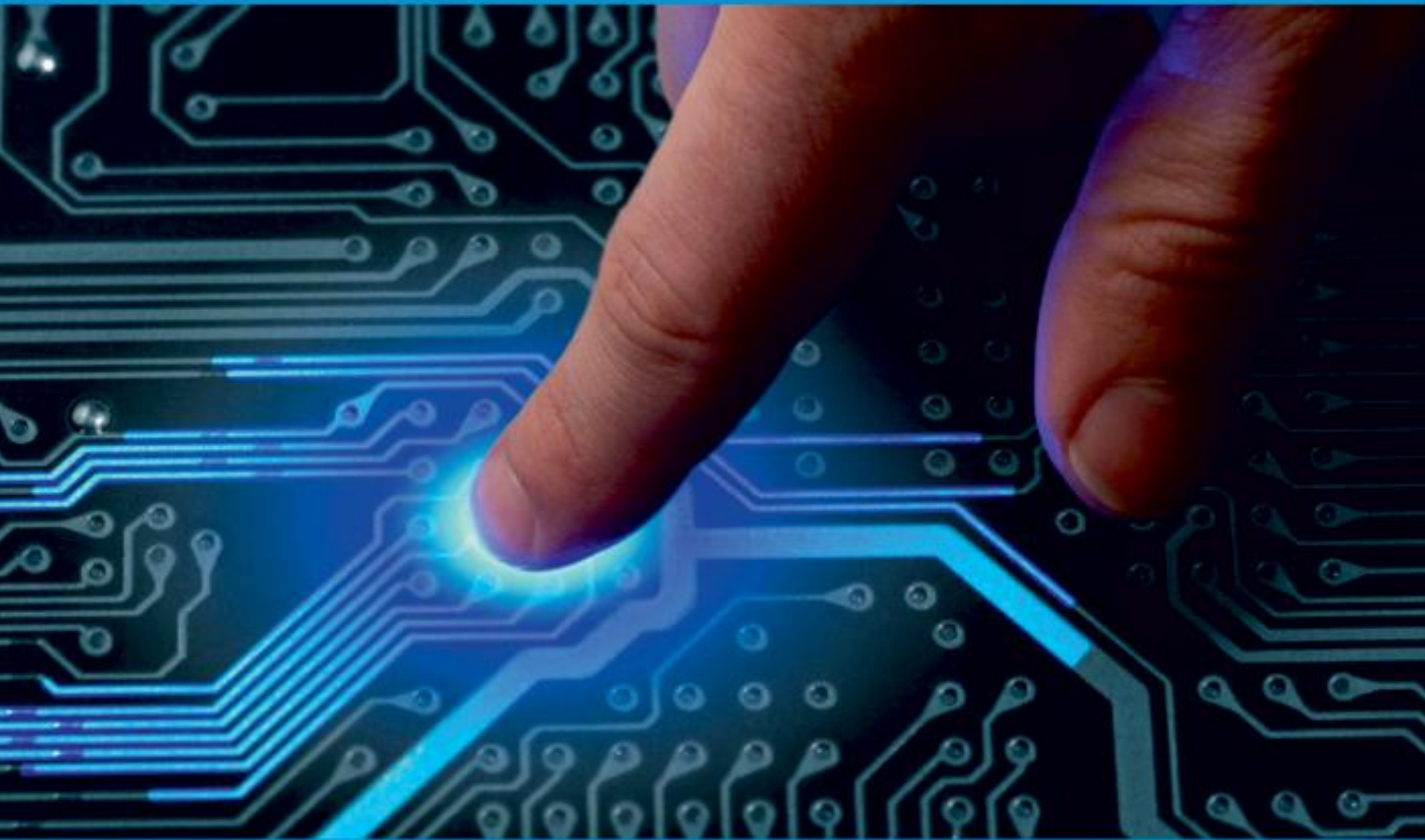
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205224 |

