An improved Machine Learning Approach to Classify Sleep Stages and Apnea Events

Swapna PREMASIRI[1], Jayasanka RANAWEERA[1], Lalith B. GAMAGE[2,1], Clarence W. DE SILVA[1]
([1]*The University of British Columbia, Vancouver Canada*;
[2]*Sri Lanka Institute of Information Technology, Malabe, Sri Lanka*)

Abstract: Sleep apnea (SA) is a common sleep disorder. Identifying patients at risk by means of comprehensive monitoring that requires overnight stay at professional sleep clinics are costly and inconvenient and can lead to unreliable results in view of the unfamiliar sleep environment. Existing wearable devices for sleep monitoring, which can be used in a familiar home environment, do not provide the same comprehensive monitoring as through clinical monitoring. The larger objective of the present work is to develop a sleep monitoring system for home use, which can provide comprehensive monitoring. In the development in this paper, machine learning (ML) models are explored for the classification of SA and sleep stages using multisensory data, without neglecting any of the required signals. The data acquired through the sensors are normalized, their features are extracted using Composite Multiscale Sample Entropy (CMSE) and are standardized using a robust scaling algorithm. Processed features are classified using a Neural Network (NN) and the obtained results for the SA classification are compared with those obtained by using a Support Vector Machine (SVM) approach. The impact of neglecting signals when classifying sleep stages is analyzed as well. The results are presented in the paper and observations are made. The NN model trained with the Bayesian regularization algorithm has provided an overall average accuracy of 94.5% and performed slightly better than when trained using the scaled conjugate gradient backpropagation algorithm (93.2%). The SVMs have yielded lower accuracy levels compared to the NNs (<92%). It is observed that the use of all 14 signals for SS classification yields an overall test accuracy of 72.3%, which is higher than that when one or few signals are used. It is concluded that ML models are effective in classifying sleep data from multiple sensors. Accuracy levels are higher when fused multisensory data are used as inputs. Furthermore, NN models are found to be better suitable in practical application and can be incorporated into an inexpensive and convenient wearable device that can carry out comprehensive monitoring.

Key words: Sleep apnea, Sleep stages, Machine learning, Neural networks, Sensor fusion.

1 Introduction

Sleep Apnea (SA) is recognized as a common form of sleep disorder, which is related to abnormal breathing, since over 20% of the adult population suffers from this disorder (Sweetman, et al., 2017) (Magnusdottir & Hilmisson, 2018). Sleep, which leads to a restricted state of awareness that is essential for a healthy life, can be seriously affected by disorders such as sleep apnea. Sleep disorders may result in daytime sleepiness, fatigue, and degraded quality of life. Prolonged affliction may increase the risks of cardiovascular and metabolic diseases and stroke. Therefore, sleep monitoring is an important procedure to identify the presence of any form of sleep apnea.

Clinically, four main categories are used to identify apneic events, namely: No apneic event (NE), Central Sleep Apnea (CSA), Obstructive Sleep Apnea (OSA), andHypopnea (HYP) where symptoms of both CSA and OSA persist (Sánchez-de-la-Torre, et al., 2013) (Hawkins, 2015) (Javaheri, 2010). Identification of the sleep stages of a person is necessary to determine their quality of sleep since this may be used as an indicator to identify symptoms of sleep apnea and other underlying diseases. In order to identify the stages of sleep or the severity of SA, the current gold standard is laboratory polysomnography (PSG), which follows the standards of the American Academy of Sleep Medicine (AASM) (Berry, et al., 2017). However, PSG typically involves multiple visits to a sleep clinic and overnight monitoring. Such monitoring in an unfamiliar environment to the patient may not generate accurate data, which may lead to misdiagnosis, and also is not economically favourable. Portable devices are available such as (Sharkey, et al., 2014) (Chen, et al., 2013) (Leth, et al., 2017) (Hao, et al., 2013) for sleep monitoring in a household environment, as discussed in (Premasiri & Clarence W. de Silva, 2018). However, they cannot provide comprehensive monitoring and do not comply with the standards of the AASM. Therefore, it is recognized that there is a need for an inexpensive wearable device that can provide comprehensive monitoring at the accuracy levels of laboratory PSG.

The focus of the present paper is to develop an improved versionof (Premasiri & Clarence W. de Silva, 2018), which is an automated and enhanced sleep scoring system to classify apneic events for all types of signals used in laboratory PSG, and a model for the classification of sleep stages. This technology will be incorporated in to the wearable device, which is being developed by our group. In the present work, priority is given to achieving a high accuracy level as that of laboratory PSG and complying with the standards of the AASM.

