



Predictive Power of XGBoost_BiLSTM Model: A Machine-Learning Approach for Accurate Sleep Apnea Detection Using Electronic Health Data

Ashir Javeed^{1,2} · Johan Sanmartin Berglund² · Ana Luiza Dallora² · Muhammad Asim Saleem³ · Peter Anderberg^{2,4}

Received: 5 October 2023 / Accepted: 30 October 2023
© The Author(s) 2023

Abstract

Sleep apnea is a common disorder that can cause pauses in breathing and can last from a few seconds to several minutes, as well as shallow breathing or complete cessation of breathing. Obstructive sleep apnea is strongly associated with the risk of developing several heart diseases, including coronary heart disease, heart attack, heart failure, and stroke. In addition, obstructive sleep apnea increases the risk of developing irregular heartbeats (arrhythmias), which can lead to low blood pressure. To prevent these conditions, this study presents a novel machine-learning (ML) model for predicting sleep apnea based on electronic health data that provides accurate predictions and helps in identifying the risk factors that contribute to the development of sleep apnea. The dataset used in the study includes 75 features and 10,765 samples from the Swedish National Study on Aging and Care (SNAC). The proposed model is based on two modules: the XGBoost module assesses the most important features from feature space, while the Bidirectional Long Short-Term Memory Networks (BiLSTM) module classifies the probability of sleep apnea. Using a cross-validation scheme, the proposed XGBoost_BiLSTM algorithm achieves an accuracy of 97% while using only the six most significant features from the dataset. The model's performance is also compared with conventional long-short-term memory networks (LSTM) and other state-of-the-art ML models. The results of the study suggest that the proposed model improved the diagnosis and treatment of sleep apnea by identifying the risk factors.

Keywords Sleep apnea · Deep learning · Computer vision · Feature engineering

1 Introduction

A common sleep disorder known as sleep apnea and hypopnea syndrome (SAHS) is characterized by abnormal reductions or pauses in breathing during sleep [1]. It is estimated

to affect 2% of middle-aged women and 4% of middle-aged men. People suffering from severe sleep apnea are prone to develop coronary artery disease, congestive heart failure, and stroke [2]. Obstructive sleep apnea (OSA) is a common problem that affects at least 5% of adults [3, 4] and is linked to a higher chance of hypertension [5], coronary heart disease [6], atrial and ventricular arrhythmias [7], and mortality [8]. The physiologic linkages between obstructive airway events and cardiac pathology are multi-factorial and carefully docu-

✉ Ashir Javeed
ajv@bth.se

Johan Sanmartin Berglund
johan.sanmartin.berglund@bth.se

Ana Luiza Dallora
ana.luiza.moraes@bth.se

Muhammad Asim Saleem
asim.saleem@riphahfsd.edu.pk

Peter Anderberg
peter.anderberg@bth.se

¹ Aging Research Center, Karolinska Institutet, Tomtebodavägen 18a, Slöna 17165, Stockholm, Sweden

² Department of Health, Blekinge Institute of Technology, Valhallavägen 1, Karlskrona 37179, Blekinge, Sweden

³ Center of Excellence in Artificial Intelligence, Machine Learning and Smart Grid Technology, Chulalongkorn University, Jan Waldenströms g. 35, Bangkok 10330, Bangkok, Thailand

⁴ School of Health Sciences, University of Skövde, Högskolevägen 1, Skövde 54128, Västra Götaland, Sweden

mented in a consensus document from the American College of Cardiology [9] and the American Heart Association [10]. The association between obstructive sleep apnea and heart disease is supported by facts that treating sleep apnea lowers systolic blood pressure, improves left ventricular systolic function, and lowers platelet activation [11].

Sleep apnea can be divided into two categories: Central sleep apnea is caused by a disruption of normal communication between the brain and the respiratory muscles, while obstructive sleep apnea is caused by upper airway obstruction. Patients with both OSA and heart failure have a high co-occurrence of different conditions. Studies show a range of prevalence from 11 to 38%. In patients diagnosed with CSA, [12]. The data gives rise to speculation about how best to identify and treat co-occurring disorders. Patients with sleep apnea are more likely to be hospitalized for an exacerbation of HF [13]. There are only positive results for obstructive apnea, while studies have shown no clear benefit for central apnea. Treatment options are being explored for both [14].

Obstructive sleep apnea (OSA) is a worldwide health crisis that accompanies the global obesity epidemic. In the US, OSA affects 17% of adult women and 34% of adult men. Recent trends show that the number of people with OSA is on the rise in the US and the rest of the world [15]. Sleep apnea has been associated with metabolic syndrome features such as insulin resistance, dyslipidemia, hypertension, and central obesity [16]. The significant association between OSA and cardiovascular disease may be attributed to the metabolic syndrome and its negative effects on inflammation, oxidative stress, and endothelial dysfunction [17].

People with heart diseases are at higher risk for sleep apnea. Large prospective patient registries have shown that sleep apnea, particularly OSA, is relatively common in outpatient and inpatient cardiology settings [18, 19]. OSA has been associated with several health conditions, including hypertension [20], coronary artery disease [21], congestive heart failure [22], stroke [23]. Cardiac arrhythmias, especially atrial fibrillation (AF), [24].

Neurobehavioral disorders are associated with obstructive sleep apnea syndrome (OSAS) [25, 26], heart disease [27, 28], poor quality of life [28, 29], and more physician visits [30], demonstrating the importance of detecting and treating this condition. Therefore, the American Academy of Pediatrics (AAP) recommends screening for OSAS during regular doctor visits. Children with typical symptoms (such as snoring, restless sleep, and daytime hyperactivity) or risk factors (such as craniofacial, neurological, or genetic disorders) should be considered for diagnosis. Overnight polysomnography may confirm the diagnosis.

Over 34 million people worldwide and at least 3 million Americans have AF [31]. Atrial fibrillation is thought to be caused by abnormal atrial tissue substrates and triggers of abnormal impulses, which often originate in the pulmonary

venous ostia[32]. However, the mechanisms by which AF arises are not fully understood. Because of its immediate effects on intrathoracic pressure and autonomic tone and its ability to drive long-term changes in the underlying atrial tissue substrate, OSA may favor the development of AF [33]. Obstructive sleep apnea is more common than usual in people with AF (OSA) of moderate or higher severity [34, 35], and the prevalence of AF in patients with moderate or severe OSA is comparable to that in patients with AF [36]. Machine learning can predict the likelihood of sleep apnea based on several factors, including age, gender, BMI, and other medical conditions.

In addition, the paper [37] discusses using supervised machine learning methods to predict obstructive sleep apnea (OSA). The authors used a noninvasive feature dataset of 231 records and applied common machine-learning algorithms to develop the prediction models. After reviewing the dataset and searching for hidden data, these were replaced with the average and most frequent records. Standard machine learning algorithms were used for modeling, and the overall performance of the models was evaluated using 10-fold cross-validation. The results showed that the Naive Bayes classifiers and logistic regression achieved the best predictive models with an overall AUC of 0.768 and 0.761, respectively. The SVM with a sensitivity of 93.42% and Naive Bayes with a specificity of 59.49% may be suitable for screening high-risk individuals with OSA.

Finally, [38] discusses occlusive sleep apnea syndrome, an airway sleep disorder characterized by intermittent nocturnal episodes of partial or complete upper airway obstruction. The article highlights the high prevalence of this disorder in the elderly population, with an estimated incidence between 20% and 60% in those over 65 years of age. It emphasizes the importance of diagnosing and treating this disorder in older patients as the average age of the world's population increases. The paper does not include specific results of studies or experiments.

This research work makes the following significant contributions and offers an updated overview of the topic:

1. To the best of our knowledge, this is the first study that illustrates the XGBoost_BiLSTM model successfully predicts sleep apnea using EHR.
2. To avoid the problem of model overfitting, XGBoost model is first time deployed for the selection of significant features from the dataset for the prediction of sleep apnea.
3. The XGBoost_BiLSTM approach has a lower time complexity due to the fewer features (six) it uses.
4. The newly developed XGBoost_BiLSTM model also identifies the sleep apnea risk factors, which ultimately assists us in lowering the likelihood of developing sleep apnea.

5. According to experimental findings, the proposed XGBoost_BiLSTM model outperformed other cutting-edge ML models and conventional LSTM in terms of accuracy.

This paper discusses apnea detection using machine learning with cross-domain features. Section 2 explains the concept of deep learning and the techniques that are used in this work. The previous work based on ML for sleep apnea is described in Section 3 (literature review). Section 4 provides the details of the material and methods of this study. Section 5 and Section 6 present the results and discussion of the proposed work. Section 8 concludes the study with an overview of future research.

