



Deep learning approaches for sleep disorder prediction in an asthma cohort

Dinh-Van Phan, Nan-Ping Yang, Ching-Yen Kuo & Chien-Lung Chan

To cite this article: Dinh-Van Phan, Nan-Ping Yang, Ching-Yen Kuo & Chien-Lung Chan (2020): Deep learning approaches for sleep disorder prediction in an asthma cohort, Journal of Asthma, DOI: [10.1080/02770903.2020.1742352](https://doi.org/10.1080/02770903.2020.1742352)

To link to this article: <https://doi.org/10.1080/02770903.2020.1742352>



Accepted author version posted online: 12 Mar 2020.
Published online: 18 Mar 2020.



Submit your article to this journal [↗](#)



Article views: 82



View related articles [↗](#)



View Crossmark data [↗](#)

Deep learning approaches for sleep disorder prediction in an asthma cohort

Dinh-Van Phan^{a,b,c} , Nan-Ping Yang^d , Ching-Yen Kuo^e, and Chien-Lung Chan^{a,f} 

^aDepartment of Information Management, Yuan Ze University, Taoyuan, ROC; ^bStatistics and Informatics Department, University of Economics, The University of Danang, Da Nang, Vietnam; ^cTeaching and Research Team for Business Intelligence, University of Economics, The University of Danang, Da Nang, Vietnam; ^dHualien Hospital, Ministry of Health and Welfare, Hualien, ROC; ^eDepartment of Medical Administration, Taoyuan General Hospital, Ministry of Health and Welfare, Taoyuan, ROC; ^fInnovation Center for Big Data and Digital Convergence, Yuan Ze University, Taoyuan, ROC

ABSTRACT

Objective: Sleep is a natural activity of humans that affects physical and mental health; therefore, sleep disturbance may lead to fatigue and lower productivity. This study examined 1 million samples included in the Taiwan National Health Insurance Research Database (NHIRD) in order to predict sleep disorder in an asthma cohort from 2002–2010.

Methods: The disease histories of the asthma patients were transferred to sequences and matrices for the prediction of sleep disorder by applying machine learning (ML) algorithms, including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest (RF), and deep learning (DL) models, including Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Convolution Neural Network (CNN).

Results: Among 14,818 new asthma subjects in 2002, there were 4469 sleep disorder subjects from 2002 to 2010. The KNN, SVM, and RF algorithms were demonstrated to be successful sleep disorder prediction models, with accuracies of 0.798, 0.793, and 0.813, respectively (AUC: 0.737, 0.690, and 0.719, respectively). The results of the DL models showed the accuracies of the RNN, LSTM, GRU, and CNN to be 0.744, 0.815, 0.782, and 0.951, respectively (AUC: 0.658, 0.750, 0.732, and 0.934, respectively).

Conclusions: The results showed that the CNN model had the best performance for sleep disorder prediction in the asthma cohort.

ARTICLE HISTORY

Received 6 September 2019

Revised 20 December 2019

Accepted 9 March 2020

KEYWORDS

Sleep disorder; asthma; machine learning; deep learning

Introduction

In recent years, the era of the fourth industrial revolution has been touted to be upon us, with the rise in intelligent manufacturing processes and the expansion of many technologies (1–3) and scientific fields such as big data, Internet of Things (IoT), Artificial Intelligence (AI), and technologic applications in health care (4–7). AI was founded in 1956, but has been practically applied and its power developed in recent years. Its use has exploded since 2015 with the development of the graphics processing unit (GPU), which enables faster, more powerful, cheaper, and broader parallel processing (8). AI has many successful applications, such as face recognition by Facebook (9), AlphaGo of Google (10), and product recommendations by Amazon (11). With the boom in big data has arisen an increasing demand for improved analysis and predictive methods. In fact, machine learning (ML), a subfield of AI, with many algorithms such as

K-Nearest Neighbors (KNN), Support Vector Machine (SVM) and Random Forest (RF), can adjust itself through structured data toward better results without the need for human interference. These ML models have increasingly shown superiority in terms of performance. In addition, deep learning (DL), a subset of ML, consists of many neural network layers to imitate the functions of the neural network in the human brain. Models using traditional ML algorithms or deep neural networks are built through training on useful characteristics extracted from data. Furthermore, the advent of DL technology has overcome some limitations of traditional ML techniques, and DL can be deployed to build even better predictive models (12–16). ML and DL techniques have been implemented in recent years for the analysis of various types of data, such as electronic health records, imaging, -omics, sensor data, text data for automated disease detection and prediction to improve health-care, medical services, and quality of human life.

