**Detection of Sleep Apnea using Machine Learning Algorithms based on ECG Signals: A Systematic Review**

***Abstract***. Sleep Apnea (SA) is a common sleep disorder that remains unknown in many patients. Past studies have highlighted ECG analysis as a method of diagnosing SA. Because the changes caused by SA on the ECG are so subtle, the need for new methods in diagnosing the disease is required more than ever. Machine Learning (ML) techniques are recognized as one of the most successful methods of computer aided diagnosis. ML uses new methods to diagnose diseases using past clinical results. The purpose of this study is to evaluate studies using ML algorithms and based on ECG characteristics to evaluate people with SA. In this study, English articles indexed in PubMed, Scopus, Web of Science, and IEEE databases, with no lower time limit and until October 2020, were systematically reviewed. A total of 391 articles were collected from various databases, of which 48 articles were entered in the systematic review. Studies have shown that the most common features used in studies were frequency, time and statistics features. Support-Vector Machine (SVM) and Neural Networks‌ (NN) performed best in full record data detection. The highest accuracy, sensitivity and specificity reported between the selected studies was 100%, which was obtained by an SVM. In another case, the classification was based on ECG elements, and accordingly, the highest classification accuracy was observed in the residual neural network algorithm. The accuracy, sensitivity and specificity of this algorithm were reported to be 99%. In general, it can be stated that ML techniques based on ECG characteristics have a high capability in diagnosing SA. This can increase the diagnosis of patients with SA, and can potentially prevent complications of the disease at later stages.

**Keywords:**

**Introduction**

Sleep Apnea (SA) refers to the periodic cessation or reduction of airflow during sleep [1]. SA occurs due to complete or partial obstruction of the upper airways (i.e. Obstructive Sleep Apnea (OSA)), reduction or cessation of brainstem respiratory motor output (i.e. Central Apnea) or both [2]. Complete cessation of respiration (i.e. apnea) or decreased airflow (i.e. hypopnea) are two respiratory events observed in SA. These events reduce oxygen levels, hypercapnia, increase sympathetic nerve activity, and fluctuations in blood pressure and heart rate. These physiological changes also affect patients’ sleep cycle. It causes brain arousal, disruption of various stages of sleep, and sleep fragmentation [3-5]. It is estimated that about 10% of middle-aged people are affected by SA [1]. Despite the high prevalence of this disorder, most patients are unaware of the effect of SA‌ on their respiratory pattern. And because of this, patients do not seek professional treatment [4]. Many studies have examined morbidity of SA. The results of these studies show that failure to diagnose and treat SA in a timely manner can cause daily drowsiness [6], cognitive dysfunction [7], cardiovascular diseases such as hypertension [8], coronary artery disease [9], heart failure [10], stroke [11] and metabolic diseases such as diabetes [12]. Therefore, the early detection of SA is very important.

Polysomnography (PSG) is known as the standard SA diagnostic test. Accordingly, PSG examines sleep and respiration parameters using electroencephalogram, electrocardiogram (ECG), electroechogram, electromyogram, pulse oximetry, airflow measurement and respiratory effort [13, 14]. PSG has high diagnostic accuracy [15], however, factors such as high cost, patient inconvenience, long data recording and difficult interpretation of data are some of the disadvantages of this method. Moreover, the long waiting list for evaluating patients with PSG device increases the possibility of not diagnosing and treating SA in time [16]. Therefore, it is necessary to provide an alternative method to enhance patients’ convenience, and reduce costs to diagnose SA at an early stage [17].

Different strategies have been used to diagnose SA without the use of PSG [18, 19]. Among these, the use of ECG signals has received much attention [20]. ECG is not stressful for the patients compared to PSG, and uses less technical equipment. It has also been observed that the ECG with a signal strength of 1-2 mV has the best signal-to-noise ratio among all physiologic signals [21]. Also using ECG, the cyclic variation of heart rate caused by SA can be seen [22]. Moreover, Parameters extracted from the ECG signal shape allow the extraction of the effort-related curve (ECG-induced respiration or EDR) [23]. On the other hand, it has been observed that the autonomic nervous system, respiration and sleep affect the ECG. Among the parameters that affect sleep are heart rate variability (HRV), LF band power distribution (low frequency) and VF (high frequency) heart rate [21]. The LF component has been shown to affect the ECG under the influence of sympathetic nerve activity changes and the HF component under the influence of respiration and vagus nerve [24]. With these interpretations, since the changes caused by SA in the ECG are very varied and subtle, the diagnosis of SA based on ECG data is a very complex task. One of the methods that has been considered to solve this problem is the use of computer algorithms [25].

Machine Learning (ML) techniques have been considered as one of the successful methods of computer-aided diagnosis [26]. According to the pattern in Figure 1, ML is an evolving branch of the computational science designed to simulate human intelligence by learning from the environment [27]. ML is used when it is not possible to interpret a particular pattern or extract relevant information [28]. When using ML algorithms, SA automatic detection is based on a large number of pre-detected samples. In other words, ML uses data from previous examinations in which the physician has diagnosed the presence or absence of a disease [26].

Various ML techniques have been used in the diagnosis of SA. In a study by Bozkurt et al. [29], the authors used electrocardiography of 10 patients with OSA against 10 healthy controls. This study first extracted HRV from ECG, and then extracted the QRS component at different frequencies using a digital filter and then selected the feature using Principal Component Analysis (PCA). Classification was performed by kNN algorithm. The results of this study showed that when using 3 features, the classification accuracy was 82.11% and when using 13 features, this value was 85.12%. In another study [30], data collected from 86 patients were used, of which 69 were used in training and 17 in test. The Residual Neural Network (RNN) algorithm was reported to offer the highest accuracy of 99%. Moreover, the study highlighted that deep learning techniques are very useful for automatic detection of SA. In another study [31], HRV data were used to automatically detect SA‌. Then a feature selection algorithm was used to select the best features. In this study, classification was performed using support vector machine (SVM), artificial neural network (ANN) and a combination of these two algorithms. The results of this study indicated that the proposed methods have a high capability in detecting SA from the healthy state. As mentioned, different ML algorithms have been used in detecting SA. Therefore, this systematic review was designed to evaluate the ability of ML algorithms to detect SA. Furthermore, as part of our work, the methods used within selected studies were also compared with each other.



Figure 1: An overview of the implementation of machine learning techniques

**Methodology**

***Search strategy and inclusion criteria***

The protocol and reporting used in this systematic review were performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines [32]. The select relevant studies, the four databases of PubMed, Web of Science (, Scopus and IEEE were searched. All searches were performed on 24th September 2020, and were subsequently updated on 24th October 2020. Searches were performed using keywords related to sleep apnea, machine learning and electrocardiogram, and according to the search strategy for each of the databases; Table 1 outlines the keywords and search strategies used for each of database. Studies selected for this systematic review included the ones that distinguish SA from healthy states, as in a binary form. A number of diagnostic studies for apnea, hypopnea, and healthy states that followed a multiclass approach were excluded from this systematic review. Within all the included studies, various ML algorithms were used, and all classifications were based on electrocardiogram data. Exclusion criteria of this systematic review were: lack of access to the full text of articles, conference proceedings and articles written in a language other than English. Furthermore, research works that used a method other than electrocardiogram were excluded from this work.

Table 1: Search strategies, and keywords

|  |  |  |  |
| --- | --- | --- | --- |
| Database | Search strategy | Date | number |
| PubMed | (Artificial Intelligence[mesh] OR "machine learning"[tiab] OR "neural networks"[tiab] OR "Bayesian models"[tiab] OR "deep learning"[tiab] OR "dimensionality reduction"[tiab] OR "decision trees"[tiab] OR "ensemble learning"[tiab] OR "instance based models"[tiab] OR "support vector machines"[tiab]) AND ( Sleep Apnea[tiab] OR "Sleep-Disordered Breathing"[tiab] OR Sleep Apnea, Central[tiab] OR Sleep Apnea Syndrome[MESH] OR Sleep Apnea, Obstructive[mesh] OR "Sleep Apnea Hypopnea Syndrome"[tiab] OR OSA[TIAB] OR OSAHS[TIAB] OR "Sleep Apnea Syndromes"[tiab]) AND ( Electrocardiography[mesh] OR ECG[tiab] OR EKG[tiab] OR Electrocardiogram[TIAB] OR "electrocardiogram derived respiration"[TIAB]) | 24/10/2020 | 59 |
| Scopus  | TITLE-ABS-KEY("Artificial Intelligence" OR "machine learning" OR "neural networks" OR "Bayesian models" OR "deep learning" OR "dimensionality reduction" OR "decision trees" OR "ensemble learning" OR "instance based models" OR "support vector machines") AND TITLE-ABS-KEY("Sleep Apnea" OR "Sleep-Disordered Breathing" OR "Sleep Apnea, Central" OR "Sleep Apnea Syndrome" OR "Obstructive Sleep Apnea" OR "Sleep Apnea Hypopnea Syndrome" OR OSA OR OSAHS OR "Sleep Apnea Syndromes") AND TITLE-ABS-KEY(Electrocardiography OR ECG OR EKG OR Electrocardiogram OR "electrocardiogram derived respiration")  | 24/10/2020 | 210 |
| WOS | TS=("Artificial Intelligence" OR "machine learning" OR "neural networks" OR "Bayesian models" OR "deep learning" OR "dimensionality reduction" OR "decision trees" OR "ensemble learning" OR "instance based models" OR "support vector machines") AND TS=("Sleep Apnea" OR "Sleep-Disordered Breathing" OR "Sleep Apnea, Central" OR "Sleep Apnea Syndrome" OR "Obstructive Sleep Apnea" OR "Sleep Apnea Hypopnea Syndrome" OR OSA OR OSAHS OR "Sleep Apnea Syndromes") AND TS=(Electrocardiography OR ECG OR EKG OR Electrocardiogram OR "electrocardiogram derived respiration")  | 24/10/2020 | 103 |
| IEEE Explore | ("Artificial Intelligence" OR "machine learning" OR "neural networks" OR "Bayesian models" OR "deep learning" OR "dimensionality reduction" OR "decision trees" OR "ensemble learning" OR "instance based models" OR "support vector machines") AND ("Sleep Apnea" OR "Sleep-Disordered Breathing" OR "Sleep Apnea, Central" OR "Sleep Apnea Syndrome" OR "Obstructive Sleep Apnea" OR "Sleep Apnea Hypopnea Syndrome" OR OSA OR OSAHS OR "Sleep Apnea Syndromes") AND (Electrocardiography OR ECG OR EKG OR Electrocardiogram OR "electrocardiogram derived respiration") | 24/10/2020 | 29 |

***Study selection and data extraction***

After removing duplicate articles using the EndNote bibliography management software, one of the authors (HGH) reviewed the title and abstract of collected articles. Accordingly, the studies that did not meet the inclusion criteria were omitted. Two authors (HGH, MM) then reviewed the full text of the remaining articles based on the inclusion and exclusion criteria. These study assessment activities were performed independently, and using blind copies. After completing the reviews at this stage, both authors reviewed the comments made regarding the approval or rejection of the articles. In case of a disagreement, another author (NS), as a senior reviewer, made the inclusion or exclusion judgement. In cases where full text of an article was not available, a request for the full text was sent to the corresponding author via email or ResearchGate. Finally, as full text was not available or could not be secured, the study had to be excluded. The following descriptive information were extracted from the remaining studies: a) year of study, b) country, c) dataset, d) pre-processing, e) feature extraction/selection, f) ML algorithm, and g) parameters reported in the study that represent performance of the algorithms used. The collected studies and their data were reviewed systematically, yet conducting meta-analyses was not deemed appropriate in this work.