2 Deep Learning

Deep Learning (DL) is a subfield of artificial intelligence (AI) that simulates how the human brain processes data and creates patterns to enable rational decision-making. It is a branch of machine learning (ML) that provides more sophisticated tools for building models and applies a layered ANN to run ML methods. Deep learning structures, known as deep neural learning (DNL), are built from multiple interconnected layers. DL can learn from incoming data and transform it into various degrees of data abstraction [39, 40]. DL Examples include recurrent neural networks (RNNs), deep neural networks (DNNs), convolutional neural networks (CNNs), and deep belief networks (DBNs). These methods provide admirable results comparable to or superior to human reasoning, making them useful for various problems in numerous fields of study, including intrusion detection.

2.1 Recurrent Neural Networks

Due to their recurrent (circular) connection mode, recurrent neural networks (RNN), a subset of Deep Learning (DL), are a better technique for processing sequential inputs. This class of neural networks maintains hidden layers while using past outputs as inputs [41, 42]. RNNs can process inputs of arbitrary length and keep the size of a model constant as inputs grow. Unlike traditional feed-forward networks, RNNs can remember what they have learned and base their judgments on that information [43]. In other words, RNNs can recall information in addition to what they learned during training while producing output. RNNs are able to handle a variety of research problems but also have problems such as vanishing gradients [43, 44]. Because of this flaw, they are unable to acquire long-term dependencies. Hochreiter and Schmidhuber [45] introduced long short-term memory (LSTM) to solve this problem.

2.2 Long-Short-Term-Memory (LSTM)

The recurrent neural network LSTM uses a gating mechanism to learn long-term dependencies. It fixes the vanishing gradient problem that occurs in traditional RNN training. To avoid units and remember larger time steps, LSTM models use numerous switching gates [46]. The LSTM design often has a memory called cells that accepts input from the current input and the previous state. These cells decide what to keep and delete from memory before merging the previous state with the current input to produce the subsequent input. In this way, they can record long-term [47]. Due to their advantages over traditional RNNs, LSTMs have attracted much interest recently. The most pressing security challenges, such as intrusion detection, are being addressed by network security researchers with LSTMs [48, 49].

3 Literature Review

This section reviews the previous studies that presented ML methods to classify and detect sleep apnea [50]. The sleep apnea study also investigated different ML strategies with diverse types of input data [51]. The aim of the [52] and meta-analysis was to investigate the association between obstructive sleep apnea (OSA) and erectile dysfunction (ED). Masa et.al., performed a prespecified secondary analysis of the largest multicenter randomized controlled trial of OHS (Pickwick Project, $n = 221$ patients with OHS and coexisting severe obstructive sleep apnea) to compare the efficacy of 3 years of NIV and CPAP on structural and functional echocardiographic changes [53]. More importantly, [54] proposes a classification combination to further improve classification performance by using the complementary information provided by each classifier. [55] Study on automatic screening for obstructive sleep apnea using a lead-in electrocardiogram. A proposed algorithm uses a lead-in electrocardiogram (ECG) to detect OSA events.

Furthermore, the goal of [56] is to determine sleep and wakefulness with a practical and applicable method. In [57], the authors investigated an expert system for automatically detecting obstructive sleep apnea from the lead-in ECG using random undersampling boosting. The problem of automatic sleep apnea detection from lead-in ECG signals was addressed.

In addition, [58] that machine learning can automate obstructive sleep apnea (OSA) detection. With tenfold cross-validation, [59] detection achieves 88.3% accuracy for four-group classification and 92.5% in the binary classification. The goal of [60] is to analyze published research over the past decade to answer research questions, such as how to implement different deep networks. In [61], an effective, efficient, and sustainable system for automatic sleep apnea detection

using pulse oximetry (SpO₂) signals indicating the percentage of oxygen in the blood is presented.

In addition, [62] study identified obstructive sleep apnea based on sleep architecture. The patient's sleep stages and transition relationships are used as features to propose a machine learning-based OSA detection method. The proposed method can be a low-cost and reliable wearable device for monitoring sleep apnea at home and in the community [63].

Furthermore, in [64], the authors propose an efficient method to discriminate between patients with obstructive sleep apnea (OSA) and normal control subjects using EEG signals and machine learning algorithms. The delta, theta, alpha, beta, and gamma subbands of the EEG signals were separated. Energy and variance were extracted as descriptive features from each frequency band. Four machine learning algorithms were used to detect OSA: Support Vector Machines (SVM), artificial neural networks (ANN), linear discriminant analysis (LDA), and Naive Bayes (NB). The results showed that SVM achieved the best classification accuracy of 97.14% compared to the other classifiers.

In addition, the paper [37] discusses using supervised machine learning methods to predict obstructive sleep apnea (OSA). The authors used a noninvasive feature dataset of 231 records and applied common machine-learning algorithms to develop the prediction models. After reviewing the dataset and searching for hidden data, these were replaced with the average and most frequent records. Standard machine learning algorithms were used for modeling, and the overall performance of the models was evaluated using 10-fold cross-validation. The results showed that the Naive Bayes classifiers and logistic regression achieved the best predictive models with an overall AUC of 0.768 and 0.761, respectively. The SVM with a sensitivity of 93.42% and Naive Bayes with

a specificity of 59.49% may be suitable for screening high-risk individuals with OSA.

Finally, [38] discusses occlusive sleep apnea syndrome, an airway sleep disorder characterized by intermittent nocturnal episodes of partial or complete upper airway obstruction. The article highlights the high prevalence of this disorder in the elderly population, with an estimated incidence between 20% and 60% in those over 65 years of age. It emphasizes the importance of diagnosing and treating this disorder in older patients as the average age of the world's population increases. The paper does not include specific results of studies or experiments.

4 Material and Methods

4.1 Dataset

The Swedish National Study on Aging and Care (SNAC) served as the data source for this study. SNAC is a long-term consortium that collects multimodal data from Sweden's aging population to develop reliable, comparable, and durable data sets for aging research [65]. SNAC was established as a multipurpose program to study health care quality in the aging population. SNAC includes several databases with information on various topics, such as medical records, social variables, lifestyle factors, metacognitive data, and physical assessment. In Blekinge, Skne, Nordanstig, and Kungsholmen, SNAC collected data on Swedish seniors. Figure 1, provides an overview of positive and negative samples in the collected dataset. The collected dataset for this study consists of 75 features with a total sample size of 10765. Table 1 displays the nature of features as feature groups, feature names, and the total number of features in a particular feature group.

Fig. 1 Description of dataset samples

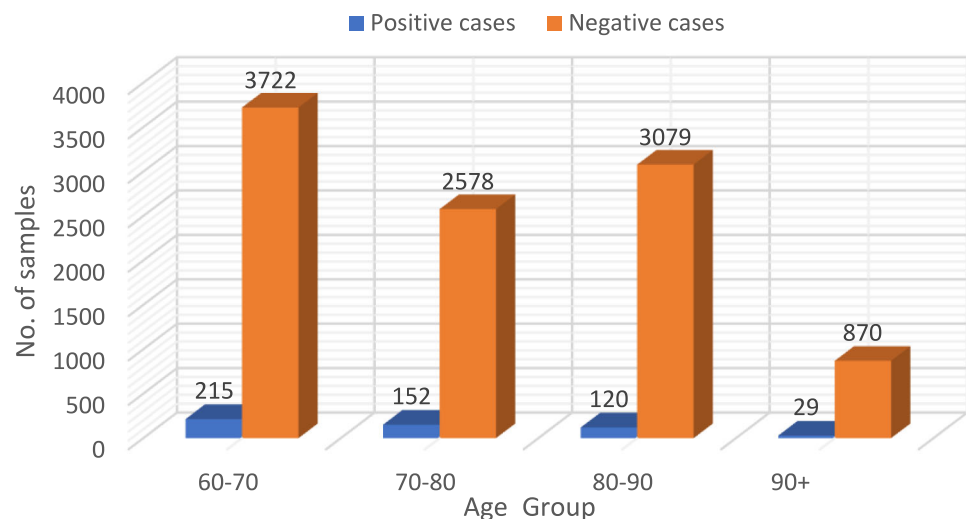


Table 1 Feature Description

| Feature_Group | Feature_Names | Total |
|----------------------|--|-------|
| Demographic | Age, Gender | 02 |
| Lifestyle | Physical-Workload, Past Smoker, Physically Demanding Activities, Present Smoker, Alcohol Quantity, Leisure Activities, Number of Cigarettes a Day, Social Activities, Work Status, Alcohol Consumption, Light Exercise | 11 |
| Social | Support Network, Education, Loneliness, Religious Belief, Social Network, Voluntary Association, Religious Activities | 07 |
| Medical History | High Blood Pressure, Myocardial Infarction, Head Trauma, Arrhythmia, Sleep Apnea, TIA/RIND, Diabetes Type 1, Stroke, Thyroid Disease, Family History of Importance, Cancer, Epilepsy, Number of Medications, Diabetes Type 2, Parkinson's Disease, Other Psychiatric Diseases, Snoring, Atrial Fibrillation, Hip Fracture, Cardiovascular Ischemia, Heart Failure, Developmental Disabilities, Depression | 22 |
| Health Instruments | Backwards Digit Span Test, Digit Span Test, Livingston Index, Sense of Coherence EQ5D Test, Instrumental Activities of Daily Living, Mini-Mental State Examination, Activities of Daily Living, Comprehensive Psychopathological Rating Scale, Physical Composite Score of the SF-12 Health Survey, Clock Drawing Test, Mental Composite Score of the SF-12 Health Survey | 12 |
| Biochemical Test | C-Reactive Protein Analysis, Hemoglobin Analysis | 02 |
| Psychological | Sense of Identity, Memory Loss, Personality Change, Memory Decline, Abstract Thinking, Memory Decline 2 | 06 |
| Physical Examination | Blood Pressure on the Right Arm, Number of Teeth, Body Mass Index (BMI), Assessment of Rising from a Chair, Heart Rate Sitting, Heart Rate Lying, Hand Strength in Right Arm in a 10s Interval, Hand Strength in Left Arm in a 10s Interval, Single-Leg Standing with Right Leg, Single Leg Standing with Left Leg, Pain in the last four weeks, Dental Prosthesis, Feeling of Safety from Rising from a Chair | 13 |

