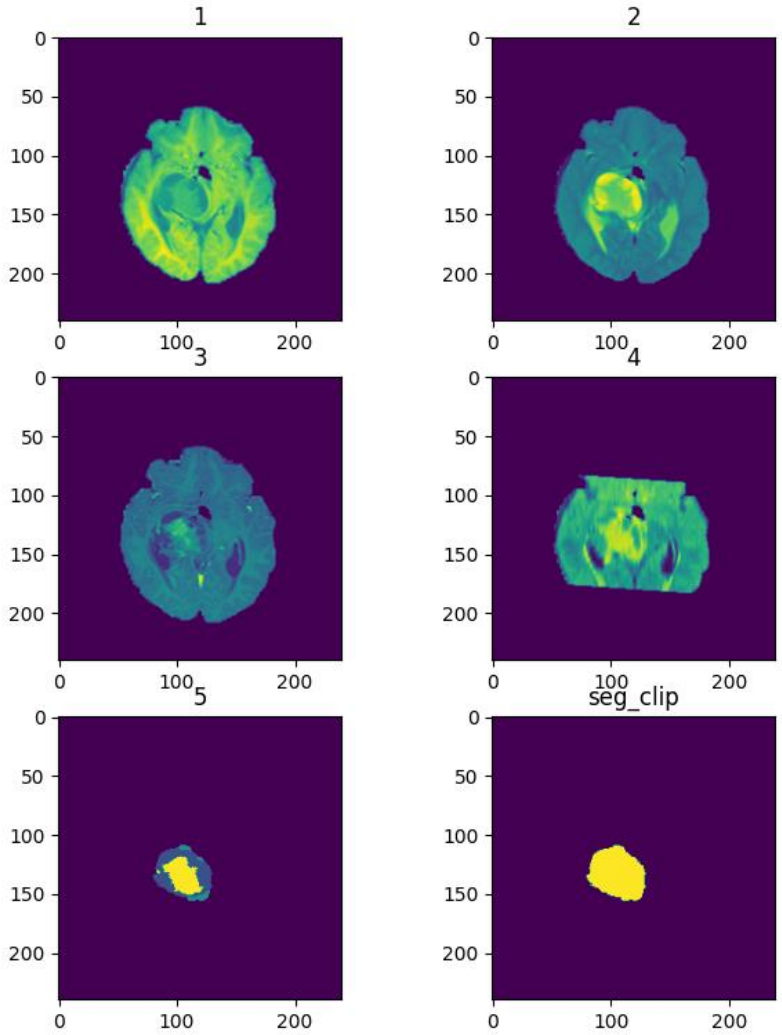
Abdur Rehman (202023019)

Assignment 2

Introduction:

Our goal in this assignment is to classify brain tumor in human brain. Glioma (brain tumor) is divided into glioblastoma (HGG) and lower grade glioma (LGG) with HGG being an aggressive and life-threating tumor. Dataset consists of tumor image, segmentation map and label. Image consists of four channels and segmentation map contain subregions, each with a different pattern on brain MRI scan. Following figure helps in visual understanding of dataset.



Assignment is divided into three parts which are as follows

1. Train a CNN-based HGG/LGG classifier and calculate its accuracy.
2. Train logistic regression, neural network and random forest using radiomics features and test their accuracy.
3. Report the difference between the deep-learning and machine learning algorithms.

Part1:

In the first part we required to train a deep learning model to classify images of HGG/LGG. During preprocessing step, we have to clip the values of segmentation map to display uniform map. Later on, this segmentation map is concatenated on the input image.

A simple neural network is used for classification problem because of small amount of training data. Architecture consists of three convolution layers of stride 2 followed by batch normalization and pooling layers with relu activation function. Dense layers are added in the end with dropout layers in between them. Here is the architecture of the training data.