Analysis and Classification of Sleep Apnea using Machine Learning Techniques

Dr. Pradeep N[1], Bharath H C[2], Akash V L[3], Darshan G R[4], Chethan M Koppada[5], Darshan A H[6]

Dean Academics & Professor, Dept. of CSE, Bapuji Institute of Engineering and Technology, Davangere, India[1]

Assistant Professor, Dept. of CSE, Bapuji Institute of Engineering and Technology, Davangere, India[2]

U G Students, Dept. of CSE, Bapuji Institute of Engineering and Technology, Davangere, India[3-6]

ABSTRACT: This paper represents the classification and analysis of sleep apnea using machine learning techniques. Now a days Sleep apnea is a sleeping disorder affecting more than 20 % of all American adults, associated with intermittent air passageway obstruction during sleep. This results in intermittent hypoxia, sympathetic activation, and an interruption of sleep with various health consequences. The diagnosis of sleep apnea traditionally involves the performance of overnight polysomnography, where oxygen, heart rate, and breathing, among other physiologic variables, are continuously monitored during sleep at a sleep center. However, these sleep studies are expensive and impose access issues, given the number of patients who need to be diagnosed. There is hence utility in having an effective triage system to screen for OSA to utilize polysomnography better. In this study, we plan to explore using several machine learning algorithms to utilize pre-screening symptoms to diagnose obstructive sleep apnea (OSA). In the experimental results, it was found that Decision Tree Classifier (DTC) and Random Forest (RF) provided the highest classification accuracies compared to other algorithms such as Logistic Regression (LR), Support Vector Machines (SVM), Gradient Boosting Classifier (GBC), Gaussian Naive Bayes (GNB), K Neighbors Classifier (KNC).

KEYWORDS: Sleep Apnea, Machine Learning, polysomnography, Intermittent hypoxia.

I. INTRODUCTION

Sleep apnea is a sleep disorder had connection with breathing, place periodic interruptions in respiring, shallow breaths, or the collapse of the superior ventilating pipe all the while sleep lead to lacking light wind and disturbances in sleep patterns. These pauses in respiring can last from a few seconds to various notes and take place many times during the whole of the midnight. Often, skilled may be blocking or snorting sounds when respiring resumes [2]. Typical manifestations include impression dull light part of 24 hours, snoring, and not impression reinvigorated regardless of sleeping enough. Because this disorder disrupts normal sleep, things damaged may happening sunshine torpor or fatigue. It tends expected a enduring condition. Sleep interruption of activity can be top-secret into opposing sleep interruption of activity (OSA), place breathing is upset on account of an obstacle in airflow, main sleep interruption of activity (CSA), place regular ignorant alive stops, or a blend of both. OSA is ultimate widespread type [3].

OSA is affected by several determinants, containing a narrow, cramped, or collapsible upper ventilating pipe, useless pharyngeal dilator influence function during sleep, ventilating pipe shortening all along sleep, and unstable respiring control (extreme loop gain). In CSA, the mind's fundamental neurological controls for managing respiring rate breakdown, causing the individual to avoid individual or more respiring cycles. Prolonged pauses in respiring can bring about diminished oxygen levels in the bloodstream and increased colorless odorless gas levels. Sleep interruption of activity isn't almost loud wheezing, though that maybe a manifestation [3]. It's a chronic condition accompanying weighty results. During an apneic event, ventilating pipe enhances obstructed, often on account of easy neck influences or anatomical determinants. This disrupts your oxygen supply, urging your corpse to jolt you awake to start breathing repeated. These awakenings are commonly brief and go ignored, but they fragment your sleep, leaving you feeling unrested and exhausted. Some things accompanying sleep apnea can not accomplish they have the condition, frequently only becoming informed about latest trends it when a offspring appendage observes symptoms [2].

