



Original Article

The application of machine learning on brain imaging features of different narcolepsy subtypes

Wei-Chih Chin^{1,2,†}, Sheng-Yao Huang^{3,‡,ID}, Feng-Yuan Liu^{4,5}, Chih-Huan Wang⁶, I Tang^{1, ID}, Ing-Tsung Hsiao^{4,5,*} and Yu-Shu Huang^{1,*}

¹Department of Child Psychiatry and Sleep Center, Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Taoyuan, Taiwan,

²College of Life Sciences and Medicine, National Tsing Hua University, Hsinchu, Taiwan,

³Department of Mathematics, Soochow University, Taipei, Taiwan,

⁴Department of Medical Imaging and Radiological Sciences, College of Medicine and Healthy Aging Center, Chang Gung University, Taoyuan, Taiwan,

⁵Department of Nuclear Medicine and Molecular Imaging Center, Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Taoyuan, Taiwan and

⁶Department of Psychology, Zhejiang Normal University, Zhejiang, China

[†]These authors contributed equally to this study.

^{*}Corresponding authors. Yu-Shu Huang, Department of Child Psychiatry and Sleep Center, Chang Gung Memorial Hospital Taoyuan, Taiwan, No.5, Fuxing St., Guishan, Taoyuan 333, Taiwan. Email: yushuhuang1212@gmail.com; Ing-Tsung Hsiao, Department of Medical Imaging and Radiological Sciences, Chang Gung University, No. 259 Wen-hua 1st Road, 333 Guishan, Taoyuan, Taiwan. Email: ihsiao@mail.cgu.edu.tw.

Abstract

Study Objectives: Narcolepsy is a central hypersomnia disorder, and differential diagnoses between its subtypes can be difficult. Hence, we applied machine learning to analyze the positron emission tomography (PET) data of patients with type 1 or type 2 narcolepsy, and patients with type 1 narcolepsy and comorbid schizophrenia, to construct predictive models to facilitate the diagnosis.

Methods: This is a retrospective and prospective case-control study of adolescent and young adult patients with type 1 or type 2 narcolepsy, and type 1 narcolepsy and comorbid schizophrenia. All participants received ¹⁸F-fluorodeoxy glucose PET, sleep studies, neurocognitive tests, sleep questionnaires, and human leukocyte antigen typing. The collected PET data were analyzed by feature selections and classification methods in machine learning to construct predictive models.

Results: A total of 314 participants with narcolepsy were enrolled; 204 had type 1 narcolepsy, 90 had type 2 narcolepsy, and 20 had type 1 narcolepsy and comorbid schizophrenia. We used three filter methods for feature selection followed by a comparative analysis of classification methods. To apply a small number of regions of interest (ROI) and high classification accuracy, the Naïve Bayes classifier with the Term Variance as feature selection achieved the goal with only three ROIs (left basal ganglia, left Heschl, and left striatum) and produced an accuracy of higher than 99%.

Conclusions: The accuracy of our predictive model of PET data are promising and can aid clinicians in the diagnosis of narcolepsy subtypes. Future research with a larger sample size could further refine the predictive model of narcolepsy.

Key words: type 1 narcolepsy; type 2 narcolepsy; PET; machine learning; feature selection

Graphical Abstract

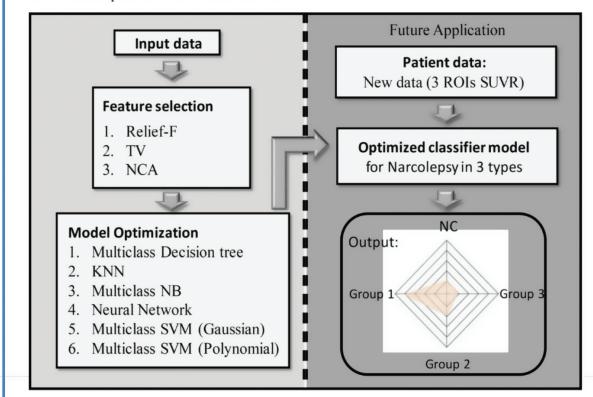
The Application of Machine Learning on Brain Imaging Features of Different Narcolepsy Subtypes

Background

Analyzing the positron emission tomography (PET) data of patients with type 1 or type 2 narcolepsy, and patients with type 1 narcolepsy and comorbid schizophrenia, to construct predictive models to facilitate the differential diagnosis.