The present paper uses neural networks (NNs) for data classification with supervised learning and compares the resulting accuracy levels when different learning algorithms are used in the model. An NN will allow the system to function through learning the classifications as done by experts. The most suitable NN model is compared with another commonly used classification technique, support vector machines (SVMs), in order to assess their ability to achieve the accuracy levels of laboratory PSG. These techniques are presented next.

2 Methodology

This section presents the steps followed in the diagnosis of apneic events and the classification of sleep stages. An overview of the process is shown in Figure 1.

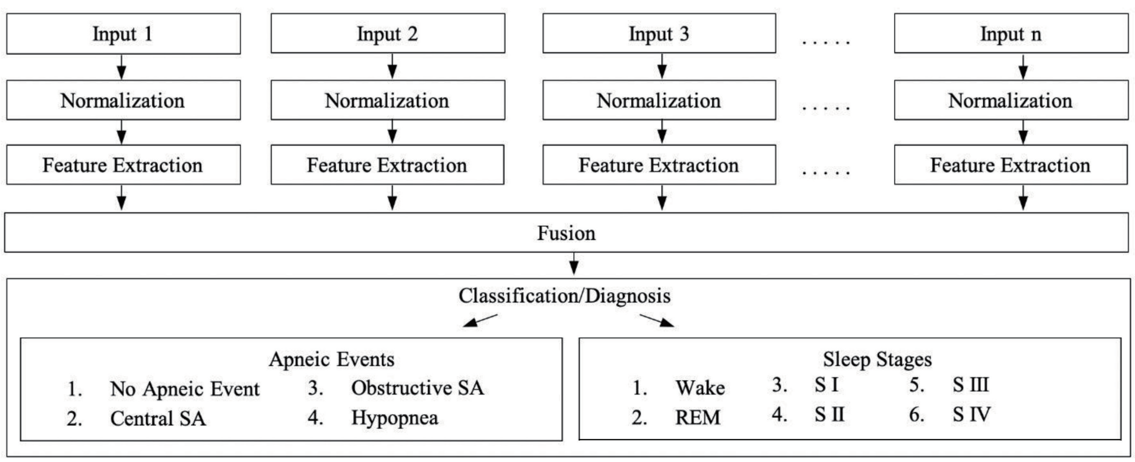

Fig. 1 Classification of the fused data from all signals.

As shown, input data that have been externally filtered, are normalized to reduce or eliminate any redundant data. Features are extracted from the filtered signals for subsequent use because too much information can undermine the effectiveness of classification. Specifically, meaningful information is extracted, which retain the key information. Fusion will allow the incorporation of information from multiple sensors to produce more consistent and accurate information. In the present application, feature level fusion is performed prior to classification. In the present work, two classification systems are used:

(1) Apneic event classification

(2) Classification of sleep stages

Each stage is discussed in detail in the subsections that follow.

According to the scoring techniques of the AASM[6], scoring is performed manually by experts for 30 second intervals of an overnight data record of a patient. To be consistent, in the automated system of the present work, the pre-processing and diagnosis are performed for the same time interval.

The data samples used in this work are the overnight polysomnography recordings provided to the Physionet database of the Massachusetts Institute of Technology (MIT) (St. Vincent´s University Hospital/University College Dublin, 2011), USA, by St. Vincent's University Hospital/University College, Dublin and Massachusetts General Hospital's (MGH) Computational Clinical Neurophysiology Laboratory (CCNL) (Massachusetts General Hospital´s (MGH), 2018) and the Clinical Data Animation Laboratory (CDAC). This database contains more patient data than in the previous database. The database is labeled only for sleep stages and arousal regions, which is not suitable for classifying sleep apneic events. Normalization, feature extraction and standardization are essential pre-processing techniques as presented in (Premasiri & Clarence W. de Silva, 2018). These methodologies are presented now.

2.1 Normalization

The original time series containing raw data are completely disparate, and their ranges vary. Therefore, prior to feature extraction, they are brought to a common basis. This normalization of the (externally) filtered data, reduces, and in some cases eliminates, the data redundancy. This is performed with respect to the mean and the standard deviation of a signal for the data points from 1 to t (here t is the number of data points in the time series), as,

$$Signal_{Normalized}(t) = \frac{Signal(t) - \mu}{\sigma} \tag{1}$$

where μ is the mean of the original time series, signal (at time t) and σ is the standard deviation of the signal.

