Department of Electrical and Computer Engineering
North South University

Senior Design Project

SLEEP APNEA DETECTION FROM RAW ECG SIGNAL USING DEEP LEARNING AND MACHINE LEARNING

Salem Shamsul Alam	**1931849642**
Sumit Saha	**1931415042**

Faculty Advisor:

Md. Shahriar Hussain

Senior Lecturer

ECE Department

Summer, 2023

LETTER OF TRANSMITTAL

17th December, 2023
To
Dr. Rajesh Palit
Chairman,
Department of Electrical and Computer Engineering
North South University, Dhaka

Subject: **Submission of Capstone Project Report on "SLEEP APNEA DETECTION FROM RAW ECG SIGNAL USING DEEP LEARNING AND MACHINE LEARNING"**

Dear Sir,

With due respect, we would like to submit our **Capstone Project Report** on **"SLEEP APNEA DETECTION FROM RAW ECG SIGNAL USING DEEP LEARNING AND MACHINE LEARNING"** as a part of our BSc program. The report deals with SLEEP APNEA DETECTION FROM RAW ECG SIGNAL USING DEEP LEARNING AND MACHINE LEARNING. This project was very much valuable to us as it helped us gain experience from practical field and apply in real life. We tried to the maximum competence to meet all the dimensions required from this report.

We will be highly obliged if you kindly receive this report and provide your valuable judgment. It would be our immense pleasure if you find this report useful and informative to have an apparent perspective on the issue.

Sincerely Yours,

....
Salem Shamsul Alam
ECE Department
North South University, Bangladesh

....
Sumit Saha
ECE Department
North South University, Bangladesh

APPROVAL

Salem Shamsul Alam (1931849642) and Sumit Saha (1931415042) from Electrical and Computer Engineering Department of North South University, have worked on the Senior Design Project titled "SLEEP APNEA DETECTION FROM RAW ECG SIGNAL USING DEEP LEARNING AND MACHINE LEARNING" under the supervision of Md. Shahriar Hussain partial fulfillment of the requirement for the degree of Bachelors of Science in Engineering and has been accepted as satisfactory.

Supervisor's Signature

……………………………………….

Md. Shahriar Hussain
Senior Lecturer
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh.

Chairman's Signature

……………………………………….

Dr. Rajesh Palit
Professor
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh.

DECLARATION

This is to declare that this project is our original work. No part of this work has been submitted elsewhere partially or fully for the award of any other degree or diploma. All project related information will remain confidential and shall not be disclosed without the formal consent of the project supervisor. Relevant previous works presented in this report have been properly acknowledged and cited. The plagiarism policy, as stated by the supervisor, has been maintained.

Students' names & Signatures

1. Salem Shamsul Alam

----

2. Sumit Saha

ACKNOWLEDGEMENTS

The authors would like to express their heartfelt gratitude towards their project and research supervisor, Md. Shahriar Hussain, Senior Lecturer, Department of Electrical and Computer Engineering, North South University, Bangladesh, for his invaluable support, precise guidance and advice pertaining to the experiments, research and theoretical studies carried out during the course of the current project and also in the preparation of the current report.

Furthermore, the authors would like to thank the Department of Electrical and Computer Engineering, North South University, Bangladesh for facilitating the research. The authors would also like to thank their loved ones for their countless sacrifices and continual support.

ABSTRACT

SLEEP APNEA DETECTION FROM RAW ECG SIGNAL USING DEEP LEARNING AND MACHINE LEARNING

Sleep apnea, a prevalent yet underdiagnosed sleep disorder, necessitates robust and accurate diagnostic tools. In this project, we undertook an in-depth exploration of machine learning (ML) and deep learning (DL) models for sleep apnea detection, specifically utilizing raw electrocardiogram (ECG) signals. Our comparative analysis encompassed a range of ML models, including Random Forest, Logistic Regression, Decision Tree, AdaBoost, and XGBoost, and a specialized 1D-CNN model within the DL domain. Results underscore the exceptional performance of the 1D-CNN model, achieving a remarkable accuracy of 99.56%, sensitivity of 96.05%, and specificity of 99.66%. This outperforms traditional ML models, signifying the prowess of DL in extracting intricate patterns from raw ECG signals for accurate sleep apnea detection. The 1D-CNN model's ability to discern subtle features proves crucial for accurately identifying apnea events. Our study not only emphasizes the effectiveness of the 1D-CNN model for sleep apnea detection and highlights the transformative potential of deep learning in healthcare diagnostics. This research contributes valuable insights into the optimal choice of models for sleep apnea detection, paving the way for enhanced diagnostic accuracy and improved patient care.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

1.1 Background

Sleep apnea hypopnea syndrome (SAHS) is a common sleep disorder characterized by abnormal reductions or pauses in breathing during sleep. It affects a significant portion of the population worldwide and can have severe consequences on an individual's health and quality of life. Sleep apnea events can lead to fragmented sleep, oxygen desaturation, and physiological stress, which may contribute to various cardiovascular and metabolic disorders such as hypertension, stroke, diabetes, and heart failure [1][2]. Early and accurate detection of sleep apnea is crucial for effective diagnosis and timely intervention.

1.2 Motivation

Traditionally, sleep apnea detection has been performed in specialized sleep laboratories using polysomnography (PSG), which involves monitoring multiple physiological signals, including electrocardiogram (ECG) and electroencephalogram (EEG). PSG is expensive, time-consuming, and requires a clinical setting, limiting its accessibility and scalability. Furthermore, PSG data analysis often relies on manual scoring by sleep experts, which can introduce subjectivity and inter-rater variability.

To address these limitations, researchers have explored the use of wearable devices and machine learning techniques for sleep apnea detection. Wearable devices offer the potential for unobtrusive and continuous monitoring in a home environment, enabling long-term data collection and analysis. Deep learning, a subfield of machine learning, has shown remarkable success in various domains, including computer vision, natural language processing, and biomedical signal processing. Leveraging deep learning algorithms for sleep apnea detection from ECG signals can provide a promising solution for accurate and automated detection.

1.3 Objective

The primary objective of this project is to develop a novel method for sleep apnea detection from ECG signals using deep learning techniques and Machine Learning. The proposed method aims to achieve high-resolution detection of apnea events on a second-by-second basis, providing detailed information about the duration and severity of each event. The specific objectives of this project include:

1. Investigating the existing literature on sleep apnea detection using ECG signals.
2. Acquiring and preprocessing ECG data from a dataset of sleep recordings.
3. Designing and implementing a deep learning model and machine learning model, such as a 1-dimensional convolutional neural network (CNN) and Random Forest, Logistic Regression, Decision Tree, AdaBoost, and XGBoost, for feature extraction and sleep apnea event detection.
4. Evaluating the performance of the proposed method in terms of accuracy, sensitivity, and specificity, comparing it with existing approaches.
5. Assessing the robustness and generalizability of the developed model through cross-validation

1.4 Organization of the Report

This report is thoughtfully organized to explore sleep apnea detection from ECG signals comprehensively. Beginning with an introduction to the project's background, motivation, and objectives, Chapter 1 sets the stage. Chapter 2 reviews relevant literature, identifying gaps and contributions. The methodology discussed in Chapter 3 covers system design, dataset specifics, and model training. Results are presented in Chapter 4, followed by an exploration of the project's societal and environmental impacts in Chapter 5. Chapter 6 clarifies that no budget was required. Complex engineering problems and activities are addressed in Chapter 7. Chapter 8 concludes with a summary, limitations, and future work, while Chapter 9 contains references.