**Results**

***Study selection***

After searching the databases, a total of 391 studies were collected. Subsequently, duplicates were omitted and only a copy was retained. Then, the remaining 318 articles were evaluated based on the inclusion and exclusion criteria, leaving 68 studies. These 68 articles were reviewed for eligibility, and finally 48 articles were included in the systematic review (Figure 2). By applying different methods, feature extraction was performed from ECG signals. These features were then used to construct training and test sets and to classify data. In these studies, SA was detected from healthy controls based on ‘per record’, ‘per segment’, or both. The studies performed on the studies showed that in 23 studies the diagnosis was made based on per record, and in 33 studies the diagnosis was made based on per segment. In the record mode, a complete ECG strip was analyzed to distinguish SA from healthy states. In another case (per segment), the ECG strip was first divided into smaller one-minute pieces and named by experts as apnea and healthy parts. Classification was then performed based on the characteristics extracted from these components.



Figure 2: PRISMA flow diagram for study selection

***Dataset***

The datasets used in most studies overlapped [18, 20, 31, 33-62]. These studies used samples from the Physionet Apnea-ECG Database [63]. Nonetheless, among the remaining studies, one study used data from YILDIRIM BEYAZIT Instructional [64] and another study used samples from Sleep laboratories at Chest Diseases Clinics in Sakarya, Turkey [29]. Referrals to Sultan Qaboos University Hospital in Oman, Samsung Medical Center in South Korea, and University of Heidelberg Hospital in Jordan were also examined in other studies. Two other research works used data from the Sleep Heart Health Study cohort [67, 68]. In 6 studies, data from more than one dataset were examined. In 4 studies of data from both Physionet Apnea-ECG Database and St. Vincent’s University Hospital/University College Dublin were used [18, 69-71]. One other research work examined data from Physionet and University Hospital Leuven [72], and one study examined 3 different datasets [73]. The remaining two studies [74, 75] had not provided clear information in relation to the use of specific databases.

***Pre-processing***

In general, the feature extraction process was very different in the studies. In almost all studies, pre-processing was performed with the aim of breaking down the ECG waves into smaller sections and clearing them of junk data. R peaks are known as one of the common features in the diagnosis of apnea, which was detected by different algorithms. In two studies, the diagnosis was made based on the opinion of specialists. One of the most common methods in studies for ECG signal analysis was the Pan Tompkins algorithm [35, 47, 54, 55, 59, 65, 67, 72, 74]. This algorithm uses the amplitude, slope, and width of an integrated window to distinguish the peaks of R from the QRS complex. Pan Tompkins is known as the QRS detection algorithm in real-time approaches [76]. Other studies [46, 48, 49, 73] have used real-time approaches algorithms. Yet the specific type of the algorithms used were not reported.

Due to the oscillating nature of the ECG signal, wavelet-based algorithms and analyzes were used in several studies. TQWT is known as one of the wavelet analyzes that was adopted in four research works [42, 44, 51, 69]. This method breaks down the ECG signal into a number of sub-bands signal to extract features. Daubechies (Db) wavelet is another method used in some works. In this method, the ECG signal is decomposed into several segments. Studies have shown that studies have used 4Db [36, 53], 6Db [33, 62] and 14Db [46, 73] to analyze or extract features. In another study [71], the continuous wavelet transfer method was applied to ECG signals to detect R Wave and R peaks. In two other studies [56, 57], the optimal biorthogonal antisymmetric wavelet filter bank was used to differentiate SA from healthy. These studies first decomposed the ECG signal into 5 levels using wavelet decomposition‌ and then extracted the feature.

Studies have used other algorithms to segment ECG signals. In two studies [52, 64], due to the non-linear nature and energy of ECG waves, the Teager Energy Operator (TEO) method was used. Fourier decomposition was another method used in ECG analysis [41]. A study [43] adopted Empirical Mode Decomposition (EMD) to analyze signals. EMD is known as a data-adaptive signal processing method that performs highly localized time-frequency estimations. Another study [70] using the Hilbert algorithm identified the QRS complex‌. Variational Mode Decomposition [75], dynamic autoregressive (AR) representation model [60] and Iterated Cumulative Sums of Squares (ICSS) [38] were other methods used to detect the RR interval. In another study, the BIOSIG-toolbox was used to detect RRI and EDR [58]. In two studies, the diagnosis of QRS complex [54] and EDR [55] was performed based on Hermite algorithm.

The filters have been applied to ECG waves with the aim of clearing the ECG from noise, ground drift and baseline drift, and detecting HRV and EDR. Moving average filter was applied with the aim of limiting and adjusting the lower and upper limits of waves and eliminating false and unexplained points [20, 46, 65, 73]. This filter was also used to detect EDR [58], RR distance [38] and mean RR value‌ [48, 49]. One study [59] used a bandwidth filter to detect EDR. Low pass filter band was adopted in a number of other studies. This particular filter was used to remove noise from the ECG strip in a study [33]. In a study [74] low pass and high pass filter and in another study [35] low pass, high pass and band pass filters were used to eliminate noise. Chebyshev bandpass filter type I and type II [29, 66, 70], Butterworth pass filter [37], FIR band pass [18, 30], ‌ powerline filter‌ [72] and Savitzky–Golay filter [52] have been used to clean signals from junk data. In one article, the type of filter used to delete junk data was not known [68].