2.2 Feature extraction

Feature extraction allows the mining of useful information without impairing key information of a signal, as needed in a classification model. Sleep disorder monitoring involves information derived from several sensors over time. Hence, multivariate time series have to be taken into account in the process of information fusion. Although the signals comprise repetitive patterns, there may exist unexpected events depending on the subject who is being monitored and their medical conditions; specifically, the signals are time- and person-variant. Also, presenting a system with original signals (time series) as the inputs for processing or classification is undesirable since excess information will reduce the effectiveness of the data. It follows that feature extraction is necessary. Inspired by the work in (Wu, et al., 2013) and (Begum, et al., 2014), it was considered that composite multiscale sample entropy (CMSE) is a suitable approach for feature extraction, specifically for classifying apneic events in the present biomedical application. CMSE (Premasiri & Clarence W. de Silva, 2018) (Wu, et al., 2013) involves coarse graining of the acquired signals according to,

$$CS_j = \frac{1}{S} \sum_{i=(j-1)S+1}^{jS} E_i \tag{2}$$

where C is the coarse-grained time series (CGTS) with scale factor S, E is the original time series element, and the indices (positions) of the original and CGTS are denoted by i and j, with $1 \leq j \leq n/S$ where n is the total number of elements in the original time series.

In the present work it is empirically established that S varies from 1 to 10 (Premasiri & Clarence W. de Silva, 2018). This process is followed by calculating the sample entropy value for each coarse-grained time series using

$$MSampEn(CS) = -ln\left(\frac{Am + 1(r)}{Am(r)}\right) \quad (3)$$

Here, C^S is the CGTS with respect to S, r is the threshold, and A^m and A^{m+1} are the number of repetitive sequences for templates with m and $m+1$ elements, respectively.

2.3 Scaling the extracted features

The step of scaling is desirable particularly in the presence of outliers. This is because the distributions are brought to the same scale and are also made to overlap such that outliers will not be present within the range of the bulk of the new distributions in the input data (Sarkar, et al., 2018) (Hackeling, 2017). The features of the same standard are fused into one information set, as suitable to be presented into a neural network.

This process is performed using the robust scaling algorithm

$$Input_i = \frac{Feature_i - Q_2(Feature)}{Q_3(Feature) - Q_1(Feature)} \quad (4)$$

This scales each coarse-grained signal feature-wise, given that Q_1, Q_2 and Q_3 are the 25th, 50th and 75th quantiles respectively and $Feature_i$ and $Input_i$ are the elements of the feature vector and the input vector to be fed into the neural network (NN) (Sarkar, et al., 2018) (Hackeling, 2017) (Liu, 2017).

2.4 Detection of Sleep Apnea events

The features extracted from each of the signals are fused and input to a neural network (NN) to diagnose the presence of apneic events and another NN, to identify sleep stages.

The process is performed for intervals of 30 seconds. The NN model used in this work and the training (TrD), validation (ValD) and test (TstD) data are split as presented in (Premasiri & Clarence W. de Silva, 2018). Supervised learning algorithms are used in this application since the data available in the database (Berry, et al., 2017) are scored.

In this work, Bayesian Regularization (BR), which uses a linear combination of squared errors and weights to modify the relationship in such a way that the system reaches a considerably high generalization ability once the training process is complete, is used as the learning algorithm. The algorithm is based on the original work of MacKay is used as the learning algorithm (MacKay, 1992) (Beale, et al., 2010). The results are compared with that obtained from (Premasiri & Clarence W. de Silva, 2018), where scaled conjugate gradient backpropagation (BR) (Møller, 1993) (Beale, et al., 2010) is used as the learning algorithm and the cross entropy loss as the loss function.

2.5 SVMs for Apnea Event Classification

SVMs are commonly used for classification applications (Noble, 2006) in biomedicine such as the work presented in (Hsu, et al., 2016) (Spilka, et al., 2017) (Varon, et al., 2015) for detection of SA. In this backdrop, the classification of apneic events is done using an SVM to compare the results obtained from the finalized NN, implemented for the same task. The features provided to the SVM are the same set of classes is not linearly separable. Hence, Medium Gaussian SVM with the Gaussian kernel is chosen as the classification learner. In order to determine the optimal hyperplanes for separation of the data patterns, results using both one-against-one (OVO) and one-against-all (OVA) approaches are compared to identify the more suitable approach.

2.6 Classification of Sleep Stages

For classification of sleep stages, a model similar to the one used for apnea event classification is used. Also, the BR algorithm is used based on the results from the previous task. The recordings in 30s intervals are classified in to 6 categories, namely;

Wake (W)	Stage II (SII)
Rapid Eye Movement (REM)	Stage III (SIII)
Stage I (SI)	Stage IV (SIV)

(Technically, there are "Indeterminate" and "Artifact" classes according to the annotations given in the database. However, these stages are not present in the available scored data. Thus, these two classes are excluded in the process of classification of sleep stages as performed in the present study).

According to (Berry, et al., 2017), scoring of sleep stages is predominantly based on EEG recordings and occasionally on EOG recordings. Therefore, since it is possible that the classification of sleep stages can be done based only on the features extracted from EEG signals, classification of sleep stages based on features extracted from the following signals are compared:

1 All fourteen signals used in laboratory PSG

2 Both EEG and EOG recordings

3 Only EEG recordings

4 Only EOG Recordings

Finally, the most suitable set of combined features for classification of sleep stages is identified.

