**AMPLIFIER HEALTH PROJECT REPORT**

**Tabish Saeed**

**COPD vs Healthy Binary Classification**

**1. Dataset**

The ICBHI dataset provides audio recordings of respiratory sounds captured under various conditions and acquisition modes. The dataset includes recordings labeled with diagnoses such as COPD, Pneumonia, LRTI, URTI and Healthy individuals.

**2. Data Preprocessing**

**2.1 Data Selection**

The recordings were filtered based on diagnostic information in the **diagnosis.txt** file. Only recordings labeled as COPD or Healthy were retained for further analysis.

**2.2 Standardization and Normalization**

To ensure consistency across recordings:

* **Sampling Rate**: All audio files were resampled to a common sampling rate**. 22.5 KHz**
* **Channel Consistency**: Mono channels were enforced across recordings, resolving inconsistencies from different acquisition modes.
* **Normalization**: Amplitude normalization was applied to standardize the volume levels across recordings.

**2.3 Train-Test Split and Segmentation**

* The dataset was divided into training and testing sets using an 80-20 split.
* Each recording was further split into segments of approximately 8 seconds, ensuring that each segment contained at least 2-3 full respiratory cycles.

**3. Exploratory Data Analysis (EDA)**

**3.1 Power Spectral Density (PSD) Plots**

Average Power Spectral Density (PSD) plots were generated for both COPD and Healthy classes to analyze the frequency distribution of respiratory sounds. These plots highlighted distinct frequency patterns between the two classes, providing insights into spectral differences.

A graph with a line graph

Description automatically generated

**3.2 Dimensionality Reduction with t-SNE**

Using extracted features from the audio recordings:

* **t-SNE** was applied to visualize feature distributions in 2D space.
* The plots revealed clustering tendencies, offering a visual understanding of class separability.

A graph with red and blue dots

Description automatically generated

**3.3 Spectrogram Analysis**

Spectrograms were created with cropped frequency limits to focus on the most informative frequency range. These spectrograms provided time-frequency representations of respiratory sounds, aiding in feature extraction and model input preparation.

* + 1. COPD Spectrogram

A blue and green lines

Description automatically generated

* + 1. Healthy Spectrogram
* A blue and green background

  Description automatically generated with medium confidence

**4. Methodology**

Two feature extraction techniques were employed in this project:

**4.1 First Technique**

Multiple features were manually extracted from the list of audio attributes, including spectral properties, MFCCs, and statistical functionals. These features provided a detailed representation of the audio characteristics.

**4.2 Second Technique**

Features were extracted using the OpenSMILE library, resulting in 6373 high-dimensional features. To reduce dimensionality and remove noise, an autoencoder was trained to compress these features into a latent space of 1024 dimensions. This process ensured computational efficiency and preserved relevant information for classification.

For both feature extraction pipelines:

* **SMOTE**: The Synthetic Minority Oversampling Technique (SMOTE) was applied to handle class imbalance by creating synthetic samples for the minority class.
* **XGBoost**: The reduced feature sets were used to train an XGBoost classifier. Hyperparameter optimization was performed to identify the best configuration for the final model.

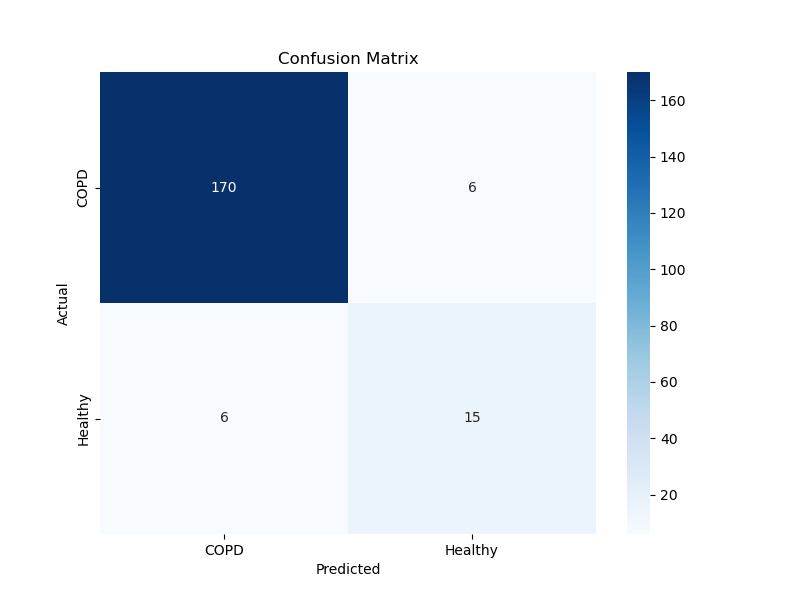
**5. Results and Evaluation**

**5.1 Classification Metrics**

The model’s performance was evaluated using:

* **Confusion Matrix**: To visualize predictions across COPD and Healthy classes.
* **Classification Report**: Providing precision, recall, F1-score, and accuracy metrics.

1. Technique 1



1. Technique 2

A screenshot of a computer screen

Description automatically generated

**Key Observation:** Technique 1 outperformed Technique 2 due to the use of smartly selected features, which were chosen based on insights from exploratory data analysis. This selective approach enabled the extraction of more relevant and impactful features, highlighting the importance of a well-designed feature engineering pipeline in improving model performance.

**6. Future Direction**

Future advancements in this work could focus on leveraging advanced transformer-based models, such as Wav2Vec2.0 or Whisper, which are well-suited for handling variable input lengths and capturing long-range dependencies in respiratory sounds through self-attention mechanisms. By using the entire ICBHI dataset and incorporating additional labels such as asthma, pneumonia, and URTI, the classification task can be expanded to a multi-class or multi-label framework, providing a more comprehensive diagnostic pipeline. These models could also enable fine-grained analysis, such as classifying wheezes and crackles or detecting their onset and offset within recordings, offering more detailed diagnostic insights.