Based on previously published research, variables were selected from SNAC database (Blekinge) and variables from eight categories were considered for this study such as demographics, social, lifestyle, medical history, physical examination, biochemical tests, psychological examination, and assessment of various health instruments [66, 67]. We obtained 10,765 data samples, of which 3461 were from SNAC-Kungsholmen and 7304 from SNAC-Blekinge. The dataset collected consists of 6816 females and 3949 males. Only 229 of the 3949 males and 287 of the 6816 females suffer from sleep apnea. Table 2 shows the statistical information for the sample population.

4.2 Proposed Model

In this study, we presented a ML model that can predict sleep apnea based on EHR. The proposed ML model is based on two components that are hybridized into a single system. The

Table 2 The summary of samples in collected dataset

| Age_Group | Female | | Male | |
|-----------|--------|------|------|------|
| | +ve | -ve | +ve | -ve |
| 60–70 | 96 | 2111 | 119 | 1611 |
| 70–80 | 98 | 1617 | 54 | 961 |
| 80–90 | 71 | 2119 | 49 | 960 |
| 90+ | 22 | 682 | 07 | 188 |
| Total | 287 | 6529 | 229 | 3720 |

first component employed the XGBoost technique to select the most significant variables from the dataset. XGBoost is used to rank the variables from the dataset, and highly ranked variables are fed into the second component for the prediction of sleep apnea. In the second component, we employed conventional LSTM and BiLSTM models. The performance

of conventional LSTM and BiLSTM based on highly ranked features from XGBoost was assessed. From the experimental results, the performance of XGBoost_BiLSTM is evident in comparison to XGBoost_LSTM. Hence, we named the newly designed model XGBoost_BiLSTM. Figure 2 presents an overview of the developed XGBoost_BiLSTM model for the prediction of sleep apnea. Data preprocessing is the first step of the proposed model because BiLSTM deals with only numeric values. Therefore, all non-numeric features in the dataset will be converted into numeric form. After successfully converting non-numeric features to numeric representations, feature scaling is the next step. The feature scaling guarantees that the dataset is normalized. Because the values of several features in the dataset have an uneven distribution, we use Min-Max scaling to scale each feature's values between 0 and 1. This ensures that our classifier does not provide biased results.

The Min-Max feature scaling equation is as follows:

$$S' = \frac{O - O_{min}}{O_{max} - O_{min}} \tag{1}$$

where S' denotes the new scaled values and O represents the original value.

The architecture of XGBoost for features ranking from the dataset, along with the intuition of a conventional LSTM architecture and details of a bidirectional LSTM (BiLSTM) architecture, is given in the below sections.

4.3 Extreme gradient boosting (Xgboost) for Feature Selection

One of the variants of the gradient boost engine is Xgboost, which is considered to be one of the best-supervised learning algorithms available on the market. The fast out-of-core execution speed of Xgboost makes it a favorite among data scientists. In addition to regression and classification problems, Xgboost can be used for feature ranking from the dataset. XGBoost is an ensemble learning algorithm that employs cacheable block structure tree learning and regularized learning. Z denotes the loss function, τ represents the t^{th} tree, and $\eta(\tau)$ is the regularized term. The second order Taylor series of Z at the t^{th} iteration is:

$$Z^t \simeq \sum_{j=1}^n \left[z \left(x_j, x_j^{(t-1)} \right) + \rho_j \tau(y_j) + \frac{1}{2} \beta_j \tau^2(y) \right] + \eta(\tau) \tag{2}$$

where g_j, β_j stand for 1^{st} and 2^{nd} order gradients. Gain is utilized to select the ideal split node throughout XGBoost training.

Gain

$$= \frac{1}{2} \left[\frac{(\sum_j \epsilon_{zZ} g_j)^2}{\sum_j \epsilon_{zZ} \beta_j + \alpha} + \frac{(\sum_j \epsilon_{zR} g_j)^2}{\sum_j \epsilon_{zR} \beta_j + \alpha} - \frac{(\sum_j \epsilon_z g_j)^2}{\sum_j \epsilon_z \beta_j + \alpha} \right] - \psi \tag{3}$$

where Z_z and Z_R represent the left and right nodes, respectively, after segmentation. $z = z_z \cup z_R$. α, ψ are penalty parameters. The average gain is used to determine the final significance value of the feature reflecting the gain for each tree split. The average gain is determined by dividing the total cumulative gain by the total cumulative number of splits for each feature. The more significant and useful a feature is, its value on the XGBoost significance scale is higher. The top features are determined in descending order of importance to describe the PPIs. In bioinformatics, the XGBoost method was used for feature selection. [68]. The loss function is binary: logistic, there are 500 boosting trees, the maximum depth is 15, and all other parameters are set to default values.

4.4 Conventional LSTM

A typical LSTM has the same control flow as a typical RNN, analyzing data and recording information as it propagates. The variations are a consequence of the LSTM's cellular activities. These characteristics allow the LSTM to ignore or store information. The different gates and cell states form the core of an LSTM. The cell state is a channel for transmitting relevant data throughout the data processing. It could be considered as the memory of the network. Different neural networks control the information that can be applied to a cell state by acting as gates. During training, the gates learn which information to store and which to forget. Three different gates control the flow of information within an LSTM cell: the input gate, the output gate, and the forget gate. The input gate determines what information is added from the current state. The output gate determines the type of hidden state that follows. The forget gate determines what must be stored from the previous state. Figure 3 represents a typical LSTM architecture. The following equations provide a mathematical description of the connection between the inputs and outputs at time τ and at time $\tau-1$.

$$\alpha_\tau = \varphi [(\omega_{\rho\alpha} * \rho_\tau) + (\omega_{v\alpha} * v_{\tau-1}) + (\omega_{\mu\alpha} * \mu_{\tau-1}) + \kappa_\alpha] \tag{4}$$

$$\beta_\tau = \varphi [(\omega_{\rho\beta} * \rho_\tau) + (\omega_{v\beta} * v_{\tau-1}) + (\omega_{\mu\beta} * \mu_{\tau-1}) + \kappa_\beta] \tag{5}$$

$$\gamma_\tau = (\beta_\tau * \gamma_{\tau-1}) + \alpha_\tau \tanh [(\omega_{v\gamma} * v_{\tau-1}) + (\omega_{\mu\gamma} * \mu_{\tau-1}) + \kappa_\gamma] \tag{6}$$

$$\phi_\tau = \varphi [(\omega_{\rho\phi} * \rho_\tau) + (\omega_{v\phi} * v_{\tau-1}) + (\omega_{\mu\phi} * \mu_{\tau-1}) + \kappa_\phi] \tag{7}$$

$$\lambda_\tau = \phi_\tau \tanh(\lambda_\tau) \tag{8}$$

where α_τ represents the input gate, ρ denotes the input vector, ϕ is the output gate, μ_τ denotes the output, and β_τ represents