3 Results

In this section, the results for each of the three modelsdiscussed in the preceding section are presented and discussed. In each case, the results are demonstrated in a confusion matrix and a table. In each confusion matrix, the lowermost row (gray) presents the prediction accuracy levels; i.e., the true positive and the false negative values. This represents the percentages of accurate classifications for each target class. The diagonal of the confusion matrices (dark blue) shows the percentage that a class is accurately classified. The rightmost column (gray) shows the accurate classification per class. The remaining cells (light blue and light purple), represent the inaccurate classification of output classes for each target class. The bottom right column (white) in each case represents the overall prediction accuracy.

3.1 Detection of Sleep Apnea events

The comparison of results when trained using BR learning algorithm in this work is compared with the results demonstrated in (Premasiri & Clarence W. de Silva, 2018) in order to identify the most suitable learning algorithm for the application of sleep apnea event detection. The results and comparisons are presented in Table 1 and the confusion matrices in Figure 2.

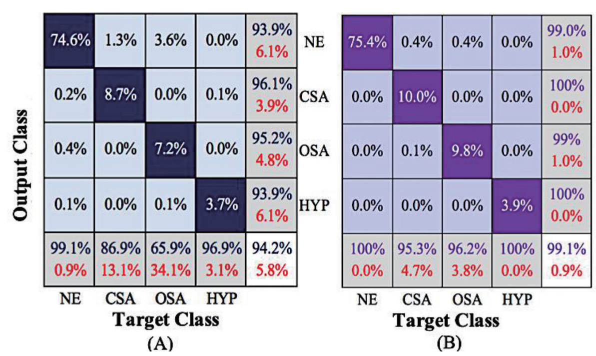

Fig. 2 Confusion Matrices for trained data using: (A) BP and (B) BR algorithms

Table 1 AASM standard compliance: PSG vs. existing wearable sleep monitoring devices

Event	Apneic Event Category								Overall Prediction Accuracy	
	NE		CSA		OSA		HYP			
	BP	BR	BP	BR	BP	BR	BP	BR	BP	BR
TrD	99.1%	100%	86.9%	95.3%	65.9%	96.2%	96.9%	100%	94.2%	99.0%
ValD	98.5%	-	90.0%	-	66.7%	-	100%	-	93.9%	-
TstD	98.5%	95.4%	66.7%	70.0%	76.5%	71.4%	100%	100%	92.2%	90.0%

For each category of SA, it can be seen that accuracy levels of classification for TrD is higher when using BR algorithm.

For TstD, the prediction accuracy when using BR is higher for cases of CSA and HYP, and lower for NE and OSA compared to when the BP algorithm is used. The overall average accuracy for TstD is only slightly different when results from the two algorithms are compared. BR algorithm does not require the use of a validation data set. Therefore, the table will not indicate the percentage values for ValD for the BR algorithm. The overall average accuracy of the final NN is 94.5%. Thus, the final system will be implemented with the NN trained using the Bayesian regularization algorithm.

For each category of SA, it can be seen that accuracy levels of classification forTrD is higher when using BR algorithm.

SVMs for Apnea Event Classification

The input features are fed in to the model and Medium Gaussian SVM with the Gaussian kernel classification learner is used for classification. The OVO and OVA approaches used for classification of data are compared as shown in Table 2 and confusion matrices in Figure 3.

Table 2 SVM results for apneic event classification using fused data

Event	Apneic Event Category								Overall Prediction Accuracy	
	NE		CSA		OSA		HYP			
	BP	BR	BP	BR	BP	BR	BP	BR	BP	BR
TrD	93.6%	93.1%	100%	100%	100%	98.6%	100%	88.1%	94.4%	93.9%
TstD	74.8%	95.4%	0.0%	70.0%	0.0%	71.4%	0.0%	100%	74.8%	90.0%

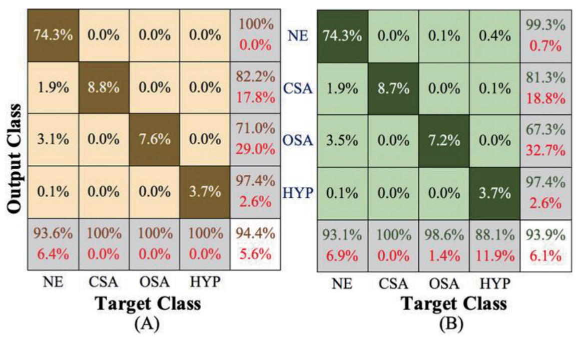

Fig. 3 Confusion matrices for SVM trained data; (A) OVO (B) OVA approach

It was observed that the SVM using OVO approach yielded slightly lower prediction accuracy levels compared to the SVM that OVA approach for training.

