Heart Disease Classification using Neural Networks

Group 1A:

Mohamed Hashi

Maziyar Babar

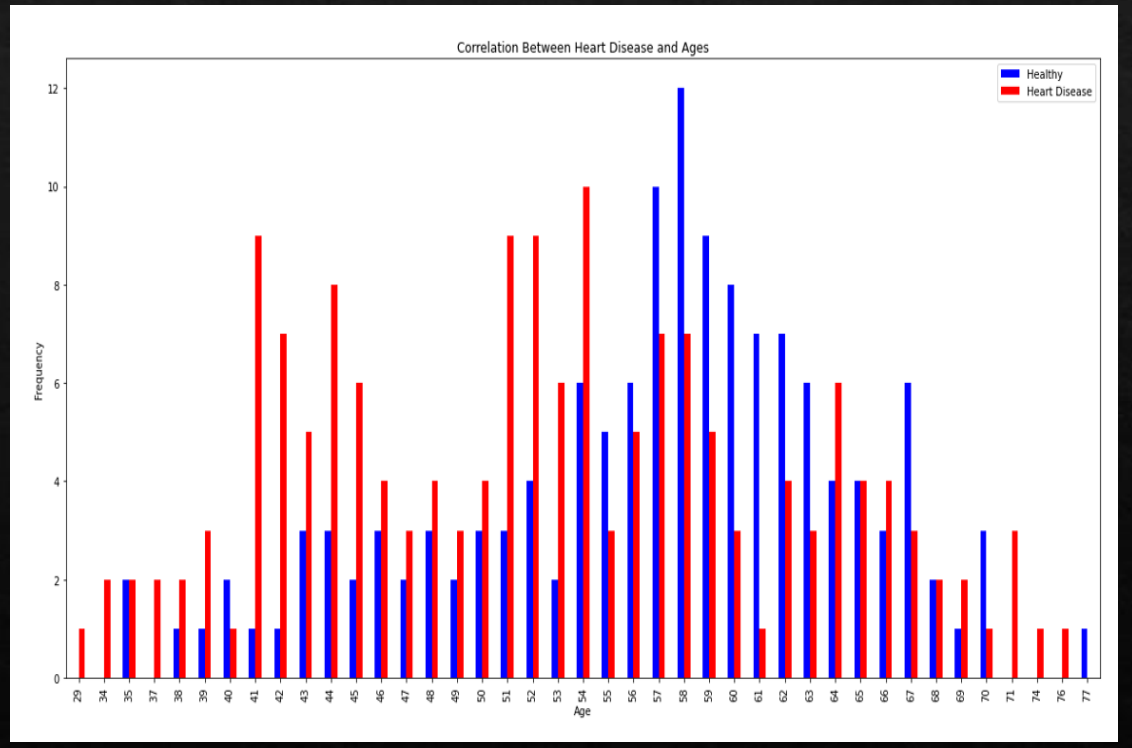
Liam Baszun

Calvin Chau

Johar Raza

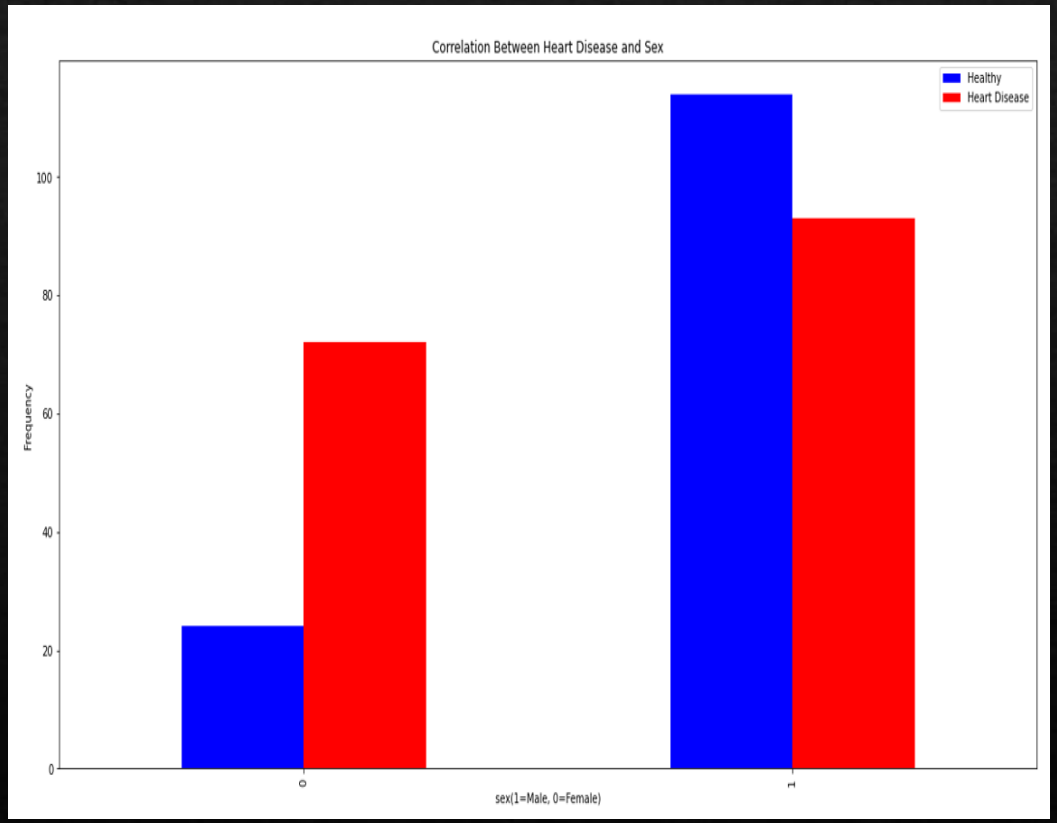
Data Exploration

- **Correlation of heart disease with age:**
  - Inconsistent correlation
  - Peaked at ages 41 and 56