Fig. 2 Overview of Proposed Model

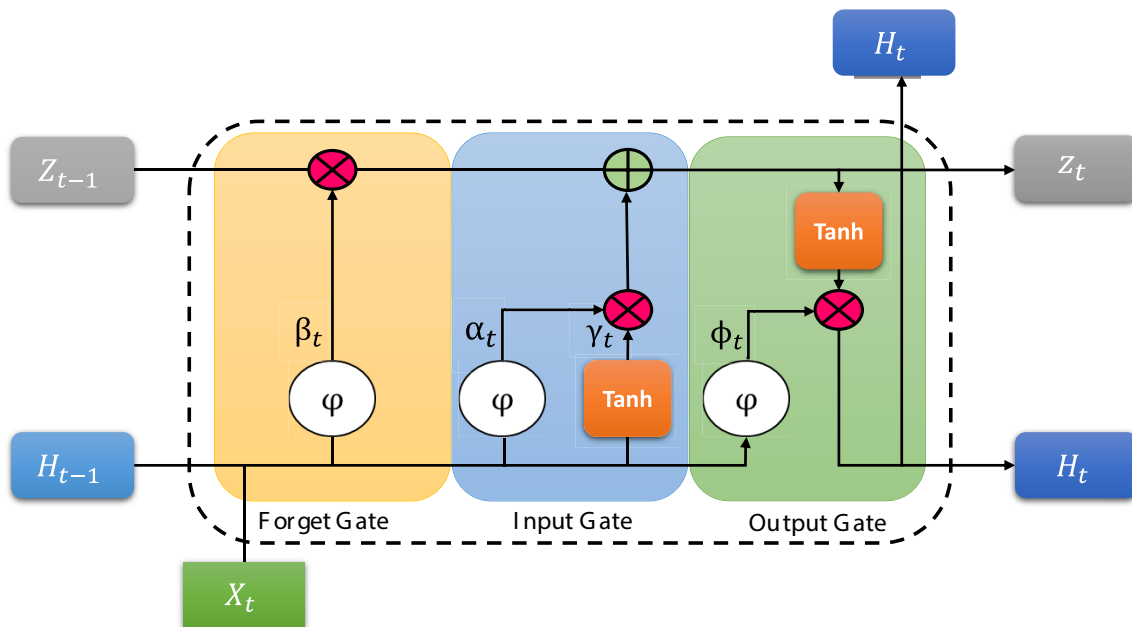
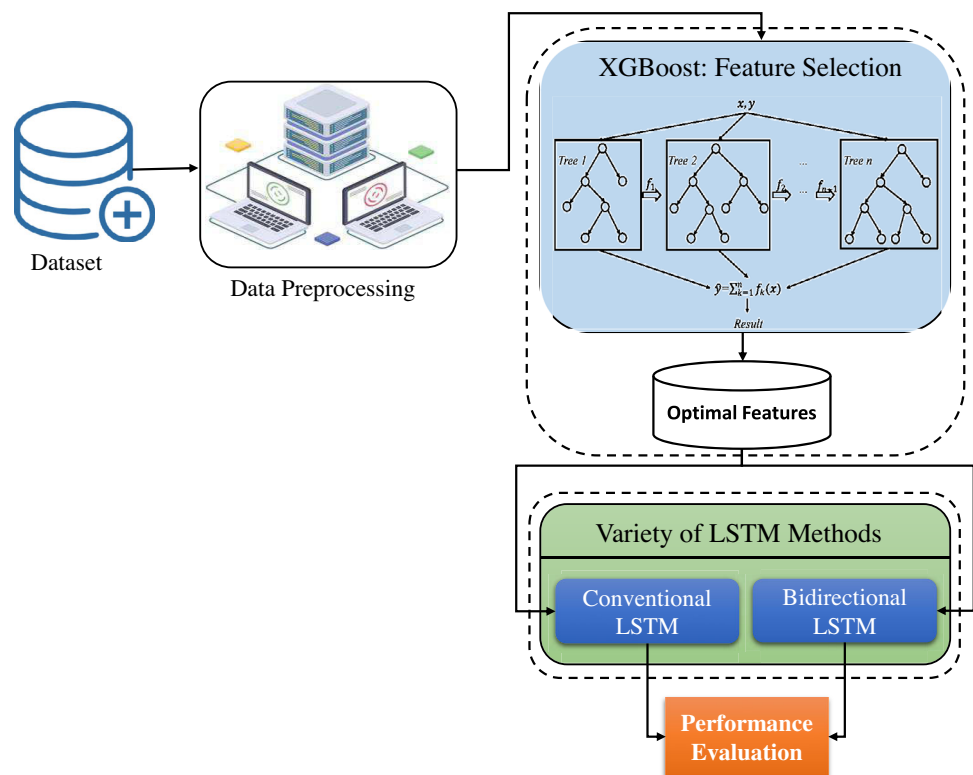


Fig. 3 Memory cell of LSTM

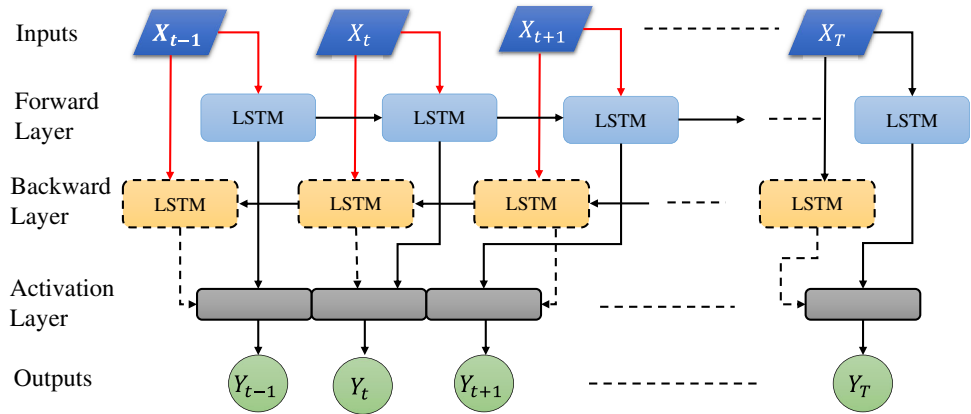
the forgetting function. The cell state is given by γ_t , where γ and κ are the weight and bias parameters, respectively.

4.5 Bidirectional LSTM (BiLSTM)

The bidirectional LSTM complements the standard LSTMs to improve the classification performance of a model. Two

LSTMs are trained based on the input data. The first LSTM was applied to the original input data, while the second was to the reverse copy. This increases the expressiveness of the network and leads to faster results. The concept underlying the BiLSTM is quite simple. It consists of duplicating the first recurrent layer of the network, passing the input data to the first layer in its original form, and then passing the input data

Fig. 4 Overview of BiLSTM architecture



to the duplicated layer in reverse order. This concept solves the problem of vanishing gradients in conventional RNNs.

The BiLSTM is trained with all available past and current input data within a specified time period. The BiLSTM uses a forward and backward layer to process input data in two directions (i.e., left-to-right and right-to-left) [69]. By accepting the initial LSTM layer as input, the Keras library in Python implements BiLSTMs via a bidirectional layer wrapper. The user can specify the fusion mode, which determines how the forward and reverse outputs are combined before being sent to the subsequent layer (Figure). In the Fig. 4, the mathematical formula for the output based on the forward hidden layer $\vec{\lambda}_\tau$, and the backward hidden layer $\overleftarrow{\lambda}_\tau$ is given as [70]:

$$\vec{\lambda}_\tau = L(\omega_{\rho \vec{l}} \rho_\tau + \omega_{\vec{l} \vec{l}} \vec{l}_{\tau-1} + \kappa_{\vec{l}}) \tag{9}$$

$$\overleftarrow{\lambda}_\tau = L(\omega_{\rho \overleftarrow{l}} \rho_\tau + \omega_{\overleftarrow{l} \overleftarrow{l}} \overleftarrow{l}_{\tau-1} + \kappa_{\overleftarrow{l}}) \tag{10}$$

$$Z_\tau = \omega_{\vec{l} z} \vec{\lambda}_\tau + \omega_{\overleftarrow{l} z} \overleftarrow{\lambda}_\tau + \kappa_z \tag{11}$$

where l represents the hidden layer and ω denotes the input weight matrices (forwards, backward) hidden weight, bias vectors for both directions are given by κ ($\kappa_{\vec{l}}$ and $\kappa_{\overleftarrow{l}}$).

5 Validation and Evaluation

To determine the efficacy of ML-based diagnostic systems, the holdout validation approach has often been used as a standard in the literature [71, 72]. However, the holdout validation scheme is inappropriate when the dataset contains imbalanced classes. Since ML models favor the majority class, we used a stratified k -fold cross-validation scheme to avoid biases caused by unbalanced classes in the collected dataset [73]. The stratified k -fold validation scheme extends the cross-validation technique by maintaining the same class ratio across K folds as the original dataset ratio. To test the efficacy of the proposed model, we used the stratified k -fold

Table 3 Performance of LSTM and BiLSTM using all features

| Model | Acc | Sens | Spec | F-Score | MCC |
|--------|-------|-------|-------|---------|--------|
| LSTM | 93.06 | 86.23 | 79.66 | 94.00 | 0.5871 |
| BiLSTM | 95.12 | 90.58 | 92.00 | 95.00 | 0.6132 |

validation with $k = 5$. Specificity, sensitivity, and accuracy are the evaluation measures used to assess the performance of the proposed model. Using a receiver operator characteristic (ROC) curve and, the Matthew correlation coefficient (MCC) and area under the curve (AUC) is calculated. These are the evaluation metrics that are mathematically specified:

$$Sensitivity = \frac{TP}{TP + FN} \tag{12}$$

$$Specificity = \frac{TN}{TN + FP} \tag{13}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

where TP stands for the number of true positives, FP for the number of false positives, TN for the number of true negatives, and FN for the number of false negatives.

$$F_score = \frac{2TP}{2TP + FP + FN} \tag{15}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{16}$$

Furthermore, a binary classification problem is subjected to statistical analysis, i.e., the F-measure is used to test the model. The F-measure ranges from 0 to 1, where 1 represents excellent predictions, and 0 represents the worst. The accuracy of a test is evaluated using MCC. MCC ranges from 1 to

