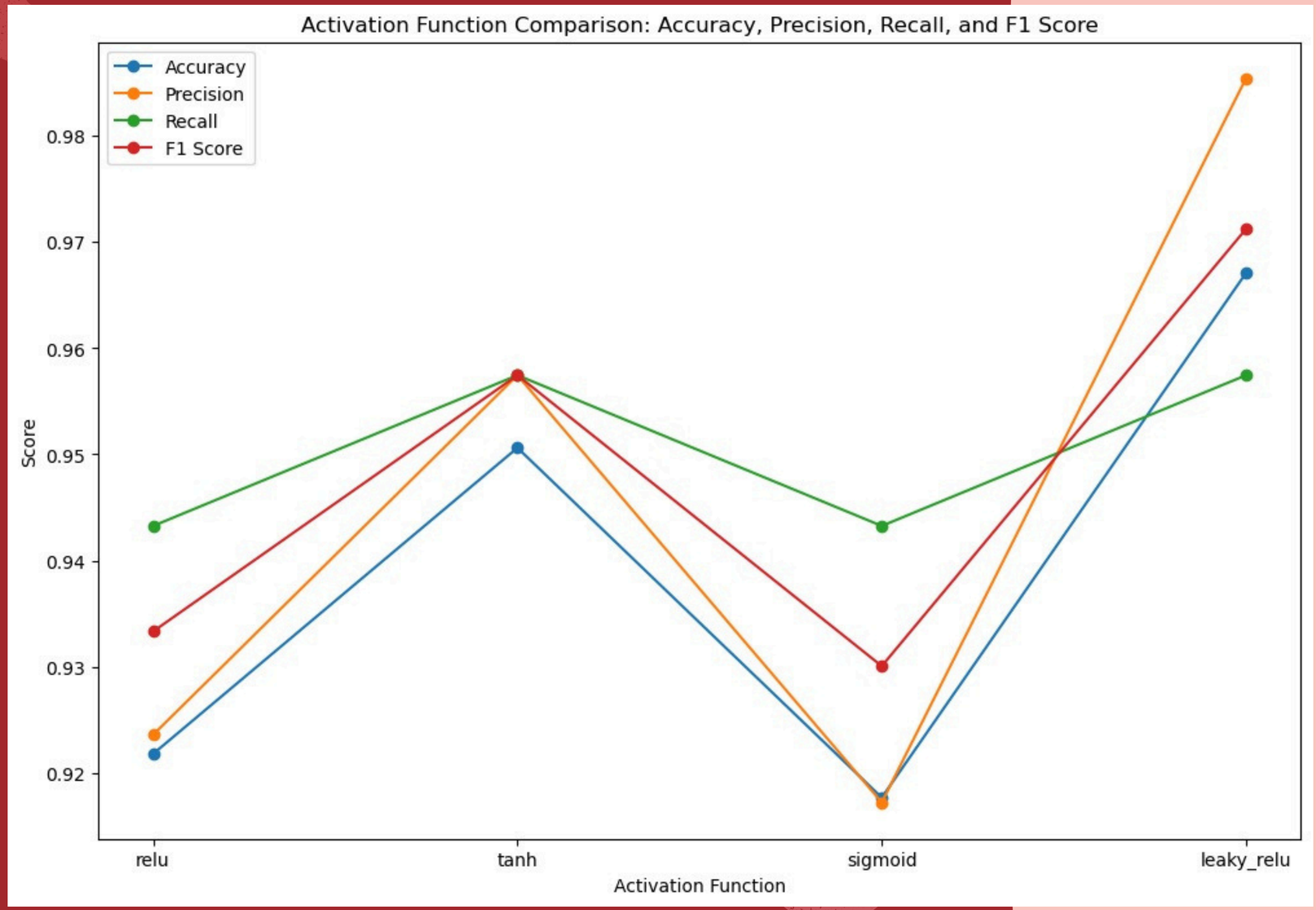
Confusion Matrix

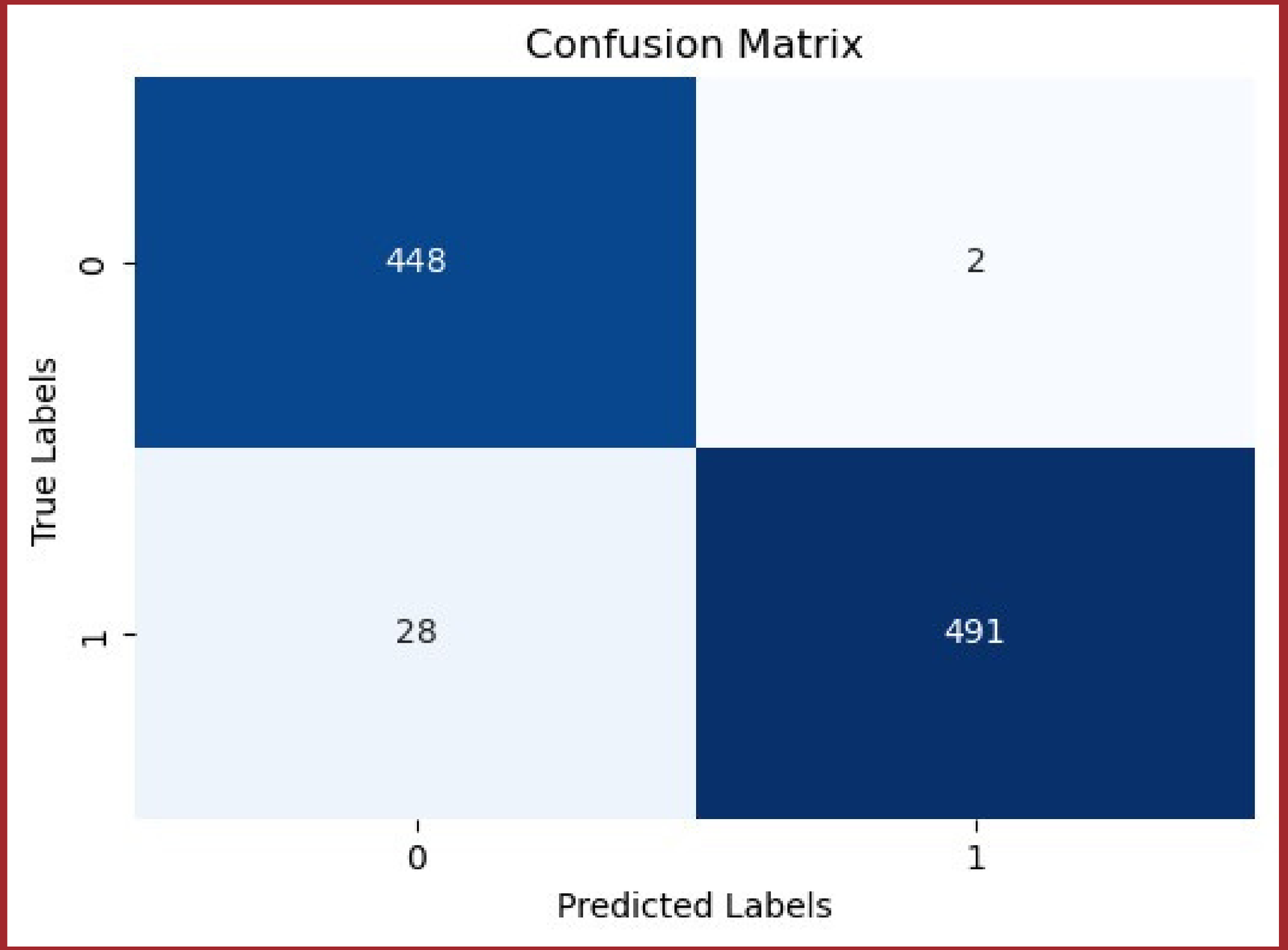
Shows the number of true positives, true negatives, false positives, and false negatives, providing insight into prediction quality.

COMPARATIVE ANALYSIS

Activation Function	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
ReLU	92.18	92.36	94.33	93.33
Tanh	95.06	95.74	95.74	95.74
Sigmoid	91.77	91.72	94.33	93.01
Leaky ReLU	96.71	98.54	95.74	97.12

GRAPHICAL ANALYSIS




Confusion Matrix
448
2
28
491
0
1
True Labels
0
1
Predicted Labels

THANK YOU