However, when the system was tested on the test data, the OVO approach did not seem to have the ability to classify all the classes efficiently; i.e., the SVM that used the OVO approach did not have the ability to classify three of the four classes at all. Thus, the approach for the SVM was chosen as OVA, since it demonstrated a prediction accuracy level of 90% once tested on the training dataset. The overall average accuracy of the finalized SVM is 91.2%. This is also consistent with the conclusion derived from the work done by Milgram et al. (Milgram, et al., 2006). When comparing the overall average classification accuracies with that of the NN model, it could be concluded that using a NN for classification of apneic events is suitable.

3.2 Detection of Sleep Stages

As mentioned in the Methodology section, two databases are used in the present work, for the detection of the sleep stages. Results obtained from each data set are presented separately in Tables 3 and 4. The data presented in Table 3 and Table 4 represents databases with 25 and 1000 patients, respectively.

As shown in Tables 3 and 4, the test data accuracies have been computed for all signals, notably, EEG signal, EOG signals, and both EOG and EEG signals. According to the results presented in the tables, it is observed that the highest overall prediction accuracy of 72.3% were obtained for all signals when compared to the cases of individual signals, EEG and EOG (with accuracies 64.9 and 53.7, respectively). The combination of EEG and EOG signals has resulted in a prediction accuracy of 66.3%. Based on these results, it is clear that in order to achieve a satisfactory performance only the EEG and EOG channels should be used at the final stage of designing a device, which requires optimization of selecting the number of channels.

However, an accuracy rate of 72.3 % cannot be considered as a substantial level of accuracy when compared with other existing models, as shown in

Table 3 Comparison of classification of sleep stages for different feature types

Event	Apneic Event Category						Overall Prediction Accuracy	
	OVO	OVA	OVO	OVA	OVO	OVA	TrD	TstD
All	98.7%	99.3%	99.4%	99.5%	96.8%	96.4%	98.9%	72.3%
EEG	98.6%	94.6%	93.1%	97.9%	80.6%	95.6%	96.1%	64.9%
EOG	96.1%	95.6%	93.9%	97.2%	85.7%	93.7%	95.5%	53.7%
EEG & EOG	98.0%	99.0%	96.2%	99.2%	96.8%	99.4%	98.3%	66.3%

Table 4 Comparison of accuracy levels of apneic event classification systems

Classification System	Overall Average Accuracy	Complies with AASM standards?
Present work	> 94 %	Yes
Premasiri et al (Premasiri & Clarence W. de Silva, 2018)	> 93 %	Yes
Varon et al. (Varon, et al., 2015)	> 85 %	No
Song et al. (Song, et al., 2016)	> 85 %	No
Lin et al. (Lin, et al., 2017)	> 70%	No

Table 5 Comparison of accuracy levels of apneic event classification systems

Proposed by	Classification Model	Signals/Channels Used	Overall Avg. Accuracy
Present work	NN	14 signals	> 85 %
Alickovic et al. (Alickovic & Subasi, 2018)	SVM	1 EEG channel	> 84 %
Gouveris et al. (Gouveris, et al., 2017)	SVM	2 EEG channels	≈ 85 %

Table 5. Hence, it is important to use a further enhanced model in order to achieve higher accuracy levels (specifically, higher than 72.3%). Therefore, it is imperative that further improvements are required to enhance the accuracy of sleep stage classification. Section 3.3 elaborates on the further investigation that is being carried out in this context.

3.3 Feature Analysis

Composite Multiscale Sample Entropy (CMSE) has been used as the feature extraction technique in the present sleep classification system. In view of its apparent shortcomings, it is necessary to further investigate the features extracted from all the patients for sleep stage classification. As shown in Figure 4, scatter plots are used to visualize and analyze the extracted features.

Since multiple features are obtained during the feature extraction, visualizing each feature using a single plot is difficult. Therefore, feature pairs are plotted in each scatter plot and it is attempted to detect which feature pairs are more suitable for clustering

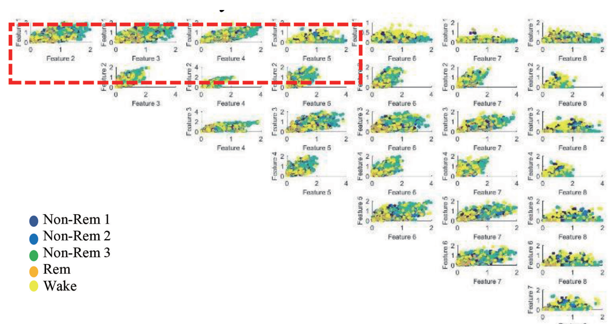

Fig. 4 Feature Scatter Plot.

different sleep stages. This method is used in the field of machine learning to visualize suitable features for use with a neural network. Results highlighted in red are presented in Figure 5. Each axis represents a feature and the sleep stages are denoted by five different colors.