Table 2: Feature extraction process, by article

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| First author, Year, | Country | Dataset, | Pre-processing: | Feature extraction/selection(FE/FS) | Features | classifier |
| Akşahin, M. 2015,[64] | Turkey | YILDIRIM BEYAZIT Instructional,20 cases | Using TEO for detection HRV series. | **FE**: Calculate CPSD by applying FT to cross correlation function **FS**: K-fold cross validation algorithm to determine training, test data | PSD Feature (LF, VLF, HF) | **Per- record**: FFNN |
| Al-Angari, H. M. 2012, [67] | America | Sleep Heart Health Study (SHHS),100 cases | Using PTA for detection RRI | **FE:** Calculate PSD by applying 256-point FFT. | PSD feature | **Per record**: SVM (linear and polynomial)**Per segment**: SVM (linear and polynomial)  |
| Ali, S. Q.2020,[65] | Oman | Sultan Qaboos University Hospital (SQUH), 80 cases | Using PTA for detection RRI,Lower limit, upper limit and medium average filters were used to clear junk data,Resampling at 1 Hz and substituting of missed peaks | **FE**: Signal breakdown into 9 and 512 sub-bands with wavelet packet decomposition. | PSD feature,Statistical feature | **Per record**: FFNN, PNN |
| Atri, R.2015[33] | Iran | PhysioNet,  35 cases | The ECG signal is parsed with a Daubechies6 (db6) wavelet to seven levels, to detect QRS complex,Low pass decomposition filter was used to clear junk data,Detection of HRV and EDR using R peaks | **FE:** Calculate PSD from HRV and EDR using FFT. | PSD feature,Bispectral feature | **Per segment**: LS-SVM |
| Babaeizadeh, S.2010,[34] | America | PhysioNet,  35 cases | detect N-N intervals,Calculate the distance of each N-N to obtain instantaneous HR | **FE**: Using the Lomb algorithm to obtain the PSD. | PSD feature | **Per segment**: QC**Per record**: QC |
| Baek, J. W.2014,[74] | South Korea | \_ | Using PTA for detection QRS complex,Low pass and high pass filter used to clear junk data, | **FE**: Calculate PSD by applying STFT | LF/HF, Average accelerometric value  | **Per record**: machine learning |
| Bali, J.2018,[35] | India | PhysioNet,  35 cases | Low pass, high pass and Band pass filter were used to clear junk data,Using PTA for detection QRS complex | **FE**: Calculate PSD by applying FFT,**FS**: Use PCA to select a feature | PSD from frequency feature,Time domain feature | **Per record**: ANN-LM, ANN-SCG |
| Bozkurt, F.2020,[29] | Turkey | Sleep laboratories at Chest Diseases Clinics of Sakarya Hendek Public Hospital,10 cases | IIR-Chebyshev Type II bandpass filter and medium average filter were used to clear junk data and detect different ECG frequency ranges | **FE**: From all 9 features identified in the previous steps, 25 features were extracted. A total of 225 features were extracted**FS**: Use PCA and Fisher score algorithm | Statistical feature | **Per segment**: DT, kNN, SVM and Ensemble Classifier |
| Bsoul, M.2011,[36] | America | Physionet Apnea-ECG Database.35 cases | Detection of QRS, P and T waves complexes using wavelet algorithm,The t-wave method was used to detect EDR,using a DB4 wavelet with eight levels for RR (m) and nine levels for EDR | **FE**: Calculate PSD by applying 256-point FFT. | Time domain and frequency | **Per segment**: SVM (RBF, polynomial, linear and MLP) |
| Chang, H. Y.2020,[37] | Taiwan | MIT PhysioNet Apnea-ECG Database. 35 cases | fourth-order Butterworth pass band filter to reduce ground drift and high frequency interference,Z score normalization | **FE**: Block diagram 1D deep CNN model | 225 features | **Per record**: CNN**Per segment**: CNN |
| Chen, L.2015[38] | China | PhysioNet Apnea-ECG Database. 35 cases | Using a medium filter, the RR distance was found in a local window,To detect potential apnea, change points, RR intervals were segmented by iterated cumulative sums of squares algorithm | **FE**: Calculate PSD by applying FFT | LF/HF | **Per record:** SVM (RBF, polynomial) |
| Kaguara, A.2014,[77] | America | PhysioNet Apnea-ECG database, 35 cases | Find all peaks within ECG recordings using slope inversion | **FE**: 11 properties were extracted from RR interval (mean and median RR interval, Standard, RMSSD, deviation of the set of RR intervals, NN50 and pNN50(for first and second RRI, SDSD, Inter-quartile range, Mean absolute deviation) | Time domain features | **Per record**: DNN |
| Dey, D.2017,[39] | India | PhysioNet Apnea-ECG database,35 cases | Marking in cases of apnea or normal by medical professionals | **FE**: Optimization of CNN parameters (ReLU, convolution pooling with stride 4 sample, convolution & ReLU, convolution pooling with stride 2 sample, fully connected layer) | \_ | **Per segment**: CNN |
| Eiseman, N. A.2012,[68] | America | SHHS cohort,4647 cases | Filter to remove junk dataresampling of a cubic spline at 2 Hz | **FE**: Calculate PSD by applying FFT | PSD from RRI and EDR | **Per record**: SVM, Naïve Bayes |
| Erdenebayar, U.2019[30] | South Korea | Samsung Medical Center (Seoul, Korea),86 cases | Use FIR pass band to eliminate noise and baseline driftApply non-overlapping ECG signals at 10-second intervals to match event-based classification | **FE**: To generate 2D input signals using Fourier transform, ECG signals were converted to 2D spectrometer images | \_ | **Per segment**: DNN, 1D CNN, 2D CNN, RNN, LSTM, GRU |
| Farouk, F. N. B. M.2019,[40] | Malaysia | Physionet apnoea-ECG dataset,35 cases | Marking in cases of apnea or normal by medical professionals | **FE**: Optimization of CNN parameters (ReLU, convolution pooling with stride 4 sample, convolution & ReLU, convolution pooling with stride 2 sample, fully connected layer) | \_ | **Per segment**: CNN |
| Fatimah, B.2020,[41] | India | MIT PhysioNet Apnea-ECG Database.35 cases | Using FDM, the ECG signal was decomposed into M FIBF | **FE**: Using statistical analysis of the extracted feature set.**FS**: A subset of statistically relevant features was selected based on their discriminatory characteristics using the Kruskal-Wallis test. | Entropy and MAD features | **Per segment**: Bagging, KNN, SVM, LogitBoost |
| Hassan, A. R.12016,[42] | Bangladesh | Physionet’s apnea-ecg Database,35 cases | ECG signal segments are decomposed into sub-bands using TQWT | **FE**: modeling of the sub-bands using symmetric NIG pdf is performed | NIG parameters | **Per segment**: RBM, SVM, Naïve Bays, ANN, RF, KNN, Bagging, LDA AdaBoost |
| Hassan, A. R. 2,2016,[43] | Bangladesh | Physionetʼs apnea-ecg Databasess,35 cases | Each section was analyzed using EMD | **FE**: compute statistical feature from IMFs**FS**: Statistical hypothesis testing is then performed for feature selection | Statistical feature  | **Per segment**: ANN, naive baye, RBM, KNN, AdaBoost, Bagging, RF, DA, ELM |
| Hassan, A. R. 3,2016,[44] | Bangladesh | Physionetʼs apnea-ecg Databasess,35 cases | \_ | **FE**: Calculate PSD by applying DFT | Statistical and spectral(PSD) feature | **Per segment**: Bagging algorithm |
| Hassan, A. R. 4,2017,[69] | Bangladesh | Physionetʼs apnea-ecg Databasess,35 cases and UCDDB, 25 case | Use of wavelet analysis (TQWT) due to the oscillating nature of ECG signals | **FE**: Different statistical moments were used as features in the proposed framework | sub-band, variance (σ2), skewness (ζ), and kurtosis (η) | **Per segment**: LS-SVM, ELM, PRAZEN-PNN, SVM, KNN, Bagging, RF, Adaboost, Rusboost |
| Jafari, A.2013,[45] | Iran | Physionet’s Apnea-ECG database: 35case | - | **FE**: 3-dimensional frequency features,6 dimensional RPS based features | NVLF, NLF, NHF, DFA, CD, 3 feature from LLEs and SE | **Per segment**: SVM(linear) |
| Khandoker, A. H.2009,[46] | Australia | physionet apnoea-ECG database, 35 cases | QRS detection times and amplitudes were determined using a real-time algorithm,A moving average method was used to remove suspicious RRs | **FE**: HRV and EDR signals were decomposed into 14 levels using daubechies Wvs with 10 vanishing moments.**FS**: Using the hill-climbing feature selection algorithm | variance of the coefficients (HRVWv, EDRWv)  | **Per-record**: SVM (polynomial, linear), LD, KNN, PNN |
| Khandoker, A. H.2.2009,[73] | Australia | SRU: 83 cases,Physionet Apnea-ECG Database: 10 cases,SVUH/UCB: 11 cases(110 case) | QRS detection times and amplitudes were determined using a real-time algorithm,A moving average method was used to remove suspicious RRs | **FE**: HRV and EDR signals were decomposed into 14 levels using daubechies Wvs with 10 vanishing moments.**FS**: Using the hill-climbing feature selection algorithm | variance of the coefficients (HRVWv, EDRWv)  | **Per-record**: SVM (polynomial, linear) |
| Li, K.2018,[47] | China | physionet apnoea-ECG database,35 case | Use the PTA to find the peaks of R,intermediate filter was used to remove Physiologically unexplained points | **FE**: Features were extracted based on DNN,Use the HMM process and finally DF to evaluate records | \_ | **Per segment**: DNN, ANN, ANN-HMM, SVM, SVM-HMM, DF**Per record**: DNN, ANN, ANN-HMM, SVM, SVM-HMM, DF |
| Lweesy, K.2011,[66] | Jordan | University of Heidelberg hospital.25 cases | band pass filter, Chebyshev filter type I and Chechyshev IIR notch filter were used to eliminate noise  | **FE**: Segmentation was performed to extract accurately the Pwaves and the time interval Tpr from an ECG signal | 3 P wave features (Tp, Pd, and Tpr) | **Per segment**: ANN |
| Mendez, M. O.2009.[48] | Italy | physionet apnea-ECG database, 35cases | QRS complex area and RR intervals were derived with real-time algorithm,The mean RR value was calculated using a moving average filter of ten multiplications | **FE**: TVAM method was used to extract the feature | time and frequency domain | **Per segment**: ANN, kNN |
| Mendez, M. O.2007.[49] | Italy | physionet apnoea-ECG database,35 cases | QRS point was performed using an automated algorithm.EDR signal was detected from the primary ECG.The base level was calculated with an average filter | **FE**: bivariate time-varying autoregressive model were used to extract PSD for RRI and DRS series**FS**: it can be evaluated by statistical analysis of features or by WRAP methods. | PSD feature | **Per segment**: kNN |
| Nakayama, C.2019[50] | Japan | physionet apnoea-ECG database,35 cases | RRI detection from ECG series,Extract HRV feature from RRI data | **FE**: time-domain HRV features,frequency-domain features | meanNN, SDNN, RMSSD, Total Power (TP), NN50, and pNN50, LF, HF, LF/HF. LFnu, and HFnu | **Per record**: RF |
| Nguyen, H. D.2014,[31] | America | Physionet Apnea-ECG database,35 cases | \_ | **FE**: The Takens time delay method, which was embedded in one phase, was used to obtain the u (t) data,Extraction of ten RQA variables from RPs,**FS**: Conditional reciprocal information was used to integrate and select the most attractive features |  (DET)), maximum diagonal line length (L), maximum vertical line length (V), (ENTR), (LAM), (MDL) (TT), (T1) and (T2) | **Per segment**: ANN, SVM and DF |
| Nishad, A.2018.[51] | India | Physionet apnea-ECG database,35 cases | Use TQWT-based bank filter to decompose ECG into sub bands | **FE**: Feature extraction was calculated using CCE. is the difference between Correntropy and Medium Correntropy | Entropy feature | **Per segment**: RF |
| Pinho, A.2019,[52] | Portugal | Physionet apnea-ECG database,35 cases | Using TEO for detection HRV seriesSavitzky-Golay filter was used to eliminate noise | **FE**: The 256-point FFT power spectral density was computed for HRV and EDR,**FS**: Measuring the quality of features using the Receiver Operating Cover curve,Discriminant Relevance method was used to rank and select the features | Extract 50 features from HRV, Extract 34 features based on EDR | **Per segment**: ANN, SVM , LDA, PLS, REG, WienerHopf, aNB, PLA |
| Rachim, V. P.2014.[53] | South Korea | Physionet apnea-ECG database,35 cases | \_ | **FE**: Dividing ECG signals into some levels at specific frequencies using wavelet analysis(db4),**FS**: Feature extraction was performed using PCA | Entropy, IQR, Variance of detail coefficient, Standard deviation, Mean Absolute Deviation | **Per segment**: SVM**Per record**: SVM |
| Rekha, B. B.2018.[70] | India | PhysioNet database,35 casesUCD database,25case. | Detect QRS using the Hilbert transform,Chebyshev Type I band pass filter with a frequency band of 6 to 18 Hz. This filter was applied both forward and reverse | **FE**: Twenty-two features were extracted in the domains of time, frequency range and statistical features,**FS**: PCA was performed on the extracted features to obtain the three principal components,CFS was applied to features to reduce dimensional complexity | domains of time, frequency range and statistical features | **Per segment**: SVM, RF |
| Sharma, H.2016,[54] | India | PhysioNet database,35 cases | Using PTA for detection RRI,the Hermite decomposition of QRS complexes was performed using 15 lower order Hermite basis functions | **FE**: Mean R-R (mRR) distances, standard deviation(std) R-R(sRR) distances are used in the feature vector**FS**: Hill climbing algorithm  | \_ | **Per segment:** KNN, MLPNN, LS-SVM and SVM**Per record**: KNN, MLPNN, LS-SVM and SVM |
| Sharma, H.2020,[55] | India | PhysioNet database,35 cases | Using PTA for detection RRI,The EDR signal was detected by Hermite analysis,Two medium filters were used to clear noise | **FE**: HRV and EDR signals were analyzed based on VMD,**FS**: PCA method was used to reduce the features,Feature selection was performed via PCs using a 10-fold validation scheme | spectral entropies, interquartile range, and energy from HRV and EDR, Standard deviation of EDR and RR interval,Standard deviation of Hermite coefficients | **Per segment:** LR, LDA, BDT, ADT, KNN, ANN, LS-SVM, SVM**Per record:** LR, LDA, BDT, ADT, KNN, ANN, LS-SVM, SVM |
| Sharma, M.2018,[56] | India, Singapore and Malaysia  | PhysioNet database,35 cases  | Apnea sections of ECG signals were detected using BAWFB system | **FE**: Extraction of fuzzy entropy from WSBsLog-energy (LE) extraction for SBs | \_ | **Per segment**: Weighted KNN, CT, LD, LR, LS-SVM(kernel) |
| Sharma, M.2019,[57] | India, Singapore and Malaysia | PhysioNet database,35 cases | Apnea sections of ECG signals were detected using OWFB designed | **FE**: Extraction of fuzzy entropy from WSBsLog-energy (LE) extraction for SBs | - | **Per segment**: KNN, CT, LD, LR, Gaussian SVM |
| Smruthy, A.2017,[75] | India | \_ | Decomposition of ECG signals using VMD algorithm into different variational mode functions | **FE**: Use FFT to extract attributes from each VMF |  mean energy The standard deviation of the peak-to-peak distance  | **Per record**: SVM |
| Song, C.2016,[58] | China and America | PhysioNet database,35 cases | Detection of peaks of QRS complexes was performed using BIOSIG toolbox. EDR detection was performed by applying two medium filtersmedium filter was used to eliminate abnormal infrastructure | **FE**: 24 features from RR intervals and 8 features from EDR signals, were extracted**FS**: Apply Hidden Markov Model on features,using a step-by-step design with integrated validation (LOOCV) | Time and frequency domain | **Per segment**: SVM**Per record**: SVM |
| Travieso, C. M.2014[71] | Spain and Colombia | PhysioNet database,35 cases.UCDDB, 25 cases | Use a continuous wavelet transform based algorithm to detect RRI | **FE**: Use Cepstrum power parameters as a feature.Apply Hidden Markov Model on features | \_ | **Per segment**: SVM-HMM |
| Tripathy, R. K.2018,[59] | India | PhysioNet database,35 cases | Using PTA for detection RRI,bandwidth filters were used to detection of EDR | **FE**: extraction of features from the intrinsic band functions (IBFs) of both EDR and HRV signals.Extraction of frequency-based properties using Fourier spectrum**FS**: PCA based algorithm was used for the extraction of EDR signal,  | Frequency feature, Fuzzy entropy | **Per record**: SVM |
| Varon, C.2015.[72] | Belgium | PhysioNet database,35 cases,UZ Leuven10 cases | Power line interference at 50 Hz was filtered using a gap filter and the average signals were removed,Sampling of all ECG signals was performed using cubic spline interpolation at 250 Hz,Using PTA for detection RRI | **FE**: EDR detection was performed by applying two medium filters removing the QRS complex, R-waves and Pi waves. Then the found baseline was used for EDR detection**FS**: Use of PCA on EDR signals | Frequency feature | **Per segment**: SVM, LDA, LS-SVM |
| Wang, L.2019.[60] | China | PhysioNet database,35 cases | RR distance detection,Remove RRIs that are 20% of the average distance | **FE**: Display RR distances using dynamic autoregressive representation model (DARRM)Optimization of CNN parameters | \_ | **Per segment**: CNN, residual network  |
| Wang, T.2019.[18] | China | PhysioNet database,35 cases | Hamilton algorithm was used to detect R peaks,FIR pass band filter and median filter were used to clear junk data | **FE**: Calculate PSD by applying FFT | Time domain and frequency feature | **Per segment**: LR, LDA, SVM, MLP, TW-MLP**Per record**: LR, LDA, SVM, MLP, TW-MLP |
| Wang, T.2019.[20] | China | PhysioNet database,35 cases.UCD database,25 cases | Hamilton algorithm was used to find the peaks of R.R peaks were used to calculate RR distances and amplitude extractionMedium filter was used to remove Cubic interpolation was used due to unequal time interval | **FE**: A simple CNN run, LeNet-5, was used to build the diagnostic model Calculate PSD by applying FFT | Frequency and amplitude feature  | **Per segment**: SVM, LR, KNN, MLP, LET-NET5 CNN**Per record**: SVM, LR, KNN, MLP, LET-NET5 CNN |
| Wang, X. W.2020.[61] | China | PhysioNet database,35 cases | \_ | The model designed in this study consists of three convolution layers (the first two convolution layers are followed by batch normalization and maximization layers. The third convolution layer is followed by three fully connected layers) | \_ | **Per record**: LET-NET5 CNN |
| Yildiz, A.2011.[62] | Turkey | PhysioNet database,35 case | multi-resolution decomposition of ECG signal was performed with Daubechies 6 (db6) wavelet, Detection of EDR and HRV using QRS complex | **FE**: PSD characteristics of HRV and EDR signals were extracted using FFT-based method**FS**: The hill-climbing method was used to select the feature | Frequency feature | **Per record**: LS-SVM |

