**VISVESVARAYA TECHNOLOGICAL UNIVERSITY**

“Jnana Sangama”, Belagavi-18, Karnataka, India



**INTERNSHIP REPORT**

**BACHELOR OF ENGINEERING**

*in*

**MEDICAL ELECTRONICS ENGINEERING**

**INTERNSHIP MONTHLY REPORT**

**during academic year 2024 – 2025**

*By*

|  |  |
| --- | --- |
| Adwin Justa Manuel | 1DS21MD004 |



**Department of Medical Electronics Engineering**

**Accredited by National Board of Accreditation (NBA)**

**DAYANANDA SAGAR COLLEGE OF ENGINEERING**

Shavige Malleshwara Hills, Kumaraswamy Layout, Bangalore-560078

An Autonomous Institute affiliated to VTU, Approved by AICTE and UGC, Accredited by NAAC with ‘A’ Grade & ISO 9001:2015 Certified Institution

***INTERNSHIP REPORT***

**Introduction:**

The Breathing dataset was classified into 3 types:

1. Normal Breathing
2. Endurance Breathing
3. Recovery Breathing

* Normal Breathing – Slow and regular, breathing in and out through the nose only is called Normal Breathing For an adult normal breathing is 12 to 20 breaths per minute.

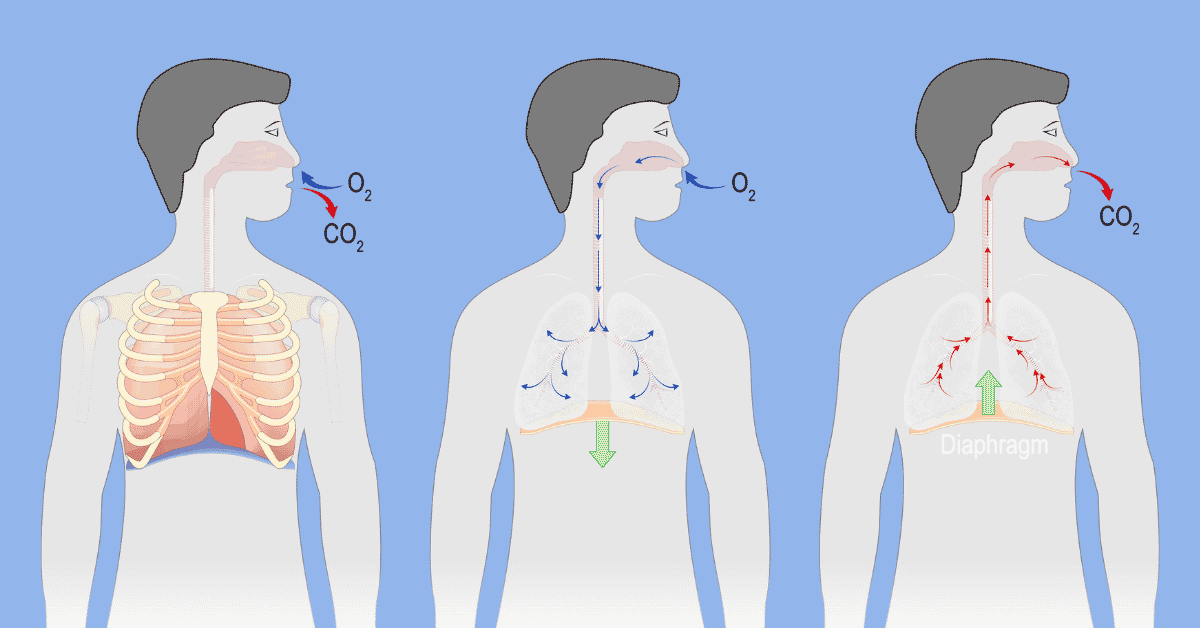


Fig1: Normal Breathing (Source: Respiratory Therapy Zone)

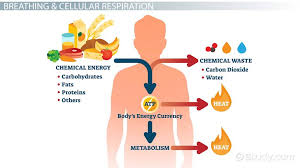
* Endurance Breathing - The ability of the respiratory system to efficiently deliver oxygen to working muscles and remove carbon dioxide during physical activity like exercise,

Fig2: Endurance Breathing(Source:Study.com)

* Recovery Breathing: This helps the body return to resting state and regulate breathing after intense physical activity, often involves slow, deep breaths to increase oxygen intake and promotes relaxation also called as box

Fig3:Recovery Breathing(Source: Medical News Today)

**Objectives:**

* To remove baseline wandering and outliers from the signal and store this as cleaned dataset.
* To Train this dataset in models such as RNN (Recurrent Neural Network), LSTM (Long Short Term Memory) and GRU (Gated Recurrent unit) and to analyse which is the best model for this particular dataset from ROC (Receiver Operating Characteristics) curve and AUC(Area Under Curve).

**Software Used:** Python.

**IDE used:** Jupyter Notebook.

**Methodology:**

1. Combining all the .csv files into one single file called as Main\_Data.csv file and adding a separate column as label where the Activity is labelled as follows:
   * + - * N 🡪 Normal Breathing.
         * E 🡪 Endurance Breathing.
         * R 🡪 Recovery Breathing.
2. Removal of outliners and training the data using simple RNN , LSTM & GRU models
3. Removal of baseline wandering and training the data using simple RNN, LSTM & GRU models.
4. Removal of both baseline wandering and outliners of the dataset and training the data using simple RNN , LSTM & GRU models.

**Removal of outliners and training the data using simple RNN, LSTM and GRU models.**

* **Simple RNN:**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove outliners interquartile method is implemented.
* Step 3 🡪 After removal of outliners this data is saved as .csv file where the name is Main\_Data\_outliners\_Removed.csv.
* Step 4 🡪 Build simple RNN.
* Step 5 🡪 Use step 3 dataset to train these models ,while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss , accuracy , f1 score and specificity.

**Basic Definitions**:

1. Precision: It measures the proportion of true positives among all predicted positives.

Formula:

1. Accuracy: it’s the fundamental metric for evaluating the performance of a classification model, providing an overview of how well the model is performing in terms of correct predictions. This is calculated as the ration of correct predictions to the total number of inputs samples.
2. Recall: Measures the proportion of true positives among all actual positives..