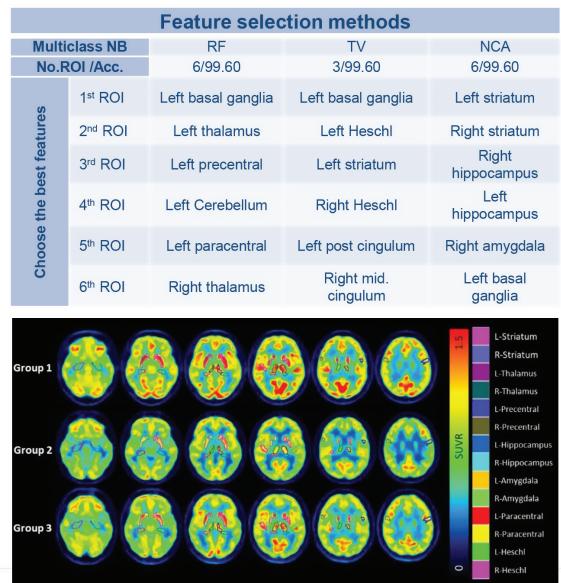
Methods

314 patients with type 1 or type 2 narcolepsy, and type 1 narcolepsy and comorbid schizophrenia. PET data were analyzed by feature selections and classification methods to construct predictive models.



Results

The Naïve Bayes classifier with the Term Variance : with only three ROIs (Lt basal ganglia, Lt Heschl, Lt striatum), accuracy > 99%.



Promising results of our predictive model from PET images. Facilitate diagnosis of narcolepsy subtypes

Statement of Significance

Narcolepsy is a central hypersomnia disorder, and differential diagnoses between its subtypes can be difficult. This study applied machine learning to analyze the positron emission tomography data of patients with type 1 or type 2 narcolepsy, and patients with type 1 narcolepsy and comorbid schizophrenia, to select features of different subtypes and construct predictive models to facilitate the diagnosis. To apply a small number of regions of interest (ROIs) and high classification accuracy, the Naïve Bayes classifier with the Term Variance as feature selection achieved the goal with only three ROIs (left basal ganglia, left Heschl, and left striatum) and produced an accuracy of higher than 99%. These results can aid clinicians in the diagnosis of narcolepsy subtypes and further treatment.

Introduction

Narcolepsy is a central hypersomnia disorder, characterized by excessive daytime sleep and disturbed nighttime sleep. Symptoms include cataplexy, sudden and transient muscle weakness, and rapid eye movement-related symptoms such as hypnagogic/hypnopompic hallucinations, sleep attacks, and sleep paralysis [1]. It was classified into two subtypes by the International Classification of Sleep Disorders, second edition, (ICSD-2) according to the presence or absence of cataplexy [2], and in recent years, the ICSD-3 divided narcolepsy into type 1 (NT1) and type 2 (NT2): NT1 has the presence of cataplexy and/or abnormally low cerebrospinal fluid (CSF) hypocretin levels, while NT2 lacks cataplexy and/or has normal CSF hypocretin levels [3].

Differential diagnosis between the two subtypes of narcolepsy is not always easy. Cataplexy can be as trivial as grimacing or eyelid weakness and is therefore difficult to detect in many patients. Its presentation and severity change during the course of the

disease, and some patients with type 1 develop cataplexy several years after the onset of the disease. It can also be mistakenly labeled as a seizure, clumsiness, or even attention-seeking behavior in children. Previous studies have found that type 1 narcolepsy is caused by the loss of hypocretin neurons [4–6], and patients with type 2 have normal hypocretin levels [7]. Although hypocretin is a fundamental marker between the two subtypes, the examination of CSF hypocretin levels is not available in most countries, and lumbar puncture raises many concerns for patients and families.

Furthermore, psychiatric comorbidities of narcolepsy can also increase the difficulty of diagnosing narcolepsy. Although it remains controversial, studies of narcolepsy and comorbid schizophrenia (SCZ) have been published. Schizophrenia is characterized by hallucinations, delusions, disorganized speech/behavior, and negative symptoms [8]. High rates of psychiatric comorbidity were reported in patients with narcolepsy [9], and the prevalence of schizophrenia was 3.4%, a four-fold increase

compared to those without narcolepsy. Douglass et al. [10] reported that 7% of their refractory schizophrenia patients have variants of human leukocyte antigen (HLA)-associated narcolepsy [10]. Clinically, it can be difficult to distinguish true hallucinations from hypnopompic or hypnagogic hallucinations, and cataplexy can mimic disorganized behavior. Besides, patients with narcolepsy can display disorganized speech when they are very sleepy and become withdrawn, secondary to narcoleptic symptoms, and similar to the presentation of the negative symptoms of schizophrenia. Narcolepsy medication, including Modafinil and Methylphenidate, can also lead to side effects of psychotic symptoms such as hallucination and delusion, and differentiation can sometimes be challenging.