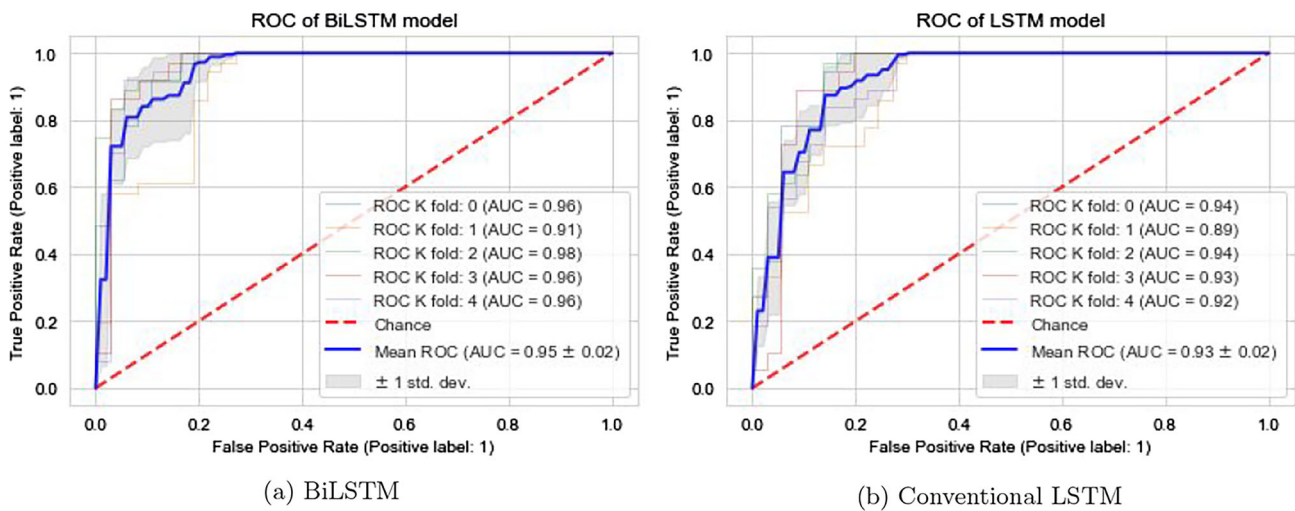


Fig. 5 Performance comparison of BiLSTM with conventional LSTM model using all 75 features

-1, where 1 represents a perfect prediction, and -1 represents the worst prediction.

6 Results

6.1 Comparison of LSTM and BiLSTM Using All Features

In the first phase of experiments, we examined the performance of conventional LSTM and BiLSTM models using all 75 features of the dataset. The performance of both conventional LSTM and BiLSTM models is evaluated using the stratified *K*-fold cross-validation method, with the value of *k* set to 5. The performance of the conventional LSTM and BiLSTM models is validated using various evaluation metrics, i.e., accuracy, sensitivity, specificity, F1 score, and MCC which are given in Table 3, with BiLSTM achieving the highest test accuracy of 95.12% compared with conventional LSTM models of 94.56%.

We also evaluate the performance of BiLSTM and conventional LSTM models while using all features from the dataset based on the ROC curve. The graph with a larger area under the curve (AUC) is considered more accurate. From Fig. 5, it can be seen that the BiLSTM has a larger AUC in comparison to the conventional LSTM. Hence, BiLSTM is more efficient than conventional LSTM.

6.2 Bidirectional LSTM Using Xgboost Feature Selection Module

In this experiment, we hybridized the XGBoost model with the BiLSTM model, using the XGBoost model to rank the

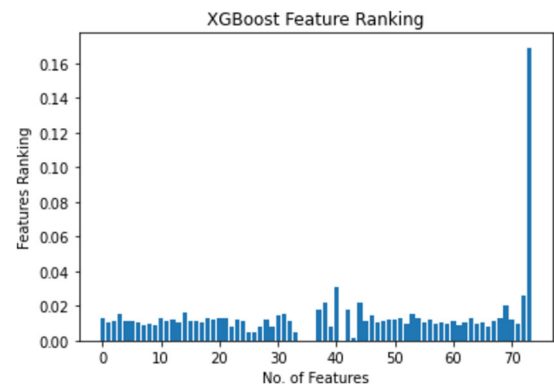


Fig. 6 Features ranking based on their importance using XGBoost

features in the dataset. All 75 features in the dataset are ranked according to their importance, as shown in Fig. 6.

After ranking the features, we set a threshold to select the best features. The selected features are then fed into the BiLSTM model for classification. The proposed model (XGBoost_BiLSTM) was evaluated using evaluation metrics such as training accuracy, test accuracy, sensitivity, specificity, and MCC. The results of the proposed model are given in Table 4 along with the number of features selected using the XGBoost module. Table 4 shows that the proposed model (XGBoost_BiLSTM) achieves the highest test accuracy of 97.00% while using the six best features of the dataset. In contrast, using the six best features of the dataset.

In addition, we used a ROC curve to validate the performance of the proposed model (XGBoost_BiLSTM). The model ML with a large area under the curve (AUC) is considered more efficient. Therefore, we tested the performance of the proposed model based on the ROC curve using the six best features from the dataset. Also, the conventional LSTM model was tested based on the ROC -curve using the six best

Table 4 Performance of BiLSTM Model based on Xgboost feature ranking

| SF ¹ | Feture_Code | Acc | Sens | Spec | F-Score | MCC |
|-----------------|--------------------------|-------|-------|-------|---------|--------|
| 01 | [73] | 95.66 | 95.00 | 90.44 | 85.26 | 0.5488 |
| 02 | [40, 73] | 97.14 | 95.25 | 95.97 | 90.00 | 0.5700 |
| 03 | [40, 72, 73] | 95.98 | 95.67 | 92.74 | 85.00 | 0.5825 |
| 04 | [40, 44, 72, 73] | 96.00 | 95.66 | 98.59 | 92.00 | 0.5874 |
| 05 | [38, 40, 44, 72, 73] | 96.53 | 96.15 | 88.76 | 90.21 | 0.6145 |
| 06 | [38, 40, 44, 69, 72, 73] | 98.75 | 97.00 | 89.26 | 95.19 | 0.6232 |

¹ Selected Features

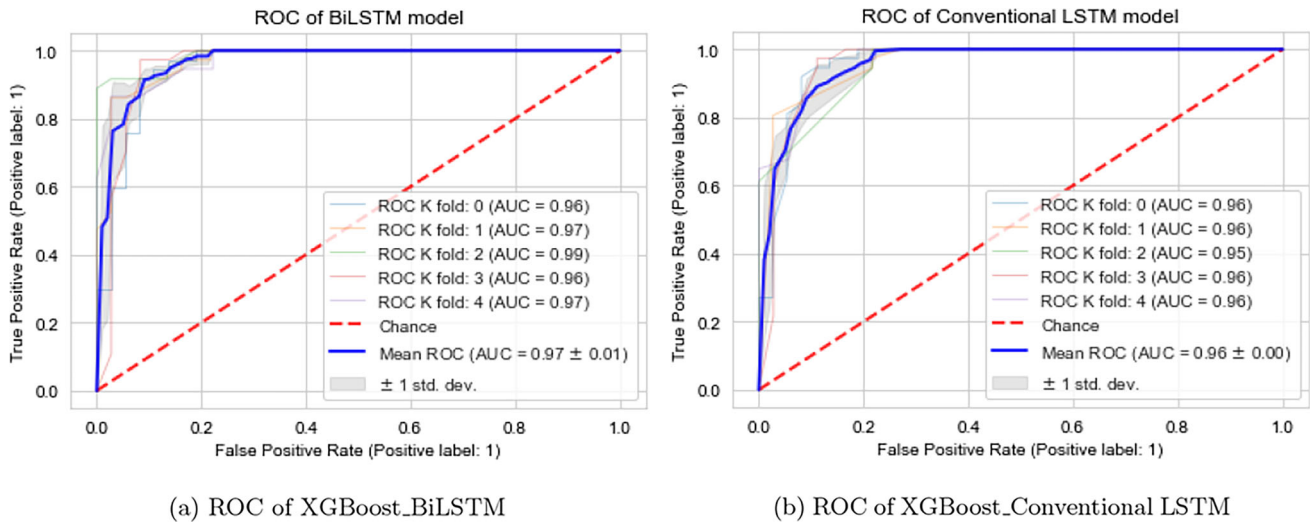


Fig. 7 Performance comparison of BiLSTM with conventional LSTM model using top 6 ranked features

features from the dataset. From Fig. 7, it can be seen that the ML (XGBoost_BiLSTM) model has a large area under the curve compared to the (XGBoost_Conventional LSTM) model.