Formula:

1. F1 score: It’s the harmonic medium between precision and recall,providing a good balance between the trade-oof between precision and recall.
2. Loss: A numerical metric that describes how wrong a models predictions are. It describes the distance between the models predictions and the actual labels.
3. True Positive Rate: TPR is considered as a portion of positive data points that are correctly considered as positive, with respect to all data points that are ppositive.
4. True Negative Rate: its also know as specificity, it quantifies the portion of actual negatives that are correctly identified by a classification model.

1. False positive rate: its actually the proportion of actual positives that are incorrectly identified as negatives

**Result Obtained:**

* Confusion Matrix:

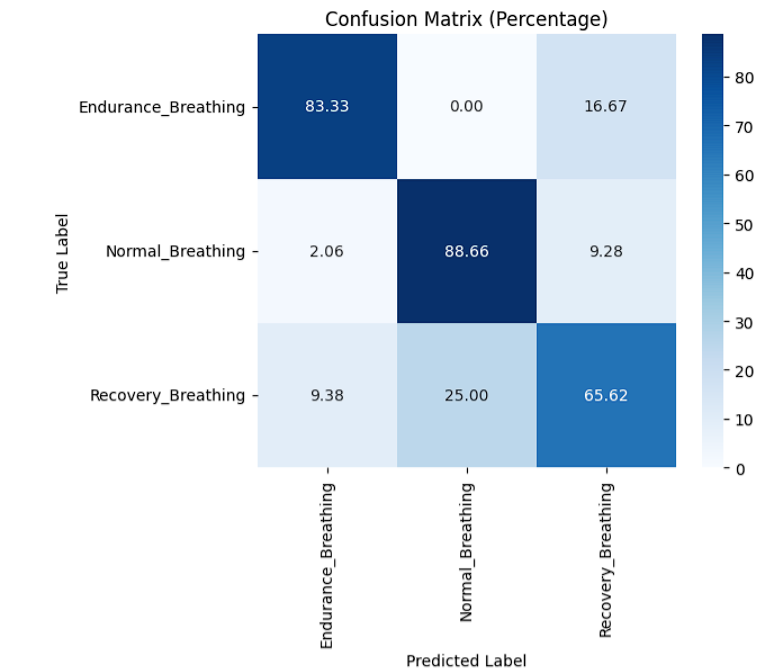


Fig 4: Confusion Matrix for the simple RNN.

1. Endurance Breathing :
   * + 83.33% correctly classified as Endurance breathing.
     + 0% misclassified as Normal breathing.
     + 16.67% misclassified as Recovery Breathing.
2. Normal Breathing:
   * + 88.66% correctly classified as Normal breathing
     + 2.06% misclassified as Endurance Breathing.
     + 9.28% misclassified as Recovery Breathing.
3. Recovery Breathing
   * + 65.62% correctly classified as Recovery breathing.
     + 9.38% misclassified as Endurance breathing.
     + 25% misclassified as Normal breathing.

**Calculations:**

1. Endurance Breathing Class:

* True Positive (TP) = 83.33% (From the diagonal of Confusion Matrix).
* False Positive (FP) = 2.06% + 9.38 % = 11.44% (other classes misclassified as Recovery Breathing).
* False Negative (FN) = 16.67% (Endurance class misclassified aas Recovery Breathing).
* True Negative (TN) =188.5% (correctly classified other class).
* Precision = 87.92%.
* Recall = 83.33%.
* F1 score = 85.58%.
* Specificity= 94.27%.

1. Normal Breathing:

* TP = 88.66%
* FP = 25%
* FN = 11.34%
* TN = 175%

(From the formulas given in Basic definitions):

* Precision = 77.99%
* Recall = 88.66%
* F1 Score = 83%
* Specificity = 87.50%

1. Recovery Breathing :

* TP = 65.62%
* FP = 25.95%
* FN = 34.38%
* TN = 174.05%
* Precision = 71.67%
* Recall = 65.62%
* F1-Score = 68.51%
* Specificity = 87%

**Model Performance of RNN**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 87.92% | 83.33% | 94.27% | 85.58% |
| Normal Breathing | 77.99% | 88.66% | 87.50% | 83% |
| Recovery Breathing | 71.67% | 65.62% | 87% | 68.51% |

ROC Curve:

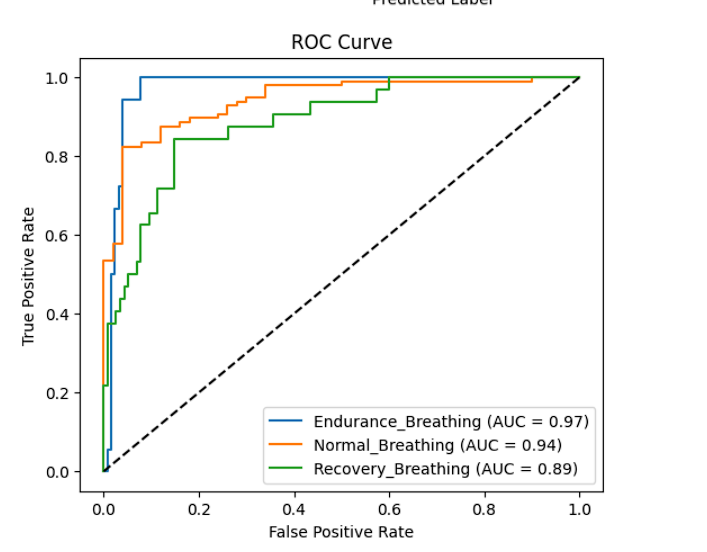


Fig 5: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from RNN Model.

* The closer the curve is to the top left corner, better the model is at differentiating between the classes.
* Higher AUC values indicate better classification performance.
* Endurance Breathing – AUC is 0.97 which indicates that it’s the best classification performance.
* Normal Breathing – AUC is 0.94 which indicates its the Good Classification.
* Recovery Breathing- AUC IS 0.89 which indicates Slightly lower performance.
* The dashed diagonal line represents a random classifier AUC is 0.5, if the model is significantly above this line, indicates strong predictive performance.
* **LSTM Model:**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove outliners interquartile method is implemented.
* Step 3 🡪 After removal of outliners this data is saved as .csv file where the name is Main\_Data\_outliners\_Removed.csv.
* Step 4 🡪 Build LSTM Model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity.