  - Ages between 41 and 56 have low heart disease with some outliers



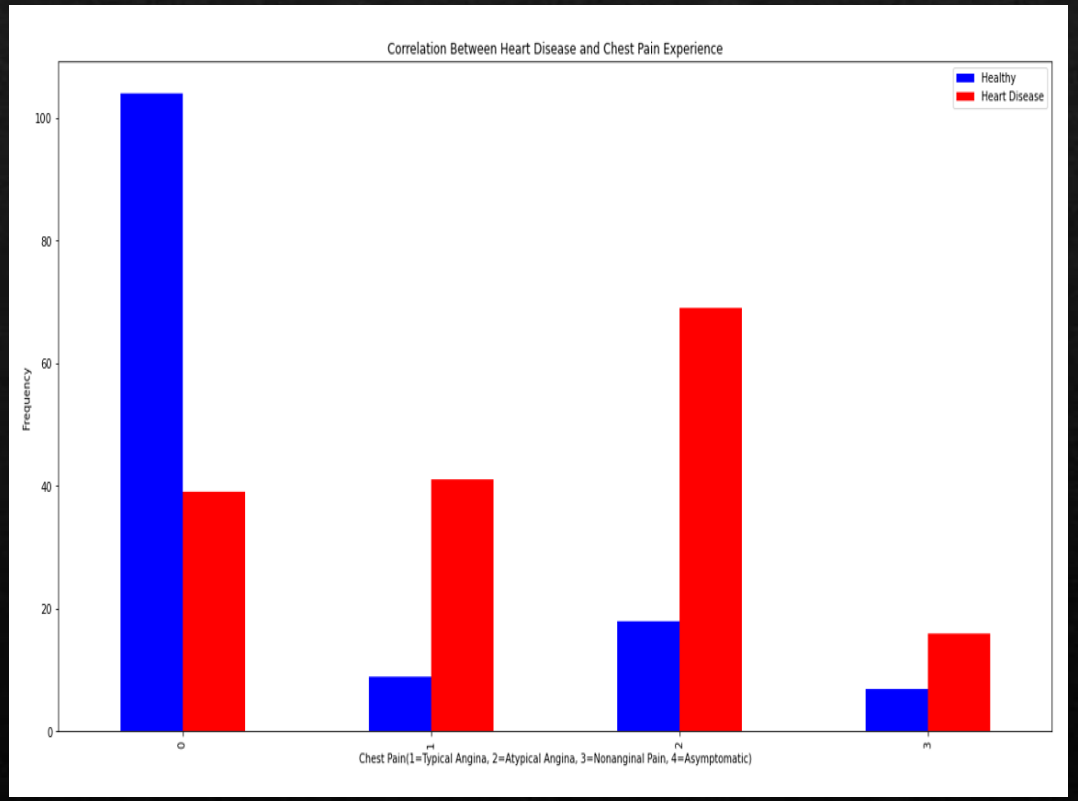
Data Exploration

- **Correlation of heart disease with sex:**
  - Male patients are twice as likely to have heart disease compared to females
  - Given sample size, females are perceived with heart disease rather than healthy