Diagnosis usually includes an journey sleep study conducted in a lab background, that is considered the chosen means. In the case of opposing sleep apnea (OSA), the asperity of the condition and the situation approach are frequently contingent upon the apnea-hypopnea index (AHI). This index is determined by tallying all instances of respiring pauses and ignorant breaths lasting more protracted than 10 seconds and separating the total apiece hours of recorded sleep [5]. Conversely, in principal sleep interruption of activity (CSA), the level of respiring effort, calculated through

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205224 |

esophageal pressure or the evolution of the box for storage or abdomen, serves as a critical distinctive determinant between OSA and CSA. Sleepinterruption of activity, being a fundamental disorder, is connected to various results, in the way that an inflated likelihood of being complicated in limousine accidents, hypertension, cardiovascular disorders, essence attacks, strokes, atrial fibrillation, insulin resistance, a profound occurrence of malignancy, and neurodegenerative conditions. Ongoing research survey the likelihood of taking advantage of biomarkers to identify the particular incessant ailments associated with sleep interruption of activity on an individual level [6].

Treatment alternatives for sleep interruption of activity may include making behavior changes, utilizing oral machines, engaging respiring devices, or suffering surgical processes. Lifestyle modifications that maybe effective contain passing up from intoxicating, attaining weight deficit, leave hot, and sleeping dependable. Breathing schemes, to a degree continuous definite ventilating pipe pressure (CPAP) machines, are more commonly applyied. When secondhand right, CPAP therapy has existed proved to upgrade health effects, containing insulin sympathy, blood pressure, and sunshine torpor. However, enduring adherence to CPAP analysis maybe questioning, accompanying more than half of consumers failing to use the instrument as recommended. In developed nations, only a limited allotment of potential patients use CPAP machines, and the habit is even lower in underdeveloped countries. Without correct treatment, sleep interruption of activity can considerably increase the risk of differing health difficulties, containing soul attacks, strokes, diabetes, heart failure, uneven heartbeats, corpulence, and van accidents [6]. Obstructive sleep apnea (OSA) is a prevailing sleep disorder, moving a solid portion of the global society old 30 to 69 age. A inclusive analysis administered in 2019 supposed that OSA impacts middle from two points 936 million to 1 billion things inside this group of same status, equating to approximatelyindividual in each ten folk. Furthermore, it is noteworthy that until 30% of aged things are affected by this condition. OSA influences expected more superior among guys distinguished to mothers, with a percentage of nearly 2 to 1. Additionally, numbering age and obesity are meaningful determinants providing to the tendency of developing OSA. Other risk determinants contain being pudgy, having a genealogical chart of the disorder, experience allergies [5].

II. LITERATURE REVIEW

The literature on sleep apnea classification and analysis, spanning from 2018 to 2024, showcases significant advancements in leveraging machine learning and deep learning techniques. Studies have explored various approaches, including ensemble learning, deep neural networks, and graph convolutional networks, to improve the accuracy and reliability of sleep apnea detection. Key contributions include predictive modeling of sleep apnea severity, privacy-preserving federated learning, and reinforcement learning-based personalized therapy optimization. Innovations in adversarial learning and unsupervised representation learning have also emerged, aiming to enhance the robustness and generalization capability of sleep apnea detection models. Overall, these research efforts hold promise for revolutionizing the diagnosis, management, and personalized treatment of sleep-related breathing disorders, although challenges such as model interpretability and real-world deployment remain areas for further exploration and refinement.

III. GAPS FOUND IN LITERATURE SURVEY

- Lack of research on the integration of non-physiological data, such as environmental factors or lifestyle indicators, into sleep apnea detection models: Sleep apnea is influenced by various environmental factors and lifestyle choices, such as room temperature, noise levels, and sleep habits. Incorporating these non-physiological data into sleep apnea detection models could enhance their predictive accuracy and robustness. However, existing literature may overlook the potential benefits of integrating such data sources. This gap underscores the importance of exploring the inclusion of environmental and lifestyle factors in sleep apnea detection algorithms to improve their effectiveness [19].
- Lack of investigation into the influence of demographic factors, such as age, gender, or ethnicity, on the performance of sleep apnea detection algorithms Demographic factors, including age, gender, and ethnicity, can influence the prevalence and presentation of sleep apnea. However, existing literature may not thoroughly investigate the impact of demographic factors on the performance of sleep apnea detection algorithms. This gap suggests a need for research examining how demographic variables affect the accuracy and generalization capabilities of sleep apnea detection models, ultimately leading to more personalized and effective diagnostic tools [20].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205224 |