ECG: Electrocardiogram, TEO: Teager Energy Operator, PTA: Pan–Tompkins Algorithm, FT: Fourier Transform, FFT: Fast Fourier Transform, STFT: Short-Time Fourier Transform, FDM: Fourier Decomposition Method, DFT: Discrete Fourier Transform, PCA: Principal Component Analysis, CCE: Centered Correntropy, HRV: Heart Rate Variability, RRI: RR Interval, RQI: Recurrence Quantification Analysis, BAWFB: Biorthogonal Antisymmetric Wavelet Filter Bank, OWFB: Orthogonal Wavelet Filter Banks, EDR: ECG-derived respiratory, CPSD: Cross Power Spectrum Density, EMD: Empirical Mode Decomposition, VMD: Variational Mode Decomposition, DARRM: Dynamic Autoregressive Representation Model , VMF: Variational Mode Functions, NIG: normal inverse Gaussian, TQWT: tunable-Q factor wavelet transform, PSD: Power Spectral Density, LF: Low Frequency, VLF: Very Low Frequency, HF: High Frequency, FFNN: Feed Forward Neural Network, SVM: Support Vector Machine, LS-SVM: least-square support vector machine, QC: quadratic classifier, ANN: Artificial Neural Network, ANN-SCG: Artificial Neural Network- Scaled Conjugate Gradient, ANN-LM: Artificial Neural Network- Levenberg-Marquardt algorithm, CNN: Convolutional neural network, DNN: Deep Neural Network, RNN: Recurrent Neural Networks DT: Decision tree, kNN: k-nearest neighbors algorithm, GRU: Gated Recurrent Unit, RBF: Restricted Boltzmann Machine, LDA: linear Discriminant Analysis, Bagging: Bootstrap Aggregating, Adaboost: Adaptive boosting, HMM: Hidden Markov Model, PLS: Partial Least Squares Regression, REG: Regression Analysis, aNBC: Augmented Naive Bayesian Classier, PLA: Perceptron Learning Algorithm, LMS: Least Mean Square, WienerHopf: Wiener–Hopf equation, BDT: Bagged Decision Tree, CT: Complex Tree, LR: Logistic Regression,

***Feature Extraction/Selection***

IN the selected studies, after ECG analysis, features were extracted from different sections such as RRI, HRV, EDR, R wave and P wave. Fourier transform has been one of the most common methods for feature extraction. This method was used in most studies to extract Power Spectral Density (PSD) of signals, and the frequency feature. PSD shows energy changes as a function of frequency. In a number of studies [18, 35, 36, 49, 50, 52, 58, 70], in addition to the frequency feature, Fourier transform was used to extract the time domain feature. Another study [65] extracted the frequency feature using wavelet packet decomposition. The Lomb algorithm was also used to obtain the PSD of signals [34]. In one article, the Hilbert algorithm was used to extract the time and frequency features [70].

In 6 studies [18, 35, 36, 48, 50, 58], frequency and time domain features were extracted from ECG signals simultaneously. Eight other studies [34, 38, 49, 62, 64, 67, 68, 72] adopted the frequency feature only for classification of SA and healthy individuals. One research work [33] also used only the time domain feature for classification. In another work [33], PSD was extracted along with bispectral feature. Another article [74] that used a portable accelerometer with 3 electrodes used the frequency along with the data accelerometer as a feature. Another work [45] adopted the frequency feature along with the RPS based feature for classification. One piece of research [18] also used the frequency feature along with the amplitude feature. Frequency and statistical feature were also extracted in one article [44]. Moreover, 7 studies [29, 31, 43, 46, 69, 70, 73] considered statistical features, 9 studies [41, 51, 53, 55-57, 59, 66, 75] the characteristics related to wave energy and entropy, and One study [66] used the properties of P and T wave analyses. Four articles [31, 59, 71, 72] adopted various features for classification, yet the specific features used in these studies were not reported. Also, a number of other studies [20, 30, 37, 39, 40, 47, 60, 61] prepared data for classification by applying different layers within the Neural Network (NN).

***Per record classification***

In 23 studies, data were classified as per record. The data used in the studies were between 20 and 4647 ECG strips. Classification was performed based on 15 different ML algorithms. The most commonly used algorithm was SVM, which was used in RBF, linear and polynomial form; The accuracy of this algorithm in detecting SA from healthy individuals was reported to be 65.4% to 100%. The sensitivity and specificity of this algorithm were also reported (43.4%-100%) and (36.4%-100%) respectively. Moreover, kNN algorithm was used in 7 studies; the lowest accuracy of this algorithm was reported 77.3%, and the highest accuracy was 97.14%. In all of these studies, the classification sensitivity was equal to or greater than 80%. The specificity of the kNN algorithm also showed that, with the exception of one work [20], in other studies at least half of the healthy individuals were correctly classified, thus the specificity was 50%. LR was another common algorithm for classifying data per record. The accuracy of this method in diagnosing SA was reported to be in the range of 74.3% to 97.14%. The classification sensitivity of this algorithm was reported to be 100% in all studies. However, the specificity of the LR algorithm was very scattered and was reported in the range of 18.2% to 90.91%.

Other methods used in the studies include neural network algorithms. These algorithms include FFNN two studies [64, 65], PNN in three studies [46, 54, 65], CNN 2 studies [20, 37], DNN [77] and ANN [35] each in one study. The results showed that neural network algorithms perform classification with very high accuracy. Classification accuracy was reported in these studies (80%-99%). The sensitivity and specificity of these algorithms were investigated in 7 studies. Sensitivity greater than 85% and specificity greater than or equal to 80% were reported in all, but one study [46]. Among the other techniques used Naive Bayes, LD, LS-SVM, LDA, BDT, ADT, MLP and QC could be mentioned mentioned. The minimum classification accuracy was obtained in the study of Eiseman et al. [68] using Naive Bayes for which the value was reported to be 63.02%. Study of Song et al. [58] also reported SA detection accuracy based on ANN algorithm as 68.6%. The results of other studies reported a high accuracy of 75% in the classification based on per record data (Table 3).

***Per segment classification***

Classification was performed per segment in 33 articles. These studies used from 1500 to 43522 segments for their classification. In 9 studies [20, 31, 42-44, 51, 53, 69, 70], the number of segments examined was not reported. However, in all of these studies, it was stated that one-minute ECG components were labeled by experts, indicating a classification based on one-minute components in these studies. Furthermore, SVM was used as the most common algorithm. In total, among the selected studies 31 different ML algorithms were used to classify SA and healthy individuals.

SVM algorithm has been used in linear, polynomial, RBF, and nonlinear forms. The accuracy of this algorithm was reported in the range of 59.22% to 93.91%. The sensitivity of the algorithm was also reported in the range of 32.7% to 95.2%. Specificity studies of SVM algorithm showed that the minimum specificity was 47.32%, and the maximum was 95.42%. In 3 articles, Hidden Markov Model (HMM) was applied to SVM algorithm, leading to the accuracy of the algorithm in these 3 studies to be between 80% and 100%. Another LS-SVM algorithm was used in a number of research works. The lowest accuracy reported was in the study of Hassan et al. [69], which was 31.88%. Other studies reported an accuracy of more than 70% for the LS-SVM algorithm. The highest accuracy was reported in the study of Atri et al. [33] with 95.57%. The sensitivity and specificity of this algorithm in this study were 98.64% and 92.51% respectively.

Neural Network (NN) algorithms have also been used in several related research works. CNN, DNN, RNN, ANN and PNN were the specific algorithms used. Eight studies [31, 42, 43, 47, 48, 52, 55, 66] adopted the ANN algorithm. The highest accuracy was reported at 92.3% and the lowest at 68.52%. CNN was another algorithm that was examined in 6 studies [20, 30, 37, 39, 40, 60]. The accuracy reported in these studies was 78.2% to 98.91%. Sensitivity and specificity were also calculated to be more than 80% in most studies. In only one article [20] the reported classification sensitivity was 26.6%. This study had the lowest accuracy among all CNN algorithms. Specificity reported 86.9%. DNN and RNN‌ algorithms were applied to the data in the study of Erdenebayar et al. [30]. The results of this study showed that the accuracy of the DNN algorithm in detecting and classifying SA and healthies was ‌ 93.1%. Moreover, in this study, the accuracy of the RNN algorithm was 99%. The sensitivity and specificity of DNN and RNN algorithms were above 90%. PNN was also investigated in a study with an accuracy of 60.95%. This work [69] did not report sensitivity and specificity for the PNN algorithm.

RF, LR, LDA and kNN were 4 other algorithms used to classify data. kNN was applied to selected datasets in 12 studies [20, 29, 41-43, 48, 54-58, 69]. Examination of the results of these studies showed that the accuracy of the kNN algorithm is in the range of 66.1% to 90.57%. In these research works, with increasing accuracy, sensitivity and specificity also increased. The accuracy of RF, LR and LD‌ algorithms were reported in the ranges of 79.26%-92.78%, 66%-85.6% and 62.93%-83.72% respectively, which shows the better performance of RF algorithm in data classification. Information on other algorithms used are provided in Table 3.