Data Exploration

- **Correlation of heart disease with chest pain:**
  - Most significant type being nonanginal chest pain
  - Each chest pain type associating with heart diseases is evident



Neural Network Architecture

- Imported from Keras, a sequential model is made after splitting the dataset into train and test.
- A second hidden layer was made before the output layer, which are connected using “Dense”

```
Network architecture is defined!
Model: "sequential"
_________________________________________________________________
 Layer (type)                 Output Shape              Param #
=================================================================
 dense (Dense)                (None, 11)                209

 activation (Activation)      (None, 11)                0

 dense_1 (Dense)              (None, 11)                132

 activation_1 (Activation)    (None, 11)                0

 dense_2 (Dense)              (None, 1)                 12

=================================================================
Total params: 353
Trainable params: 353
Non-trainable params: 0
```

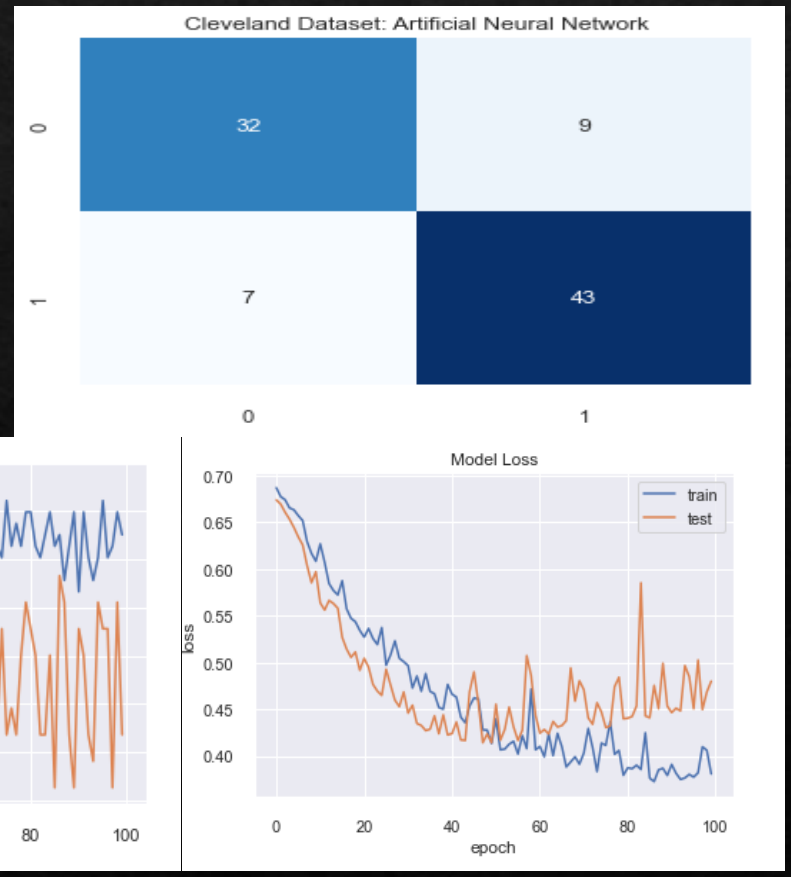
Neural Network Architecture

- Using model compile, the model was compiled with parameters including:
  - Loss, which was set to categorical_crossentropy
  - Metrics, which was set to accuracy
  - Optimizer, being set to adam
- At the end, validation accuracy is 0.8242 and accuracy is 0.8066

```
Epoch 98/100

22/22 [==============================] - 0s 3ms/step - loss: 0.3810 -
accuracy: 0.8396 - val_loss: 0.4602 - val_accuracy: 0.7802

Epoch 99/100

22/22 [==============================] - 0s 3ms/step - loss: 0.3781 -
accuracy: 0.8349 - val_loss: 0.5041 - val_accuracy: 0.7473

Epoch 100/100

22/22 [==============================] - 0s 3ms/step - loss: 0.3868 -
accuracy: 0.8066 - val_loss: 0.4563 - val_accuracy: 0.8242

model is fit!
```

Neural Network Architecture

- Lastly, a confusion matrix was made to determine the accuracy of the model which was found to be 79.12%
- The accuracy was then defined with a model accuracy graph
- Furthermore, a model loss graph was made to describe the NNA