When human health is being considered, we need to assess physical activities, sleep quality, and diseases, as many studies have indicated that physical activities and sleep affect physical and mental health (17–21). Sleep is typically a natural action of humans, comprising two states: rapid eye movement (REM) and non-REM (NREM) (22). When people sleep, they will not respond to the environment, a reversible behavioral state of perceptual disengagement (23), and sufficient sleep is required to recover and maintain health. Sleep deprivation may negatively impact cognitive health functioning as well as physical health (24), and in more serious cases sleep disorder is diagnosed. Previous studies indicated that there are many internal and external factors that affect sleep quality, such as sleep environment, events, and diseases. One disease that affects sleep quality is asthma, as attacks incite breathlessness, wheezing and obstructive sleep apnea, resulting in difficulty in getting to sleep (25,26). Therefore, this study applied ML and DL for the prediction of sleep disorder in an asthma cohort based on their disease history from 2002–2010. When sleep disorder is predicted in patients with asthma at an early stage, better treatment choices are available, and health can be improved.

Materials and methods

Sleep disorder

Sleep disorder is a disease related to problems such as the timing, quantity, and quality of sleep. Patients with sleep disorder have difficulty falling asleep, with no apparent cause, and the disorder may disturb normal physical, mental, social and emotional functioning (27). This disease may affect safety, as well as quality of life, including the patient's ability to drive safely, and may increase the risk of development of other health problems (28,29).

Recurrent neural network

Recurrent Neural Networks (RNNs) are a type of artificial neural network, the essence of which is that they employ recurrence, basically using information from a previous forward pass over the neural network. RNNs are flexibly applied for speech recognition (30), language modeling (31), language translation (32), image captioning (33) and healthcare (34).

Some special kinds of RNN are Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), which were introduced in 1997 (35,36) and 2014 (37,38), respectively. LSTM can remember information for long periods of time, while RNN may not be able to

remember for such a long time (39). While standard RNNs have a simple structure, such as a tanh layer, LSTM has four different structures that interact in a special way. LSTM has the ability to remove or add information to the cell state by gates. Gates are composed of a sigmoid layer (forget gate layer) and a pointwise multiplication operation. GRU is a variation of LSTM in which the forget gate is combined with an input gate to form a single “update gate”. The cell state is also combined with a hidden state. GRU often yields better performance for analysis of smaller datasets (40).

Convolutional neural network

Convolution Neural Network (CNN) is a type of artificial neural network that has been applied to many areas, such as computer vision, recommender systems, and natural language processing. A CNN consists of a number of convolutional and subsampling layers, optionally followed by fully-connected layers (41). Convolution is a mathematical operation to merge two sets of information. In this case, the convolution is applied to the input data using a convolution filter (kernel) to produce a feature map. The kernel is slid on input data with element-wise matrix multiplication, and the result is summed at every respective location. After a convolution operation, an m subsampling operation via a pooling layer is performed to reduce the number of parameters. The most common type of pooling layer is max pooling, which means taking the max values in each pooling window. The convolution and pooling layers extract features and reduce the number of parameters from the input data. After that, fully-connected layers are applied, and outputs corresponding to the number of classes that it needs to classify are generated (42–44).

Data source

This study used a Taiwan National Health Insurance Research Database (NHIRD) dataset that included 1,000 000 random samples from approximately 27.38 million contained in the Registry of Beneficiaries (ID) in 2010. No significant differences in gender distribution were noted between the population data and the sampling data (45–49) (data approval number: NHIRD-104–183). Retrospective records of the one-million samples obtained in 2010 were retrieved from 2000 using the Longitudinal Health Insurance Database (LHID).

The aim of the study was to build a sleep disorder prediction model for patients for use from the first time they were diagnosed with asthma. We therefore analyzed asthma subjects that were newly-diagnosed in