Table 3: Classification result based on per record or per segment data

|  |
| --- |
|  Per segment  |
| First author, Year, | **NO, of segments** | **classification** | **ACC%** | **Other parameters** |
| Al-Angari, H. M. 2012, | 39575 segments | SVM (linear) C= 10 | 68.8 | Sen: 51.6, spec: 81.4 |
| SVM (polynomial) C= 10 | 69.5 | Sen: 54.8, spec: 80.1 |
| Atri, R. 2015 | 16479 segments | LS-SVM | 95.57 | Sen: 98.64, Spec: 92.51 |
| Babaeizadeh, S. 2010, | 17010 training. 17268 test sesments  | QC | 84.7 | Sen: 76.7, Spec: 89.6, PPV: 81.8, NPV:86.3 |
| Bozkurt, F. 2020, | 2460 segments (1242 A, 1218 N),1230 training segments,1230 test segments | DT (113 feature) | 79.11 | Sen: 76, sp:82 |
| kNN (90 feature) | 82.2 | Sen: 78, sp:75 |
| SVM (113 feature) | 84.15 | Sen: 76, sp:82 |
| Ensamble (113 feature) | 85.12 | Sen: 76, sp:82 |
| Bsoul, M. 2011, | 14700 one-minute segments | SVM (linear) C=32  | 91.16 | Sen:89.12, Spec: 92.35, F: 90.70 |
| SVM (poly, d=2) C=0.5, γ= 0.5 | 89.85 | Sen:88.25, Spec: 88.25, F: 90.82 |
| SVM (RBF) C=2, γ= 0.5 | 90.86 | Sen: 89.02, Spec: 91.94, F: 90.46 |
| SVM (MLP) C=0.5, γ= 0.5 | 80.45 | Sen:74.66, Spec: 83.96, F: 79.04 |
| Chang, H. Y. 2020, | 34230 segments, 17234 test segments | CNN | 87.9 | Sen:81.1, Spec:92 |
| Dey, D. 2017, | 10787 segments, (4987A, 5800 N) | CNN | 98.1 | Sen:97.82, Spec: 99.2, PPV: 99.06, NPV: 98.14 |
| Erdenebayar, U. 2019 | 43522 segments,37338 training segments,6184 test segments (1623A, 4561N) | DNN | 93.1 | Sen: 93 Spec:94 |
| 1DCNN | 98.5 | Sen: 99 Spec:99 |
| 2DCNN | 95.9 | Sen: 96 Spec:96 |
| RNN | 85.4 | Sen: 97 Spec:87 |
| RNN(LSTM) | 98 | Sen: 98 Spec:98 |
| RNN(GRU) | 99 | Sen: 99 Spec:99 |
| Farouk, F. N. B. M. 2019, | 17402 segments, 10787 test segments | CNN | 98.91 | Sen:97.82, Spec:99.20 PPV:99.06, NPV:98.14 |
| Fatimah, B.2020, | 17010 segments (6514A, 10496N) | Bagging | 91.44 | Sen:91.61 Spec: 92.52 Pre: 88.20 |
| KNN | 90.57 | Sen: 89.13 Spec: 91.56 Pre: 86.90 |
| SVM | 92.59 | Sen: 89.70 Spec: 94.67 Pre: 91.27 |
| LogitBoost | 85.84 | Sen: 79.17 Spec: 89.97 Pre: 83.04 |
| Hassan, A. R.1 2016, | \_ | RBM | 38.79 | AdaBoost:Sen: 81.99, Spec: 90.72 |
| SVM | 59.22 |
| Naïve Bays  | 62.15 |
| ANN | 81.37 |
| RF | 82.70 |
| KNN | 83.32 |
| Bagging | 83.33 |
| LDA | 83.72 |
| AdaBoost | 87.33 |
| Hassan, A. R. 2, 2016, | - | ANN  | 68.52  | ELM:Sen: 85.20, Spec: 82.79 |
| NBC | 39.47 |
| RBM | 61.20 |
| KNN | 69.72 |
| AdaBoost | 80.07 |
| Bagging | 79.82 |
| RF | 79.26 |
| DA | 64.60 |
| ELM | 83.77 |
| Hassan, A. R. 3, 2016, | - | Bagging | 85.97 | Sen:84.14, Spec: 86.83 |
| Hassan, A. R. 4, 2017, | - | LS-SVM | 31.88 | RusboostSen: 87.58, Spec: 91.49 |
| ELM | 53.02 |
| PRAZEN-PNN | 60.95 |
| SVM | 72.4 |
| KNN | 79.77 |
| Bagging | 84.29 |
| RF | 84.49 |
| Adaboost | 86.94 |
| Rusboost | 88.88 |
| Jafari, A. 2013, | 16711 segments, 6711 test segments | SVM | 94.80 | Sen:94.16, Spec:95.42 |
| Li, K. 2018, | 17122 test segments (6517A, 10605N) | ANN | 78.3 | Sen:66.6 Spec:85.4 |
| ANN-HMM | 83 | Sen:91.5 Spec:77.7 |
| SVM | 78.6 | Sen:66.5 Spec:86.1 |
| SVM-HMM | 84.7 | Sen:68.8 Spec:94.5 |
| Decision fusion | 84.7 | Sen:88.9 Spec:82.1 |
| Lweesy, K. 2011, | 1500 segments, 224 test segments | ANN | 92.3 | Sen:90.1, Spec:94.4, |
| Mendez, M. O. 2009. | 24432 segments,12077 training segments, 12355 test segments | KNN Feature=10 | 88 | Sen:86, Spec: 87 |
| ANN Feature=10 | 88 | Sen: 89, Spec:86 |
| Nguyen, H. D. 2014, | - | ANN | 83.23 | Sen:85.57, Spec:79.09 |
| SVM | 84.14 | Sen:93.72, Spec:65.88 |
| Decision fusion | 85.26 | Sen:86.37, Spec:83.47 |
| Nishad, A. 2018. | - | RF | 92.78 | Sen: 90.95, Spec:93.91 |
| Pinho, A. 2019, | 5671 test segments | ANN Features: 20 | 82.12 | Sen:88.41, Spec:72.29 |
| SVM Features: 70 | 75.18 | Sen:86.79, Spec:56.45 |
| LDA Features: 20 | 62.93 | Sen:83.98, Spec:28.40 |
| PLS Features: 20 | 64.49 | Sen:57.78, Spec:66.05 |
| REG Features: 20 | 65.13 | Sen:62.23, Spec:65.65 |
| WienerHopf Features: 20 | 64.05 | Sen:58.14, Spec:65.07 |
| aNBC Features: 44 | 62.12 | Sen:0, Spec:62.12 |
| PLA Features: 6 | 61.36 | Sen:36.84, Spec:61.70 |
| LMS Features: 84 | 61.72 | Sen:28.70, Spec:62.35 |
| Rachim, V. P. 2014 | - | SVM (RBF), C=10, Ϭ=0.5, PCA=5 | 93.91 | Sen: 95.20, Spec: 92.65 |
| Rekha, B. B. 2018. | \_ | SVM (without feature reduction) | 91 | Sen: 90.38, Spec: 91.54 |
| RF (with feature reduction) | 94.32 | Sen:92.98, spec: 94.77 |
| Sharma, H. 2016, | 32727 segments,16845 training segments, 15873 test segments | KNN | 73.3 | Sen:72.5, Spec:73.8 AUC:73.8 |
| MLPNN | 81.2 | Sen:77.5, Spec:83.4 AUC:80.7 |
| LS-SVM | 82.6 | Sen:76.7, Spec:88.2 AUC:82 |
| SVM | 83.8 | Sen:79.5, Spec:88.4 AUC:83.4 |
| Sharma, H. 2020, | 34313 set,17045 training, 17268 test segments | LR | 85.6 | Sen:82.1, SP:90.2 AUC: 0.93 |
| LDA | 82.5 | Sen: 73, Spec: 88.5 AUC: 0.89 |
| BDT | 84.7 | SEN:79.5, Spec: 87.5 AUC: 0.92 |
| ADT  | 81.6 | Sen: 74, Spec: 86.3 AUC: 0.88 |
| KNN | 87.5 | Sen: 84.9, Spec: 88.2 AUC: 0.93 |
| ANN | 86.1 | Sen: 84, Spec: 86.9, AUC: 0.82 |
| LS-SVM | 86.2 | Sen: 81.3, Spec: 87.7 AUC: 0.91 |
| SVM | 85.3 | Sen: 82.5, Spec: 88 AUC: 0.91 |
| Sharma, M. 2018, | 16993 test segments (6513A, 10480N) | Weighted KNN  | 89.1 | Sen: 91.8, Spec: 84.9, PPV: 90.36, NPV:87 |
| CT | 83.5 | Sen: 86.3, Spec:78.9, PPV: 87.1, NPV:77.8 |
| LD | 66.1 | Sen:70.9, Spec: 56.7, PPV: 76.6, NPV: 49.8 |
| LR | 66.5 | Sen: 70, Spec: 58.1 PPV:80, NPV: 44.8 |
| LS-SVM(kernel)  | 90.11 | Sen: 90.9, Spec:88.9 PPV: 93, NPV: 85.8 |
| Sharma, M. 2019, | 16993 test segments (6513A, 10480N) | KNN  | 90.3 | Sen: 91.5, Spec: 88.5, PPV: 92.8, NPV:86.6 |
| CT | 82.9 | Sen: 91, Spec:76.8, PPV: 85.8, NPV:78.1 |
| LD | 68.9 | Sen:86.9, Spec: 40, PPV:70, NPV: 65.4 |
| LR | 71.1 | Sen: 86, Spec: 47.2 PPV:72.4, NPV: 67.7 |
| Gaussian SVM | 90.87 | Sen: 92.43, Spec:88.33 PPV: 92.8, NPV: 88.3 |
| Song, C. 2016, | 17268 test segments (6550A, 10718N) | SVM | 81.2 | Sen: 75.7, Spec: 84.7 |
| SVM-HMM  | 86.2 | Sen: 82.6, Spec: 88.4  |
| LR | 81.2 | Sen: 74.4, Spec: 85.4 |
| LR+HMM | 86.2 | Sen: 80, Spec: 89.9 |
| LDA | 80.5 | Sen: 83.1, Spec:78.9 |
| LDA+HMM | 85.3 | Sen: 77.5, Spec: 90.1 |
| KNN | 80.7 | Sen:75.3, Spec: 83.9 |
| KNN+HMM | 84.5 | Sen: 74, Spec: 90.8 |
| Travieso, C. M. 2014 | 11266 test segments (4375A, 6891N) | SVM-HMM | 99.2 | Sen: 98.8, Spec: 99.5, PPV: 99.2, NPV:99.3 |
| Varon, C. 2015. | 32 477 segments,6000 training segments, 26477 test segments | LDA | 71.43 | Sen: 71.74, Spec:71.2  |
| SVM(LIN) | 71.16 | Sen: 74.63, Spec: 68.51 |
| SVM(POLY) | 72.6 | Sen: 78.36, Spec: 68.22 |
| SVM(RBF) | 73.86 | Sen: 78.2, Spec: 70.55 |
| LS-SVM(LIN) | 71.78 | Sen: 71.48, Spec: 76.26 |
| LS-SVM(POLY) | 73.43 | Sen: 73.4, Spec:73.43 |
| LS-SVM(RBF) | 84.74 | Sen: 84.71, Spec: 84.69 |
| Varon, C.2 2015. | 5205 segments,3000 training segments, 2205 test segments | LDA  | 79.86 | Sen: 81.22, Spec: 79.71 |
| SVM(LIN) | 75.83 | Sen: 86.46, Spec: 74.6 |
| SVM(POLY) | 77.78 | Sen: 88.21, Spec: 76.57 |
| SVM(RBF) | 81.9 | Sen: 82.97, Spec: 81.78 |
| LS-SVM(LIN) | 80.05 | Sen: 80.35, Spec: 80.01 |
| LS-SVM(POLY) | 81.68 | Sen: 81.66, Spec: 81.7 |
| LS-SVM(RBF) | 83.95 | Sen: 78.81, Spec: 84.56 |
| Wang, L. 2019. | 16,988 test segments,6496A, 10492A | CNN | 90.97 | Sen: 83.43, Spec: 85.5 |
| residual network | 94.39 | Sen: 93.04, Spec: 94.95 |
| Wang, T. 2019. | 34,313 segments, 17045 training segments, 17268 test segments. | LR | 81.5 | Sen: 72, Spec: 87.4 |
| LDA | 81.8 | Sen: 70.9, Spec: 88.4  |
| SVM | 80.6 | Sen: 72.1, Spec: 85.6 |
| MLP | 81.4 | Sen: 74.3, Spec: 85.7 |
| TW-MLP | 87.3 | Sen: 85.1, Spec: 88.7 |
| Wang, T.2 2019. |  | SVM | 81.4 | Sen: 76.9, Spec: 84.3 |
| LR | 80.8 | Sen: 75.7, Spec: 84 |
| KNN | 77.5 | Sen: 68.1, Spec: 83.4 |
| MLP | 81.1 | Sen: 71.3, Spec: 87.2 |
| LET-NET5 CNN | 87.6 | Sen: 83.1, Spec: 90.3 |
| Wang, T.3 2019. |  | SVM | 70.6 | Sen: 32.7, Spec: 83.3 |
| LR | 69.6 | Sen: 34.7, Spec: 81.3 |
| KNN | 66.1 | Sen: 38.1, Spec: 75.4 |
| MLP | 67.2 | Sen: 38.5, Spec: 76.8 |
|  CNN | 71.2 | Sen: 26.6, Spec: 86.9 |
| Per record |
| First author, Year, | **No, of records** | **ML algorithm** | **ACC%** | **Other parameters** |
| Akşahin, M. 2015, | 20test, (10 A, 10 N) | FFNN | 99 | \_ |
| Al -Angari, H. M. 2012, | 100 records50 apnea, 50 normal | SVM(linear) C=5 | 79 | Sen: 79.6, spec: 78.4 |
| SVM(polynomial) C=5, 10 | 78 | Sen: 67.3, spec: 88.2 |
| Ali, S. Q. 2020, | - | FFNN | 87.5 | Sen:86.67, spec: 90 |
| PNN | 85 | Sen:86.67, spec: 80 |
| Babaeizadeh, S. 2010, | 70 records, 30 test segments (20 A, 10 N) | QC | 100 | Sen: 100, Spec: 100, PPV: 100, NPV: 100 |
| Baek, J. W. 2014, | 20 records (10A, 10N), | ML | 85 | Sen: 90, spec: 80 |
| Bali, J. 2018, | 70 records,35 training segments, 35 test segments (23A, 12N) | ANN-LM | 91 | Sen:91, spec: 92, PR:95 |
| ANN-SCG | 95 | Sen:94, spec: 91, PR: 96 |
| Chang, H. Y. 2020, | 35 training segments, 35 test segments | CNN | 97.1 | Sen:95.7, spec:100 |
| Chen, L. 2015 | 90 subjecta,59 training, 31 test segments (23A, 8N) | SVM (RBF kernel), C=3 | 97.41 | Sen:98.99, Spec: 92.87 |
| SVM(polynomial) C=5, Order=1 | 97.03 | Sen:99.16, Spec: 90.91 |
| Kaguara, A. 2014, | 70 records, (35 training segments, 35 test segments) | DNN (fold 4) | 91 | \_ |
| Eiseman, N. A. 2012, | 4647 records,2090 A, 2557 N | SVM | 65.4 | Sen: 43.4, Spec: 83.5 PPV: 68.3, NPV: 64.4 |
| Naive Bayes | 63.02 | Sen: 39, Spec: 82.7 PPV: 64.8 NPV: 62.4 |
| Khandoker, A. H. 2009, | 60 records,30 training segments, 30 test segments (20A, 10N) | SVM(Poly) Feature:5,6,7or8, C=0.1,1,1 | 100 | Sen: 100, Spec: 100 |
| SVM(Linear) Feature: 6 or7, C=10 | 100 | Sen: 100, Spec: 100 |
| LD | 90 | Sen: 100, Spec: 93 |
| KNN (K =1) | 80 | Sen: 90, Spec: 83 |
| PNN (Ϭ =0.5) | 80 | Sen: 50, Spec:70 |
| Khandoker, A. H.2. 2009, | 125 records,83 training segments, 42 test segments | SVM(Polynomial) D= 3, C =0.8 | 100 | Sen: 100, Spec: 100  |
| SVM(Linear) C =10 | 98.8 | Sen: 100, Spec: 94.44 |
| SVM(RBF)Ϭ =0.5, C= 10 | 96.39 | Sen: 98.46, Spec: 88.89 |
| Mendez, M. O. 2007. | 25 training segments, 25 test segments | KNN | 85 | Sen: 83.90, Spec:88.50 |
| Nakayama, C. 2019 | 61 test segments (25A, 36N) | RF | 85 | Sen: 76, Spec: 92 |
| Rachim, V. P. 2014 | 35 test segments (22A,13N) | SVM | 94.3 | Sen: 100, Spec: 81.3 |
| Sharma, H. 2016, | 70 records,35 training segments, 35 test segments | KNN | 77.3 | Sen:100, Spec:63.6, AUC:77.3 |
| MLPNN | 93.4 | Sen:95.8, Spec:90.9, AUC:93.4 |
| LS-SVM(RBF) | 97.8 | Sen:95.8, Spec:100 AUC:97.8 |
| SVM(RBF) | 97.8 | Sen:95.8, Spec:100 AUC:97.8 |
| Sharma, H. 2020, | 70 records,35 training segments, 35 test segments | LR | 97.14 | Sen:100, Spec:90.91AUC:97.8 |
| LDA | 100 | Sen:100, Spec:100 AUC:0.95 |
| BDT | 97.14 | Sen:100, Spec:90.91 AUC:1 |
| ADT | 94.28 | Sen:91.67, Spec:100 AUC: 0.95 |
| KNN | 97.14 | Sen:100, Spec: 90.91 AUC: 0.95 |
| ANN | 97.14 | Sen:100, Spec: 90.91 AUC: 0.95 |
| LS-SVM | 94.28 | Sen:100, Spec: 90.91 AUC: 0.95 |
| SVM | 97.14 | Sen:95.8, Spec: 90.91 AUC: 0.93 |
| Smruthy,A. 2017 | 40 test segments | SVM | 97.5 | Sen:95.45, spec: 100, PPV: 100, NPV: 94.7 |
| 25 test segments | SVM | 95 | Sen: 100, Spec: 80, PPV: 94.12, NPV:1 |
| Song, C. 2016, | 35 test segments | SVM  | 80 | Sen: 100, Spec: 36.4 |
| SVM-HMM | 97.1 | Sen: 95.8, Spec:100 |
| LR  | 74.3 | Sen: 100, Spec: 18.2 |
| LR+HMM  | 97.1 | Sen:95.8, Spec: 100 |
| LDA | 68.6 | Sen:100, Spec: 0 |
| LDA+HMM | 97.1 | Sen: 95.8, spec: 100 |
| KNN | 91.4 | Sen: 100, Spec: 72.7 |
| KNN+HMM | 91.4 | Sen: 87.5, Spec: 100 |
| Tripathy, R. K. 2018, | 31 test segments | KELM (RBF) K=5 | 78.71 | Sen:83.45, Spec:73.27 |
| KELM (LINEAR) K=10 | 75 | Sen:91.26, Spec:58.19 |
| KELM (POLY) K=10 | 83.46 | Sen: 85.6, Spec: 81.30 |
| KELM (CWK) K=6 | 78.71 | Sen: 79.06, Spec: 78.33 |
| Wang, T. 2019. | 35 test segments | LR  | 91.4 | Sen: 100, Spec: 75 |
| LDA | 88.6 | Sen: 100, Spec: 66.7 |
| SVM | 82.9 | Sen: 100, Spec: 50 |
| MLP | 82.9 | Sen: 100, Spec: 50 |
| TW-MLP | 97.1 | Sen: 100, Spec: 91.7 |
| Wang, T.2 2019. | 35 test segments | SVM | 88.6 | Sen: 100, Spec: 66.7 |
| LR | 88.6 | Sen: 100, Spec: 66.7 |
| KNN | 82.9 | Sen: 100, Spec: 50 |
| MLP | 85.7 | Sen: 95.7, Spec: 66.7 |
| LET-NET5 CNN | 97.1 | Sen: 100, Spec: 91.7 |
| Wang, T.3 2019. | 25 test segments | SVM | 92.3 | Sen: 100, Spec: 50 |
| LR | 84.6 | Sen: 100, Spec: 50 |
| KNN | 84.6 | Sen: 90.9, Spec: 0 |
| MLP | 92.3 | Sen: 100, Spec: 50 |
| LET-NET5 CNN | 92.3 | Sen: 90.9, Spec: 100 |
| Yildiz, A. 2011. | 35 test segments | LS-SVM(RBF) | 83.3 | Sen: 95, Spec: 60 |
| LS-SVM(POLY) | 76.7 | Sen: 85, Spec:60 |
| LS-SVM(LIN) | 86.7 | Sen: 90, Spec: 80 |

