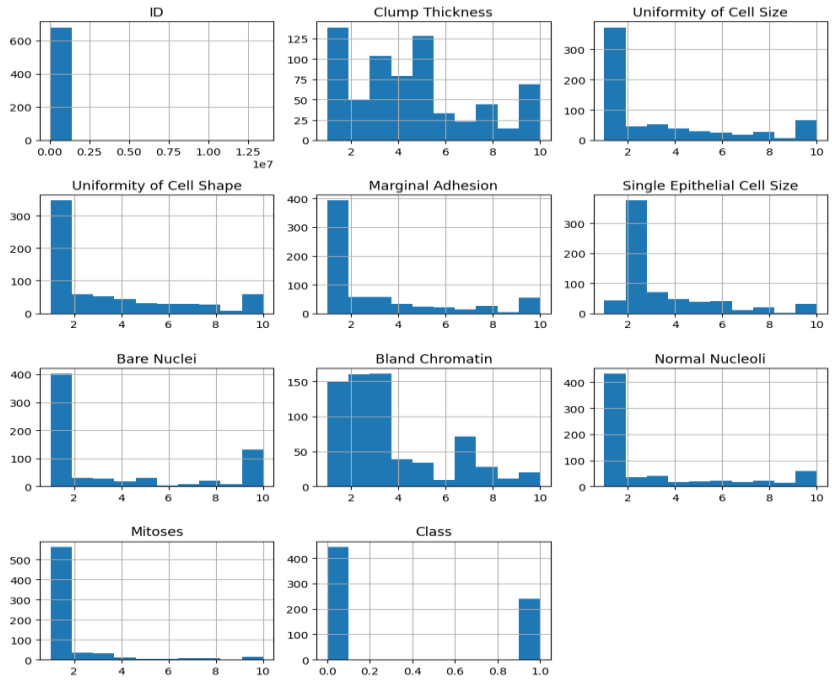
Machine Learning Assignment #4
Breast Cancer Malignancy Prediction Report

Name: Ruthvik Reddy Anugu
Student Id: 001096522

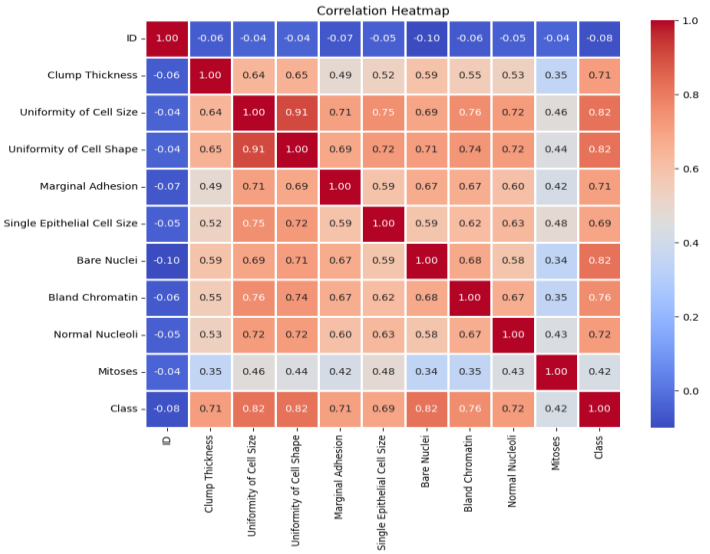
Visualization of Dataset:

The dataset includes a number of characteristics linked to tumors of breast cancer. These are some examples of visualizations that show the features of the data:

Histogram of Features:
Each feature in the collection is represented as a histogram, which sheds light on its ranges and distributions.



Scatter Plot for Correlation:
The pairwise correlations between characteristics are shown visually in the correlation scatter plot, which aids in determining any possible connections between them.