On observing the distribution of the sleep stages, it is clear that they are heavily overlapped, and this has caused difficulty in clustering or categorizing the sleep stages using a basic neural network with a single hidden layer. Feature selection is very important

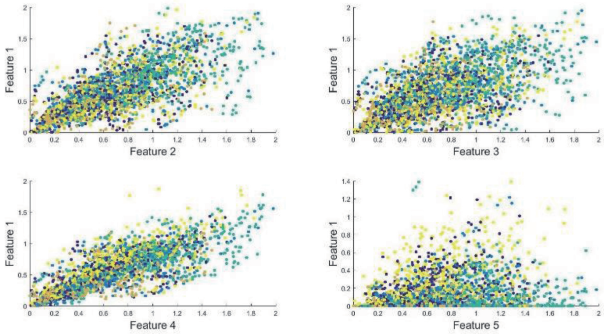

Fig. 5 Feature Scatter Plot (Extracted).

in a classification application. Poor feature selection will lead to poor performance. Therefore, further investigation is needed on the selection of features and enhanced methods will help better feature selection. With an improved feature selection procedure, better accuracies can be obtained from the classifier.

4 Conclusion

In this work, three main tasks have been presented; implementation of a NN for classification of apneic events, comparing the results of the finalized NN with an optimized SVM for the same task, and classification of sleep stages using a NN. Filtered raw monitored signals were normalized, features were extracted using the CMSE algorithm, and the extracted features were standardized prior to being used as inputs to the NN or SVM. For the classification of apneic events using NNs, appropriate analysis and comparison have been made for different learning algorithms (specifically, scaled conjugate gradient back propagation and Bayesian regularization). The results and comparison show that the neural network has a higher prediction accuracy when trained using the Bayesian regularization algorithm rather than when the scaled conjugate gradient backpropagation algorithm is used as the learning algorithm. Hence, the final NN comprises a single hidden layer, while Bayesian regularization is used as the learning algorithm

A Gaussian kernel is used for the SVM for classification of apneic events. It was shown that, the overall average accuracy was higher when the one versus all approach was used, compared to when the one versus one approach was used. This conclusion that is based on results seems to be consistent with other studies that involve similar comparisons such as (Milgram, et al., 2006). However, the finalized NN, using the BR algorithm yielded a higher overall accuracy compared to the SVM. Hence, it was decided that the NN model was more suitable for the task.

Based on the results obtained for the NN designed for classification of sleep stages, it is observed that the accuracy levels are the highest, i.e., 72.3%, when all features are used. However, when features extracted from EEG and EOG signals together are used, the resulting overall average accuracy is only 6% less than the highest accuracy. Hence, given that a situation arises where all channels are not used in the final device that will be implemented, the latter is also an option.

The models used for classification of apneic events and sleep stages are compared with existing models or proposed systems as shown in Table 4 and Table 5 respectively. It can be observed from this comparison that the models presented in this work have higher accuracies predicting apneic events while complying with the standards of AASM. On the other hand, sleep stage classification system has less accuracy levels compaired to the exsisting models. This highlights the importance of further investigation on new features and methods, to improve the current system for classifying sleep stages.

The limitations of this work are mostly based on the limitations of the available data. This is because the database contains scored sleep data of only patients who suffer from sleep apnea as the patient would have visited the sleep clinic for overnight recordings only in the presence of symptoms of sleep apnea. Therefore, recordings from healthy subjects who do not suffer from sleep apnea are not readily available. Also, the occurrence of certain apneic events is uncommon within the database. In such cases, when the data is split into training, validation and test data, the features corresponding to the uncommon classes may not be adequately available within the datasets. This can be overcome by including recordings of more patients.

Future Work

Future research based on this work may be focused on investigating new features and methods to improve the sleep stage classification system. The use of an improved model associated with a health quality index, would provide more meaningful interpretation of the severity of sleep apnea and the quality of sleep. The present work used the open access database, available in (St. Vincent´s University Hospital/University College Dublin, 2011) and (Massachusetts General Hospital´s (MGH), 2018). Therefore, the developed models must be optimized as suited for the signals obtained from the sensors that are being developed for a wearable sleep monitoring device. Thereafter, the accuracy and completeness of the data needs to be estimated and the confidence level of the final decision needs to be interpreted. These tasks maybe achieved by comparing and determining the percentage errors between the outputs by the implemented system with scoring done by several experts of sleep scoring.

ACKNOWLEDGMENT

This work is funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada through the strategic partnership grants project STPGP 493908 "Research in Sensory Information Technologies and Implementation in Sleep Monitoring."

