

Neural Network

- **Data Normalization:**

Z-score normalization is used to standardize the dataset for both training and testing phases.

- **One-Hot Encoding:**

Converts labels into a one-hot encoded format for multi-class classification tasks.

- **Activation Functions:**

- **Sigmoid:** Outputs probabilities for binary classification. Derivatives are used for backpropagation.
- **ReLU:** Introduces non-linearity by replacing negative values with zero. Used in hidden layers.
- **Softmax:** Converts outputs to probabilities for multi-class classification tasks.

- **Forward Propagation:**

- Layers are built using a **dense** function, applying weights, biases, and activation functions sequentially.
- Multiple layers (hidden and output) are computed to propagate the input forward.

- **Output Layers:**

- **Binary Classification:** Sigmoid activation is used in the output layer.
- **Multi-class Classification:** Softmax activation is used in the output layer for probabilities.

- **Backward Propagation:**

- Gradients are computed for each layer to minimize the error.
- Weight and bias updates use the **Gradient Descent** algorithm with the specified learning rate.

- **Prediction:**

- For binary classification, outputs are thresholded (e.g., ≥ 0.5) for class labels.
- For multi-class classification, the class with the highest probability is selected.

- **Loss Function:**

- Binary Cross-Entropy: For binary classification tasks (commented in the code).
- Cross-Entropy Loss: For multi-class classification tasks (commented in the code).

- **Accuracy Calculation:**

Computes the percentage of correct predictions.