Pre-Processing Steps:

The dataset underwent the following pre-processing procedures before the model was trained:

Missing Values Handling: To maintain data integrity, rows with missing values were removed.
Class Label Encoding: Binary encoding was used to encode the class labels, with 0 denoting benign tumors and 1 denoting malignant tumors.
Feature Scaling: To make sure that every feature contributes equally to model training, features were standardized using the Standard Scaler.

Model Architecture:

A straightforward neural network model was used for this prediction task:

Model of Neural Networks:
Input Layer: A single dense layer containing a sigmoid activation function and a single neuron.

Model Training:

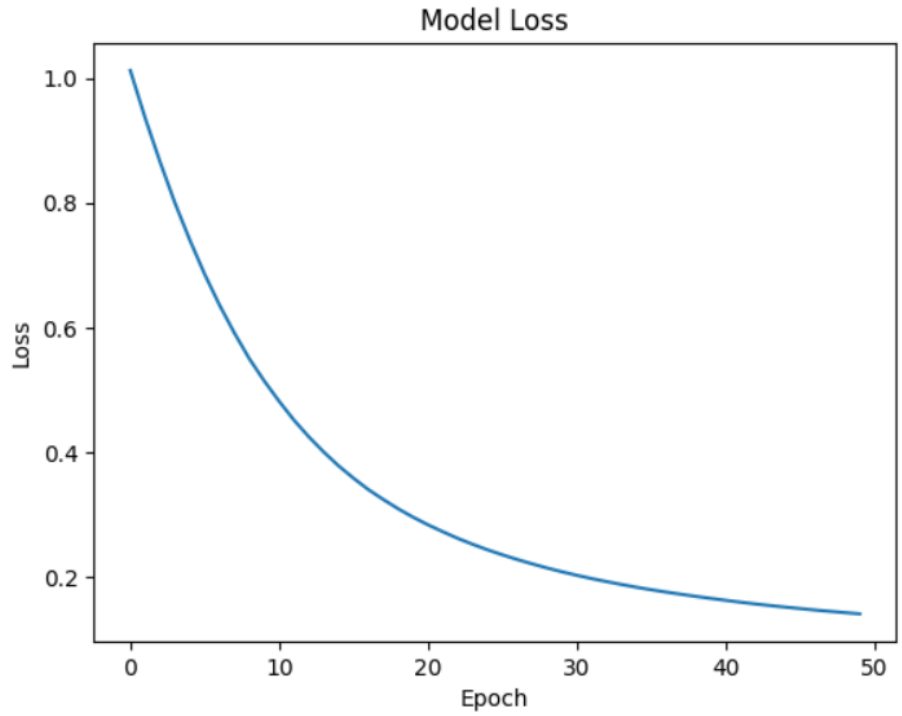
The preprocessed data was used to train the neural network model. The model used backpropagation to minimize the binary cross entropy loss function during training. To guarantee convergence and avoid overfitting, we kept an eye on the training process across several epochs.

Training Progress:
Throughout several epochs, the model's training progress was tracked, and its goal was to minimize the binary cross entropy loss function. The drop in loss values throughout epochs is depicted in the following plot:

Results:

After training was finished, a number of indicators were used to assess the model's performance:
Total Counts of Model Parameters:

```
3/3 [==============================] - 0s 3ms/step
Confusion Matrix:
[[39  0]
 [ 2 28]]
Accuracy: 0.9710144927536232
Precision: 1.0
Recall: 0.9333333333333333
F1 Score: 0.9655172413793104
Total Parameters: 10
```



The neural network model has Ten parameters in total.

Conclusion:

In summary, based on tumor features, our machine learning technique shows promising results in predicting the aggressiveness of breast cancer. Through the use of a neural network model and thorough data preparation, we were able to identify tumors as benign or malignant with high accuracy. By assisting medical practitioners in the early identification and diagnosis of breast cancer, this predictive model may eventually improve patient outcomes and survival rates. It is advised to do more research and validation to improve the model and guarantee its dependability in clinical situations.