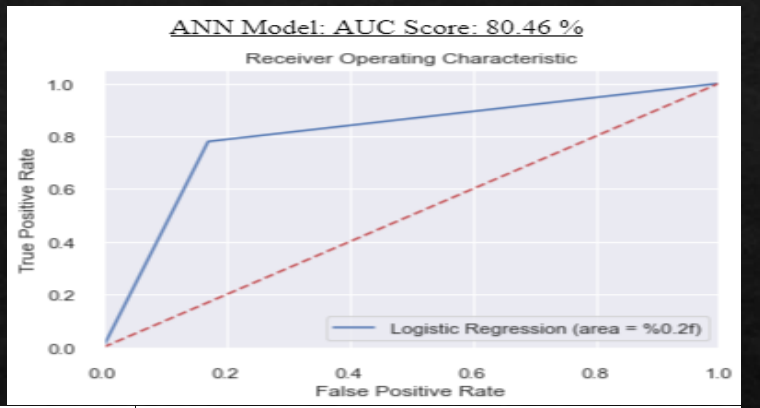
Parameter Sensitivity Analysis

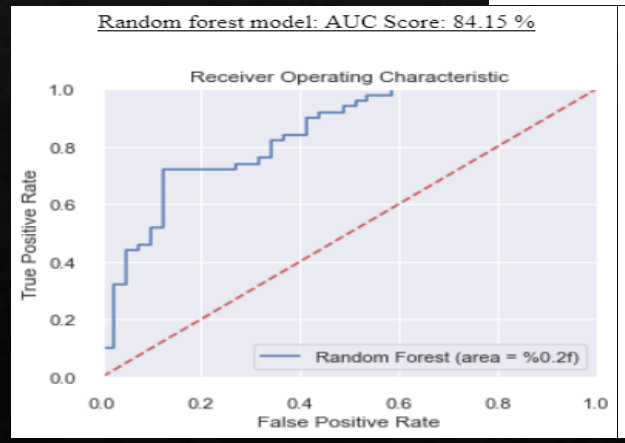
- After reexamining, the optimal number of epoch is found to be 7 where accuracy peaks at 0.9
- After epoch 7 the accuracy drops to approximately 0.5 but peaks again at 25
- Looking over, epoch 50 has a decreased accuracy, which keeps declining the more it is used even when the epoch is significantly more than the usage

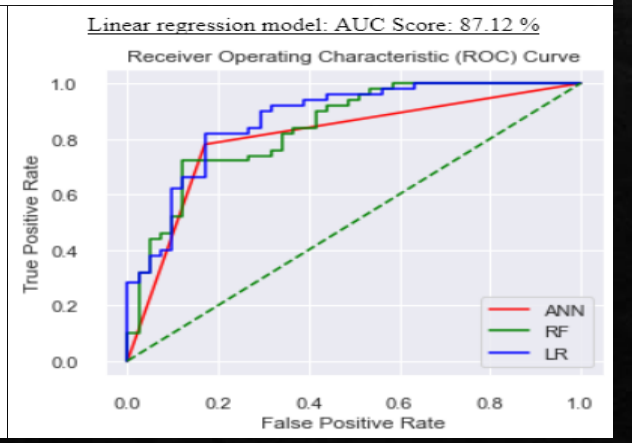
```
Epoch 7/50
22/22 [==============================] - ETA: 0s - loss: 2.3842e-08 -
accuracy: 0.90 - 0s 2ms/step - loss: 6.6352e-08 - accuracy: 0.5896 - val_loss:
6.1570e-08 - val_accuracy: 0.5714
.
.
.
Epoch 25/50
22/22 [==============================] - ETA: 0s - loss: 4.7684e-08 -
accuracy: 0.60 - 0s 2ms/step - loss: 6.6352e-08 - accuracy: 0.5896 - val_loss:
6.1570e-08 - val_accuracy: 0.5714
.
.
.
Epoch 50/50
22/22 [==============================] - 0s 3ms/step - loss: 6.6352e-08 -
accuracy: 0.5425 - val_loss: 6.1570e-08 - val_accuracy: 0.5714
```

AUC – ROC Analysis

- **ROC Analysis performed on:**
  - ANN model which had an AUC score of 80.46%
  - Random Forest model which had an AUC score of 84.15%
  - Linear Regression model which had an AUC score of 87.12%
- Therefore, the models have a high percentage of correct predictions







Conclusion

- Our findings from the dataset determines the important variables to predict heart diseases
- In our model, sex was a major factor in predicting a possible heart disease
- Even though, the older a patient gets the higher chance of heart disease, our data contradicts with this fact
- Lastly, the type of chest pain was also a factor in determining a heart disease with the main types being atypical angina and nonanginal pain