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1         [-1, 16, 111, 111]             736
       BatchNorm2d-2         [-1, 16, 111, 111]              32
              ReLU-3         [-1, 16, 111, 111]               0
         MaxPool2d-4           [-1, 16, 55, 55]               0
            Conv2d-5           [-1, 32, 27, 27]           4,640
       BatchNorm2d-6           [-1, 32, 27, 27]              64
              ReLU-7           [-1, 32, 27, 27]               0
         MaxPool2d-8           [-1, 32, 13, 13]               0
            Conv2d-9             [-1, 64, 6, 6]          18,496
      BatchNorm2d-10             [-1, 64, 6, 6]             128
             ReLU-11             [-1, 64, 6, 6]               0
        MaxPool2d-12             [-1, 64, 3, 3]               0
          Dropout-13                  [-1, 576]               0
           Linear-14                   [-1, 10]           5,770
             ReLU-15                   [-1, 10]               0
          Dropout-16                   [-1, 10]               0
           Linear-17                    [-1, 1]              11
          Sigmoid-18                    [-1, 1]               0
================================================================
Total params: 29,877
Trainable params: 29,877
Non-trainable params: 0
----------------------------------------------------------------
```

Adam optimizer and binary cross entropy loss was used for training. Without any regularization model quickly overfits on data. After adding dropout layers accuracy improves a bit but model still overfits and achieves an accuracy of **~80%**. Here are the accuracy graphs:

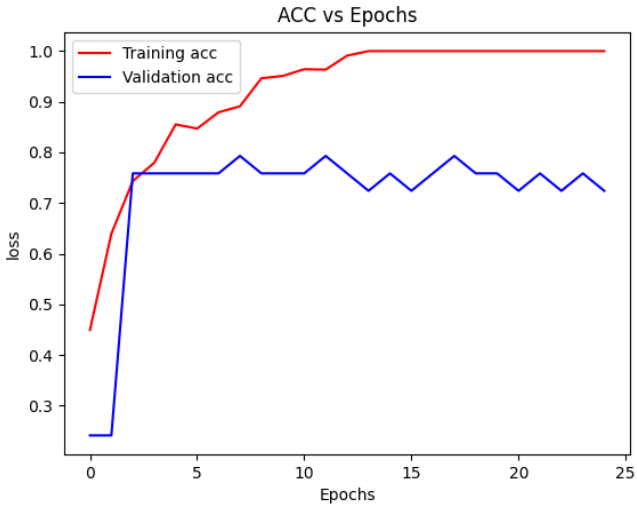


Figure 1: Without dropout

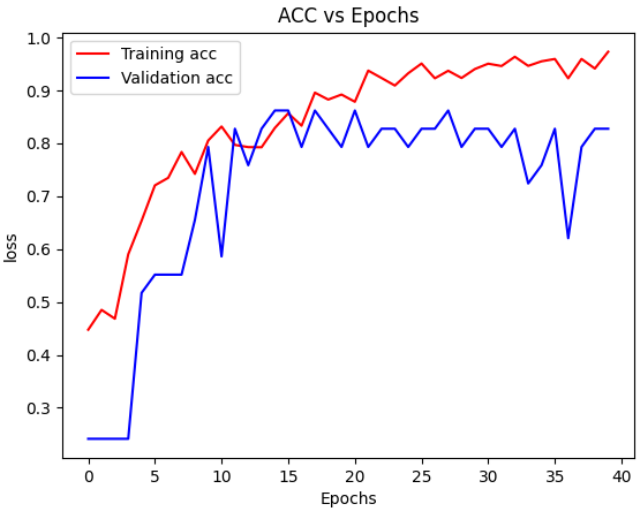


Figure 2: With dropout

Part2:

In this part we have to train logistic regression, neural network(MLP) and random forest on radiomics features. Radiomics is a method that extracts a large number of features from medical images using data-characterization algorithms. These features, termed radiomic features, have the potential to uncover tumoral patterns and characteristics that fail to be appreciated by the naked eye. We use Pyradiomics library developed by Artificial Intelligence in Medicine (AIM) Harvard University. Using Pyradiomics library we can extract several features from an image. These features can be used to train the above-mentioned algorithms. Details of these algorithm are provided below.

Logistic regression:

Logistic regression is used for modeling the probability of a certain class or event. We can use it for the given classification task. We achieve a classification accuracy of **76.19%** and an F1 score of **0.85**. Here is the confusion matrix of the test data.

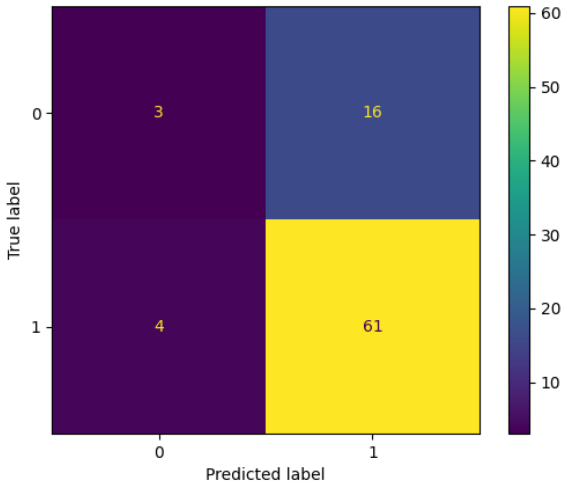


Figure 3: Logistic Regression Confusion Matrix

Neural Network (MLP):

Neural network (NN) or multilayer perceptron can also be used for classification task. In this assignment, we used a three-layer NN of size (32). Results of NN were quite inconsistent and accuracy fluctuates between **30%-70%**. Here is the confusion matrix of the test data.

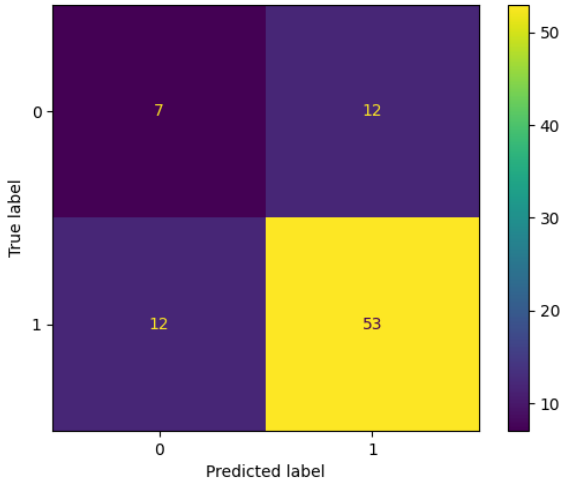


Figure 4: Neural Network Confusion Matrix

Random Forest:

Random forest is an ensemble learning method for classification and regression tasks by constructing a multitude of decision trees at training time. For classification task, the output of random forest is the class selected by most trees.

Random forest achieved the best accuracy out of all methods. We achieved a test data accuracy of **85.71%** with F1 score of **0.91**.

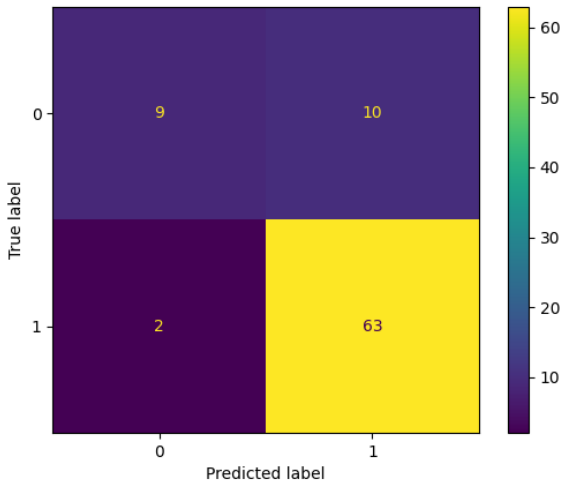


Figure 5: Random Forest Confusion Matrix

Here is the summary of all methods described in part2 of this assignment:

	Accuracy	**F1 Score**
Logistic Regression	76.19	0.859
Neural Network	73.80	0.849
Random Forest	85.71	0.913

Part 3:

After trying both deep learning and machine learning methods, it was observed that deep learning method was on par with machine learning methods and random forest achieves the highest accuracy.

Low accuracy of deep learning method can be attributed to low amount of training data. Deep learning models out perform traditional methods when training data is abundant which is not true in our case. Another problem that I observed was data imbalance. It arises when one class has a lot more data than

the other class. During training model learns to predict only one class which results in higher accuracy but model is not learning anything. This problem was particularly observed in machine learning algorithms (logistic regression and neural networks) and random forest was found to be somewhat immune to this problem sue to the fact that it uses ensemble learning to make prediction.

Based on the observation made during this assignment, it is better to use random forest when we have small imbalanced training data and use deep learning with regularization when training data is large.