Machine Learning Assignment #4
Breast Cancer Malignancy Prediction Report

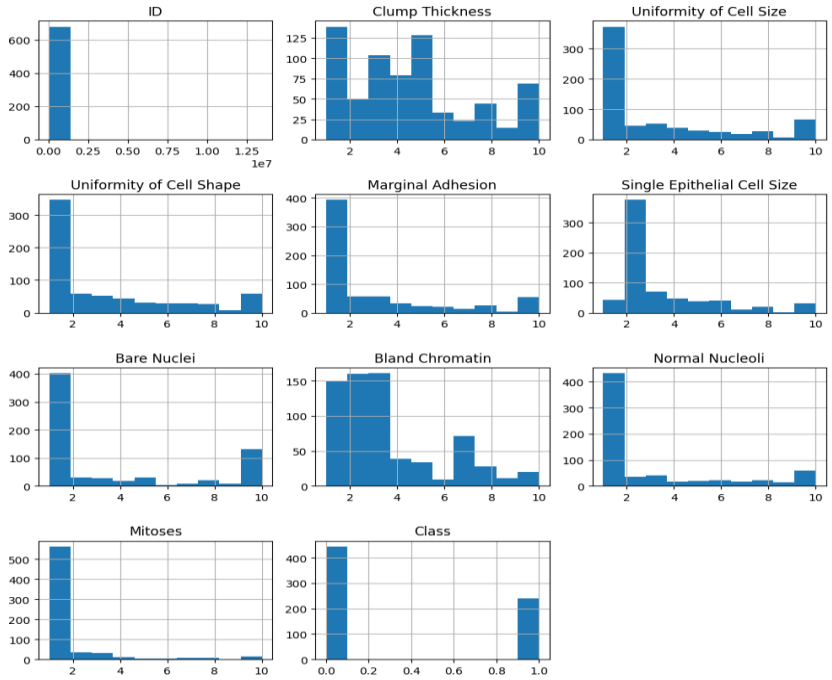
Name: Ruthvik Reddy Anugu
Student Id: 001096522

Visualization of Dataset:

The dataset contains several features associated with breast cancer tumors. The following are a few instances of data feature visualizations:

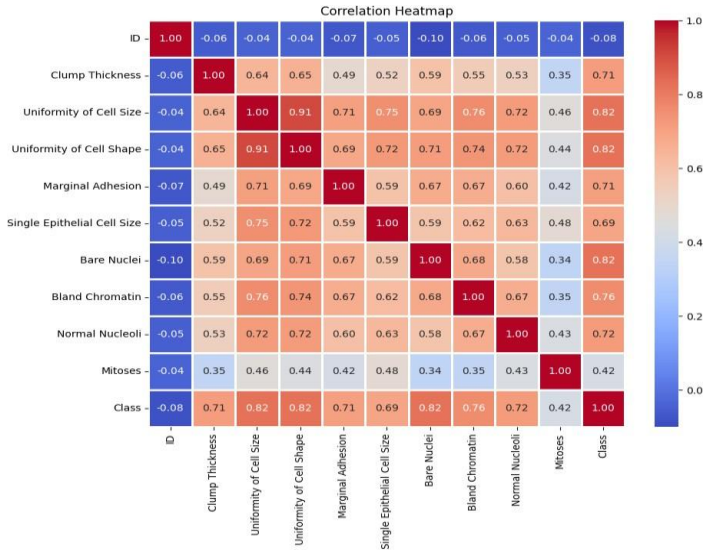
Histogram of Features:

A histogram is used to illustrate the ranges and distributions of each characteristic in the collection.



Scatter Plot for Correlation:

The correlation scatter plot graphically represents the pairwise correlations between the parameters and helps to identify potential relationships between them.



Pre-Processing Steps:

The dataset underwent the following pre-processing procedures before the model was trained:
Missing Values HandlingRows with missing values were eliminated in order to preserve data integrity.

Class Label Encoding: Using binary encoding, the class labels were encoded, where 0 represented benign tumors and 1 represented malignant tumors.
Feature Scaling: Using the Standard Scaler, features were standardized to ensure that each feature contributes equally to model training.

Model Architecture:

This prediction task employed a simple neural network model:

The neural network model
One dense layer with a single neuron and a sigmoid activation function is the input layer.

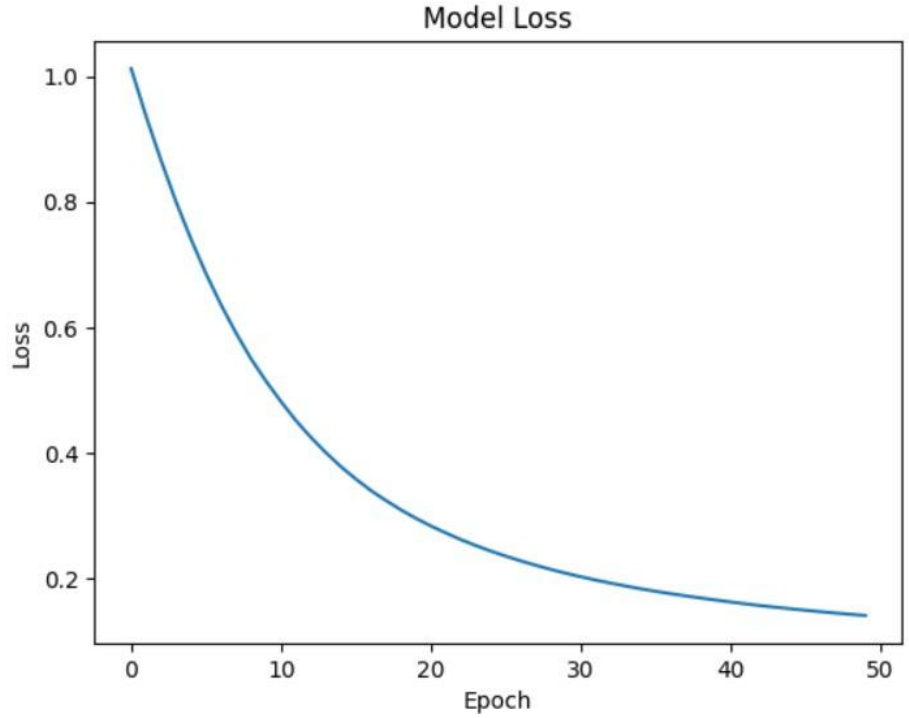
Model Training:

The neural network model was trained using the preprocessed data. During training, the model minimized the binary cross entropy loss function via backpropagation. We monitored the training process across many epochs to ensure convergence and prevent overfitting.
Training Advancement:
The model was trained across many epochs with the objective of minimizing the binary cross entropy loss function. The graphic below shows the decline in loss values throughout several epochs:

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 conclusion, our machine learning approach demonstrates encouraging outcomes in predicting the aggressiveness of breast cancer based on tumor traits. Using a neural network model and careful data preparation, we achieved great accuracy in classifying tumors as benign or malignant. This prediction model has the potential to eventually enhance patient outcomes and survival rates by aiding medical professionals in the early detection and diagnosis of breast cancer. Further study and validation are recommended in order to enhance the model and ensure its reliability in clinical settings.