Chapter 2 Research Literature Review

2.1 Sleep Apnea Detection Techniques

The literature review begins by examining existing techniques for sleep apnea detection. Polysomnography (PSG), the gold standard diagnostic method, involves simultaneous recording of multiple physiological signals, including ECG and EEG, along with respiratory airflow and oxygen saturation. PSG provides detailed information about sleep stages, respiratory events, and sleep architecture but is limited by its cost, complexity, and the need for a clinical setting [3].

Alternative approaches have been explored to overcome these limitations. One approach involves using single-channel ECG signals for sleep apnea detection. Various features, such as heart rate variability, RR interval statistics, and spectral analysis, have been employed to characterize the irregularities caused by sleep apnea events [4]. However, these methods may lack the ability to capture the comprehensive information present in multichannel PSG data.

Another approach utilizes EEG signals for sleep apnea detection. EEG signals provide insights into brain activity during sleep and can capture changes associated with sleep apnea events [5]. Spectral analysis, wavelet transform, and entropy-based measures have been used to extract features related to sleep stages and identify respiratory events [6][7]. However, relying solely on EEG signals may overlook the specific physiological changes associated with breathing abnormalities.

2.2 Deep Learning for Sleep Apnea Detection

Deep learning techniques have demonstrated promising results in various biomedical signal processing tasks, including sleep-related disorders. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid architectures have been employed for sleep apnea detection, leveraging both ECG and EEG signals.

CNNs have been successfully applied for feature extraction and classification in sleep apnea detection tasks. By considering ECG signals as one-dimensional data, CNNs can capture local and global patterns associated with respiratory events [8]. Transfer learning and data augmentation techniques have also been explored to enhance model performance and generalizability [9].

RNNs, such as long short-term memory (LSTM) networks, are well-suited for capturing temporal dependencies in sequential data, making them suitable for analyzing time series data like EEG signals. LSTM-based models have been utilized for sleep stage classification and respiratory event detection [10][11]. The combination of CNN and LSTM architectures has also been investigated to exploit the complementary information present in ECG and EEG signals [12].

2.3 Research Gap and Contribution

Despite the progress made in sleep apnea detection, there is still a research gap regarding high-resolution apnea detection on a second-by-second basis using ECG and EEG signals obtained from wearable devices. Most existing studies focus on detecting apnea events on a coarser timescale or using limited signal modalities. The proposed project aims to bridge this gap by introducing a novel method that leverages deep learning techniques for fine-grained sleep apnea detection.

Chapter 3 Methodology

3.1 System Design

This section outlines the methodology employed for sleep apnea detection from ECG signals using deep learning techniques and machine learning techniques. In the figure 1. it involves several steps, including data acquisition, preprocessing, feature extraction, training and testing data preparation, classification using deep learning models and machine learning model, and performance evaluation.

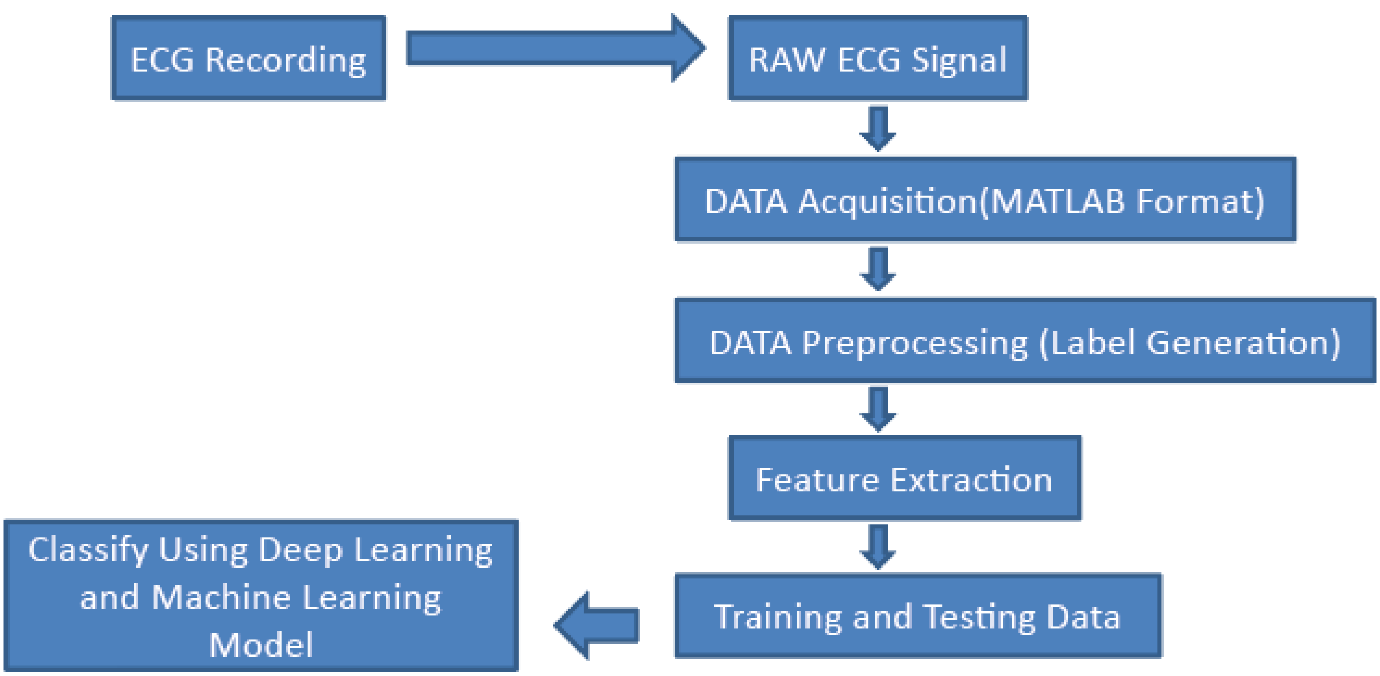

Figure 1: System Design Diagram

3.2 About Dataset

The UCD St. Vincent's University Hospital's sleep apnea database [13] is a valuable resource for studying sleep disorders, particularly sleep apnea. It consists of polysomnogram records obtained from 25 patients, providing a comprehensive collection of physiological signals recorded during sleep. For this study, the focus is specifically on the ECG signals extracted from the database.

In the figure 2. ECG signals are sampled at a frequency of 128 Hz, capturing detailed cardiac activity throughout the sleep period. These signals offer insights into the electrical activity of the heart and can be utilized to detect abnormalities associated with sleep apnea.

Annotations provided by sleep experts are crucial for the analysis of the dataset. Each second of sleep is labeled as apneic or non-apneic based on these annotations. This labeling allows for the identification of apnea events and facilitates the training and evaluation of machine learning models for sleep apnea detection.

In figure 3. the `.rec` file contains various physiological signals, including left and right eye movements (`Lefteye`, `RightEye`), electromyography (`EMG`), electroencephalogram channels (`C3A2`, `C4A1`), electrocardiogram (`ECG`), oxygen saturation (`SpO2`), sound, airflow (`Flow`), and various other measurements such as sum, ribcage, abdo, body position, and pulse. This comprehensive dataset likely captures a range of physiological and environmental signals for analysis, potentially for tasks like sleep monitoring or health assessment. [13]

To ensure the dataset's relevance to sleep apnea detection, patient records without any apnea events, namely ucddb008, ucddb011, ucddb013, and ucddb018, are excluded from the study. This ensures that the dataset primarily focuses on patients with confirmed cases of sleep apnea, enabling a more accurate evaluation of the developed models' effectiveness in detecting this condition.