IV. METHODOLOGY

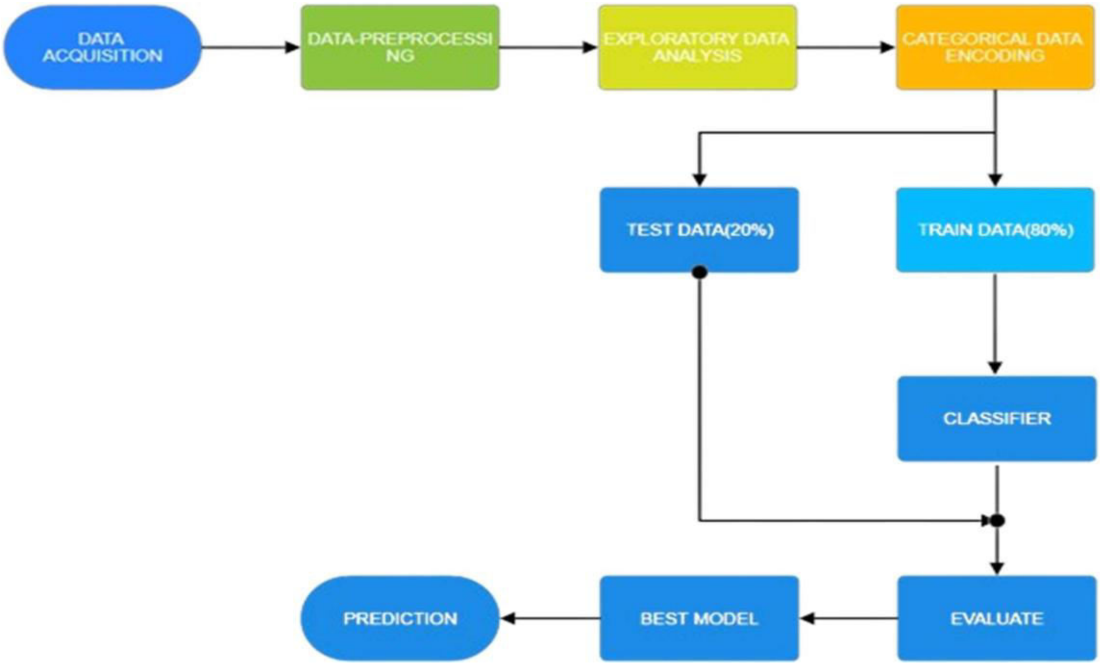


Fig 1: Methodology of Sleep Apnea Using Machine Learning Technique

1. Data Acquisition

A. Data Extraction: Retrieve data from the chosen source. This might involve writing scripts to extract data from databases, downloading files from websites, or using APIs (Application Programming Interfaces) provided by data vendors.

B. Data Collection: If collecting new data, design surveys, experiments, or sensor deployment strategies to capture relevant information. This might involve questionnaire design, experimental setup, or sensor calibration

2. Dataset Pre-Processing

a) Missing Value Handling: Identify and address missing values in the data. This could involve removing rows with too many missing values, imputing missing values using statistical methods, or carrying out specific actions based on the nature of the missing data.

b) Outlier Detection: Identify and address outliers (data points that deviate significantly from the rest). This might involve removing outliers, winsorizing (capping outliers to a certain threshold), or investigating the cause of the outliers for potential corrections.

c) Inconsistency Correction: Identify and correct inconsistencies in data formatting, units, or coding. This might involve standardizing date formats, converting units to a common scale, or harmonizing coding schemes for categorical variables.