**Confusion Matrix:**

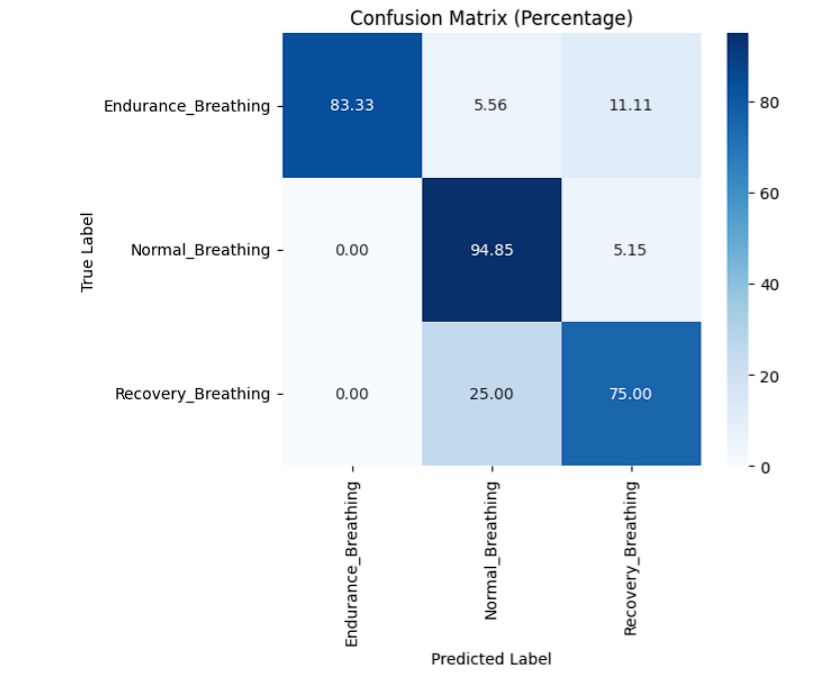


Fig 6: Confusion Matrix for the LSTM model.

**Observations:**

1. Endurance Breathing (Actual):

* 83.33% correctly classified as Endurance Breathing. (TP)
* 5.56% misclassified as Normal Breathing. (FP)
* 11.11% misclassified as Recovery Breathing. (FN)

1. Normal Breathing (Actual):

* 94.85% correctly classified as Normal Breathing (TP).
* 0% misclassified as Endurance Breathing (FN).
* 5.15% misclassified as Recovery Breathing (FN).

1. Recovery Breathing (Actual):

* 75 % correctly classified as Recovery Breathing (TP).
* 25 % misclassified as Normal Breathing (FN).
* 0 % misclassified as Endurance Breathing (FN).

**Model Performance of LSTM**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 100% | 83.33 | 94.62% | 90.91% |
| Normal Breathing | 75.63% | 94.85% | 89.74% | 84% |
| Recovery Breathing | 82.18% | 75% | 87.81% | 78.43% |

**ROC curve:**

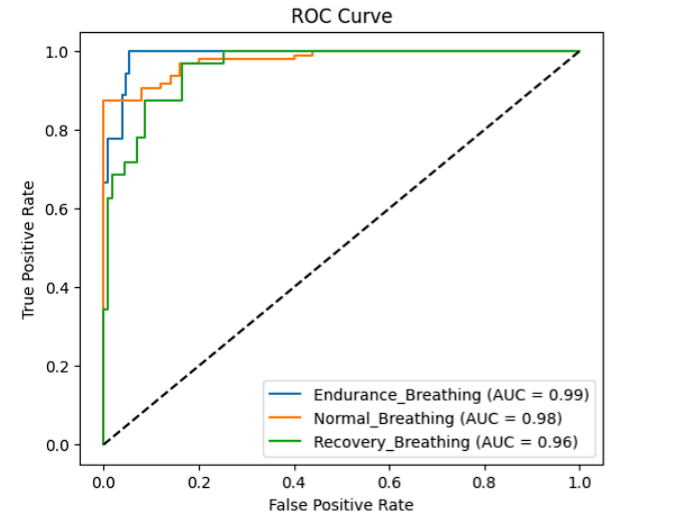


Fig 7: ROC curve of Endurance Breathing ,Normal Breathing ,Recovery Breathing from LSTM Model.

**Observations:**

* Endurance Breathing – AUC score is 0.99, Excellent performance, nearly perfect classification. The curve is closest to the top left corner, showing a **high true positive rate** with a **low False Positive rate**.
* Normal Breathing – AUC score is 0.98, Very strong performance ,slight misclassification but overall robust. There are minor misclassifications.
* Recovery Breathing- AUC score is 0.96, Strong Performance, but slightly lower than the other two. The model finds it a bit harder to differentiate Recovery Breathing from Other classes.

**GRU MODEL:**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove outliners interquartile method is implemented.
* Step 3 🡪 After removal of outliners this data is saved as .csv file where the name is Main\_Data\_outliners\_Removed.csv.
* Step 4 🡪 Build GRU Model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity

**Confusion matrix:**

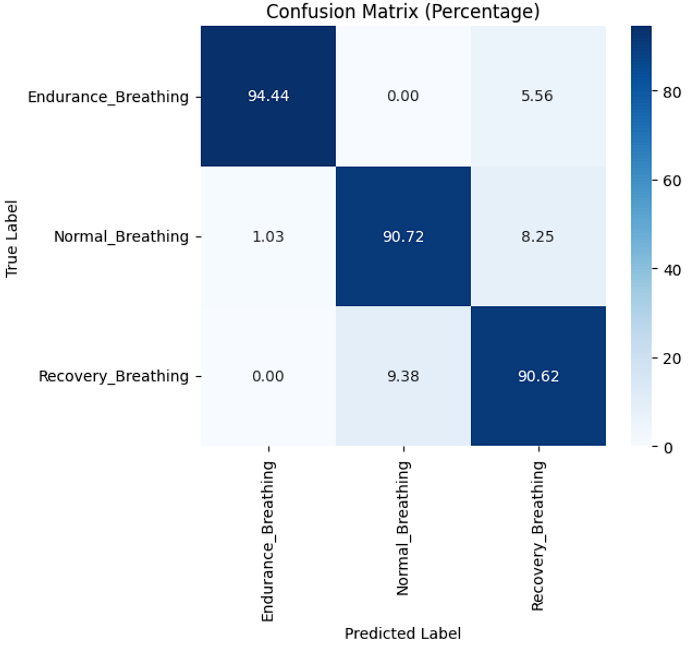


Fig 8: Confusion Matrix for the GRU model.