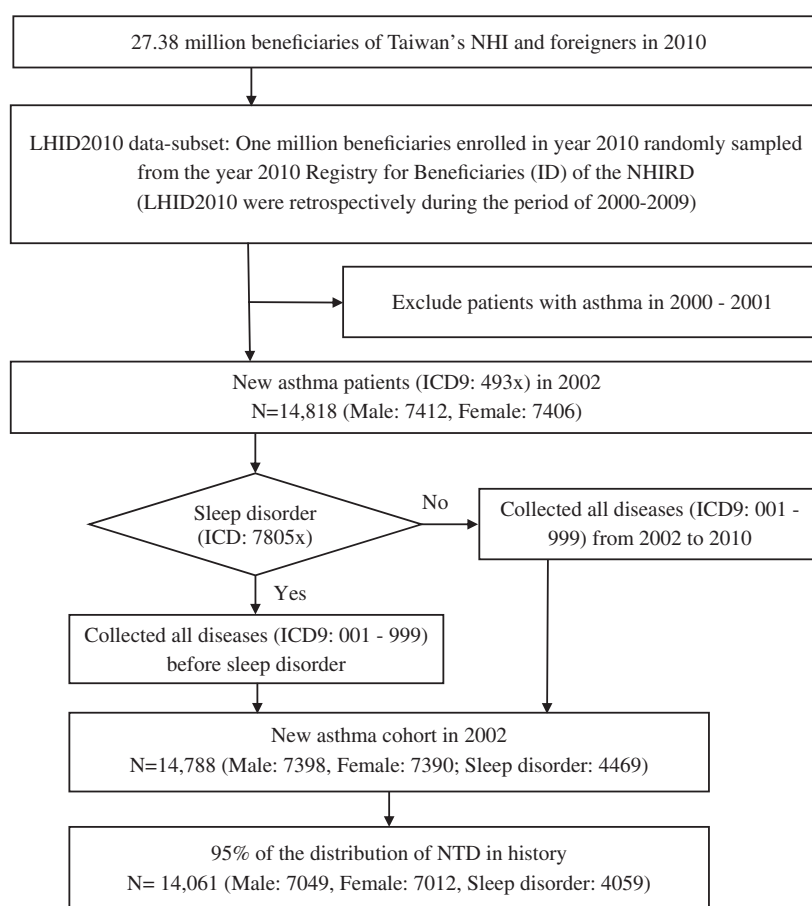


Figure 1. Study sample flowchart.

2002, but excluded asthma subjects already diagnosed in 2000 and 2001. These new asthma subjects were studied prospectively during the period of 2002–2010 to form a new asthma fixed cohort for the prediction of sleep disorder. Asthma and sleep disorder were identified according to the International Classification of Diseases Revision 9 (ICD-9) codes 493x and 7805x, respectively. For asthma patients diagnosed with sleep disorder, the history of diseases included all diseases (ICD-9: 001–999) diagnosed prior to sleep disorder, while for asthma patients not diagnosed with sleep disorder, all diseases diagnosed from 2002 to 2010 were included in the disease history (Figures 1 and 2).

Data processing

Transforming disease histories to sequences

Patients were diagnosed with diseases by doctors based on their symptoms and from the results of examinations such as ultrasound, x-ray and CT. Sometimes, it is necessary to consider a patient's disease history in order to diagnose new diseases. Many diseases are life-threatening, but are very difficult to treat, or cause severe sequelae such as stroke and cancer; however, if

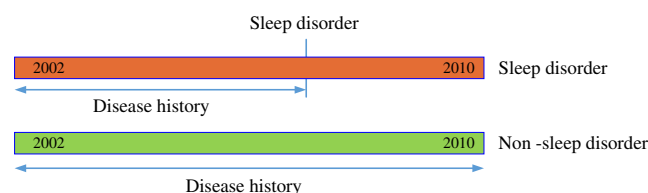


Figure 2. History of disease.

these diseases are predicted at an early stage, and a warning is given, the chances of prevention or successful treatment are increased. Indeed, there are many methods by which to predict the onset of a disease, such as examinations, prognosis, and diagnosis, especially based on the disease history of the patient. Therefore, in this study, we proposed a prognostic approach based on patient disease history using NHIRD data and a deep neural network.

In the NHIRD, the International Classification of Diseases, ninth revision, Clinical Modification (ICD-9-CM) codes for diagnosis are used, which classify by disease groups 001 to 999, V01–V91, and E000–E999 (50). We used the disease group including 001 to 999, with a total of 999 disease groups. The disease codes of each patient were reorganized into a

Table 1. A sample of diagnostic histories in NHIRD.

ID	FUNC_DATE	ACODE_ICD9_1	ACODE_ICD9_2	ACODE_ICD9_3
P1	20020208	36813		
P1	20020402	4659		
P1	20020408	36813		
P1	20020430	5719	7805	
P1	...	...	...	...
P1	20020715	36813		
P2	20020230	6829		
P2	20020324	61611		
P2	...	...	...	...
P2	20021230	4659	5369	7031

sequence for application of the RNN. Table 1 presents the examples of patients P1 and P2, who were diagnosed with diseases over a period of time with three main disease codes (ACODE_ICD9_1, ACODE_ICD9_2, ACODE_ICD9_3), along with the dates on which the patients were diagnosed (FUNC_DATE).

The disease histories of the patients were sorted by date (FUNC_DATE), and the numbers of the first three disease codes were transformed into a sequence. The order of the disease codes was ACODE_ICD9_1, ACODE_ICD9_2, ACODE_ICD9_3, following the date values in the FUNC_DATE column (P1: 368 465 368 571, label: 1; P2: 682 616 465 536 ... 703, label: 0).

Transforming disease histories into matrices

NHIRD data were embedded into 2D matrices $n \times m$, in which n was the number of months from January 2002 to December 2010, and m was the number of the ICD-9 code third group (50) based on three digits from 001 to 999. The values of data $x_{t,i}$ ($t \in n$, $i \in m$) were calculated as 0 or 1 based on the diagnosis of disease i at time t .

$$x_{t,i} = \begin{cases} 1 & \text{if the patient was diagnosed with disease } i \text{ at time } t \\ 0 & \text{if the patient was not diagnosed with disease } i \text{ at time } t \end{cases} \quad (1)$$

For example, a patient was diagnosed with a diabetic disease (ICD-9 = 250x) on 2002/01/06 and 2005/02/17. Here, we used the monthly time dimension, so the cell at row 1 (2002/01) and row 48 ($12 \times 3 + 2$) of column 250 (diabetic disease code) will be 1.