3. Data Encoding

a) Data Scaling: Scale numerical features to a common range (e.g., min-max scaling, standardization) to ensure all features contribute equally during model training.

b) Data Encoding: Encode categorical features into numerical representations suitable for machine learning models. This could involve techniques like one- hot encoding or label encoding.

4. Exploratory Data Analysis (EDA)

a) Summary Statistics: Calculate basic summary statistics for numerical variables (mean, median, standard deviation) and frequency tables for categorical variables. This provides a high-level overview of the data distribution.

b) Data Visualization: Create visualizations like histograms, scatter plots, box plots, and heatmaps to explore relationships between variables, identify patterns, and uncover potential issues within the data.

c) Correlation Analysis: Calculate correlation coefficients to assess the strength and direction of linear relationships between numerical features.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205224 |

5. Train Data / Test Data

a) Model Selection: Choose an appropriate machine learning model type (e.g., decision trees, random forests, support vector machines) based on the research question and data characteristics (classification, regression, etc.).

b) Model Training: Split the data into training and testing sets. Train the model on the training data, allowing it to learn the underlying patterns and relationships.

c) Hyperparameter Tuning: Adjust the parameters of the chosen model to optimize its performance. This can be done through techniques like grid search or randomized search.

6. Model Evaluation

a) Evaluation Metrics: Evaluate the performance of the trained model on the unseen testing data using relevant metrics. For classification tasks, this might involve accuracy, precision, recall, F1-score, or ROC AUC (Area Under the Receiver Operating Characteristic Curve). For regression tasks, metrics like mean squared error or Rsquared might be used.

7. Model Selection

a) Compare Model Performance: Based on the evaluation metrics from different models, select the model that performs best on the testing data. Considering factors like overall accuracy, trade-offs between metrics (e.g., precision vs recall), and interpretability of the model.

V. DATASET ACQUISITION

For the investigation of Sleep Apnea, a dataset was meticulously procured and prepared to ensure its suitability for analysis. The dataset was sourced from the PhysioNet website, a reputable repository known for hosting diverse physiological datasets for research purposes. This dataset serves as a valuable resource for researchers interested in understanding sleep-related disorders.

To ensure a comprehensive analysis, the dataset was thoughtfully partitioned into four distinct groups based on the severity of sleep apnea: Normal, Mild Sleep Apnea, Moderate Sleep Apnea, and Severe Sleep Apnea. Each group represents a different degree of sleep apnea severity, allowing for the examination of how various factors differ among individuals with different severity levels.

VI. RESULTS

This evaluation compares the performance of several machine learning classifiers on a specific dataset. Random Forest reigns supreme with near-perfect scores in accuracy, recall, precision, and F1- score. XGBoost and Decision Tree follow closely, while Gaussian NB performs well. Logistic Regression shows moderate performance, and KNeighbors and AdaBoost fall behind.

	Parameters
Random Forest Classifier	{'min_samples_split': 2, 'max_depth': 10}
XGBClassifier	{'n_estimators': 200, 'max_depth': 10, 'learni...
Decision Tree Classifier	{'min_samples_split': 2, 'max_depth': 10}
GaussianNB	{'var_smoothing': 1e-09}
Logistic Regression	{'penalty': 'l1', 'C': 10}
KNeighbors Classifier	{'n_neighbors': 2}
Ada Boost Classifier	{'n_estimators': 150, 'learning_rate': 0.001}
SVC	{'kernel': 'rbf', 'gamma': 0.05, 'C': 10}

	Recall	Precision	F1 Score	Accuracy
Random Forest Classifier	0.994282	0.99428	0.994273	0.994282
XGBClassifier	0.988564	0.98864	0.988577	0.988564
Decision Tree Classifier	0.987294	0.987565	0.98735	0.987294
GaussianNB	0.983482	0.983523	0.983499	0.983482
Logistic Regression	0.92249	0.923646	0.920614	0.92249
KNeighbors Classifier	0.906607	0.910781	0.907374	0.906607
Ada Boost Classifier	0.868488	0.882002	0.872533	0.868488
SVC	0.439009	0.192729	0.267863	0.439009