* Endurance Breathing –
* Correctly classified as 94.44% (TP).
* Misclassified as Recovery Breathing.5.56% (FN).
* Normal Breathing –
* Correctly classified as 90.72%.
* 8.25% were classified as Recovery Breathing.
* 1.03% were classified as Endurance Breathing.
* Recovery Breathing –
* Correctly classified as 90.62%.
* 9.38 % were misclassified as Normal Breathing.

**Model Performance of LSTM**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 98.92% | 94.44% | 99.25% | 96.62% |
| Normal Breathing | 90.63% | 90.72% | 94.28% | 90.67% |
| Recovery Breathing | 91.09% | 90.62% | 94.73% | 90.86% |

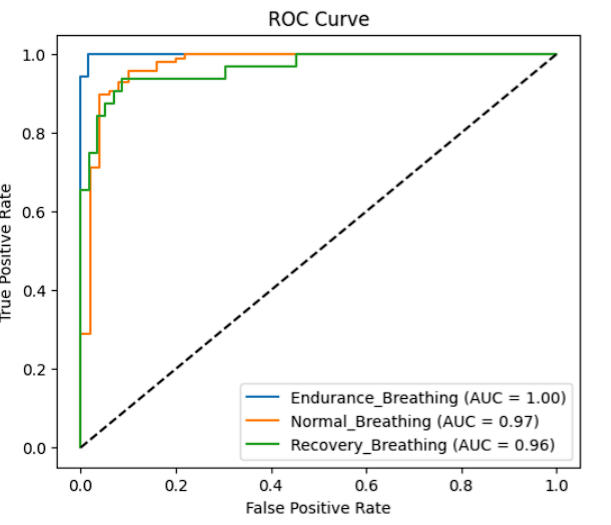
**ROC curve:**

Fig 9: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from GRU Model.

* Endurance Breathing – The curve reaches the top left corner, indicating perfect classification with an AUC of 1.00. This means the model never makes false predictions for Endurance Breathing.
* Normal Breathing-The AUC score is 0.97. The curve is slightly below the perfect classification line , which means some false positive or false negatives may occur.
* Recovery Breathing – The AUC of 0.96 shows that it’s a highly reliable classification, but it is slightly lower than the normal and endurance Breathing, There are slightly more misclassifications compared to Endurance Breathing.

**Removal of Baseline Wandering and training the data using simple RNN, LSTM and GRU models.**

**LSTM Model:**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove Baseline Wandering using High pass filter.
* Step 3 🡪 After removal of Baseline Wandering this data is saved as .csv file where the name is Main\_Data\_BW\_Removed.csv.
* Step 4 🡪 Build simple LSTM model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity.

**Confusion matrix:**

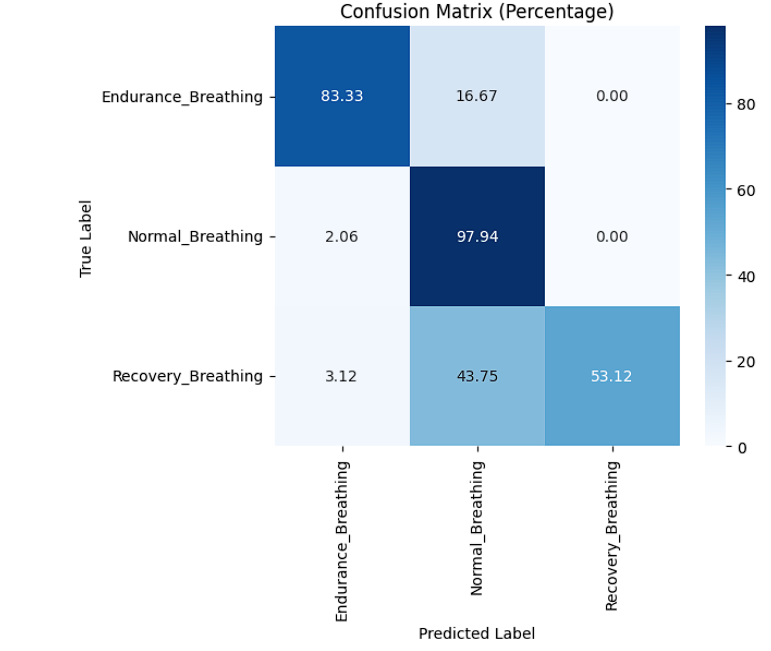
****

Fig 10: Confusion Matrix for the Simple LSTM model.

* Endurance Breathing:
* 83.33% of Endurance Breathing samples were correctly classified.
* 16.67% were misclassified as Normal Breathing.
* Normal Breathing:
* 97.94% OF Normal Breathing samples were correctly classified.
* 2.06% were misclassified as Endurance Breathing.
* Recovery Breathing:
* 53.12% of Recovery Breathing samples were correctly classified.
* 43.75% were misclassified as Normal Breathing, showing a high level of confusion between these two classes.
* 3.12% were misclassified as Endurance Breathing.

**Model Performance of LSTM**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 94.21% | 83.33% | 98.97% | 88.42% |
| Normal Breathing | 85.71% | 97.94% | 77.85% | 91.44% |
| Recovery Breathing | 100% | 53.12% | 76.22% | 69.39% |

**ROC curve:**

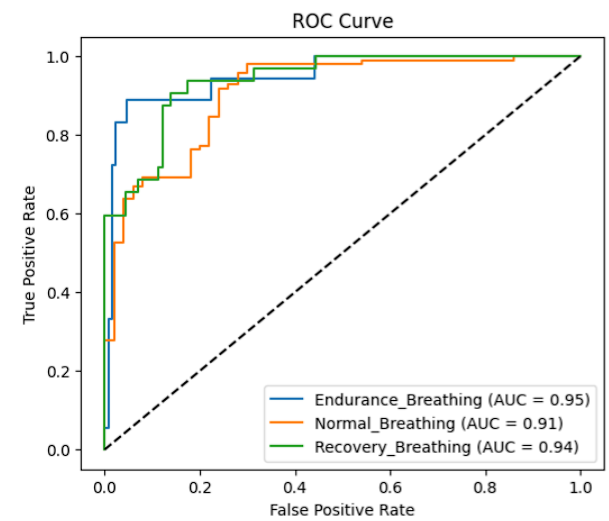
****

Fig 11: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from LSTM model.