References

[1] Alickovic, E. & Subasi, A., 2018. Ensemble SVM Method for Automatic Sleep Stage Classification. *IEEE Transactions on Insturmentation and Measurement*, June, 62(6), pp. 1258-1265.

[2] Beale, M. H., Hagan, M. T. & Demuth, H. B., 2010.*Neural Network Toolbox 7: User´s Guide*, : Mathworks.

[3] Begum, S., Barua, S. & Ahmed, M. U., 2014. Physiological Sensor Signals Classification for Healthcare Using Sensor Data Fusion and Case-Based Reasoning.*Sensors*, 3 July, 2014(14), pp. 11770-11785.

[4] Berry, R. B. et al., 2017.*The AASM Manual for the Scoring of Sleep Associated Events - Terminology and Technical Specifications*, Darien: American Academy of Sleep Medicine.

[5] Chen, Z. et al., 2013.*Unobtrusive Sleep Monitoring using Smartphones.* Venice, Institute for Computer Science, Social-Informatics and Tlecommunication Engineering (ICST), pp. 145-152.

[6] Gouveris, H. et al., 2017. Sleep stage classification using spectral analyses and support vector machine algorithm on C3- and C4-EEG signals.*Sleep Medicine*, December, 40(1), p. e116.

[7] Hackeling, G., 2017. *Mastering Machine Learning with scikit-learn.* Birmingham: PACKT.

[8] Hao, T., Xing, G. & Zhou, G., 2013.*iSleep: Unobtrusive Sleep Quality Monitoring using Smartphones.* Rome, s.n., pp. 11-15.

[9] Hawkins, T., 2015.*Apnea, Apnea No More: Easy Ways Out Of Sleep.* s.l.:JNR Publishing.

[10] Hsu, W.-C.et al., 2016. EEG Classification of Imaginary Lower Limb Stepping Movements Based on Fuzzy Support Vector Machine with Kernel-Induced Membership Function. *International Journal of Fuzzy Systems*, 19(2), pp. 566-579.

[11] Javaheri, S., 2010. Central Sleep Apnea. *Clinics in Chest Medicine*, June, 31(2), pp. 235-248.

[12] Leth, S., Hansen, J., Nielsen, O. W. & Dinesen, B., 2017. Evaluation of Commercial Self-Monitoring Devicesfor Clinical Purposes: Results from the Future Patient Trial, Phase I. *Sensors*, January, 17(1), pp. 1-11.

[13] Lin, Y.-Y.et al., 2017. Sleep Apnea Detection Based on Thoracic and Abdominal Movement Signals of Wearable Piezoelectric Bands. *IEEE Journal of Biomedical and Health Informatics*, November, 21(6), pp. 1533-1544.

[14] Liu, Y., 2017.*Python Machine Learning By Example.* Birmingham: Packt Publishing.

[15] MacKay, D. J. C., 1992. A Practical Bayesian Framework for backropagation Networks.*Neural Computation*, 4(3), pp. 448-472.

[16] Magnusdottir, S. & Hilmisson, H., 2018. Ambulatory screening tool for sleep apnea: analyzing a single-

lead electrocardiogram signal (ECG). *Sleep Breath*, 2018(22), pp. 421-429.

[17] Massachusetts General Hospital´s (MGH), 2018. *You Snooze You Win - The PhysioNet Computing in Cardiology Challenge* 2018, s.l.: physionet.mit.edu.

[18] Milgram, J., Cheriet, M. & Sabourin, R., 2006. "*One against one*" *or* "*one against all*": *Which one is better for handwriting recognition with SVMs*?. La Baule, Suvisoft.

[19] Møller, M. F., 1993. A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. *Neural Networks*, 6(4), pp. 525-533.

[20] Noble, W. S., 2006. What is a support vector machine?. *Nature Biotechnology*, December, 24(12), pp. 1565-1567.

[21] Premasiri, S. & Clarence W. de Silva, L. B. G., 2018. *A Multi-Sensor Data Fusion Approach for Sleep Apnea Monitoring Using Neural Networks.* Anchorage, s.n.

[22] Sánchez-de-la-Torre, M., Campos-Rodriguez, F. & Barbé, F., 2013. Obstructive sleep apnoea and cardiovascular disease. *The Lancet*: *Respiratory Medicine*, March, 1(1), pp. 61-72.

[23] Sarkar, D., Bali, R. & Sharma, T., 2018. *Practical Machine Learning with Python.* Karnataka: Apress.

[24] Sharkey, K. M. et al., 2014. Validation of the Apnea Risk Evaluation System (ARES) Device Against Laboratory Polysomnography in Pregnant Women at Risk for Obstructive Sleep Apnea Syndrome. *Journal of Clinical Sleep Medicine*, May, 10(5), pp. 497-502.