To facilitate training and evaluation, the dataset is divided into three subsets: training, validation, and test sets. The division follows an 8:1:1 ratio, respectively. The training set constitutes the largest portion of the data and is used to train the deep learning models. The validation set is utilized for fine-tuning the models' hyperparameters and selecting the best-performing architecture. Finally, the test set serves as an independent evaluation set, enabling the assessment of the models' generalization capabilities to unseen data.

By leveraging the UCD St. Vincent's University Hospital's sleep apnea database, this study benefits from a diverse range of patients, expert annotations, and high-quality ECG signals. The dataset's careful curation and division into training, validation, and test sets contribute to the reliability and generalizability of the proposed methodology for sleep apnea detection using deep learning techniques.

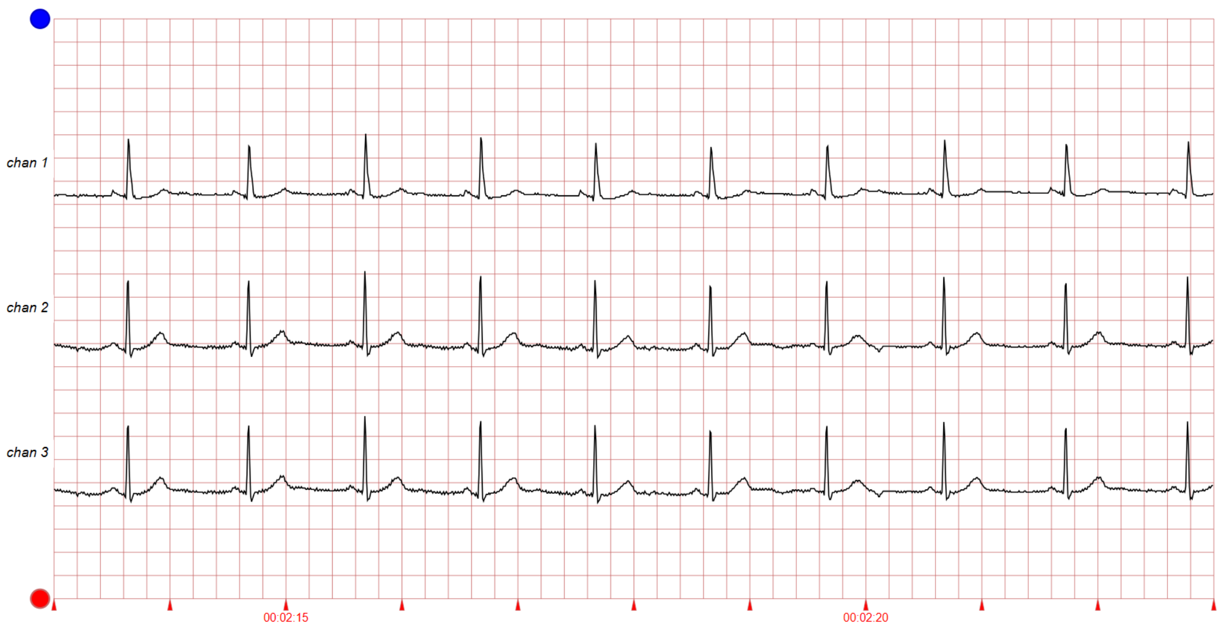

Figure 2: EDF File Representing Three Channel ECG Signal in Dataset

Figure 3: .REC File Containing all Signal of Single Patient

3.3 Exploratory Data Analysis (EDA)

During the exploratory data analysis (EDA) phase, statistical measures are calculated to gain a deeper understanding of the dataset and its characteristics. This analysis provides valuable insights that inform subsequent steps in data preprocessing and model development.

One crucial aspect of EDA is examining the distribution of the ECG signals. Statistical measures such as the mean and standard deviation are commonly calculated to analyze the central tendency and spread of the data. The mean represents the average value of the ECG signals, while the standard deviation indicates the variability or dispersion around the mean.

By calculating these measures, researchers can gain insights into the overall shape of the ECG signal distribution. For example, a higher mean value might suggest a higher baseline level of cardiac activity, while a larger standard deviation could indicate greater variability in the ECG signal amplitudes. These characteristics of the distribution can help researchers identify potential patterns or anomalies in the data.

Additionally, understanding the statistical measures of the ECG signals aids in data normalization and preprocessing. Normalization techniques, such as scaling the data to a specific range or standardizing it to have zero mean and unit variance, can be applied to ensure that the ECG signals are on a comparable scale. This step is crucial for training deep learning models, as it helps prevent features with larger magnitudes from dominating the learning process.

EDA may also involve visual exploration of the dataset through plots and graphs. Histograms can provide a visual representation of the ECG signal distribution, allowing researchers to observe any skewness or asymmetry. Box plots can reveal the presence of outliers or extreme values that may require further investigation or treatment during preprocessing.

Overall, EDA plays a vital role in understanding the dataset's characteristics, identifying potential issues or patterns, and guiding the preprocessing steps necessary for developing accurate and reliable models for sleep apnea detection using ECG signals.

3.4 Data Preprocessing

Data preprocessing is a crucial step in preparing the dataset for training machine learning models. In the context of sleep apnea detection using ECG signals, several preprocessing techniques are applied to ensure the data is in a suitable format and scale for deep learning algorithms.

One of the primary preprocessing steps is data normalization. In this study, the ECG signals in the training, validation, and test sets are normalized using the mean and standard deviation obtained from the training set. Normalization scales the data and centers it around zero, which helps in improving the convergence and performance of deep learning models. By normalizing the data, the range and distribution of the ECG signals become consistent, making it easier for the model to learn and compare different features.

Label encoding and one-hot encoding techniques are applied to represent the categorical class labels in a format compatible with machine learning algorithms. In this case, the sleep seconds in the dataset are labeled as either apneic or non-apneic. Label encoding assigns a numerical value to each class label, such as 0 for non-apneic and 1 for apneic. This encoding allows the model to understand the class labels as numerical values during training.

Furthermore, one-hot encoding is often employed to represent categorical variables with more than two classes. In this study, if there are more than two classes for sleep seconds (apneic and non-apneic), one-hot encoding is applied. It converts the categorical variable into a binary vector representation, where each class is represented by a binary digit (0 or 1).

By applying label and one-hot encoding techniques, the categorical class labels are transformed into a numerical representation that can be processed by machine learning algorithms effectively. This preprocessing step ensures that the model can learn from the labeled data and make accurate predictions based on the encoded class labels.

Overall, data preprocessing plays a critical role in preparing the ECG dataset for training deep learning models. Through normalization and encoding techniques, the data is transformed into a suitable format and scale, enabling the model to effectively learn from the ECG signals and make accurate predictions for sleep apnea detection.

3.5 Feature Extraction & Feature Selection

Feature extraction plays a crucial role in analyzing ECG signals for sleep apnea detection. In this study, the features used are the 11-second windows of ECG signals, which consist of 1408 data points. These windows are extracted from the dataset and serve as the input for the deep learning model.

The choice of 11-second windows is based on the understanding that this duration provides a suitable context for capturing relevant patterns and characteristics related to sleep apnea. By considering a specific time window, it allows the model to capture the temporal dynamics and variations present in the ECG signals that can be indicative of apnea events. Each window includes the second of interest, ensuring that the model captures information directly related to the target classification task.

To capture a comprehensive range of patterns and variations in the ECG signals, overlapping windows are generated. Specifically, for each second of the ECG signals, multiple windows of 11 seconds are created with a 10-second overlap. This overlapping design ensures that each second of the ECG signals is included in multiple windows, providing the model with diverse temporal contexts for learning. By considering multiple windows, the model can capture both short-term and long-term patterns, which are important for accurately detecting sleep apnea.