6.3 Performance of ML_Models Using All Features

In this experiment, all 75 features of the dataset were used to evaluate the performance of several modern ML models. The performance of the ML models was evaluated using the following evaluation metrics: training accuracy (Acc.Train), test accuracy (Acc.Test), sensitivity (Sens.), specificity (Spec.), F1 score, and Matthew’s correlation coefficient (MCC) based on a holdout validation scheme with 70% data for the training of the model and 30% data for the testing of ML models. From Table 5, it can be seen that the model RF achieves the highest test accuracy of 83.40 compared to the other models from ML. We also used ROC curves to test the effectiveness of the ML models. The ML model is more accurate and reliable because it has a larger area under the curve (AUC). From Fig. 8, it can be seen that the performance of the RF model is much better than the other ML models, which have an AUC of 83.40%.

Table 5 ML models performance with all features used

| Model | Acc | Sens | Spec | F-Score | MCC |
|----------|-------|-------|-------|---------|--------|
| LR | 88.53 | 56.21 | 66.45 | 82.00 | 0.4521 |
| GNB | 70.10 | 58.51 | 74.45 | 65.00 | 0.3864 |
| LDA | 82.00 | 85.63 | 70.47 | 82.00 | 0.4600 |
| KNN | 68.00 | 72.11 | 59.45 | 62.00 | 0.3656 |
| DT | 61.44 | 44.33 | 88.00 | 57.00 | 0.3317 |
| RF | 92.54 | 90.23 | 81.54 | 84.00 | 0.4851 |
| SVM | 76.63 | 78.47 | 68.14 | 71.00 | 0.4287 |
| AdaBoost | 90.62 | 82.36 | 78.94 | 83.00 | 0.4700 |

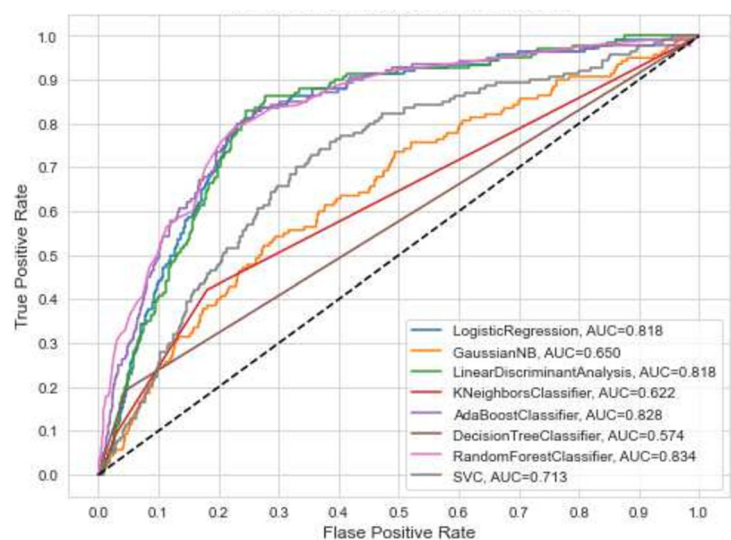
In this study, we developed a hybrid machine learning model combining XGBoost and BiLSTM to detect risk factors and diagnose sleep apnea. The proposed model consists of two modules: The first module classifies the most important features from the dataset, and the second module performs sleep apnea classification. To evaluate the performance of our model, we used a k-fold cross-validation scheme (k=5). We compared it with eight other state-of-the-art machine learning models, including a conventional LSTM model. Our proposed model achieved 97% accuracy

using only the top six features in the dataset. These six features, including type 2 diabetes, external injuries, mental and behavioral disorders, psychological stress and emotions, a comprehensive psychopathology rating scale, and respiratory system diseases, are the main risk factors for sleep apnea in older adults. Overall, our proposed model outperformed the conventional LSTM and other modern machine learning models, demonstrating its potential for early detection and diagnosis of sleep apnea in older adults

7 Discussion

Using machine learning and deep learning techniques, we identified the six major factors contributing to sleep apnea in older adults. The proposed sleep apnea prediction model consists of two modules, the first of which ranks features by importance. We used a machine learning model (XGBoost) to rank the features in the dataset. After ranking the features, we used deep learning techniques such as LSTM and BiLSTM. The performance of the LSTM and BiLSTM models was tested using various evaluation metrics. According to the experimental results, BiLSTM outperformed the LSTM model and several other state-of-the-art machine learning models. BiLSTM achieved the highest accuracy of 97% compared to the accuracy of the LSTM model, while using the same six main features. These (38, 40, 44, 69, 72, 73) six features are critical for the development of sleep apnea in older adults. Table 6 provides information regarding the feature code (F_Code), feature label (F_Label) selected by the proposed model (XGBoost_BiLSTM) and describes the six main risk factors for sleep apnea in older adults. If we avoid these risk factors, we can improve the health of older adults and reduce the risk of sleep apnea.

Fig. 8 ROC curve analysis of ML models



It is also important to mention the limitations of this study so that future researchers can benefit from it. The proposed study used EHR for experimental purposes, and the sample size of the dataset was modest (10,765). But deep learning methods work well for larger datasets. Therefore, in further research, researchers should collect datasets with a larger sample size. Although sleep apnea is a rare disease, the number of positive cases of sleep apnea is far less in comparison to healthy subjects. The machine learning models tend to bias toward the majority class in the dataset; therefore, a balanced dataset should be collected in the future. To avoid the problem of bias in the proposed model, we used a cross-validation scheme with several evaluation metrics to truly validate the performance of the developed algorithm. Moreover, instead of using a single modality, a multimodal dataset should be used for the prediction of sleep apnea. For clinical setup, a huge amount of medical data is generated these days. The generated data can be utilized for the improvement of the health conditions of adults. Based on the given scenario, the proposed model used the EHR of the patients to predict sleep apnea. The proposed model also identified the risk factors for sleep apnea in the patient. Through identification of the risk factors, medical practitioners can advise the patient to alter their lifestyle so that the development of sleep apnea can be avoided.

8 Conclusion

In this study, we developed a hybrid ML model (XGBoost_BiLSTM) to diagnosis of sleep apnea. The proposed model consists of two modules: The first module ranks the most important features from the feature space, and the second module used for the classification of sleep apnea. To evaluate the performance of (XGBoost_BiLSTM) model,

Table 6 Description of top six risk factors of sleep apnea

| F_Code ¹ | F_Label ² | Description |
|---------------------|----------------------|--|
| 38 | K70 | Diabetes type 2 |
| 40 | K76 | Injury by external causes (external causes: S00-T98) |
| 44 | K91 | Mental and behavioural disorders (F00-F99) |
| 69 | O199 | Psychological distress and emotions |
| 72 | CPRS (Sum) | Comprehensive Psychopathological Rating Scale |
| 73 | K57 | Diseases of the respiratory system (J00-J99) |

¹ Features Code² Features Label

we used a k-fold cross-validation scheme ($k = 5$). The performance of (XGBoost_BiLSTM) model is compared with other state-of-the-art ML models, including a conventional LSTM model. The newly constructed (XGBoost_BiLSTM) model achieved the highest accuracy of 97% while using only the top six features from feature space. The selected top six features consist of; type 2 diabetes, external injuries, mental and behavioral disorders, psychological stress and emotions, a comprehensive psychopathology rating scale, and respiratory system diseases, are the main risk factors for sleep apnea in older adults. Overall, the proposed (XGBoost_BiLSTM) model outperformed the conventional LSTM and other ML models by demonstrating its potential for early detection and diagnosis of sleep apnea. In this study, the dataset used for the experimental purpose was based on the EHR, but in further research, researchers can employ a multimodality-based dataset. Furthermore, deep learning algorithms perform well when the number of samples in the dataset is large; therefore, we need to collect a dataset with a larger sample size.

Supplementary information Detail of list of variables used in this study is provided along with the upload material.

Acknowledgements The first author's learning process was supported by the National E-Infrastructure for Aging Research (NEAR), Sweden. NEAR is working on improving the health condition of older adults in Sweden

Author Contributions Conceptualization by Johan Sanmartin Berglund, Data curation by Ana Luiza Dallora, Formal analysis by Liaqat Ali, Write up and Methodology by Ashir Javeed, Proofread by Muhammad Asim Saleem, Supervised by Peter Anderber.

Funding Open access funding provided by Blekinge Institute of Technology.

Availability of data and materials Data will be provide on demand from the first author.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethics approval This study was carried out in accordance with the Declaration of Helsinki and was approved by the Research Ethics Committee at Blekinge Institute of Technology (BTH).