* Endurance Breathing – The AUC score is 0.95, The model is highly confident in distinguishing endurance breathing from other classes. This curve remains close to the top left corner, indicating low false positives and high true positives.
* Recovery Breathing – The AUC score is 0.94, The model performs well but may have some slight confusion with other breathing types. The curve is still close to the ideal classification boundary.
* Normal Breathing – The AUC score is 0.91, The lowest AUC of the three, indicating some overlap with other classes. The model is still very good but slightly less accurate in distinguishing Normal Breathing.

**Simple RNN MODEL:**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove Baseline Wandering using High pass filter.
* Step 3 🡪 After removal of Baseline Wandering this data is saved as .csv file where the name is Main\_Data\_BW\_Removed.csv.
* Step 4 🡪 Build simple RNN model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity.

**Confusion Matrix:**

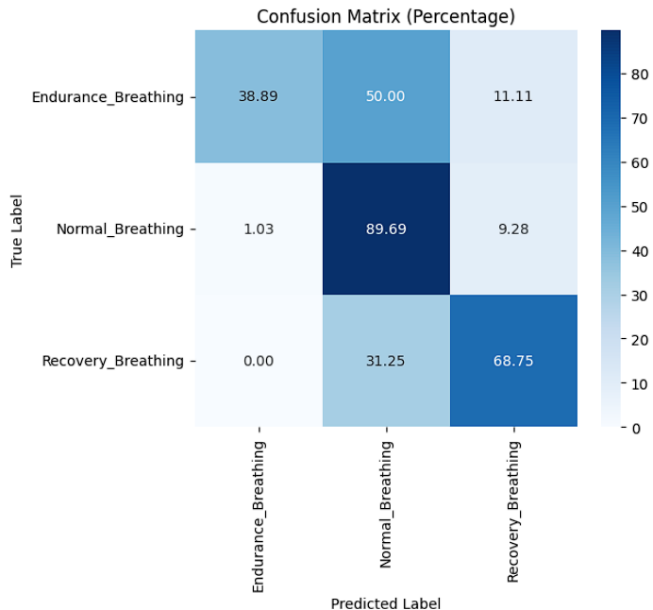
****

Fig 12: Confusion Matrix for the Simple RNN model.

* Endurance Breathing:
* 38.89% correctly classified as Endurance Breathing.
* 50% misclassified as Normal Breathing.
* 11.11% misclassified as Recovery Breathing.
* Normal Breathing:
* 89.69% is correctly classified as Normal breathing.
* 1.03% is misclassified as Endurance Breathing.
* 9.28% is misclassified as Recovery Breathing.
* Recovery Breathing:
* 68.75% is correctly classified as Recovery breathing.
* 31.25% is misclassified as Normal Breathing.

**Model Performance of Simple RNN**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 97.42% | 38.89% | 99.49% | 55.59% |
| Normal Breathing | 52.47% | 89.69% | 59.38% | 66.21% |
| Recovery Breathing | 77.13% | 68.75% | 89.81% | 72.70% |

**ROC curve:**

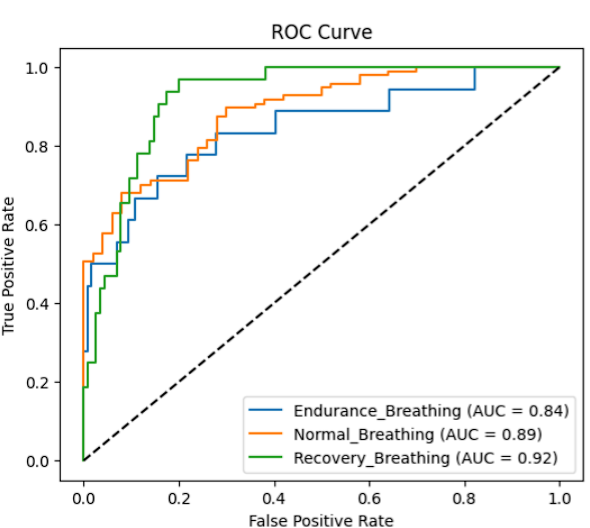
****

Fig 13: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from RNN model.

* Endurance Breathing: The AUC score is 0.84 which is the lowest, indicating that the distinguish between the other classes is not accurate.
* Normal Breathing: The AUC score is 0.89 which shows a good balance between sensitivity and specificity.
* Endurance Breathing : The AUC score is 0.92 which indicates that the mode, struggles more with distinguishing it from the other classes.

**GRU Model:**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove Baseline Wandering using High pass filter.
* Step 3 🡪 After removal of Baseline Wandering this data is saved as .csv file where the name is Main\_Data\_BW\_Removed.csv.
* Step 4 🡪 Build simple GRU model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity.

**Confusion Matrix:**

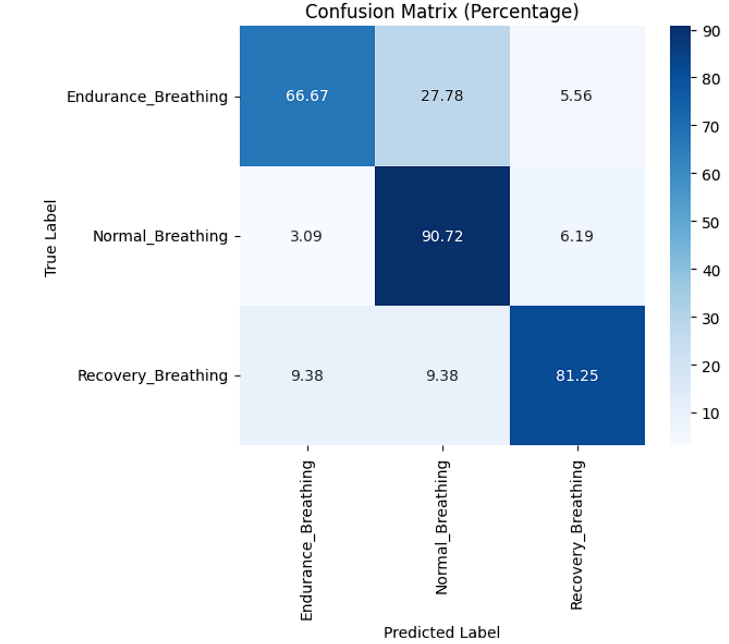


Fig 14: Confusion Matrix for the Simple GRU model.