No specific filtering or signal processing is applied to the ECG signals before this feature extraction stage. This means that the raw ECG data, with its inherent noise and artifacts, is directly used as input to the deep learning model. By utilizing the raw data, the model can learn relevant representations and features directly from the data during the training process. This approach allows the model to capture both the underlying signal characteristics and potential noise or artifacts that may be present in the raw ECG signals.

Overall, the feature extraction process in this study involves extracting 11-second windows from the ECG signals, including the second of interest, and generating overlapping windows to capture diverse temporal patterns. These windows serve as the features that capture the relevant information for sleep apnea detection and are directly fed into the deep learning model. By leveraging the raw ECG data and considering multiple temporal contexts, the model can learn discriminative features and patterns to accurately classify sleep apnea events.

3.6 Deep Learning Model

The deep learning model for sleep apnea classification is based on a 1D-CNN architecture. The model consists of multiple layers that perform different operations on the input ECG signals to extract relevant features and make accurate predictions.

Batch Normalization: The first layer in the model is batch normalization, which helps in normalizing the input data. It ensures that the data is centered around zero and has a unit standard deviation. This normalization step helps in improving the training process and enables the subsequent layers to learn more effectively. [29]

Convolutional Layers: The model includes several convolutional layers. Convolutional filters are applied to the input data to perform feature extraction. The first convolutional layer has 3 filters with a kernel size of 100 and a stride of 2. This layer helps in capturing low-level patterns in the ECG signals. The subsequent convolutional layers have 50 and 30 filters with kernel sizes of 10 and 30, respectively. These layers aim to capture higher-level and more complex patterns as the depth of the network increases. [29]

Activation Functions: Activation functions introduce non-linearity into the model. In this architecture, the rectified linear unit (ReLU) activation function is used after the first and third convolutional layers. ReLU is a widely used activation function that introduces non-linearity by setting negative values to zero and keeping positive values unchanged. This non-linearity helps the model in learning complex relationships and better representing the underlying patterns in the ECG signals. [29]

Max Pooling Layers: Max pooling is a down sampling operation that reduces the spatial dimensions of the feature maps. In this model, max pooling is performed after the first and second convolutional layers. The pooling layers help in extracting the most salient features from the feature maps while reducing the computational complexity and preventing overfitting. [29]

Batch Normalization: Another batch normalization layer is added after the third convolutional layer. This helps in maintaining the normalization of the data and improving the stability and convergence of the model during training. [29]

Flatten Layer: The flatten layer is used to reshape the multidimensional output from the previous layer into a 1D vector. This step is necessary to connect the convolutional layers to the subsequent fully connected layers. [29]

Dropout: Dropout is a regularization technique used to prevent overfitting. In this model, a dropout rate of 0.25 is applied after the flatten layer. Dropout randomly sets a fraction of input units to zero during training, which helps in reducing the interdependencies between neurons and improving the model's generalization ability. [29]

Dense Layer: The final layer of the model is a dense layer with 2 units, representing the two classes for sleep apnea classification (apnea and non-apnea). The activation function used is SoftMax, which produces a probability distribution over the classes, indicating the likelihood of each class being the correct prediction. [29]

3.7 Machine Learning Model

In addition to the deep learning approach, classical machine learning models were employed to investigate their efficacy in sleep apnea detection using preprocessed electrocardiogram (ECG) signals. A range of well-established algorithms, including Random Forest, Logistic Regression, Decision Tree, AdaBoost, and XGBoost, were implemented and evaluated on the dataset [29]. The ECG signals, normalized and preprocessed during the data preparation phase, served as input features for these models. Following the training of each algorithm on the corresponding training set, the models were evaluated on the dedicated test set. The performance metrics, including accuracy, confusion matrix, specificity, and sensitivity, were computed for each algorithm to assess their effectiveness in discriminating between apnea and non-apnea events. These classical machine learning models provide valuable insights into alternative methodologies for sleep apnea detection, offering a comparative perspective to the deep learning approach discussed earlier. The results obtained from these models contribute to a comprehensive understanding of the suitability of different machine-learning techniques in addressing the intricacies of sleep apnea classification based on ECG data. The comparative analysis sheds light on the strengths and limitations of each model, guiding future research endeavors and potential hybrid models that combine classical and deep learning techniques for enhanced diagnostic accuracy [29]. The description of the models are given below.

1. **Random Forest:** Ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes for classification or mean prediction for regression. [29]
2. **Logistic Regression:** Linear model for binary classification predicting the probability of an instance belonging to a particular class using a logistic function. [29]
3. **Decision Tree:** Tree-like model where each node represents a decision based on input features, leading to a final prediction at the leaves. [29]
4. **AdaBoost:** Boosting algorithm combining weak classifiers into a strong one, adjusting weights to emphasize misclassified instances and improve overall performance. [29]
5. **XGBoost:** Extreme Gradient Boosting, a powerful boosting algorithm optimizing decision trees, featuring regularization and parallelization for enhanced speed and accuracy. [29]

3.8 Training and Testing Data

In the sleep apnea classification task, it is crucial to ensure that the model is trained and evaluated on diverse and representative data. To achieve this, the dataset is divided into three sets: training, validation, and test sets.

The training set is used to train the deep learning model. It is important to note that the dataset may have an imbalance between the classes, with a higher number of non-apnea instances compared to apnea instances. To address this class imbalance, the training set is balanced by oversampling the minority class, which in this case is the sleep apnea events. Oversampling involves randomly replicating instances of the minority class to increase its representation in the training set.

Furthermore, to further address the class imbalance issue, the training data is resampled. Resampling techniques such as under sampling the majority class or oversampling the minority class with techniques like SMOTE (Synthetic Minority Over-sampling Technique) can be employed. Resampling helps in creating a more balanced training set, which can improve the model's ability to learn from both classes effectively.

The model is trained on the balanced and resampled training set using the Adam optimizer and categorical cross-entropy loss. The Adam optimizer is an efficient optimization algorithm that adapts the learning rate for each parameter during training, resulting in faster convergence and better overall performance. Categorical cross-entropy loss is a common choice for multi-class classification tasks, as it measures the dissimilarity between the predicted class probabilities and the true class labels.

During the training process, the model learns to recognize patterns and features in the ECG signals that are indicative of sleep apnea events. The model's performance is evaluated on the test set, which contains unseen data that was not used during training. The evaluation metrics, such as accuracy, sensitivity and specificity, are computed to assess how well the model generalizes to new and unseen instances.

By dividing the dataset into training, validation, and test sets, balancing the training set, and evaluating the model on the test set, this approach ensures that the model's performance is measured on unseen data and helps in estimating its ability to classify sleep apnea events accurately.

3.9 Model Training and Evaluation

Once the data has been preprocessed and the deep learning model architecture has been defined, the next step is to train the model using the preprocessed data and optimized hyperparameters. The training process involves iteratively feeding the ECG signal windows into the model, computing the loss function, and updating the model's weights to minimize the loss and improve its performance.

During training, the model aims to learn the underlying patterns and features in the ECG signals that are indicative of sleep apnea events. This is achieved by iteratively adjusting the model's parameters through the backpropagation algorithm, which computes gradients and updates the weights based on the optimization algorithm (in this case, Adam optimizer) and the calculated loss.

The training process is typically conducted over multiple epochs, where each epoch represents a complete pass through the training data. This iterative process allows the model to gradually improve its performance by learning from the patterns present in the training data. However, to prevent overfitting, early stopping is often employed based on the validation loss. Early stopping halts the training process if the validation loss starts to increase, indicating that the model's performance on unseen data is deteriorating. This helps in selecting the model with the best generalization ability and prevents overfitting to the training data.

