Section III: Results

The seizure detection model was evaluated using several metrics such as training and validation accuracy, training and validation loss, precision, recall, and F1-score. As seen in Figure 1, both the validation loss and training loss are decreasing at about the same rate across all epochs with it beginning to plateau around epoch twenty-two. Inversely, as seen in Figure 2, both the validation accuracy and training accuracy are increasing at roughly the same rate but similarly, they began to plateau at around epoch twenty-two. Much like the validation and training accuracy, the F1-score also increased over time with the epochs. As seen in Figure 3 and as discussed earlier, there is a positive linear correlation between F1-score and epoch number. There is also statistically significant evidence to suggest this since the p-value is less than 0.05. In Figure 4, it is seen that the model's recall and precision slightly increase over epochs while maintaining a higher recall than precision – prioritizing detecting seizures.

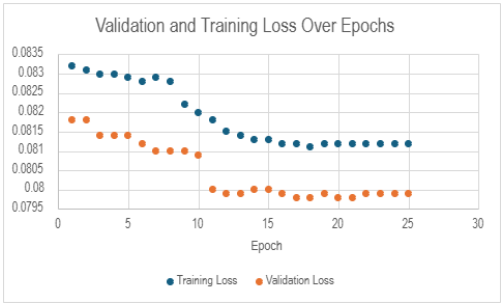

Figure 2: Validation and Training Loss Over Epochs. Twenty-five epochs were run when training, validating, and testing the LSTM. Starting values of 0.0832 training loss and 0.0818 validation loss. Ending values of 0.0812 and 0.0799

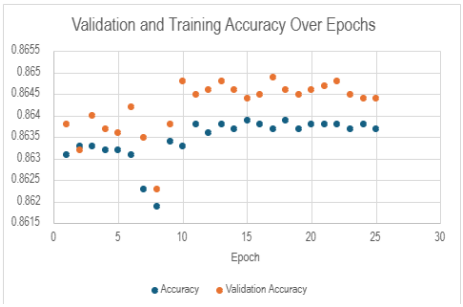

Figure 1: Validation and Training Accuracy Over Epochs. Starting values of 0.8631 training accuracy and 0.8638 validation accuracy. Ending values of 0.8637 and 0.8644.

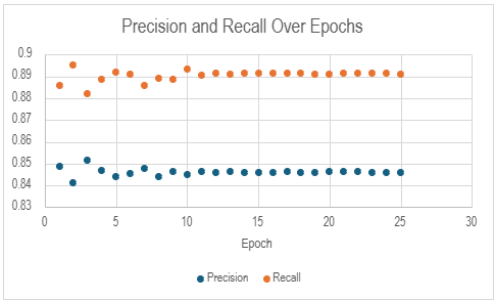

Figure 4: Precision and Recall Over Epochs. Starting recall of 0.8857 and 0.8486 precision. Ending values of 0.8912 and 0.8459

Figure 3: F1-score Over Epochs. Starting F1-score of 0.8663 and ending F1-score of 0.8674