* Endurance Breathing- 66.67% correctly classified, Endurance Breathing misclassified as normal breathing which is 27.78% and Endurance Breathing is misclassified as Recovery breathing which is about 5.56%.
* Normal Breathing- 90.72% correctly classified , Normal Breathing misclassified as Endurance Breathing about 3.09%and Normal Breathing misclassified as Recovery Breathing about 6.19%.
* Recovery Breathing- 81.25% correctly classified, Recovery Breathing misclassified as Endurance Breathing about 9.38%, Recovery Breathing is misclassified as Normal Breathing as 9.38%..

**Model Performance of Simple RNN**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 78% | 67% | 85% | 74% |
| Normal Breathing | 92% | 91% | 95% | 91% |
| Recovery Breathing | 87% | 81% | 91% | 85% |

**ROC curve:**

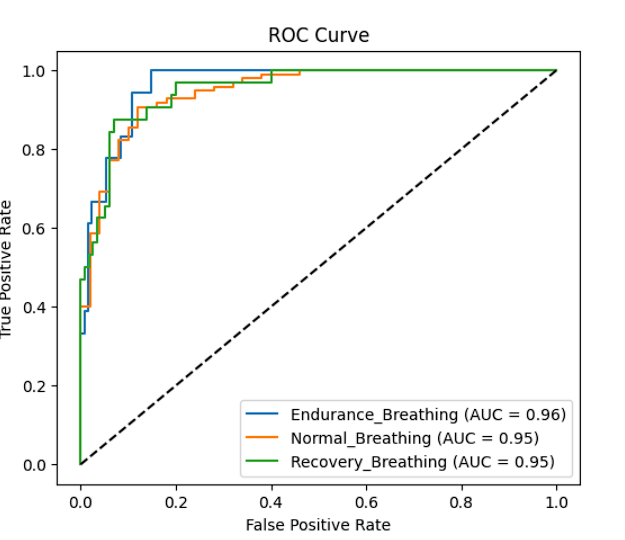
****

Fig 15: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from GRU model

* Endurance Breathing-The AUC score is 0.96, this model can identify endurance breathing with high accuracy and minimal misclassification.
* Normal Breathing – The AUC Score is 0.95, the classifier performs well in distinguishing normal breathing from the other categories.
* Recovery Breathing – The AUC score is 0.95, The model effectively classifies recovery breathing, with very few false positives.

**Removal of Baseline Wandering and Removal of Outliers training the data using simple RNN, LSTM and GRU models**

**Simple RNN Model:**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove Baseline Wandering using High pass filter.
* Step 3 🡪 After removal of Baseline Wandering and Outliners this data is saved as .csv file where the name is Main\_Data\_BW \_&\_outliners\_Removed.csv.
* Step 4 🡪 Build simple RNN model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity

**Confusion Matrix:**

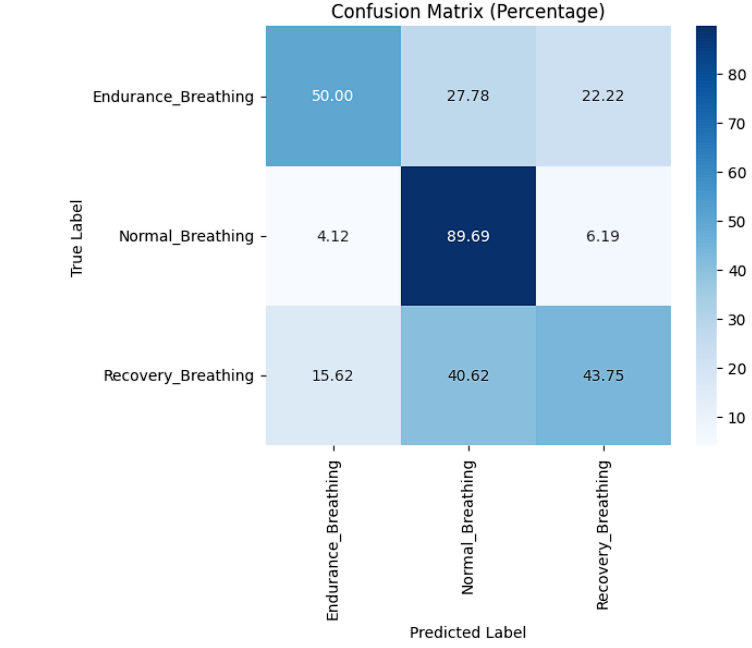
****

Fig 16: Confusion Matrix for the Simple RNN model.

* Endurance Breathing- 50% of endurance breathing samples were correctly classified, 27.78% were misclassified as normal breathing, 22.22% were misclassified as recovery breathing.
* Normal Breathing- 89.69% of normal breathing samples were correctly classified, 4.12% were misclassified as endurance breathing, 6.19% were misclassified as recovery breathing.
* Recovery Breathing- 43.75% of recovery breathing samples were correctly classified,40.62% were misclassified as normal breathing,15.62% were misclassified as endurance breathing.