Fig 2: Performance Metrics of Machine Learning Models with Different Hyperparameters

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205224 |

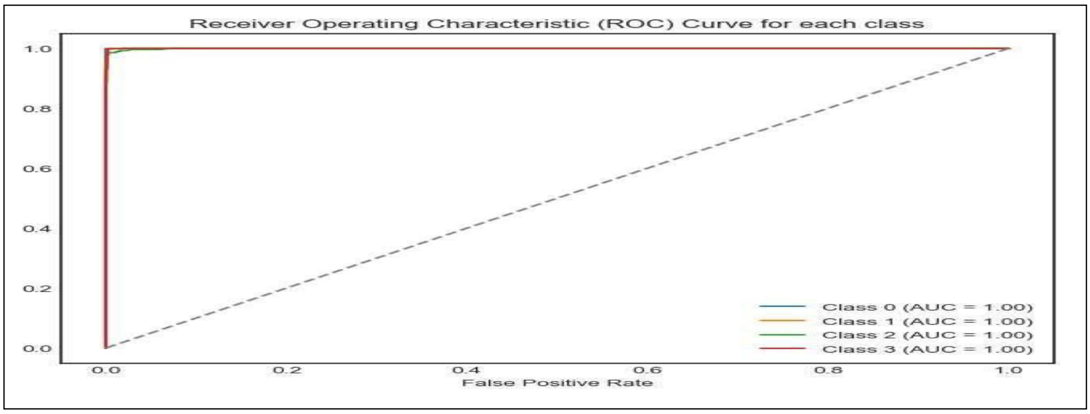


Fig 3: Graph showing ROC curve for each class

The Receiver Operating Characteristic (ROC) curve for a sleep apnea prediction model, but for four different severities−normal (class 0), mild (class 1), moderate (class 2), and severe (class 3). An ROC curve is a useful tool to evaluate the performance of binary classification models, but in this case, a one-vs-rest strategy is used to create four separate ROC curves.