Statistical analysis

Besides using SPSS Ver.22 to perform statistical methods for the analysis of data, we also used ML methods to evaluate the performance of the models. Some standard coefficients were employed, including accuracy (Acc), sensitivity (Sen), specificity (Spec), Matthews's correlation coefficient (MCC), precision (Pre), F_1 Score

(F_1), area under the curve (AUC), true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (3)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FN)(TN + FN)}} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$F1 - \text{Score} = \frac{2TP}{2TP + FP + FN} \quad (7)$$

There are different ways in which to implement a DL model. However, building a DL model with a large training dataset requires a high-performance system and takes a long time. In this study, the DL model was implemented using Python 3.6 (open-source projects) with Anaconda 1.9.2, TensorFlow 1.11, Keras (51) and GPU: ROG-STRIX-GTX1080TI-O11G-GAMING, CPU: Intel Core i7 Processor 7700 in Ubuntu 18.04.1 LTS operation system environment. The dataset was partitioned into the training part (67%) and testing part (33%) using scikit-learn 0.21.2.

Ethical approval and consent for participation and publication

This study was approved by the Institutional Review Board (IRB) of Taipei Hospital, Taiwan, certificated by the Ministry of Health & Welfare, Taiwan (IRB approval number: TH-IRB-0015-0003).

Results

Demographic characteristics of the new asthma cohort in 2002

Of the one-million random samples in the NHIRD in 2010, there were 14 818 new asthma subjects (asthma

Table 2. Demographic characteristics of new asthma cohort in 2002.

New asthma patients in the year 2002					
Asthma			Non- Asthma		Chi-Square Test
	N	%	N	%	
Subject	14,818	100.00	741,135	100.00	
Gender					
Female	7,406	49.98	386,146	52.10	0.000
Male	7,412	50.02	354,989	47.90	
Age					
0–15	4,523	30.52	166,242	22.43	0.000
16–30	2,376	16.03	191,898	25.89	
31–45	2,133	14.39	190,415	25.69	
46–60	2,416	16.30	122,481	16.53	
>=61	3,370	22.74	70,099	9.46	

subjects who were diagnosed in 2002, excluding those already diagnosed in 2000 and 2001) and 741 135 non-asthma subjects in 2002. Among this new asthma cohort, there was a significantly higher proportion (30.52%, $p < 0.001$) of children, but there was no significant difference in gender (50.02% male, 49.98% female, $p < 0.001$) (Table 2).

In the new asthma cohort of 2002, 4469 subjects were diagnosed with sleep disorder from 2002 to 2010, increasing year after year from 8.29% in 2002 to 10.21% in 2010 (Table 3). In particular, a greater proportion of female asthma patients (61.58%) were diagnosed with sleep disorder. However, after the data were filtered by 95% of the distribution of the number of times disease was recorded (NTD) from 2002–2010, 4059 (female, 62.16%) asthma patients were diagnosed with sleep disorder (Table 4).

Sleep disorder prediction in the new asthma cohort from 2002–2010

Distribution of sleep disorder in the new asthma cohort

The new asthma cohort was used to predict sleep disorder using the DL model. However, the distribution statistic of this group showed many outlier values. Accordingly, we filtered the dataset by using 95% of the distribution of data for elimination of outlier values. After filtering out 95% of the distribution of the NTD, the new asthma cohort contained 4059 asthma subjects with sleep disorder and 10 002 asthma subjects without sleep disorder (Table 5). The NTD according to records of disease history ranged from 8 to 723 diseases (8–722 for sleep disorder and 8–723 for non-sleep disorder), with a mean of 140.21 (SD: 134.68) and a median of 97 for sleep disorder, and a mean of 224.99 (SD: 145.16) and a median of 191 for the non-sleep disorder group (Table 5).

Sleep disorder prediction in the new asthma cohort based on the sequence of disease history

The datasets for ML and RNN, LSTM and GRU were built from the disease history of the asthma patients as a sequence of ICD-9-CM codes. For sleep disorder subjects, the data included all diseases prior to the diagnosis of sleep disorder. Contrarily, for subjects without sleep disorder, the data included all ICD-9-CM codes for diseases. The disease history of each new asthma patient was transferred to a sequence of text, with a label value of 0 (non-sleep disorder) or 1 (sleep disorder). The sequence of ICD-9-CM codes was split into 2 parts: 67% for the training, 33% for the test. The KNN (number of neighbors $k=30$), SVM (kernel = 'rbf', $\gamma = 0.001$), and RF (number of estimator trees $n=50$) algorithms indicated an accuracy of 0.798, 0.793, and 0.813, respectively; while the AUC values were 0.737, 0.690, and 0.719, respectively (Table 6).

The sequence of disease history of the new asthma cohort from 2002–2010 was transformed into a vector with 300 dimensions by embedding a layer of Keras before being placed into the RNN, LSTM, and GRU layers. The results of these three models showed the accuracies of the RNN, LSTM, and GRU to be 0.744, 0.815, and 0.782, respectively (MMC: 0.343, 0.533, 0.469; AUC: 0.658, 0.750, 0.732, respectively) (Table 6).