FFNN: Feed Forward Neural Network, SVM: Support Vector Machine, LS-SVM: least-square support vector machine, QC: quadratic classifier, ANN: Artificial Neural Network, ANN-SCG: Artificial Neural Network- Scaled Conjugate Gradient, ANN-LM: Artificial Neural Network- Levenberg-Marquardt algorithm, CNN: Convolutional neural network, DNN: Deep Neural Network, RNN: Recurrent Neural Networks DT: Decision tree, kNN: k-nearest neighbors algorithm, GRU: Gated Recurrent Unit, RBF: Restricted Boltzmann Machine, LDA: linear Discriminant Analysis, Bagging: Bootstrap Aggregating, Adaboost: Adaptive boosting, HMM: Hidden Markov Model, PLS: Partial Least Squares Regression, REG: Regression Analysis, aNBC: Augmented Naive Bayesian Classier, PLA: Perceptron Learning Algorithm, LMS: Least Mean Square, WienerHopf: Wiener–Hopf equation, BDT: Bagged Decision Tree, CT: Complex Tree, LR: Logistic Regression

**Discussion**

This systematic review was conducted to investigate the applications, and current performance of ML techniques in the diagnosis of SA. ML is used when storing, efficiently managing and extracting information from data are challenging tasks [78]. ML includes a set of artificial intelligence algorithms that enable computers to think and learn on their own. These algorithms are used to achieve more accuracy [79, 80]. It has always been observed that providing accurate diagnosis based on experimental data is ambiguous and difficult [81]. Therefore, in recent years, the use of ML algorithms in the healthcare sector has been increasing rapidly [82].

Finding evidence on the high accuracy of ML algorithms in providing medical diagnoses is an important objective, since providing an accurate diagnosis is recognized as one of the major challenges facing global health systems [83]. Nonetheless, the high accuracy of ML techniques generally has led to their successful in many medical disciplines [84]. One of the features of ML techniques is that it allows the analysis of larger and more complex data. On the other hand, ML models extract features that are not normally extractable by human [85].

In this study, different ML algorithms that are used in the diagnosis of SA were reviewed. Our study shows that SA diagnosis using both per record and per segment methods have high accuracy, sensitivity and specificity. Within the reviewed articles, various features of the ECG were also extracted and used in the classification. From the medical perspective, hormonal and neurological changes are observed automatically and frequently during sleep. One of the changes that is commonly seen during sleep is a decrease in Heart Rate (HR) [34]. When apnea stops breathing, the oxygen level in the blood decreases rapidly, which causes a rapid increase in HR. Therefore, Heart Rate Variability (HRV) is one of the cases used in the diagnosis of apnea [86]. EDR is another feature that was commonly monitored. This feature is important because the electrodes on the surface of the body move due to the filling and emptying of the lungs. This movement shifts the axis relative to the heart, and a continuous ECG-derived respiration is extracted for each normal QRS complex [87]. As a result, many classification studies were performed based on EDR and HRV characteristics.

The number and selection of appropriate attributes is crucial for the success of classification [88]. PSD signals was one of the most common features used in most research works standalone, or in combination with other features. Time domain and statistical feature were other features that were used for classification. In general, our review showed that all the features used in the selected studies were applicable to the diagnosis of SA. However, it was not possible to determine the best feature to diagnose SA. This is due to the high accuracy of classification in various studies. Moreover, the combined use of the features in some research works prevents their detailed study.