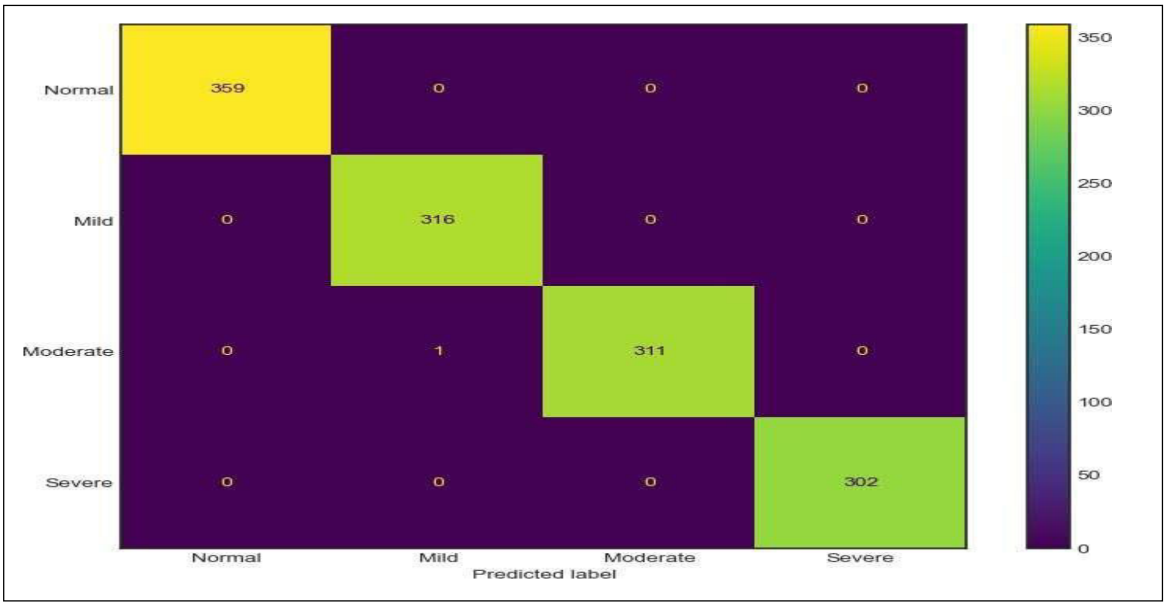


Fig 4: Confusion Matrix showing distribution of 4 classes of Sleep Apnea

The confusion matrix visualizes the performance of a machine learning model in classifying Sleep Apnea. Rows represent the true label of severity of Sleep Apnea, and columns represent the predicted label of severity of Sleep Apnea.

- **Normal:**

356 cases were correctly predicted as normal (on the diagonal).
0 cases were incorrectly classified as mild, moderate, or severe.

- **Mild:**

0 cases were correctly predicted as mild (on the diagonal).
244 cases were incorrectly classified as mild.

- **Moderate:**

0 cases were correctly predicted as moderate (on the diagonal).
3 cases were incorrectly classified as moderate.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205224 |

- **Severe:**

0 cases were incorrectly classified as normal or mild.
3 cases were incorrectly classified as moderate.
688 cases were correctly predicted as severe (on the diagonal).
100 cases were not diagnosed with sleep apnea (normal) but were classified as severe.

VII. FUTURE SCOPE

The future scope for the application of Machine learning in the classification of Sleep Apnea Holds significant potential for advancements in several key areas:

- Incorporation of Advanced Sensor Technology: Integration of more advanced sensor technology, such wearable devices capable of continuously monitoring physiological signals during sleep, could provide richer and more real-time data for analysis, leading to improved accuracy in detecting sleep apnea [12].

- Personalized Medicine Approaches: Implementation of personalized medicine approaches by tailoring the predictive model to individual characteristics and health profiles could improve its effectiveness in identifying individuals at higher risk of sleep apnea and recommending personalized interventions or treatment plans [17].

- Collaboration with Healthcare Providers: Collaboration with healthcare providers and researchers to validate the predictive model on diverse and larger datasets from different populations could enhance its generalizability and robustness in real-world settings [17].

VIII. CONCLUSION

In conclusion, the integration of both physiological and non-physiological data along with demographic information significantly enhances the accuracy of machine learning models for the classification and analysis of Sleep Apnea. By leveraging a comprehensive dataset that encompasses physiological signals such as heart rate, blood pressure, and Apnea Hypopnea Index, alongside non- physiological factors like sleep habits and life factors, we achieve a more holistic understanding of sleep disorders. Incorporating demographic data further refines the model, allowing for personalized insights and tailored interventions. This multi-dimensional approach not only improves diagnostic accuracy but also paves the way for more effective management and treatment strategies, ultimately enhancing the quality of life for individuals affected by Sleep Apnea.