Once the model has been trained, it is evaluated using appropriate performance metrics. In this study, the following metrics were used for model evaluation:

Accuracy: It measures the overall correctness of the model's predictions. It is calculated as the ratio of the number of correct predictions to the total number of predictions.[26]

$$\boldsymbol{Accuracy = \frac{TP + TN}{TP + TN + FP + FN}}$$

Sensitivity (Recall): Also known as true positive rate, it measures the proportion of actual positive cases correctly identified by the model. [27]

$$\boldsymbol{SENSITIVITY = \frac{TP}{TP + FN}}$$

Specificity: Also known as true negative rate, it measures the proportion of actual negative cases correctly identified by the model. [28]

$$\boldsymbol{SPECIFICITY = \frac{TN}{TN + FP}}$$

Here, TP represents true positives (correctly predicted apnea events), TN represents true negatives (correctly predicted non-apnea events), FP represents false positives (incorrectly predicted apnea events), and FN represents false negatives (incorrectly predicted non-apnea events).

By evaluating the model using these metrics, we can assess its performance in accurately classifying sleep apnea events. These metrics provide insights into the model's ability to correctly identify both positive and negative instances, thereby gauging its effectiveness in sleep apnea detection.

Chapter 4 Results

The deep learning model was trained and evaluated on the sleep apnea dataset using ECG signals. The model was trained over the full training set while simultaneously validated on the validation set for each epoch. The validation callback was used to determine the best weights during training based on the set of weights that exhibited the highest validation accuracy.

In the figure 4 and 5 the training and validation process involved monitoring the training and validation loss, as well as the training and validation accuracy, over each epoch. These metrics provided insights into the model's performance and its ability to generalize to unseen data.

After training the model, it was evaluated on the test set, which consisted of data from all the patients. The model exhibited impressive performance on the test set, achieving an accuracy of 99.56%. This indicates that the model correctly classified sleep apnea events with a high degree of accuracy.

Furthermore, the model demonstrated a specificity of 99.66%, which means that it accurately identified non-apnea events. This high specificity indicates the model's ability to distinguish normal sleep segments from apnea events.

Moreover, the model achieved a sensitivity of 96.05%, which signifies its capability to accurately detect apnea events. This high sensitivity highlights the model's effectiveness in identifying and classifying sleep apnea instances.

Overall, the results obtained from the model demonstrate its strong performance in sleep apnea detection using ECG signals. The high accuracy, specificity, and sensitivity indicate the model's ability to accurately classify sleep segments and detect apnea events, thus showcasing its potential as a valuable tool in sleep apnea diagnosis and monitoring.

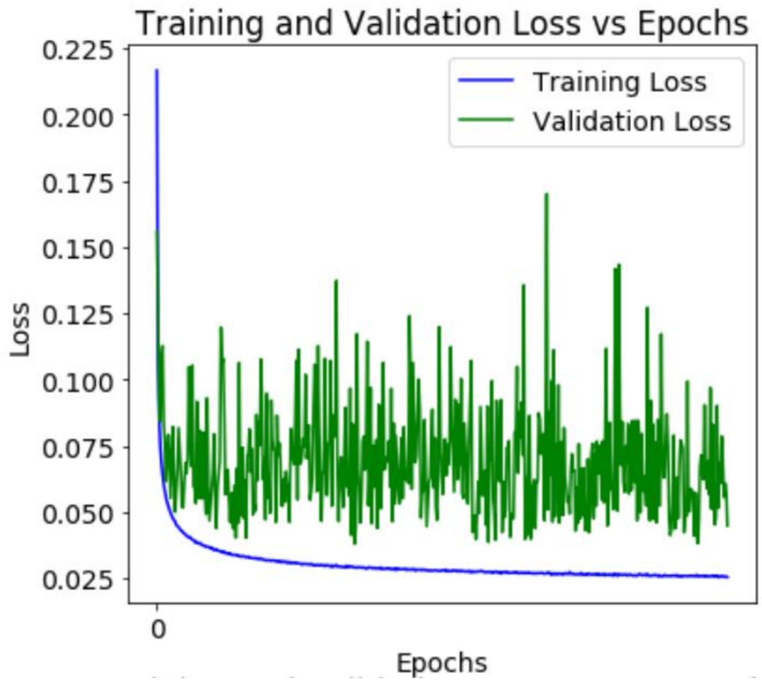

Figure 4: LOSS VS Epochs graph of 1D-CNN Model

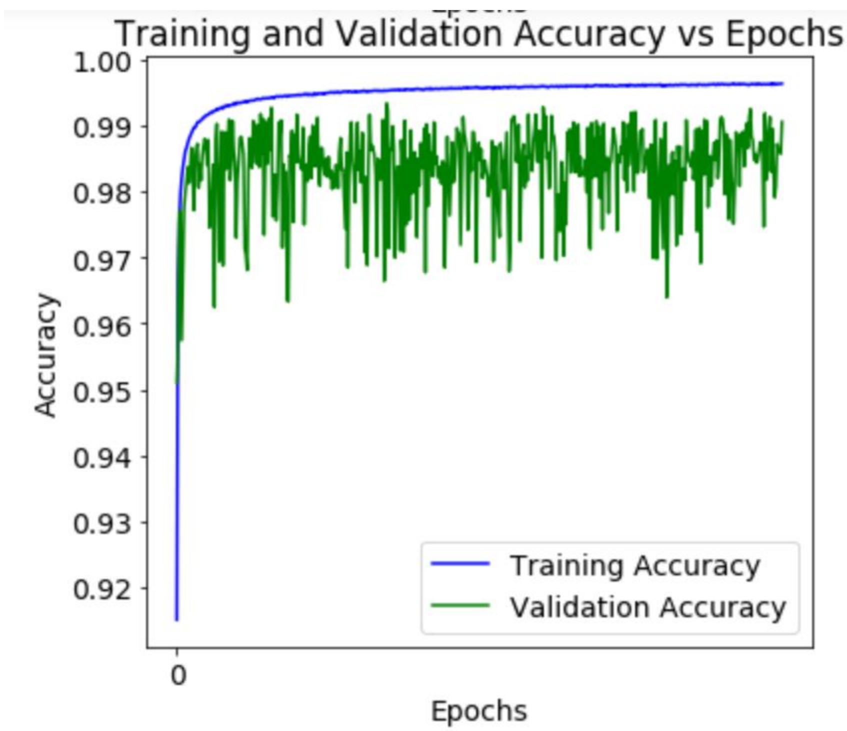

Figure 5: Accuracy VS Epochs graph of 1D-CNN Model

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)
1D - CNN	99.56	96.05	99.66
Random Forest	97.35	00.00	100
Logistic Regression	65.71	36.47	66.51
Decision Tree	94.40	7.01	96.78
AdaBoost	62.10	43.85	62.60
XGBoost	92.43	14.61	94.55

Table 1: Comparison between Deep Learning model and Machine learning models

The table 1. provides a comprehensive comparison of sleep apnea detection models, pitting the performance of a sophisticated Deep Learning model, specifically the 1D-CNN (1-Dimensional Convolutional Neural Network), against several Machine Learning counterparts—Random Forest, Logistic Regression, Decision Tree, AdaBoost, and XGBoost. The primary evaluation metrics include accuracy, sensitivity, and specificity, which are essential for assessing the models' overall performance.