The highest classification accuracy (100%) was observed when performing the analysis based on a complete ECG record. The study by Babaeizadeh et al. [34] adopted the QC algorithm to classify data. This study, which reported 100% accuracy, sensitivity and specificity, extracted PSD features from HR. However, none of the studies used the QC algorithm, therefore, it is not possible to accurately outline the performance of this algorithm. Other studies such as Khandoker [46, 73] and Sharma et al. [55] also reported 100% accuracy, sensitivity and specificity. In the first two studies which used a wavelet-based analysis, HRV and EDR waves were decomposed into 10 different layers. Then, using hill climbing algorithm, the best features were selected and classified accordingly. In both studies, the SVM algorithm performed a more accurate classification compared to the other algorithms. The better performance of SVM than other algorithms has also been reported in some other articles [54, 58]. Although the SVM algorithm was highly accurate in the study of Sharma et al. [55], the highest classification accuracy was observed using the LDA algorithm. In this study, classification was performed based on energy, entropy and standard deviation characteristics in HRV and EDR. Overall, existing research works have shown that classifying data per record is a very good way to diagnose and separate SA patients from healthy individuals. Classification using full ECG tape seems to provide more accurate diagnoses of SA than ECG components. However, this does not mean that segmentation per segment performs poorly.

Unlike per record data classification, none of the classifications provided based on ECG components had 100% accuracy, sensitivity and specificity. However, a review of all the presented results shows that more than 80% of the algorithms applied to the ECG components have a high accuracy that is in the range of 70% to 99%. This indicates that the use of small ECG components also performs very well in the diagnosis and classification of SA. Our review also highlighted that the best algorithms in segment classification were Neural Network, RF and SVM. Erdenebayar et al. [30] used neural network algorithms in their work. The results of their study showed that the best algorithm in segmentation is RNN. DNN and CNN algorithms were also highly accurate. Such high accuracies were also observed in a number of other studies [20, 37, 39, 40, 60, 66].

In general, the results of selected studies indicate that ML techniques are successful in diagnosing SA. Unlike the present study, which only diagnoses SA from binary healthy individuals, some studies have classified patients with SA based on the severity of the disease, the results of which again showed the high accuracy of ML models. Another approach was to evaluate the ability of models to detect the exact value of Apnea Hypopnea Index (AHI), and the results of these research works were also of acceptable accuracy [89].

With these interpretations, ML still faces challenges that may affect the results obtained. One of the issues with ML algorithms is that they use random models to train their data. This means that if the same model is retrained with the same data, different values of the parameters may be reported. In other words, the reproducibility of the models is one of the issues that should be considered [90].

Another factor that may affect the results of this study is the possibility of bias. In one piece of research, the possibility of unintended racial bias within healthcare algorithms has been raised [91]. A 2015 study was conducted to investigate the impact of ethnicity on the prognosis of cardiovascular disease. The results of this study showed that predicting the risk of cardiovascular disease for non-whites was associated with bias. In fact, the risk of cardiovascular disease for this group of people was reported either more or less than normal [91, 92]. Furthermore, when different causes of a disease are provided, ML techniques will be unable to diagnose the cause of the disease [83]. Also, most of the selected studies in our review used fixed datasets, and ethnic and genetic differences were not measured. This issue can cause challenges in using ML methods in diagnosing SA.

***Limitations***

One of the limitations of this study is that most studies had used the same datasets. This made it impossible to review data from the same area, so the impact of environmental and individual factors on the incidence of SA was not measured. The small samples used in these datasets is another factor that can affect the generalizability of the results. Another limitation observed was the lack of reporting true positive, true negative, false positive and false negative in most studies. Accordingly, it was not possible to perform meta-analysis on the data.

**Conclusion**

Studies have confirmed the effectiveness of ML techniques in diagnosing SA. Since neurological, hormonal, and respiratory changes are quite effective on the ECG, HRV and EDR were the most common ECG features extracted for classification. The features of PSD waves obtained from ECG analysis appear to be very useful in the diagnosis of SA. It was also observed that SVM and Neural Network algorithms are highly accurate in detecting SA. Additionally, other ML techniques such as KNN, RF, and LR performed well in classifying SA-related data. There was no significant difference between the parameters related to the classification based on the complete ECG record and ECG components. However, the ML performance seems to have been better in the full record classification. In future studies, more up-to-date datasets can be used to classify SA. Furthermore, datasets with higher number of records, and the use of samples in various geographical areas are other things that can be examined.

**References**

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM: **Increased prevalence of sleep-disordered breathing in adults**. *American journal of epidemiology* 2013, **177**(9):1006-1014.

2. Dempsey JA, Veasey SC, Morgan BJ, O'Donnell CP: **Pathophysiology of Sleep Apnea**. *Physiological Reviews* 2010, **90**(1):47-112.

3. Parati G, Lombardi C, Narkiewicz K: **Sleep apnea: epidemiology, pathophysiology, and relation to cardiovascular risk**. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 2007, **293**(4):R1671-R1683.

4. Veasey SC, Rosen IM: **Obstructive Sleep Apnea in Adults**. *New England Journal of Medicine* 2019, **380**(15):1442-1449.

5. Young T, Skatrud J, Peppard PE: **Risk Factors for Obstructive Sleep Apnea in Adults**. *JAMA* 2004, **291**(16):2013-2016.

6. Kainulainen S, Töyräs J, Oksenberg A, Korkalainen H, Sefa S, Kulkas A, Leppänen T: **Severity of desaturations reflects OSA-related daytime sleepiness better than AHI**. *Journal of Clinical Sleep Medicine* 2019, **15**(8):1135-1142.

7. Ferini-Strambi L, Baietto C, Di Gioia M, Castaldi P, Castronovo C, Zucconi M, Cappa S: **Cognitive dysfunction in patients with obstructive sleep apnea (OSA): partial reversibility after continuous positive airway pressure (CPAP)**. *Brain research bulletin* 2003, **61**(1):87-92.

8. Torres G, Sánchez-de-la-Torre M, Barbé F: **Relationship between OSA and hypertension**. *Chest* 2015, **148**(3):824-832.

9. Torres-Alba D, Gemma D, Armada-Romero E, Rey-Blas JR, López-de-Sá E, López-Sendon JL: **Obstructive sleep apnea and coronary artery disease: from pathophysiology to clinical implications**. *Pulmonary medicine* 2013, **2013**.

10. Wang H, Parker JD, Newton GE, Floras JS, Mak S, Chiu K-L, Ruttanaumpawan P, Tomlinson G, Bradley TD: **Influence of obstructive sleep apnea on mortality in patients with heart failure**. *Journal of the American College of Cardiology* 2007, **49**(15):1625-1631.

11. Dyken ME, Im KB: **Obstructive sleep apnea and stroke**. *Chest* 2009, **136**(6):1668-1677.

12. Kendzerska T, Gershon AS, Hawker G, Tomlinson G, Leung RS: **Obstructive sleep apnea and incident diabetes. A historical cohort study**. *American journal of respiratory and critical care medicine* 2014, **190**(2):218-225.

13. Rundo JV, Downey R: **Chapter 25 - Polysomnography**. In: *Handbook of Clinical Neurology. Volume 160*, edn. Edited by Levin KH, Chauvel P: Elsevier; 2019: 381-392.

14. Gottlieb DJ, Punjabi NM: **Diagnosis and Management of Obstructive Sleep Apnea: A Review**. *JAMA* 2020, **323**(14):1389-1400.

15. Ali SQ, Khalid S, Belhaouari SB: **A Novel Technique to Diagnose Sleep Apnea in Suspected Patients Using Their ECG Data**. *IEEE Access* 2019, **7**:35184-35194.

16. Portier F, Portmann A, Czernichow P, Vascaut L, Devin E, Benhamou D, Cuvelier A, Muir JF: **Evaluation of home versus laboratory polysomnography in the diagnosis of sleep apnea syndrome**. *American journal of respiratory and critical care medicine* 2000, **162**(3):814-818.

17. Pombo N, Silva BMC, Pinho AM, Garcia N: **Classifier Precision Analysis for Sleep Apnea Detection Using ECG Signals**. *IEEE Access* 2020, **8**:200477-200485.

18. Wang T, Lu CH, Shen GH: **Detection of Sleep Apnea from Single-Lead ECG Signal Using a Time Window Artificial Neural Network**. *BioMed research international* 2019, **2019**.

19. Bozkurt S, Bostanci A, Turhan M: **Can Statistical Machine Learning Algorithms Help for Classification of Obstructive Sleep Apnea Severity to Optimal Utilization of Polysomnography Resources?** *Methods of information in medicine* 2017, **56**(4):308-318.

20. Wang T, Lu CH, Shen GH, Hong F: **Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network**. *PeerJ* 2019, **7**.

21. Kesper K, Canisius S, Penzel T, Ploch T, Cassel W: **ECG signal analysis for the assessment of sleep-disordered breathing and sleep pattern**. *Medical & Biological Engineering & Computing* 2012, **50**(2):135-144.

22. Hayano J, Watanabe E, Saito Y, Sasaki F, Fujimoto K, Nomiyama T, Kawai K, Kodama I, Sakakibara H: **Screening for obstructive sleep apnea by cyclic variation of heart rate**. *Circ Arrhythmia Electrophysiol* 2011, **4**(1):64-72.

23. Janbakhshi P, Shamsollahi M: **Sleep apnea detection from single-lead ECG using features based on ECG-derived respiration (EDR) signals**. *Irbm* 2018, **39**(3):206-218.

24. Bailón R, Sornmo L, Laguna P: **A robust method for ECG-based estimation of the respiratory frequency during stress testing**. *IEEE transactions on biomedical engineering* 2006, **53**(7):1273-1285.

25. FAUST O, ACHARYA UR, NG EYK, FUJITA H: **A REVIEW OF ECG-BASED DIAGNOSIS SUPPORT SYSTEMS FOR OBSTRUCTIVE SLEEP APNEA**. *Journal of Mechanics in Medicine and Biology* 2016, **16**(01):1640004.

26. Li M, Zhou Z: **Improve Computer-Aided Diagnosis With Machine Learning Techniques Using Undiagnosed Samples**. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 2007, **37**(6):1088-1098.

27. El Naqa I, Murphy MJ: **What is machine learning?** In: *machine learning in radiation oncology.* edn.: Springer; 2015: 3-11.

28. Dey A: **Machine learning algorithms: a review**. *International Journal of Computer Science and Information Technologies* 2016, **7**(3):1174-1179.

29. Bozkurt F, Ucar MK, Bozkurt MR, Bilgin C: **Detection of Abnormal Respiratory Events with Single Channel ECG and Hybrid Machine Learning Model in Patients with Obstructive Sleep Apnea**. *Irbm* 2020, **41**(5):241-251.

30. Erdenebayar U, Kim YJ, Park JU, Joo EY, Lee KJ: **Deep learning approaches for automatic detection of sleep apnea events from an electrocardiogram**. *Computer methods and programs in biomedicine* 2019, **180**:105001.

31. Nguyen HD, Wilkins BA, Cheng Q, Benjamin BA: **An online sleep apnea detection method based on recurrence quantification analysis**. *IEEE journal of biomedical and health informatics* 2014, **18**(4):1285-1293.

32. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement**. *PLoS medicine* 2009, **6**(7):e1000097.

33. Atri R, Mohebbi M: **Obstructive sleep apnea detection using spectrum and bispectrum analysis of single-lead ECG signal**. *Physiological measurement* 2015, **36**(9):1963-1980.

34. Babaeizadeh S, White DP, Pittman SD, Zhou SH: **Automatic detection and quantification of sleep apnea using heart rate variability**. *Journal of electrocardiology* 2010, **43**(6):535-541.

35. Bali J, Nandi A, Hiremath PS, Patil PG: **Detection of sleep apnea in ecg signal using pan-tompkins algorithm and ann classifiers**. *Compusoft* 2018, **7**(11):2852-2861.

36. Bsoul M, Minn H, Tamil L: **Apnea MedAssist: Real-time Sleep Apnea Monitor Using Single-Lead ECG**. *Ieee Transactions on Information Technology in Biomedicine* 2011, **15**(3):416-427.