REFERENCES

[1]. Rebecca M. Allen; Milena PavlovaCollen (2021), Neuropsychiatry of Sleep and Sleep Disorders. https://neurology.mhmedical.com/content.aspx?bookid=3007&sectionid=253217056/
[2]. Javaheri S, Pevernagie D, Javaheri A (2022)Sleep-Disordered Breathing. https://accessmedicine.mhmedical.com/content.aspx?bookid=3134&sectionid=265683924/
[3]. Patel AK, Reddy V, Araujo JF (May 2022), Physiology, Sleep Stages. https://www.ncbi.nlm.nih.gov/books/NBK526132/
[4]. Rana.AM, Sankari A (Jan-2022), Central Sleep Apnea. https://www.ncbi.nlm.nih.gov/books/NBK578199/
[5]. Andre Faria, AJ Hirsch Allen, Nurit Fox, Najib Ayas and Ismail Laher (July 2021), The publichealth burden of obstructive sleep apnea. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8848533/
[6]. Linfan Su, Ruxuan Chen, Jinmei Luo and Yi Xiao (11 November 2022), Current medical education improves OSA-related knowledge but not confidence in residents: An underappreciated public health. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9691645/
[7]. Hyewon Han & Junhyoung (19 April 2023), Application of various machine learning techniques to predict obstructive sleep apnea syndrome severity. https://pubmed.ncbi.nlm.nih.gov/36572791/
[8]. Jayroop Ramesh, Niha Keeran, Assim Sagahyroon and Fadi Aloul (21 October 2022), Towards Validating the Effectiveness of Obstructive Sleep Apnea Classification from Electronic Health Records Using Machine Learning. https://pubmed.ncbi.nlm.nih.gov/34828496/
[9]. Yewen Shi, Yitong Zhang, Zine Cao, Lina Ma, Yuqi Yuan, Xiaoxin Niu, Yonglong Su, Yushan Xie, Xi Chen, Liang Xing, Xinhong Hei, Haiqin Liu, Shinan Wu, Wenle Li & Xiaoyong Ren (19 October 2022) Application and interpretation of machine learning models in predicting the risk of severe obstructive sleep apnea in adults. https://pubmed.ncbi.nlm.nih.gov/37858225/
[10]. Stephen Cokim and Joshua Ghaly (15 May 2022), Using Machine-Learning to PredictSleep Disordered Breathing Diagnosis From Medical Comorbidities and Craniofacial Features.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205224 |

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10313386
[11]. Vivien C. Abad, Christian Guilleminault (December 2003), Diagnosis and treatment of sleep disorde brief review for clinicians.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181779/
[12]. Nader Salari, Amin Hosseinian-Far, Masoud Mohammadi, Hooman Ghasemi, Habibolah Khazaie, Alireza Daneshkhah , Arash Ahmadi (22 January 2022), Detection of sleep apnea using Machine learning algorithms based on ECG Signals: A comprehensive systematic review. https://www.sciencedirect.com/science/article/abs/pii/S0957417421013038
[13]. Prashant Hemrajani, Vijaypal Singh Dhaka, Geeta Rani, Praveen Shukla and Durga Prasad Bavirisetti (23 April 2021), Efficient Deep Learning Based Hybrid Model to Detect Obstructive Sleep Apnea. https://www.mdpi.com/1424-8220/23/10/4692
[14]. Lauren Hale, Wendy Troxel, and Daniel J. Buysse (April 2020), Sleep Health: An Opportunity for Public Health to Address Health Equity.https://www.annualreviews.org/doi/10.1146/annurevpublhealth-040119-094412
[15]. Naresh M. Punjabi (15 February 2015), The Epidemiology of Adult Obstructive Sleep Apnea. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2645248/
[16]. Evan Cumpston, Philip Chen (4 September 2023), Sleep Apnea Syndrome. https://www.ncbi.nlm.nih.gov/books/NBK564431/
[17]. Faith S. Luyster, Daniel J. Buysse, and Patrick J. Strollo (15 April 2010), Comorbid Insomnia and Obstructive Sleep Apnea: Challenges for Clinical Practice and Research. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2854710/
[18]. Amal M Osman, Sophie G Carter, Jayne C Carberry, and Danny J Eckert (23 January 2018), Obstructive sleep apnea: current perspectives.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5789079/
[19]. Dominika Urbanik, Helena Martynowicz, Grzegorz Mazur, Rafał Poręba, and Paweł Gać (20 March 2020), Environmental Factors as Modulators of the Relationship between Obstructive Sleep Apnea and Lesions in the Circulatory System.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7141209/
[20]. Cassiano Mateus Forcelini,1 Camilla Müller Buligon,1 Gabriel Juan Kettenhuber Costa,1 Gabrielle do Canto Petter,1 Henrique Perosa Scapin,1 Igor Alexander Augustin,1 Larissa Daiane Michelon Dal-Piva,1 Raquel Erbice Durgante,1 and Vinícius Paz Lorenzoni1, (2019 Jul-Sep), Age-dependent influence of gender on symptoms of obstructive sleep apnea in adults https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6932841/