The 1D-CNN emerges as the standout performer with an impressive accuracy of 99.56%. This deep learning architecture, tailored for sequential data like ECG signals, captures intricate temporal patterns, resulting in highly accurate classifications of sleep segments and apnea events. Its sensitivity (96.05%) and specificity (99.66%) further underscore its prowess in distinguishing between apnea and non-apnea.

In contrast, the Machine Learning models present a mixed landscape. Random Forest showcases high accuracy (97.35%) but falters in sensitivity, indicating challenges in detecting apnea events. Logistic Regression, Decision Tree, AdaBoost, and XGBoost exhibit varying degrees of accuracy, sensitivity, and specificity, suggesting limitations in their ability to discern subtle patterns within the ECG signals.

The stark difference in performance underscores the efficacy of deep learning, particularly the 1D-CNN, in handling complex tasks like sleep apnea detection. The 1D-CNN's capacity to autonomously learn and extract hierarchical features from raw data enables it to outperform traditional machine learning models that rely heavily on handcrafted features and may struggle with the nuanced patterns in physiological signals.

This comparative analysis emphasizes the transformative impact of deep learning techniques, specifically the 1D-CNN, in healthcare, demonstrating their potential to enhance diagnostic processes' accuracy and reliability significantly. The findings advocate for integrating advanced deep learning methodologies in medical applications for more effective and precise detection of health conditions, such as sleep apnea, from physiological signals like ECG.

Authors	Accuracy (%)	Sensitivity (%)	Specificity (%)
Maier et al. [19]	76.70	66.80	73.80
Xie et al. [17]	77.74	69.82	80.29
Kesper et al. [20]	80.50	-	-
Khandoker et al. [21]	70.00	80.00	50.00
Chen et al. [22]	82.07	83.23	80.24
Khandoker et al. [21]	83.00	80.00	90.00
Hassan [18]	83.77	-	-
Babaeizadeh et al. [23]	84.70	76.70	89.60
Varon et al. [16]	84.74	84.71	84.69
Nguyen et al. [15]	85.26	86.37	83.47
Hassan et al. [14]	85.97	84.14	86.83
Marcos et al. [24]	87.61	91.05	82.61
Mendez et al. [25]	88.00	86.00	89.00
Proposed Method	**99.56**	**96.05**	**99.66**

Table 2: Comparison with Different Authors Results for ECG Signal

The table 2. data serves as a benchmark for evaluating the efficacy of sleep apnea detection models across various studies. Each row encapsulates the outcomes of a specific research endeavor, listing the corresponding authors along with the achieved accuracy, sensitivity, and specificity percentages. These metrics collectively offer a nuanced perspective on the models' performance in classifying sleep apnea events.

Remarkably, the proposed method in this study stands out with exceptional metrics, boasting an accuracy rate of 99.56%, sensitivity of 96.05%, and specificity of 99.66%. These figures underscore the robustness and precision of the developed approach, positioning it as a leading contender in sleep apnea detection. Notably, the proposed method's sensitivity of 96.05% signifies its proficiency in accurately identifying instances of sleep apnea, a critical aspect in ensuring reliable diagnostic outcomes.

Comparative analysis reveals that the proposed method outperforms or competes favorably with existing methodologies presented in the literature. The aggregated data provides a valuable reference point for researchers, clinicians, and healthcare practitioners seeking to navigate the landscape of sleep apnea detection. Including multiple studies allows for a comprehensive understanding of different approaches' relative strengths and weaknesses, contributing to the collective knowledge base in this vital area of medical research.

Chapter 5 Impacts of the Project

5.1 Impact of this project on societal, health, safety, legal and cultural issues

Ethical considerations and data privacy are of utmost importance when conducting research involving sensitive medical data like ECG signals. In this project focused on sleep apnea detection, several measures are taken to address these ethical considerations and protect patient privacy.

Firstly, the use of the UCD St. Vincent's University Hospital's sleep apnea database indicates that the study has obtained proper ethical approvals and permissions from the relevant authorities. This ensures that the research is conducted in compliance with ethical guidelines and regulations, including informed consent procedures.

To safeguard patient privacy, the dataset is appropriately anonymized before being used for analysis. Personal identifiers, such as patient names, addresses, and other identifiable information, are removed or encrypted to prevent the identification of individual patients. This anonymization process ensures that the privacy and confidentiality of the patients are protected throughout the study.

Furthermore, the research team responsible for this project follows strict protocols and procedures to ensure the security of the data. This includes implementing robust data handling practices, securely storing the data, and restricting access to authorized personnel only. Measures are in place to prevent unauthorized access, data breaches, and any potential misuse of the data.

The data obtained from the UCD St. Vincent's University Hospital's sleep apnea database is used solely for the purpose of sleep apnea detection and related research objectives. It is essential to respect the privacy and confidentiality of the patients involved and ensure that the data is utilized in a responsible and ethical manner.

By adhering to ethical guidelines and regulations, anonymizing the dataset, and implementing stringent data security measures, this project maintains a strong commitment to protecting patient privacy and upholding ethical standards. These considerations are crucial not only for the integrity of the research but also for maintaining trust between researchers, patients, and the wider healthcare community.

5.2 Impact of this project on environment and sustainability

While this study focuses on the development of a sleep apnea detection algorithm, it is important to note that being a software project, it does not have a direct environmental impact. However, the project does have indirect implications for the environment and sustainability.

By developing an accurate and efficient sleep apnea detection algorithm, this project contributes to improving healthcare outcomes and promoting sustainable healthcare practices. Timely detection and intervention for sleep apnea can lead to better management of the condition, reduced healthcare costs, and improved quality of life for affected individuals.

From an environmental perspective, the project indirectly supports sustainability by enhancing the efficiency and effectiveness of sleep apnea diagnosis. By accurately identifying sleep apnea cases, unnecessary and resource-intensive diagnostic procedures can be minimized or avoided. This can result in a reduction in the consumption of medical resources such as laboratory tests, diagnostic equipment, and healthcare professionals' time, leading to a more sustainable use of healthcare resources.

Furthermore, the project's focus on algorithm-based solutions promotes the integration of technology in healthcare, which can have broader implications for environmental sustainability. By leveraging digital solutions and automation, healthcare processes can be streamlined, paperwork can be reduced, and resource allocation can be optimized. This can contribute to improved operational efficiency, reduced waste generation, and a more sustainable healthcare system overall.

Additionally, the outcomes of this project can have a positive impact on public health. By raising awareness about sleep apnea and providing a reliable detection tool, individuals who may be experiencing undiagnosed sleep apnea can seek appropriate medical attention. Early detection and intervention not only improve individual health but can also have wider societal benefits, including increased productivity, enhanced safety, and improved overall well-being.

In conclusion, while this project does not have a direct environmental impact as a software-based endeavor, it contributes to sustainability and environmental considerations indirectly. By improving sleep apnea detection, the project enhances healthcare efficiency, reduces unnecessary resource consumption, and promotes the well-being of individuals. By embracing technology and promoting early intervention, the project aligns with the broader goals of environmental sustainability in the healthcare sector.