[25] Song, C. et al., 2016. An Obstructive Sleep Apnea Detection Approach Using a Discriminative Hidden Markov Model From ECG Signals. *IEEE Transactions on Biomedical Engineering*, July, 63(7), pp. 1532-1542.

[26] Spilka, J. et al., 2017. Sparse Support Vector Machine for Intrapartum Fetal Heart Rate Classification. *IEEE Journal of Biomedical and Health Informatics*, 21(3), pp. 664-671.

[27] St. Vincent´s University Hospital/University College Dublin, 2011. *Sleep Apnea Database Sleep Apnea Database*, Belfield: physionet.mit.edu.

[28] Sweetman, A. M. et al., 2017. Developing a successful treatment for co-morbid insomnia and sleepapnoea. *Sleep Medicine Reviews*, 2017(33), pp. 28-38.

[29] Tan, H.-L. et al., 2014. Overnight Polysomnography versus Respiratory Polygraphy in the Diagnosis of Pediatric Obstructive Sleep Apnea. *Sleep*, January, 37(02), pp. 255-261.

[30] Varon, C., Caicedo, A., Testelmans, D. & Huffel, B. B. a. S. V., 2015. A Novel Algorithm for the Automatic Detection of Sleep Apnea From Single-Lead ECG. *IEEE Transactions on Biomedical Engineering*, September, 62(9), pp. 2269-2278.

[31] Wu, S.-D. et al., 2013. Time Series Analysis Using Composite Multiscale Entropy. *Entropy*, 18 March, 15(3), pp. 1069-1084.

Authors' Biographies

Swapna PREMASIRI received the B. Eng. degree in Electronic Engineering and Postgraduate Diploma in Telecommunication and Electronic Engineering from the Sheffield Hallam University in 2011 and 2016, respectively. She completed her MASc degree at the University of British Columbia, Vancouver, Canada, in 2018, under the supervision of Prof. Clarence W. de Silva and co-supervision of Prof. Lalith B. Gamage. Currently she is a PhD student at McGill University, Montreal, Canada.

Jayasanka RANAWEERA received the BSc Engineering (Hons) degree in Electrical & Electronic Engineering from Sri Lanka Institute of Information Technology in 2016. He is currently reading for his MASc degree at the University of British Columbia, Vancouver, Canada, under the supervision of Prof. Clarence W. de Silva and co-supervision of Prof. Lalith B. Gamage.

Lalith B. GAMAGE received the BSc. degree in Electronic and Telecommunication Engineering from the University of Moratuwa, Sri Lanka, in 1985, the M.Sc. in Information Technology from De Montfort University, UK, in 1987, and the Ph. D. from the University of British Columbia, Vancouver, British Columbia, Canada, in 1993. He has served the University of Moratuwa from 1985 to

1999 as a Lecturer, Senior Lecturer and the Director of Computing Services Centre. He has been the Chairman, Arthur C Clarke Institute for Modern Technologies from 2010 to 2013, Chairman, Lanka Software Foundation 2012 to 2018 and President Sri Lanka Association of Non-State Higher Education Institutes from 2016 to 2018. Currently, he is the President/Vice Chancellor/CEO of the Sri Lanka Institute of Information Technology and an Affiliate Professor at the University of British Columbia, Vancouver, Canada.

Clarence W. de Silva received the Ph. D. degrees from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1978, and from the University of Cambridge, Cambridge, U.K., in 1998, and the honorary D.Eng. degree from the University of Waterloo, Waterloo, ON, Canada, in 2008. He is a Professor of mechanical engineering at The University of British Columbia, Vancouver, BC, Canada. He is a fellow of: IEEE, ASME, Canadian Academy of Engineering, and Royal Society of Canada. His appointments include the Tier 1 Canada Research Chair in Mechatronics and Industrial Automation, a Professorial Fellow, the Peter Wall Scholar, the Mobil Endowed Chair Professor, and the NSERCBC Packers Chair in Industrial Automation. His recent books published are Modeling of Dynamic Systems-With Engineering Applications (Taylor & Francis/CRC, 2018), Sensor Systems (Taylor & Francis/CRC, 2017), Sensors and Actuators-Engineering System Instrumentation, 2nd Edition (Taylor & Francis/CRC, 2016), Mechanics of Materials (Taylor & Francis/CRC, 2014), Mechatronics-A Foundation Course (Taylor & Francis/CRC, 2010), Modeling and Control of Engineering Systems (Taylor & Francis/CRC, 2009), VIBRATION-Fundamentals and Practice, 2nd Edition (Taylor & Francis/CRC, 2007), and Soft Computing and Intelligent Systems Design- Theory, Tools, and Applications (with F. Karray, Addison Wesley, 2004.