**Model Performance of Simple RNN**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 71.7% | 50% | 90.1% | 59.0% |
| Normal Breathing | 56.8% | 89.7% | 65.8% | 69.2% |
| Recovery Breathing | 60.6% | 43.8% | 85.8% | 50.9% |

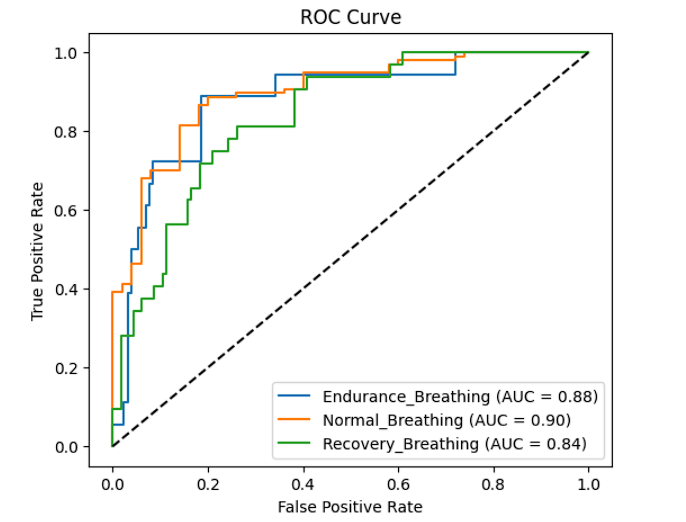
**ROC curve:**

Fig 17: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from simple RNN model.

* Normal Breathing – The AUC score is 0.90, this indicates a highly effective classification.
* Endurance Breathing- The AUC score is 0.88, this indicates it’s performance is slightly lower than normal breathing.
* Recovery Breathing – The AUC score is 0.84, this indicates more classification errors compared to other two classes.

**LSTM Model**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove Baseline Wandering using High pass filter.
* Step 3 🡪 After removal of Baseline Wandering and Outliners this data is saved as .csv file where the name is Main\_Data\_BW \_&\_outliners\_Removed.csv.
* Step 4 🡪 Build LSTM model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity

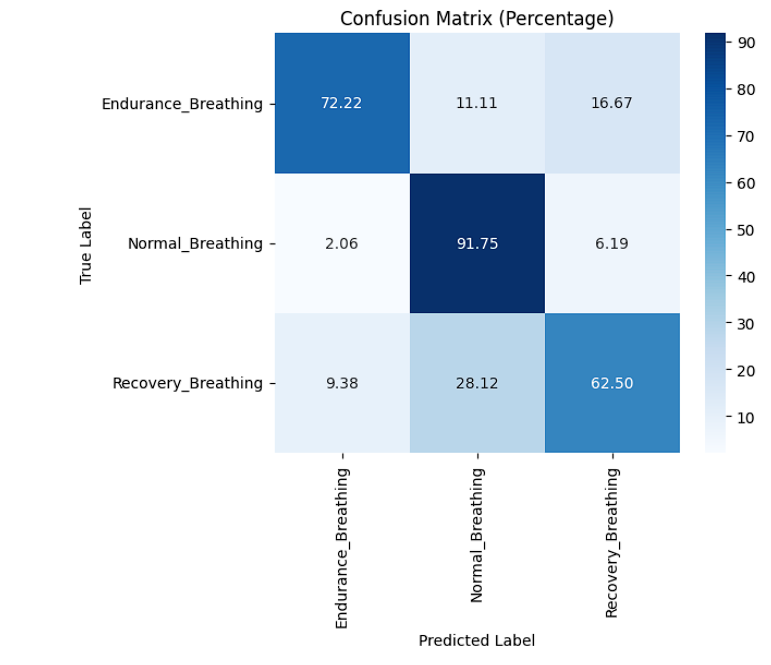
**Confusion Matrix**

Fig18: Confusion Matrix for the Simple RNN model

* **Endurance Breathing**
* **72.22% correctly classified** as Endurance Breathing.
* 11.11% misclassified as Normal Breathing.
* 16.67% misclassified as Recovery Breathing.
* **Normal Breathing (Best Performance)**
* 91.75% correctly classified.
* 2.06% misclassified as Endurance Breathing.
* 6.19% misclassified as Recovery Breathing**.**
* **Recovery Breathing-**
* 62.50% correctly classified.
* 28.12% misclassified as Normal Breathing.
* 9.38% misclassified as Endurance Breathing.

**Model Performance of Simple RNN**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 86.7% | 72% | 94.5% | 78.7% |
| Normal Breathing | 70.2% | 92% | 80.4% | 79.7% |
| Recovery Breathing | 72.9% | 62.26% | 88.5% | 67.4% |

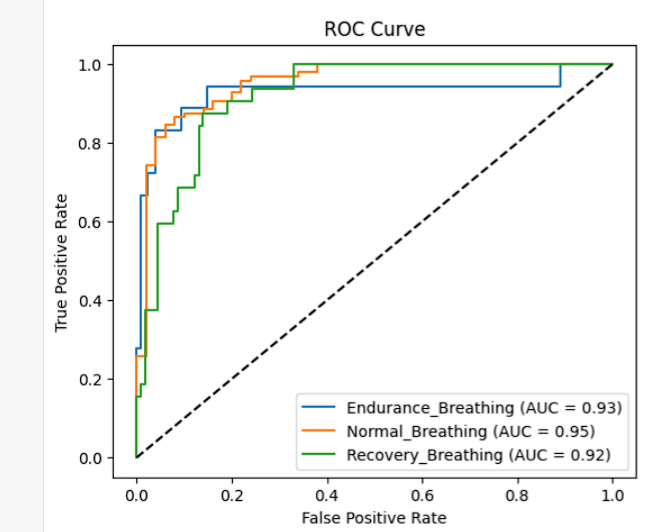
**ROC curve:**

Fig 19: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from simple LSTM model.

* Endurance Breathing – The AUC score is 0.93, indicating high classification accuracy for endurance breathing.
* Normal Breathing – The AUC score is 0.95, The model is highly effective in detecting normal breathing instances with minimal misclassification.
* Recovery breathing-The AUC score is 0.92, While slightly lower than the other two classes, the curve remains well above the diagonal line (random guessing). Some overlap with other classes might occur, leading to a higher false positive rate compared to normal breathing.