Sleep disorder prediction in the new asthma cohort based on the matrix

The new asthma cohort data were preprocessed before being placed into the CNN model. Each subject was transformed into a 108×999 matrix (108 months, 999 ICD-9-CM codes). The labels were defined by [0, 1]; value 0 represented asthma patients without sleep disorder from 2002 to 2010, while value 1 represented asthma patients with sleep disorder in the same period. When a patient was diagnosed with sleep disorder, the label value would be set to 1, and all the disease history would be set to 0. All 2D matrices of all asthma patients were converted into a 1D matrix (108×999) and stored in a file for training and testing of the CNN models. We employed three models CNN1, CNN2, CNN3, with accuracies of 0.943, 0.951, and 0.923, respectively, and AUC values of 0.926 (95% CI: 0.920, 0.938), 0.934 (95% CI: 0.925, 0.942), and 0.885 (95% CI: 0.874, 0.897), respectively (Table 6). The CNN2 model had the highest performance, with a structure consisting of an input layer (108×999), three 2D convolution (Conv2D) layers with ReLU activation function, kernels $((5 \times 5) - (3 \times 3) -$

Table 3. The number of sleep disorder subjects in new asthma cohort from 2002–2010.

Year	2002	2003	2004	2005	2006	2007	2008	2009	2010
Asthma	14,818	14,601	14,592	14,570	14,414	14,396	13,712	14,341	14,323
Sleep disorder (%)	1,229 (8.29)	1,174 (8.04)	1,302 (8.92)	1,351 (9.27)	1,336 (9.27)	1,344 (9.34)	1,328 (9.68)	1,434 (10.00)	1,462 (10.21)

Table 4. Sleep disorder in new asthma cohort from 2002 to 2010.

	Sleep disorder in the new asthma cohort from 2002–2010		95% of the distribution of NTD from 2002–2010	
	N	%	N	%
Subject	4469	100.00	4059	100
Female	2752	61.58	2523	62.16
Male	1717	38.42	1536	37.84

NTD: times a disease was recorded.

Table 5. Description statistic of NTD in new asthma cohort.

	NTD in new asthma cohort before sleep disorder			
	All (N = 14,788)		95% of the distribution (N = 14,061)	
	Sleep disorder (N = 4469)	Non-Sleep disorder (10,319)	Sleep disorder (N = 4059)	Non-Sleep disorder (10,002)
Minimum	1	2	8	8
Maximum	1762	2365	722	723
Mean	140.26	247.12	140.21	224.99
Median	88	197	97	191
Standard Deviation	165.86	144.10	134.68	145.16

NTD: times a disease was recorded.

Table 6. Result of machine learning algorithms and deep learning models.

Model	Accuracy	Sensitivity	Specificity	Precision	F1-score	MCC	AUC
KNN	0.798	0.595	0.880	0.669	0.630	0.493	0.737
SVM	0.793	0.489	0.933	0.732	0.556	0.453	0.690
Random forest	0.813	0.497	0.941	0.775	0.606	0.551	0.719
RNN	0.744	0.453	0.862	0.580	0.508	0.343	0.658
LSTM	0.815	0.593	0.907	0.725	0.725	0.533	0.750
GRU	0.782	0.614	0.851	0.631	0.622	0.469	0.732
CNN1	0.943	0.885	0.966	0.913	0.899	0.859	0.926
CNN2	0.951	0.893	0.974	0.933	0.913	0.879	0.934
CNN3	0.923	0.796	0.975	0.927	0.856	0.807	0.885

KNN: K-Nearest Neighbors; SVM: Support Vector Machine; RNN: Recurrent Neural Network; LSTM: Long Short-Term Memory network; GRU: Gate Recurrent Unit; CNN: Convolution Neural Network; MCC: Matthews's correlation coefficient; AUC: area under the curve.

(3 × 3)), max-pooling (MP) layers ((4 × 4) – (4 × 4) – (2 × 2)), dropout layers (rate: 0.25, 0.25, 0.4, 0.3), Flatten, Dense layers, and Adam optimization (52). These types of layer (output shape) were built as follows: Conv2D (none, 104, 995, 128), MP (none, 26, 248, 128), Dropout, Conv2D (none, 24, 246, 62), MP (none, 6, 61, 62), Dropout, Conv2D (none, 4, 59, 62), MP (none, 2, 29, 62), Dropout, Flatten (none, 3596), Dense (none, 32), Dropout, Dense (none, 2).