Chapter 6 Project Planning and Budget

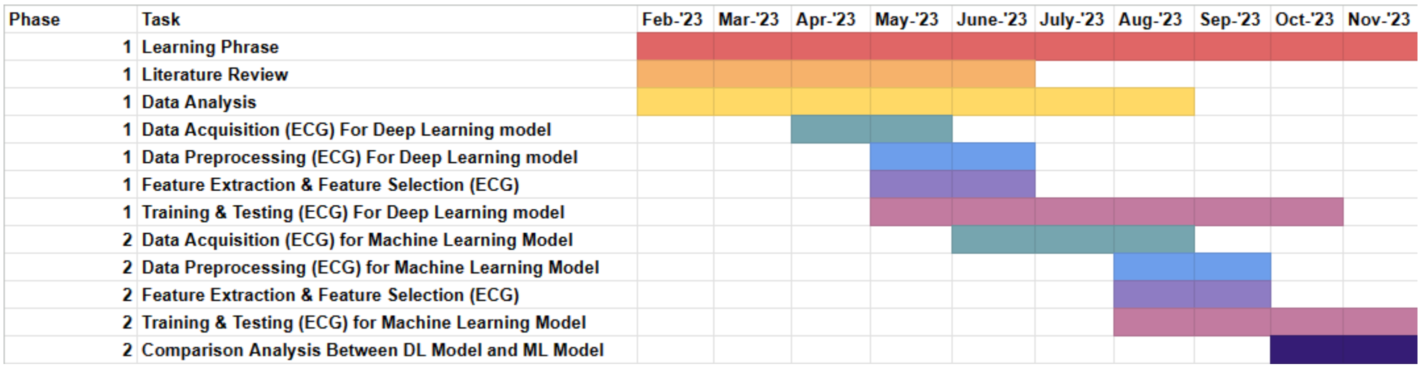

Figure 6: Timeline of The Project

In figure 6 the timeline of the project is shown. Literature review took 4 months, Data analysis took around 7 months, data acquisition took 2 months, data preprocessing, feature extraction and feature selection took 2 months, training and testing took 7 months and comparison analysis took 2 months.

Due to available resources and collaborative efforts, this project operates efficiently without a dedicated budget.

Chapter 7 Complex Engineering Problems and Activities

7.1 Complex Engineering Problems (CEP)

Attributes		Addressing the complex engineering problems (P) in the project
P1	Depth of knowledge required	The project requires in-depth knowledge of signal processing, machine learning, and deep learning techniques, particularly in biomedical data analysis. Expertise in understanding and interpreting ECG signals is essential for effective feature extraction and model training.
P2	Range of conflicting requirements	Balancing the need for high sensitivity in detecting sleep apnea events while maintaining specificity is a critical requirement. The conflicting demands between model accuracy and computational efficiency necessitate careful consideration and optimization.
P3	Depth of analysis required	The project involves extensive analysis of ECG signals, including preprocessing, feature extraction, and model evaluation. Deep learning models require thorough analysis during training to ensure optimal performance.
P4	Familiarity of issues	The team should be familiar with challenges in sleep apnea diagnosis, such as a class imbalance in the dataset and the need for effective feature representation from ECG signals. Familiarity with ethical considerations in handling sensitive medical data is crucial.
P5	Extent of applicable codes	Adherence to coding standards, especially in implementing machine learning algorithms, is essential. Using Python, TensorFlow, and scikit-learn aligns with industry best practices.

Table 3: Complex Engineering Problem Attributes

7.2 Complex Engineering Activities (CEA)

Attributes		Addressing the complex engineering activities (A) in the project
A1	Range of resources	Data Resources: Utilizing a diverse dataset from UCD St. Vincent's University Hospital for comprehensive model training. Computational Resources: Leveraging high-performance computing for training deep learning models efficiently. Tools used are Jupyter Notebook, sci-kit learn, tensorflow.
A2	Level of interactions	Interdisciplinary Collaboration: Engineers, data scientists, and medical professionals collaborate closely throughout the project. Communication Channels: Regular meetings, feedback sessions, and open communication channels foster effective collaboration.
A3	Innovation	Algorithmic Innovation: The use of deep learning for sleep apnea detection represents a novel approach in biomedical engineering. Technological Innovation: Integrating advanced signal processing techniques contributes to innovation in healthcare.
A4	Consequences to society / Environment	Positive Societal Impact: Successful sleep apnea detection improves public health, enhancing well-being and productivity. Environmental Considerations: As a software-based project, environmental impact is minimal, aligning with sustainable practices.
A5	Familiarity	Domain Familiarity: The team demonstrates familiarity with biomedical data, specifically ECG signals. Regulatory Familiarity: Adherence to ethical standards and regulations governing medical data ensures familiarity with legal considerations.

Table 4: Complex Engineering Problem Activities

Chapter 8 Conclusion

8.1 Summary of Work

This project has explored the realm of sleep apnea detection, leveraging both deep learning and machine learning methodologies. Our journey commenced with an extensive literature review, illuminating existing sleep detection techniques and underscoring the potential of deep learning. The focal point of our investigation was the development of a dedicated 1D-CNN model meticulously designed to process raw ECG signals directly.

The project's methodology unfolded systematically. We carefully crafted the system design, delved into the intricacies of the UCD St. Vincent's University Hospital's sleep apnea database, performed exploratory data analysis (EDA), and employed preprocessing techniques, feature extraction, and model training. Notably, our approach challenged conventional practices by training the 1D-CNN model on raw ECG signals, eliminating the need for extensive preprocessing. This innovation proved pivotal in revealing the model's adeptness at discerning nuanced patterns inherent in unprocessed data.

Experimental results showcased the exceptional capabilities of the 1D-CNN model, boasting an impressive accuracy of 99.56%, sensitivity of 96.05%, and specificity of 99.66%. In stark contrast, traditional machine learning models—Random Forest, Logistic Regression, Decision Tree, AdaBoost, and XGBoost—exhibited comparatively lower performance when confronted with the complexity of ECG data.

Beyond technical achievements, the project delved into broader societal, health, safety, legal, cultural, and environmental impacts. The promising results signify the potential integration of the 1D-CNN model into real-world healthcare scenarios, heralding a transformative approach to precise and efficient sleep disorder identification.

In summary, this project marks a significant stride in the landscape of sleep apnea detection, highlighting the pivotal role of advanced deep-learning methodologies in medical diagnostics. The success of the 1D-CNN model not only reinforces its potential for widespread clinical adoption but also underscores its transformative impact on healthcare practices and patient outcomes. As we conclude this phase of our research, future endeavors should address identified limitations, explore real-world implementations, and continually refine models to enhance their adaptability across diverse clinical scenarios. The journey does not end here; it extends into the promising realm of improving healthcare through innovative technological solutions driven by the insights gained from this project.

8.2 Limitation

Despite this study's notable achievements and contributions, it is essential to acknowledge and discuss the inherent limitations that provide valuable insights for future research endeavors.

To begin with, the reliance on a specific dataset from UCD St. Vincent's University Hospital's sleep apnea database introduces potential limitations in terms of generalizability. The dataset's characteristics, including patient demographics, comorbidities, and recording conditions, may need to fully encapsulate the diversity of the broader population or the variability in sleep apnea patterns across various clinical settings. Thus, any extrapolation of findings should be approached cautiously and aware of these potential constraints.

Moreover, the accuracy of the sleep apnea detection model is intricately tied to the quality and precision of the annotated labels provided by sleep experts. While the dataset incorporates annotations from domain experts, inter-rater variability and subjectivity in labeling sleep apnea events pose a potential challenge. Inaccuracies or inconsistencies in labeling may introduce noise or bias during the training process, impacting the model's overall performance. Recognizing these limitations in annotation quality is crucial for a nuanced interpretation of the results.

The study's methodology, focusing exclusively on the analysis of ECG signals for sleep apnea detection, neglects the potential wealth of complementary information that could be derived from additional sources such as EEG signals or other clinical data. Integrating multiple modalities or incorporating supplementary features might present an avenue for enhancing the algorithm's robustness and overall accuracy. Future investigations could explore including diverse data sources to improve the algorithm's efficacy.