**GRU Model :**

**Algorithm:**

* Step 1 🡪 Using the dataset which is combined into one .csv file.
* Step 2 🡪 To remove Baseline Wandering using High pass filter.
* Step 3 🡪 After removal of Baseline Wandering and Outliners this data is saved as .csv file where the name is Main\_Data\_BW \_&\_outliners\_Removed.csv.
* Step 4 🡪 Build GRU model.
* Step 5 🡪 Use step 3 dataset to train these models, while training the data is segmented to 150 samples per segment because we take time as 10 sec and frequency as 15 Hz.
* Step 6 🡪 Display Confusion Matrix and ROC curve.
* Step 7 🡪 Print model performance table such as loss, accuracy, f1 score and specificity

**Confusion Matrix:**

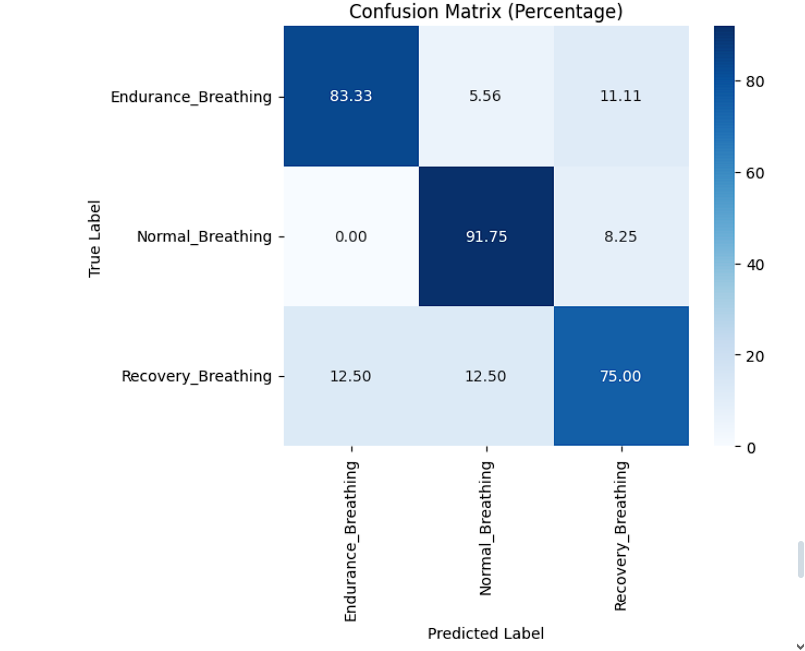
****

Fig 20: Confusion Matrix for the GRU model.

* Endurance Breathing –
* **83.33%** of Endurance Breathing samples were correctly classified.
* **5.56%** were misclassified as Normal Breathing.
* **11.11%** were misclassified as Recovery Breathing.
* Normal Breathing –
* **91.75%** of Normal Breathing samples were correctly classified.
* **0.00%** were misclassified as Endurance Breathing (perfect separation).
* **8.25%** were misclassified as Recovery Breathing.
* Recovery Breathing-
* **75.00%** of Recovery Breathing samples were correctly classified.
* **12.50%** were misclassified as Endurance Breathing.
* **12.50%** were misclassified as Normal Breathing.

**Model Performance of GRU Model**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | Specificity | F1-score |
| Endurance Breathing | 87% | 83.3% | 93.8% | 85.1% |
| Normal Breathing | 83.6% | 91.8% | 91% | 87.5% |
| Recovery Breathing | 79.5% | 75% | 90.3% | 77.2% |

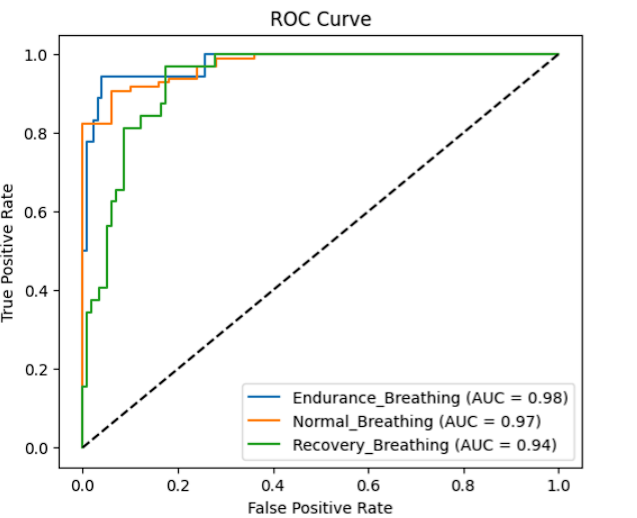
**ROC Curve:**

Fig 21: ROC curve of Endurance Breathing, Normal Breathing, Recovery Breathing from simple LSTM model.

* Endurance Breathing- The AUC score is 0.98, meaning it is the best at distinguishing between endurance breathing and the other classes.
* Normal Breathing – The AUC score, slightly below the endurance breathing curve, meaning it is slightly less effective at distinguishing normal breathing but still highly reliable.
* Recovery Breathing- The AUC score is 0.94, indicating a slightly higher misclassification rate compared to the other two. The curve still remains well above the diagonal, indicating good classification ability.

**Conclusion:**

The dataset was classified into three types of breathing (Normal, Endurance, and Recovery) using deep learning models (RNN, LSTM, GRU) and evaluating their performance after preprocessing steps (Outliner removal, Baseline Wandering correction and combination of both), so based on this comparison study is as follows:

1. **Model Performance:**

* **GRU** consistently outperformed other models, achieving the highest AUC scores (e.g.**0.98–1.00** for Endurance Breathing) and robust F1-scores across classes, especially when both preprocessing steps were applied.
* **LSTM** demonstrated strong performance, closely trailing GRU, with high AUC values (e.g., **0.92–0.99**).
* **Simple RNN** lagged behind, struggling with misclassifications (e.g., Recovery Breathing AUC dropped to **0.84**), likely due to its limitations in handling sequential dependencies.

1. **Impact of Preprocessing:**

* Removing **outliers** improved precision and recall for all models.
* **Baseline wandering removal** enhanced specificity, particularly for Normal and Recovery Breathing.
* Combining **both preprocessing steps** yielded the most balanced results, with GRU achieving **AUC > 0.94** for all classes.

1. **Class-Specific Challenges:**

* **Recovery Breathing** was the hardest to classify, often confused with Normal Breathing (e.g., 43.75% misclassification in RNN).
* **Endurance Breathing** was the most distinguishable, with near-perfect AUC scores in GRU and LSTM models.