Discussion

A study applied CNN (53) for the detection of diabetic retinopathy and diabetic macular edema based on 128 175 retinal images of 69 573 patients in the

United States and India. The model was validated by two image data sets, EyePACS-1, which consisted of 9,963 retinal fundus images from 4,997 patients, and Messidor-2, which included 1748 retinal fundus images from 874 patients. The results showed that the prediction models were of high sensitivity and specificity (for EyePACS-1: 97.5%, 93.4%; for Messidor-2: 96.1%, 93.9%, respectively) (54). Our study also used CNN models to predict sleep disorder, with high accuracy (> 92.3%). The daily activities of elderly people in their home were converted into sequences of activities, which were then used to build a predictive model based on a non-linear autoregressive network with exogenous (NARX) RNN, with excellent results in terms of identifying the temporal relationships of the input data (55). In particular, Electronic Health Record (EHR) data represent another type of data to which DL can be applied for building prediction models. Edward Choi proposed the Med2Vec model, which can represent the sequences of both visit times and medical codes in EHR data. This model showed significant improvement in both predictive accuracy and clinical interpretation (56). In addition, EHR data were represented for every patient as a temporal matrix, which has a time dimension and event dimension. Then, CNN was applied to build a model with four layers for the extraction of phenotypes and performance of chronic disease prediction (57). The RNN model with the GRU was used to detect relations among event time series and predict heart failure

based on EHR data. The results were compared for multilayer perceptron (MLP), SVM, and KNN with 12-month and 18-month observation windows. According to the results, the DL model attained the highest AUC as compared with the other models for the detection of heart failure with both a 12-month and an 18-month observation window (0.764 and 0.883, respectively) (58). DL models using EHR data can support decision-making and advance clinical research. A study of data of 700 000 patients obtained from the Mount Sinai data warehouse also employed DL, and that model improved the clinical prediction performance for diabetes, schizophrenia, and various cancers (59). In this study, we converted the disease history of a new asthma patient cohort from 2002–2010 into sequences of text and then utilized ML algorithms (KNN, SVM, RF) and deep learning (RNN, LSTM, GRU, and CNN) techniques to predict sleep disorder, with a high performance. ML algorithms and DL were also applied to a Hospital Information System (HIS) and an electronic medical record (EMR) database in a Chinese study, and the results showed that the deep neural network model achieved a higher performance than previous predictive models (60). Our experimentally-confirmed results showed that DL models attained a better performance than traditional ML models.

Conclusions

Several ML algorithms and some DL models were applied for the prediction of sleep disorder in a new asthma patient cohort from 2002 to 2010. The results showed that CNN models had the best performance ($\text{Acc} > 92.3\%$, $\text{MCC} > 0.80$ and $\text{AUC} > 0.88$). In our next study, we will continue to apply DL models for the prediction of other serious diseases, such as cancer and stroke; nevertheless, this method may not work for other diseases. In addition, we will use DL models to build decision support systems to predict some diseases at an early stage.

Acknowledgement

The authors acknowledge support from the Taiwan Typhoon and Flood Research Institute.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

This study was supported by the Ministry of Science and Technology (<http://www.most.gov.tw/>), MOST 105–2221-E-155–041-MY3, CLC received the funding.

ORCID

Dinh-Van Phan  <http://orcid.org/0000-0002-7015-1432>
 Nan-Ping Yang  <http://orcid.org/0000-0001-5675-5937>
 Chien-Lung Chan  <http://orcid.org/0000-0002-7486-7075>

References

1. Li C, Konomis D, Neubig G, Xie P, Cheng C, Xing E. Convolutional neural networks for medical diagnosis from admission notes. *arXiv Preprint. arXiv:171202768*. 2017.
2. Bengio Y, Ducharme R, Vincent P, Jauvin C. A neural probabilistic language model. *J Mach Learn Res*. 2003; 3:1137–55.
3. Bottomly D, McWeeney SK, Wilmot B. HitWalker2: visual analytics for precision medicine and beyond. *Bioinformatics*. 2016;32(8):1253–5. doi:10.1093/bioinformatics/btv739.
4. Tomita K, Nagao R, Touge H, Ikeuchi T, Sano H, Yamasaki A, Tohda Y. Deep learning facilitates the diagnosis of adult asthma. *Allergol Int*. 2019;68(4): 456–61. doi:10.1016/j.alit.2019.04.010.
5. Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang GZ. Big data for health. *IEEE J Biomed Health Inform*. 2015;19(4):1193–208. doi:10.1109/JBHI.2015.2450362.
6. Zhou K, Taigang L, Lifeng Z, editors. Industry 4.0: Towards future industrial opportunities and challenges. 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD); 2015. 15–17 Aug. 2015. doi:10.1109/FSKD.2015.7382284.
7. Lee I, Lee K. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Bus Horiz*. 2015;58(4):431–40. doi:10.1016/j.bushor.2015.03.008.
8. Copeland M. What's the difference between artificial intelligence, machine learning, and deep learning? 2016. Available from: <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/> [last accessed 06 May 2018].
9. Becker BC, Ortiz EG, editors. Evaluation of face recognition techniques for application to facebook. 2008. 8th IEEE International Conference on Automatic Face & Gesture Recognition; 2008 17–19 Sept. 2008.
10. Chang HS, Fu MC, Hu J, Marcus SI. Google Deep Mind's AlphaGo. *OR/MS Today*. 2016;43(5):24–9.
11. Schafer JB, Konstan J, Riedl J, editors. Recommender systems in e-commerce. *Proceedings of the 1st ACM conference on Electronic commerce*; 1999. ACM. doi: 10.1145/336992.337035.
12. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw*. 2015;61:85–117. doi:10.1016/j.neunet.2014.09.003.