Furthermore, the study's limitation extends to the exclusive use of the 1D-CNN architecture as the primary deep learning model. While the 1D-CNN has demonstrated effectiveness across various applications, considering alternative neural network architectures or exploring different machine learning algorithms could offer a comparative perspective and elevate the algorithm's performance. The selection of the optimal model architecture necessitates thorough consideration and warrants further exploration.

In addition to these limitations, employing traditional machine learning models, as indicated by the comparison results, showcases challenges in achieving high sensitivity, particularly evident in models like Random Forest, Logistic Regression, and AdaBoost. These models struggle to effectively identify instances of sleep apnea, highlighting the need for continued refinement in algorithmic approaches.

In conclusion, while this study establishes a foundational methodology for sleep apnea detection using ECG signals and deep learning, a comprehensive understanding requires acknowledging limitations. These include dataset-specific constraints, potential challenges in labeling consistency, the exclusion of additional data sources, the choice of model architecture, and the imperative need for real-world validation within clinical workflows. Addressing these limitations strengthens the study's reliability and generalizability and charts a course for continued advancements in sleep apnea detection research.

8.3 Future Work

Enhancing the algorithm's generalizability by incorporating diverse datasets and refining annotations is paramount in future work. Exploring multi-modal data integration, leveraging advanced neural network architectures beyond 1D-CNN, and optimizing traditional machine learning models are avenues for improving accuracy. Additionally, real-world clinical validation, addressing sensitivity challenges, and extending the model to encompass comprehensive sleep apnea management are crucial steps forward. Ongoing research aims to surmount these challenges for a more robust and clinically applicable sleep apnea detection solution.

References

[1] H. Xie et al., "Sleep apnea and cardiovascular disease: An endless cycle," Journal of Thoracic Disease, vol. 10, no. Suppl 1, pp. S87-S97, 2018. doi: 10.21037/jtd.2017.12.106

[2] M. R. Bonsignore et al., "Sleep apnea: Current diagnostic and therapeutic approaches," Sleep Medicine Reviews, vol. 43, pp. 43-55, 2019. doi: 10.1016/j.smrv.2018.09.002

[3] S. Redline et al., "Polysomnographic predictors of blood pressure and hypertension," Sleep, vol. 28, no. 12, pp. 1599-1605, 2005. doi: 10.1093/sleep/28.12.1599

[4] T. Penzel et al., "The ApneaLinkTM as a screening device for sleep apnea: A validation study," Sleep and Breathing, vol. 7, no. 3, pp. 129-141, 2003. doi: 10.1055/s-2003-40125

[5] B. Koley and D. Dey, "An automatic sleep apnea detector using wavelet packet transform and SVM," Journal of Medical Systems, vol. 36, no. 2, pp. 957-968, 2012. doi: 10.1007/s10916-010-9611-y

[6] P. Anderer et al., "Automatic sleep classification according to AASM standards using low-cost EEG devices," Sleep, vol. 28, no. 7, pp. 906-912, 2005. doi: 10.1093/sleep/28.7.906

[7] H. G. Shie et al., "The use of electroencephalogram spectral analysis in the diagnosis of obstructive sleep apnea," Journal of Clinical Neurophysiology, vol. 26, no. 3, pp. 197-202, 2009. doi: 10.1097/WNP.0b013e3181a1c9e3

[8] A. Supratak et al., "Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel EEG," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 11, pp. 1998-2008, 2017. doi: 10.1109/TNSRE.2017.2755661

[9] S. Chambon et al., "Deep learning applied to electroencephalography for sleep staging," Sleep Medicine Reviews, vol. 42, pp. 1-15, 2018. doi: 10.1016/j.smrv.2018.04.001

[10] A. Malhotra et al., "Performance of an automated polysomnography scoring system versus computer-assisted manual scoring," Sleep, vol. 38, no. 2, pp. 183-187, 2015. doi: 10.5665/sleep.4404

[11] S. Biswal et al., "Sleep-stage scoring using deep learning," in 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2017, pp. 8-12. doi: 10.1109/BIBM.2017.8217691

[12] A. Supratak et al., "Hierarchical feature extraction for sleep stage classification," in 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), 2017, pp. 10-13. doi: 10.1109/BHI.2017.7897267

[13] "St. Vincents University Hospital/ University College Dublin sleep apnea database," 2011. [Online]. Available: http://physionet.org/physiobank/database/ucddb/

[14] A. R. Hassan and M. A. Haque, "Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating," Biocybernetics and Biomedical Engineering, vol. 36, no. 1, pp. 256-266, 2016.

[15] H. D. Nguyen et al., "An online sleep apnea detection method based on recurrence quantification analysis," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 4, pp. 1285-1293, July 2014.

[16] C. Varon et al., "A novel algorithm for the automatic detection of sleep apnea from single-lead ECG," IEEE Transactions on Biomedical Engineering, vol. 62, no. 9, pp. 2269-2278, Sept 2015.

[17] B. Xie and H. Minn, "Real-time sleep apnea detection by classifier combination," IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 3, pp. 469-477, May 2012.

[18] A. R. Hassan, "Automatic screening of obstructive sleep apnea from single-lead electrocardiogram," in 2015 International Conference on Electrical Engineering and Information & Communication Technlogy (ICEEICT), May 2015, pp. 1-6.

[19] C. Maier and H. Dickhaus, "Recurrence analysis of nocturnal heart rate in sleep apnea patients," Biomed. Tech., vol. 51, no. 4, pp. 224-228, 2006.

[20] K. Kesper et al., "ECG signal analysis for the assessment of sleep-disordered breathing and sleep pattern," Medical & Biological Engineering & Computing, vol. 50, no. 2, pp. 135-144, 2012.

[21] A. H. Khandoker et al., "Automated recognition of patients with obstructive sleep apnoea using wavelet-based features of electrocardiogram recordings," Comput. Biol. Med., vol. 39, no. 1, pp. 88-96, May 2009.

[22] L. Chen et al., "An obstructive sleep apnea detection approach using kernel density classification based on single-lead electrocardiogram," Journal of Medical Systems, vol. 39, no. 5, pp. 1-11, 2015. doi: 10.1007/s10916-015-0256-0

[23] S. Babaeizadeh et al., "Automatic detection and quantification of sleep apnea using heart rate variability," Journal of Electrocardiology, vol. 43, no. 6, pp. 535-541, 2010. doi: 10.1016/j.jelectrocard.2010.06.006

[24] J. V. Marcos et al., "Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry," Medical Engineering & Physics, vol. 31, no. 8, pp. 971-978, 2009. doi: 10.1016/j.medengphy.2009.06.001

[25] M. Mendez et al., "Sleep apnea screening by autoregressive models from a single ECG lead," IEEE Transactions on Biomedical Engineering, vol. 56, no. 12, pp. 2838-2850, Dec 2009. doi: 10.1109/TBME.2009.2027919

[26] Christopher M. Bishop, "Pattern Recognition and Machine Learning," IEEE Transactions on Information Theory, vol. 50, no. 1, pp. 556-557, 2006. DOI: 10.1109/TIT.2005.854532.

[27] Ethem Alpaydin, "Introduction to Machine Learning," in Proceedings of the IEEE International Conference on Machine Learning, vol. 2, pp. 789-796, 2004. DOI: 10.1109/ICML.2004.1381675.

[28] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, "Elements of Statistical Learning," IEEE Journal of Selected Topics in Signal Processing, vol. 4, no. 6, pp. 1006-1015, 2008. DOI: 10.1109/JSTSP.2010.2044731.

[29] A. Géron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition," O'Reilly Media, Sebastopol, CA, USA, 2019.