37. Chang HY, Yeh CY, Lee CT, Lin CC: **A Sleep Apnea Detection System Based on a One-Dimensional Deep Convolution Neural Network Model Using Single-Lead Electrocardiogram**. *Sensors (Basel, Switzerland)* 2020, **20**(15).

38. Chen L, Zhang X, Song C: **An automatic screening approach for obstructive sleep apnea diagnosis based on single-lead electrocardiogram**. *IEEE Trans Autom Sci Eng* 2015, **12**(1):106-115.

39. Dey D, Chaudhuri S, Munshi S: **Obstructive sleep apnoea detection using convolutional neural network based deep learning framework**. *Biomedical Engineering Letters* 2018, **8**(1):95-100.

40. Farouk FNBM, Anwar T, Zakaria NB: **Hybrid bayesian network in neural network based deep learning framework for detection of obstructive sleep apnea syndrome**. *Int J Eng Adv Technol* 2019, **9**(1):4922-4926.

41. Fatimah B, Singh P, Singhal A, Pachori RB: **Detection of apnea events from ECG segments using Fourier decomposition method**. *Biomedical Signal Processing and Control* 2020, **61**.

42. Hassan AR: **Computer-aided obstructive sleep apnea detection using normal inverse Gaussian parameters and adaptive boosting**. *Biomedical Signal Processing and Control* 2016, **29**:22-30.

43. Hassan AR, Haque MA: **Computer-aided obstructive sleep apnea identification using statistical features in the emd domain and extreme learning machine**. *Biomed Phys Eng Express* 2016, **2**(3).

44. Hassan AR, Haque MA: **Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating**. *Biocybern Biomed Eng* 2016, **36**(1):256-266.

45. Jafari A: **Sleep apnoea detection from ECG using features extracted from reconstructed phase space and frequency domain**. *Biomedical Signal Processing and Control* 2013, **8**(6):551-558.

46. Khandoker AH, Karmakar CK, Palaniswami M: **Automated recognition of patients with obstructive sleep apnoea using wavelet-based features of electrocardiogram recordings**. *Computers in biology and medicine* 2009, **39**(1):88-96.

47. Li K, Pan W, Li Y, Jiang Q, Liu G: **A method to detect sleep apnea based on deep neural network and hidden Markov model using single-lead ECG signal**. *Neurocomputing* 2018, **294**:94-101.

48. Mendez MO, Bianchi AM, Matteucci M, Cerutti S, Penzel T: **Sleep apnea screening by autoregressive models from a single ECG lead**. *IEEE transactions on bio-medical engineering* 2009, **56**(12):2838-2850.

49. Mendez MO, Ruini DD, Villantieri OP, Matteucci M, Penzel T, Cerutti S, Bianchi AM: **Detection of sleep apnea from surface ECG based on features extracted by an autoregressive model**. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference* 2007, **2007**:6106-6109.

50. Nakayama C, Fujiwara K, Sumi Y, Matsuo M, Kano M, Kadotani H: **Obstructive sleep apnea screening by heart rate variability-based apnea/normal respiration discriminant model**. *Physiological measurement* 2019, **40**(12):125001.

51. Nishad A, Pachori RB, Acharya UR: **Application of TQWT based filter-bank for sleep apnea screening using ECG signals**. *Journal of Ambient Intelligence and Humanized Computing* 2018:1-12.

52. Pinho A, Pombo N, Silva BMC, Bousson K, Garcia N: **Towards an accurate sleep apnea detection based on ECG signal: The quintessential of a wise feature selection**. *Applied Soft Computing* 2019, **83**.

53. Rachim VP, Li G, Chung WY: **Sleep apnea classification using ECG-signal wavelet-PCA features**. *Bio-medical materials and engineering* 2014, **24**(6):2875-2882.

54. Sharma H, Sharma KK: **An algorithm for sleep apnea detection from single-lead ECG using Hermite basis functions**. *Computers in biology and medicine* 2016, **77**:116-124.

55. Sharma H, Sharma KK: **Sleep apnea detection from ECG using variational mode decomposition**. *Biomed Phys Eng Express* 2020, **6**(1).

56. Sharma M, Agarwal S, Acharya UR: **Application of an optimal class of antisymmetric wavelet filter banks for obstructive sleep apnea diagnosis using ECG signals**. *Computers in biology and medicine* 2018, **100**:100-113.

57. Sharma M, Raval M, Acharya UR: **A new approach to identify obstructive sleep apnea using an optimal orthogonal wavelet filter bank with ECG signals**. *Inform Med Unlocked* 2019, **16**.

58. Song C, Liu K, Zhang X, Chen L, Xian X: **An Obstructive Sleep Apnea Detection Approach Using a Discriminative Hidden Markov Model From ECG Signals**. *IEEE transactions on bio-medical engineering* 2016, **63**(7):1532-1542.

59. Tripathy RK: **Application of intrinsic band function technique for automated detection of sleep apnea using HRV and EDR signals**. *Biocybern Biomed Eng* 2018, **38**(1):136-144.

60. Wang L, Lin YF, Wang J: **A RR interval based automated apnea detection approach using residual network**. *Computer methods and programs in biomedicine* 2019, **176**:93-104.

61. Wang XW, Cheng MW, Wang YF, Liu SH, Tian ZH, Jiang F, Zhang HJ: **Obstructive sleep apnea detection using ecg-sensor with convolutional neural networks**. *Multimedia Tools and Applications* 2020, **79**(23-24):15813-15827.

62. Yildiz A, Akin M, Poyraz M: **An expert system for automated recognition of patients with obstructive sleep apnea using electrocardiogram recordings**. *Expert Systems with Applications* 2011, **38**(10):12880-12890.

63. Penzel T, Moody GB, Mark RG, Goldberger A, Peter JH: **The apnea-ECG database**, vol. 27; 2000.

64. Akşahin M, Erdamar A, Firat H, ArdIç S, Eroʇul O: **Obstructive sleep apnea classification with artificial neural network based on two synchronic HRV series**. *Biomed Eng Appl Basis Commun* 2015, **27**(2).

65. Ali SQ, Hossen A: **Identification of obstructive sleep apnea using artificial neural networks and wavelet packet decomposition of the HRV signal**. *J Eng Res (Oman)* 2020, **17**(1):24-33.

66. Lweesy K, Fraiwan L, Khasawneh N, Dickhaus H: **New automated detection method of OSA based on artificial neural networks using P-wave shape and time changes**. *Journal of medical systems* 2011, **35**(4):723-734.

67. Al-Angari HM, Sahakian AV: **Automated Recognition of Obstructive Sleep Apnea Syndrome Using Support Vector Machine Classifier**. *Ieee Transactions on Information Technology in Biomedicine* 2012, **16**(3):463-468.

68. Eiseman NA, Westover MB, Mietus JE, Thomas RJ, Bianchi MT: **Classification algorithms for predicting sleepiness and sleep apnea severity**. *J Sleep Res* 2012, **21**(1):101-112.

69. Hassan AR, Haque MA: **An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting**. *Neurocomputing* 2017, **235**:122-130.

70. Rekha BB, Kandaswamy A, Ramanathan RMPL: **Ensemble classification approach for screening of obstructive sleep apnoea using ECG**. *International Journal of Biomedical Engineering and Technology* 2018, **27**(1-2):139-150.

71. Travieso CM, Alonso JB, del Pozo M, Ticay JR, Castellanos-Dominguez G: **Building a Cepstrum-HMM kernel for Apnea identification**. *Neurocomputing* 2014, **132**:159-165.

72. Varon C, Caicedo A, Testelmans D, Buyse B, Van Huffel S: **A Novel Algorithm for the Automatic Detection of Sleep Apnea From Single-Lead ECG**. *IEEE transactions on bio-medical engineering* 2015, **62**(9):2269-2278.

73. Khandoker AH, Palaniswami M, Karmakar CK: **Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings**. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society* 2009, **13**(1):37-48.

74. Baek JW, Kim YN, Kim DE, Lee JH: **Computer-aided detection with a portable electrocardiographic recorder and acceleration sensors for monitoring obstructive sleep apnea**. *Sensors Transducers* 2014, **167**(3):80-87.

75. Smruthy A, Suchetha M: **Real-Time Classification of Healthy and Apnea Subjects Using ECG Signals With Variational Mode Decomposition**. *IEEE Sensors Journal* 2017, **17**(10):3092-3099.

76. Fariha MAZ, Ikeura R, Hayakawa S, Tsutsumi S: **Analysis of Pan-Tompkins Algorithm Performance with Noisy ECG Signals**. *Journal of Physics: Conference Series* 2020, **1532**:012022.

77. Kaguara A, Nam K, Reddy S: **A deep neural network classifier for diagnosing sleep apnea from ECG data on smartphones and small embedded systems**. 2014.

78. Jordan MI, Mitchell TM: **Machine learning: Trends, perspectives, and prospects**. *Science* 2015, **349**(6245):255.

79. Larrañaga P, Calvo B, Santana R, Bielza C, Galdiano J, Inza I, Lozano JA, Armañanzas R, Santafé G, Pérez A *et al*: **Machine learning in bioinformatics**. *Briefings in Bioinformatics* 2006, **7**(1):86-112.

80. Alzubi J, Nayyar A, Kumar A: **Machine Learning from Theory to Algorithms: An Overview**. *Journal of Physics: Conference Series* 2018, **1142**:012012.

81. Becker H: **Computing with words and machine learning in medical diagnostics**. *Information Sciences* 2001, **134**(1):53-69.

82. Kolachalama VB, Garg PS: **Machine learning and medical education**. *NPJ digital medicine* 2018, **1**(1):54.

83. Richens JG, Lee CM, Johri S: **Improving the accuracy of medical diagnosis with causal machine learning**. *Nature Communications* 2020, **11**(1):3923.

84. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H: **eDoctor: machine learning and the future of medicine**. *Journal of Internal Medicine* 2018, **284**(6):603-619.

85. Maity NG, Das S: **Machine learning for improved diagnosis and prognosis in healthcare**. In: *2017 IEEE Aerospace Conference: 4-11 March 2017 2017*; 2017: 1-9.

86. Bušek P, Vaňková J, Opavský J, Salinger J, Nevšímalová S: **Spectral analysis of heart rate variability in sleep**. *Physiol res* 2005, **54**(4):369-376.

87. Moody GB, Mark RG, Bump MA, Weinstein JS, Berman AD, Mietus JE, Goldberger AL: **Clinical validation of the ECG-derived respiration (EDR) technique**. *Computers in cardiology* 1986, **13**(1):507-510.

88. Foster KR, Koprowski R, Skufca JD: **Machine learning, medical diagnosis, and biomedical engineering research - commentary**. *Biomedical engineering online* 2014, **13**(1):94.

89. Mencar C, Gallo C, Mantero M, Tarsia P, Carpagnano GE, Foschino Barbaro MP, Lacedonia D: **Application of machine learning to predict obstructive sleep apnea syndrome severity**. *Health Informatics Journal* 2019, **26**(1):298-317.

90. Beam AL, Manrai AK, Ghassemi M: **Challenges to the Reproducibility of Machine Learning Models in Health Care**. *JAMA* 2020, **323**(4):305-306.

91. Char DS, Shah NH, Magnus D: **Implementing Machine Learning in Health Care - Addressing Ethical Challenges**. *N Engl J Med* 2018, **378**(11):981-983.

92. Gijsberts CM, Groenewegen KA, Hoefer IE, Eijkemans MJ, Asselbergs FW, Anderson TJ, Britton AR, Dekker JM, Engström G, Evans GW: **Race/ethnic differences in the associations of the Framingham risk factors with carotid IMT and cardiovascular events**. *PLoS One* 2015, **10**(7):e0132321.