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Xie, B., Minn, H.: Real-time sleep apnea detection by classifier combination. *IEEE Trans. Inf Technol. Biomed.* **16**(3), 469–477 (2012). <https://doi.org/10.1109/TITB.2012.2188299>
- Monahan, K., Redline, S.: Role of obstructive sleep apnea in cardiovascular disease. *Curr. Opin. Cardiol.* **26**(6), 541 (2011)
- Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S., Badr, S.: The occurrence of sleep-disordered breathing among middle-aged adults. *N. Engl. J. Med.* **328**(17), 1230–1235 (1993)
- Young, T., Peppard, P.E., Gottlieb, D.J.: Epidemiology of obstructive sleep apnea: a population health perspective. *Am. J. Respir. Crit. Care Med.* **165**(9), 1217–1239 (2002)
- Peppard, P.E., Young, T., Palta, M., Skatrud, J.: Prospective study of the association between sleep-disordered breathing and hypertension. *N. Engl. J. Med.* **342**(19), 1378–1384 (2000)
- Marin, J.M., Carrizo, S.J., Vicente, E., Agusti, A.G.: Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *The Lancet* **365**(9464), 1046–1053 (2005)
- Mehra, R., Benjamin, E.J., Shahar, E., Gottlieb, D.J., Nawabit, R., Kirchner, H.L., Sahadevan, J., Redline, S.: Association of nocturnal arrhythmias with sleep-disordered breathing: The sleep heart health study. *Am. J. Respir. Crit. Care Med.* **173**(8), 910–916 (2006)
- Yaggi, H.K., Concato, J., Kernan, W.N., Lichtman, J.H., Brass, L.M., Mohsenin, V.: Obstructive sleep apnea as a risk factor for stroke and death. *N. Engl. J. Med.* **353**(19), 2034–2041 (2005)
- Bradley, T.D., Floras, J.S.: Obstructive sleep apnoea and its cardiovascular consequences. *The Lancet* **373**(9657), 82–93 (2009)

10. Kasai, T., Bradley, T.D.: Obstructive sleep apnea and heart failure: pathophysiologic and therapeutic implications. *J. Am. Coll. Cardiol.* **57**(2), 119–127 (2011)
11. Jean-Louis, G., Zizi, F., Clark, L.T., Brown, C.D., McFarlane, S.I.: Obstructive sleep apnea and cardiovascular disease: role of the metabolic syndrome and its components. *J. Clin. Sleep Med.* **4**(3), 261–272 (2008)
12. Damy, T., Margarit, L., Noroc, A., Bodez, D., Guendouz, S., Boyer, L., Drouot, X., Lamine, A., Paulino, A., Rappeneau, S., et al.: Prognostic impact of sleep-disordered breathing and its treatment with nocturnal ventilation for chronic heart failure. *Eur. J. Heart Fail.* **14**(9), 1009–1019 (2012)
13. Tsai, M., Khayat, R.: Sleep apnea in heart failure. *Current Treatment Options in Cardiovascular Medicine* **20**(4) (2018). <https://doi.org/10.1007/s11936-018-0624-0>. Cited by: 12
14. Ishman, S.L., Maturro, S., Schwartz, S., McKenna, M., Baldassari, C.M., Bergeron, M., Chernobilsky, B., Ehsan, Z., Gagnon, L., Liu, Y.-C.C., et al.: Expert consensus statement: Management of pediatric persistent obstructive sleep apnea after adenotonsillectomy. *Otolaryngology-Head and Neck Surgery* **168**(2), 115–130 (2023)
15. Faverio, P., Zanini, U., Monzani, A., Parati, G., Luppi, F., Lombardi, C., Perger, E.: Sleep-disordered breathing and chronic respiratory infections: A narrative review in adult and pediatric population. *Int. J. Mol. Sci.* **24**(6), 5504 (2023)
16. Giampá, S.Q., Furlan, S.F., Freitas, L.S., Macedo, T.A., Lebkuchen, A., Cardozo, K.H., Carvalho, V.M., Martins, F.C., Azam, I.F., Costa-Hong, V., et al.: Effects of cpap on metabolic syndrome in patients with osa: A randomized trial. *Chest* **161**(5), 1370–1381 (2022)
17. Wang, X., Fan, J., Guo, R., Hao, W., Gong, W., Yan, Y., Zheng, W., Ai, H., Que, B., Hu, D., et al.: Association of obstructive sleep apnoea with cardiovascular events in women and men with acute coronary syndrome. *European Respiratory Journal* **61**(1) (2023)
18. Senaratna, C.V., Perret, J.L., Lodge, C.J., Lowe, A.J., Campbell, B.E., Matheson, M.C., Hamilton, G.S., Dharmage, S.C.: Prevalence of obstructive sleep apnea in the general population: a systematic review. *Sleep Med. Rev.* **34**, 70–81 (2017)
19. Tietjens, J.R., Claman, D., Kezirian, E.J., De Marco, T., Mirzayan, A., Sadroonri, B., Goldberg, A.N., Long, C., Gerstenfeld, E.P., Yeghiazarians, Y.: Obstructive sleep apnea in cardiovascular disease: a review of the literature and proposed multidisciplinary clinical management strategy. *J. Am. Heart Assoc.* **8**(1), 010440 (2019)
20. Guillot, M., Sforza, E., Achour-Crawford, E., Maudoux, D., Saint-Martin, M., Barthélémy, J.-C., Roche, F.: Association between severe obstructive sleep apnea and incident arterial hypertension in the older people population. *Sleep Med.* **14**(9), 838–842 (2013)
21. O'Connor-Reina, C., Alcalá, L.R., Ignacio, J.M., Iriarte, M.T.G., Llatas, M.C., Morente, J.C.C., Del Rey, D.P., Alvarez, I.M., Ibarburu, G.H., Baptista, P., et al.: Risk of diabetes in patients with sleep apnea: comparison of surgery versus cpap in a long-term follow-up study. *Journal of Otolaryngology-Head & Neck Surgery* **52**(1), 16 (2023)
22. Hetland, A., Vistnes, M., Haugaa, K.H., Liland, K.H., Olseng, M., Edvardsen, T.: Obstructive sleep apnea versus central sleep apnea: prognosis in systolic heart failure. *Cardiovascular Diagnosis and Therapy* **10**(3), 396 (2020)
23. Linz, B., Hohl, M., Lang, L., Wong, D.W., Nickel, A.G., De La Torre, C., Sticht, C., Wirth, K., Boor, P., Maack, C., et al.: Repeated exposure to transient obstructive sleep apnea-related conditions causes an atrial fibrillation substrate in a chronic rat model. *Heart Rhythm* **18**(3), 455–464 (2021)
24. Acharya, R., Basnet, S., Tharu, B., Koirala, A., Dhital, R., Shrestha, P., Poudel, D., Ghimire, S., Kafle, S.: Obstructive sleep apnea: risk factor for arrhythmias, conduction disorders, and cardiac arrest. *Cureus* **12**(8) (2020)
25. Lopes Costa, M., Correia-Costa, L., Santos, A.C., Azevedo, I.: Pediatric obstructive sleep apnea and metabolic syndrome—a birth portuguese cohort. Available at SSRN 4371008
26. Óskarsdóttir, S., Boot, E., Crowley, T.B., Loo, J.C., Arganbright, J.M., Armando, M., Baylis, A.L., Breetvelt, E.J., Castelein, R.M., Chadehumbe, M., et al.: Updated clinical practice recommendations for managing children with 22q11. 2 deletion syndrome. *Genetics in Medicine*, 100338 (2023)
27. Zhao, X., Xu, J., Wang, M.-Y., Hou, Z.-W., Shi, H.-S., Zhang, X.-X.: Effect of oral xiao-xian decoction combined with acupoint application therapy on pediatric adenoid hypertrophy: A randomized trial. *Medicine* **102**(5), 32804–32804 (2023)
28. Wang, H., Xu, W., Zhao, A., Sun, D., Li, Y., Han, D.: Clinical characteristics combined with craniofacial photographic analysis in children with obstructive sleep apnea. *Nature and Science of Sleep*, 115–125 (2023)
29. Xin, Y., Zhou, H., Zhao, Y., Zhao, L.: Invisalign treatment of a three-year-old child with bilateral posterior scissor bite and multisite upper airway obstruction: A case report. *J. Clin. Med.* **12**(1), 333 (2023)
30. Fujita, K., Huang, T., Rohrbaugh, T., Trott, K.: Demographic factors associated with postoperative polysomnography for pediatric patients with persistent obstructive sleep apnea. *Int. J. Pediatr. Otorhinolaryngol.* **164**, 111428 (2023)
31. Alonso, A., Alam, A.B., Kamel, H., Subbian, V., Qian, J., Boerwinkle, E., Cicek, M., Clark, C.R., Cohn, E.G., Gebo, K.A., et al.: Epidemiology of atrial fibrillation in the all of us research program. *PLoS ONE* **17**(3), 0265498 (2022)
32. Amesz, J.H., de Groot, N.M., Langmuur, S.J., Azzouzi, H.e., Tiggeloven, V.P., van Rooij, M.M., Knops, P., Bogers, A.J., Taverne, Y.J.: Biomimetic cultivation of atrial tissue slices as novel platform for in-vitro atrial arrhythmia studies. *Scientific Reports* **13**(1), 3648 (2023)
33. Marulanda-Londono, E., Chaturvedi, S.: The interplay between obstructive sleep apnea and atrial fibrillation. *Front. Neurol.* **8**, 668 (2017)
34. Soltis, M.E., Spector, A.R.: A concise review of polysomnography and obstructive sleep apnea for the neurophysiologist. *J. Clin. Neurophysiol.* **40**(3), 191–197 (2023)
35. Huang, L., Xu, Y., Gong, X., Gao, X.: Anatomical phenotype of obstructive sleep apnea (osa) patients based on cluster analysis. *Orthodontics & Craniofacial Research* (2023)
36. Khor, Y.H., Khung, S.-W., Ruehland, W.R., Jiao, Y., Lew, J., Munsif, M., Ng, Y., Ridgers, A., Schulte, M., Seow, D., et al.: Portable evaluation of obstructive sleep apnea in adults: A systematic review. *Sleep Medicine Reviews*, 101743 (2023)
37. Learning, M.U.S.M.: Obstructive sleep apnea: a prediction model using supervised machine learning method. *The Importance of Health Informatics in Public Health during a Pandemic* **272**, 387 (2020)
38. Iannella, G., Maniaci, A., Magliulo, G., Cocuzza, S., La Mantia, I., Cammaroto, G., Greco, A., Vicini, C.: Current challenges in the diagnosis and treatment of obstructive sleep apnea syndrome in the elderly. *Pol Arch Intern Med* **130**(7–8), 649–654 (2020)
39. Bengio, Y., Courville, A., Vincent, P.: Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(8), 1798–1828 (2013)
40. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *nature* **521**(7553), 436–444 (2015)
41. Berman, D.S., Buczak, A.L., Chavis, J.S., Corbett, C.L.: A survey of deep learning methods for cyber security. *Information* **10**(4), 122 (2019)
42. Kim, J., Kim, H.: Applying recurrent neural network to intrusion detection with hessian free optimization. In: *Information Security Applications: 16th International Workshop, WISA 2015, Jeju*