13. Kasun LLC, Zhou H, Huang G-B, Vong CM. Representational learning with ELMs for big data. 2013.
14. Xue-Wen C, Xiaotong L. Big data deep learning: Challenges and perspectives. *IEEE Access*. 2014;2: 514–25. doi:10.1109/access.2014.2325029.
15. Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E. Deep learning applications and challenges in big data analytics. *J Big Data*. 2015;2(1):1. doi:10.1186/s40537-014-0007-7.
16. Carson MB, Scholtens DM, Frailey CN, Gravenor SJ, Powell ES, Wang AY, et al. Characterizing teamwork in cardiovascular care outcomes: a network analytics approach. *Circ Cardiovasc Qual Outcomes*. 2016;9(6): 670–8. doi:10.1161/CIRCOUTCOMES.116.003041.
17. Phan D-V, Chan C-L, Pan R-H, Yang N-P, Hsu H-C, Ting H-W, et al. Investigating the effect of daily sleep on memory capacity in college students. *THC*. 2018; 27(2):1–12. doi:10.3233/THC-181350.
18. Candela F, Zucchetti G, Magistro D, Rabaglietti E. The effects of a physical activity program and a cognitive training program on the long-term memory and selective attention of older adults: a comparative study. *Act Adapt Aging*. 2015;39(1):77–91. doi:10.1080/01924788.2014.977191.
19. Fimm B, Blankenheim A. Effect of sleep deprivation and low arousal on eye movements and spatial attention. *Neuropsychologia*. 2016;92:115–28. doi:10.1016/j.neuropsychologia.2016.03.021.
20. Vanhelst J, Béghin L, Duhamel A, Manios Y, Molnar D, De Henauf S, Moreno LA, Ortega FB, Sjöström M, Widhalm K, et al. Physical activity is associated with attention capacity in adolescents. *J Pediatr*. 2016; 168:126–31 e2. doi:10.1016/j.jpeds.2015.09.029.
21. Phan D-V, Chan C-L, Pan R-H, Yang N-P, Hsu H-C, Ting H-W, Lai KR. A study of the effects of daily physical activity on memory and attention capacities in college students. *J Healthcare Eng*. 2018;2018:1–9. doi:10.1155/2018/2942930.
22. Carskadon MA, Dement W. Normal human sleep: an overview. *Principles Pract Sleep Med*. 2005;4:13–23.
23. Carskadon MA, Dement WC. Chapter 2—Normal human sleep: an overview. In: Kryger MH, Roth T, Dement WC, editors. *Principles and practice of sleep medicine*. 5th ed. Philadelphia: W.B. Saunders; 2011: 16–26.
24. Drummond SP, Anderson DE, Straus LD, Vogel EK, Perez VB. The effects of two types of sleep deprivation on visual working memory capacity and filtering efficiency. *PLoS One*. 2012;7(4):e35653. doi:10.1371/journal.pone.0035653.
25. Pokrzywinski RF, Meads DM, McKenna SP, Glendenning GA, Revicki DA. Development and psychometric assessment of the COPD and asthma sleep impact scale (CASIS). *Health Qual Life Outcomes*. 2009;7(1):98. doi:10.1186/1477-7525-7-98.
26. Boulet L-P, Boulay M-È. Asthma-related comorbidities. *Expert Rev Resp Med*. 2011;5(3):377–93. doi:10.1586/ers.11.34.
27. Wikipedia. Sleep disorder. 2018. Available from: https://en.wikipedia.org/wiki/Sleep_disorder. [last accessed 20 November 2018].
28. Clinic M. Sleep disorders. 2016. Available from: <https://www.mayoclinic.org/diseases-conditions/sleep-disorders/symptoms-causes/syc-20354018>. [last accessed 20 November 2018].
29. Parekh R. What Are Sleep Disorders?: American Psychiatric Association; 2017. Available from: <https://www.psychiatry.org/patients-families/sleep-disorders/what-are-sleep-disorders>. [last accessed 20 November 2018].
30. Graves A, Mohamed A, Hinton G, editors. *Speech recognition with deep recurrent neural networks*. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 2013. p. 26–31. May 2013.
31. Mikolov T, Kombrink S, Burget L, Černocký J, Khudanpur S, editors. *Extensions of recurrent neural network language model*. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2011. p. 22–7. May 2011.
32. Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. *arXiv Preprint. arXiv:14090473*. 2014.
33. Vinyals O, Toshev A, Bengio S, Erhan D, editors. *Show and tell: A neural image caption generator*. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015.
34. Choi E, Bahadori MT, Sun J, Kulas J, Schuetz A, Stewart W, editors. *Retain: An interpretable predictive model for healthcare using reverse time attention mechanism*. *Advances in Neural Information Processing Systems*; 2016.
35. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735–80. doi:10.1162/neco.1997.9.8.1735.
36. Greff K, Srivastava RK, Koutnik J, Steunebrink BR, Schmidhuber J. LSTM: A search space odyssey. *IEEE Trans Neural Netw Learning Syst*. 2017;28(10): 2222–32. doi:10.1109/TNNLS.2016.2582924.
37. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv Preprint. arXiv:14061078*. 2014.
38. Yao K, Cohn T, Vylomova K, Duh K, Dyer C. Depth-gated recurrent neural networks. *arXiv Preprint. arXiv:150803790*. 2015.
39. Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans Neural Netw*. 1994;5(2):157–66. doi:10.1109/72.279181.
40. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv Preprint. arXiv 14123555*. 2014.
41. Fonseca TL. What's happening inside the convolutional neural network? The answer is convolution. 2017; [last accessed 17 January 2019].
42. Dertat A. *Applied deep learning—Part 4: Convolutional neural networks*. 2017. Available from: <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2>. [last accessed 17 January 2019].