- Island, Korea, August 20–22, 2015, Revised Selected Papers 16, pp. 357–369 (2016). Springer
43. Bengio, Y., Simard, P., Frasconi, P.: Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Networks* **5**(2), 157–166 (1994)
 44. Intelligence, A.A.: The following abbreviations are used for frequently cited conferences and journals: Aaai proceedings of the aaai conference on artificial intelligence. *IEEE Intell. Syst.* **19**(1), 8–12 (1985)
 45. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
 46. Hochreiter, S., Schmidhuber, J.: A inips. Lstm can solve hard long time lag problems. *J Adv Neural Inform Process Syst* **9**, 473–9 (1996)
 47. Kim, J., Kim, H., *et al.*: An effective intrusion detection classifier using long short-term memory with gradient descent optimization. In: 2017 International Conference on Platform Technology and Service (PlatCon), pp. 1–6 (2017). IEEE
 48. Staudemeyer, R.C., Omlin, C.W.: Evaluating performance of long short-term memory recurrent neural networks on intrusion detection data. In: Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference, pp. 218–224 (2013)
 49. Staudemeyer, R.C.: Applying long short-term memory recurrent neural networks to intrusion detection. *South African Computer Journal* **56**(1), 136–154 (2015)
 50. Shahnaz, C., Minhaz, A.T., Ahamed, S.T.: Sub-frame based apnea detection exploiting delta band power ratio extracted from eeg signals. In: 2016 IEEE Region 10 Conference (TENCON), pp. 190–193 (2016). IEEE
 51. Guijarro-Berdinas, B., Hernández-Pereira, E., Peteiro-Barral, D.: A mixture of experts for classifying sleep apneas. *Expert Syst. Appl.* **39**(8), 7084–7092 (2012)
 52. Kellesarian, S., Malignaggi, V., Feng, C., Javed, F.: Association between obstructive sleep apnea and erectile dysfunction: a systematic review and meta-analysis. *Int. J. Impot. Res.* **30**(3), 129–140 (2018)
 53. Masa, J.F., Mokhlesi, B., Benítez, I., Mogollon, M.V., Gomez de Terreros, F.J., Sánchez-Quiroga, M.Á., Romero, A., Caballero-Eraso, C., Alonso-Álvarez, M.L., Ordax-Carbajo, E., *et al.*: Echocardiographic changes with positive airway pressure therapy in obesity hypoventilation syndrome. long-term pickwick randomized controlled clinical trial. *American Journal of Respiratory and Critical Care Medicine* **201**(5), 586–597 (2020)
 54. Baty, F., Boesch, M., Widmer, S., Annaheim, S., Fontana, P., Camenzind, M., Rossi, R.M., Schoch, O.D., Brutsche, M.H.: Classification of sleep apnea severity by electrocardiogram monitoring using a novel wearable device. *Sensors* **20**(1), 286 (2020)
 55. Hassan, A.R.: Automatic screening of obstructive sleep apnea from single-lead electrocardiogram. In: 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), pp. 1–6 (2015). IEEE
 56. Uçar, M.K., Bozkurt, M.R., Bilgin, C., Polat, K.: Automatic sleep staging in obstructive sleep apnea patients using photoplethysmography, heart rate variability signal and machine learning techniques. *Neural Comput. Appl.* **29**, 1–16 (2018)
 57. Hassan, A.R., Haque, M.A.: An expert system for automated identification of obstructive sleep apnea from single-lead ecg using random under sampling boosting. *Neurocomputing* **235**, 122–130 (2017)
 58. Prabha, A., Trivedi, A., Kumar, A.A., Kumar, C.S.: Automated system for obstructive sleep apnea detection using heart rate variability and respiratory rate variability. In: 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1303–1307 (2017). IEEE
 59. Kim, T., Kim, J.-W., Lee, K.: Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques. *Biomed. Eng. Online* **17**, 1–19 (2018)
 60. Mostafa, S.S., Mendonça, F., G. Ravelo-García, A., Morgado-Dias, F.: A systematic review of detecting sleep apnea using deep learning. *Sensors* **19**(22), 4934 (2019)
 61. Sharma, M., Kumbhani, D., Yadav, A., Acharya, U.R.: Automated sleep apnea detection using optimal duration-frequency concentrated wavelet-based features of pulse oximetry signals. *Applied Intelligence*, 1–13 (2022)
 62. Liu, J., Li, Q., Xin, Y., Lu, X.: Obstructive sleep apnea detection using sleep architecture. In: 2020 IEEE International Conference on Mechatronics and Automation (ICMA), pp. 255–260 (2020). IEEE
 63. Hafezi, M., Montazeri, N., Saha, S., Zhu, K., Gavrilovic, B., Yadollahi, A., Taati, B.: Sleep apnea severity estimation from tracheal movements using a deep learning model. *IEEE Access* **8**, 22641–22649 (2020)
 64. Almuhammadi, W.S., Aboalayon, K.A., Faezipour, M.: Efficient obstructive sleep apnea classification based on eeg signals. In: 2015 Long Island Systems, Applications and Technology, pp. 1–6 (2015). IEEE
 65. Lagergren, M., Fratiglioni, L., Hallberg, I.R., Berglund, J., Elmståhl, S., Hagberg, B., Holst, G., Rennemark, M., Sjolund, B.-M., Thorslund, M., *et al.*: A longitudinal study integrating population, care and social services data. the swedish national study on aging and care (snac). *Aging clinical and experimental research* **16**(2), 158–168 (2004)
 66. Yu, J.-T., Xu, W., Tan, C.-C., Andrieu, S., Suckling, J., Evangelou, E., Pan, A., Zhang, C., Jia, J., Feng, L., *et al.*: Evidence-based prevention of alzheimer’s disease: systematic review and meta-analysis of 243 observational prospective studies and 153 randomised controlled trials. *Journal of Neurology, Neurosurgery & Psychiatry* **91**(11), 1201–1209 (2020)
 67. Arvanitakis, Z., Shah, R.C., Bennett, D.A.: Diagnosis and management of dementia. *Jama* **322**(16), 1589–1599 (2019)
 68. Yu, J., Shi, S., Zhang, F., Chen, G., Cao, M.: Predgly: predicting lysine glycation sites for homo sapiens based on xgboost feature optimization. *Bioinformatics* **35**(16), 2749–2756 (2019)
 69. Graves, A., Mohamed, A.-r., Hinton, G.: Speech recognition with deep recurrent neural networks. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6645–6649 (2013). Ieee
 70. Mousa, A., Schuller, B.: Contextual bidirectional long short-term memory recurrent neural network language models: A generative approach to sentiment analysis (2017)
 71. Javeed, A., Ali, L., Mohammed Seid, A., Ali, A., Khan, D., Imrana, Y.: A clinical decision support system (cdss) for unbiased prediction of caesarean section based on features extraction and optimized classification. *Computational Intelligence and Neuroscience* **2022** (2022)
 72. Javeed, A., Khan, S.U., Ali, L., Ali, S., Imrana, Y., Rahman, A.: Machine learning-based automated diagnostic systems developed for heart failure prediction using different types of data modalities: a systematic review and future directions. *Computational and Mathematical Methods in Medicine* **2022** (2022)
 73. Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., Khan, J.A.: An automated diagnostic system for heart disease prediction based on chi square statistical model and optimally configured deep neural network. *Ieee Access* **7**, 34938–34945 (2019)