43. Scherer D, Müller A, Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition. In *International Conference on Artificial Neural Networks—ICANN 2010*. 2010. p. 92–101.
44. Ciresan DC, Meier U, Masci J, Gambardella LM, Schmidhuber J, editors. Flexible, high performance convolutional neural networks for image classification. *Twenty-Second International Joint Conference on Artificial Intelligence*; 2011.
45. Shyam R, Ganesh HBB, Kumar S S, Poornachandran P, Soman KP. Apache spark a big data analytics platform for smart grid. *Procedia Technol*. 2015;21:171–8. doi:[10.1016/j.protcy.2015.10.085](https://doi.org/10.1016/j.protcy.2015.10.085).
46. Wu SC, Chen WT, Muo CH, Ke TW, Fang CW, Sung FC. Association between appendectomy and subsequent colorectal cancer development: an Asian population study. *PLoS One*. 2015;10(2):e0118411. doi:[10.1371/journal.pone.0118411](https://doi.org/10.1371/journal.pone.0118411).
47. Yang NP, Chan CL, Chu D, Lin YZ, Lin KB, Yu CS, et al. Epidemiology of hospitalized traumatic pelvic fractures and their combined injuries in Taiwan: 2000–2011 National Health Insurance data surveillance. *Biomed Res Int*. 2014;2014:878601. doi:[10.1155/2014/878601](https://doi.org/10.1155/2014/878601).
48. Yang NP, Chan CL, Yu IL, Lee CY, Chou P. Estimated prevalence of orthopaedic fractures in Taiwan—A cross-sectional study based on nationwide insurance data. *Injury*. 2010;41(12):1266–72. doi:[10.1016/j.injury.2010.05.025](https://doi.org/10.1016/j.injury.2010.05.025).
49. Chan CL, Lin W, Yang NP, Huang HT. The association between the availability of ambulatory care and non-emergency treatment in emergency medicine departments: a comprehensive and nationwide validation. *Health Policy*. 2013;110(2–3):271–9. doi:[10.1016/j.healthpol.2012.12.003](https://doi.org/10.1016/j.healthpol.2012.12.003).
50. 2015 ICD-9-CM Diagnosis Codes. Available from: <http://www.icd9data.com/2015/Volume1/default.htm>. [last accessed 07 May 2019].
51. Keras. Available from: <https://keras.io/>. [last accessed 08 August 2018].
52. Kingma DP, Ba J, Adam: A method for stochastic optimization. *arXiv Preprint*. arXiv 1412.6980. 2014.
53. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z, editors. Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016.
54. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–10. doi:[10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216).
55. Mahmoud S, Lotfi A, Langensiepen C. Behavioural pattern identification and prediction in intelligent environments. *Appl Soft Comput*. 2013;13(4):1813–22. doi:[10.1016/j.asoc.2012.12.012](https://doi.org/10.1016/j.asoc.2012.12.012).
56. Choi E, Bahadori MT, Searles E, Coffey C, Thompson M, Bost J, et al. Multi-layer representation learning for medical concepts. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD16*. 2016. p. 1495–504. doi:[10.1145/2939672.2939823](https://doi.org/10.1145/2939672.2939823).
57. Cheng Y, Wang F, Zhang P, Hu J. Risk prediction with electronic health records: a deep learning approach. *Proceedings of the 2016 SIAM International Conference on Data Mining*. p. 432–40. doi:[10.1137/1.9781611974348.49](https://doi.org/10.1137/1.9781611974348.49).
58. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc*. 2017; 24(2):361–70. doi:[10.1093/jamia/ocw112](https://doi.org/10.1093/jamia/ocw112).
59. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep*. 2016;6(1):26094. doi:[10.1038/srep26094](https://doi.org/10.1038/srep26094).
60. Liang Z, Zhang G, Huang JX, Hu QV, editors. Deep learning for healthcare decision making with EMRs. 2014 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*; 2014 2–5 Nov. 2014. doi:[10.1109/BIBM.2014.6999219](https://doi.org/10.1109/BIBM.2014.6999219).