Neuroimaging has been applied in narcolepsy studies, including ¹⁸F-fluorodeoxy glucose (FDG) positron emission tomography (PET) and ^{99m}Tc-ethyl-cysteinate dimer single photon emission computed tomography methods, to evaluate the functional brain imaging of patients with narcolepsy. Our previous studies used PET and compared the differences between type 1 and type 2 narcolepsy, and the control group [11, 12]. Compared to the healthy control group, patients with narcolepsy had hypometabolism in the right mid-frontal lobe, and angular gyrus and hypermetabolism in the olfactory lobe, hippocampus, parahippocampus, amygdala, fusiform, left inferior parietal lobe, left superior temporal lobe, striatum, basal ganglia and thalamus, right hypothalamus, and pons [12]. Between the two narcolepsy subtypes, less hypermetabolism in the fusiform gyrus, striatum, hippocampus, thalamus, basal ganglia, and cerebellum was found in patients with type 2 narcolepsy, as well as less hypometabolism in the frontal lobe, posterior cingulum, angular gyrus, and parietal lobe, than those with type 1 narcolepsy [11]. In our recent study, we compared the results of PET studies between patients with type 1 narcolepsy, schizophrenia, and dual diagnoses (type 1 narcolepsy with comorbid schizophrenia) [13]. Compared to the control group, patients with dual diagnosis had significant hypometabolism in the right mid-frontal, right orbital inferior frontal, and right posterior cingulum, and significant hypermetabolism in the left amygdala, bilateral striatum, bilateral substantia nigra, bilateral basal ganglia, and bilateral thalamus. They also had significant hypometabolism in the bilateral lingual compared to those with type 1 only. These PET findings of our two studies were consistent with neurocognitive function tests [11, 13].

In recent years, more studies have started to adopt machine learning and data mining, aiming to improve the precision of diagnoses by new approaches to data analysis. After data collection, analyses can classify and regroup different factors to find the key markers to distinguish diseases. Algorithms such as decision trees and support vector machines (SVM) can construct prediction models. Once clinicians have input the clinical data, the results of these models can serve as references to support diagnoses and they also have the potential to increase the accuracy and shorten the time needed to obtain a diagnosis. Further applications may even help to reach precision medicine goals. In this study, we used machine learning to analyze the PET data of patients with type 1 and type 2 narcolepsy, and patients with narcolepsy and comorbid schizophrenia to construct predictive models to facilitate clinical differentiation and diagnosis.

Materials and Methods

This is a retrospective and prospective case-control study of adolescent and young adult patients with type 1 or type 2 narcolepsy,

type 1 narcolepsy with comorbid schizophrenia, and healthy controls. All participants received ¹⁸F-FDG PET, sleep studies, and neurocognitive tests. The collected PET data were analyzed by feature selection and classification methods in machine learning to construct predictive models for differential diagnoses between the three diagnostic groups.

Participants

The study was conducted in the sleep center of a medical center located in northern Taiwan. This study was approved by the Institutional Review Board of Chang Gung Hospital, Taiwan (IRB: 201407075A3, 201702299A3, and 201902163A3). All participants and their legal representatives signed informed consent forms before entering this study.

Narcolepsy evaluation

All participants received a thorough clinical evaluation, including the patient's history and a physical examination. Demographic data were collected, and questionnaires were completed, including the Stanford narcolepsy questionnaire, the Epworth Sleepiness Scale, Pediatric Daytime Sleepiness Scale, and Visual Analog Scales, to evaluate excessive daytime sleep and sleep diaries for at least 14 consecutive days. HLA typing was performed, and some participants also underwent a cerebrospinal fluid hypocretin level test. The presence of different narcoleptic symptoms was recorded.

Actigraphy, polysomnography (PSG), and multiple sleep latency tests (MSLT) were performed on each participant. Participants wore an actigraph for 2 weeks to record the quantity of sleep during the night and day. During nighttime PSG, variables were recorded, including electroencephalography, electrooculogram, chin and leg electromyography, electrocardiography, body-position sensor, nasal cannula/pressure transducer, mouth thermistor, thoracic and abdominal plethysmography bands, neck microphone, and finger pulse oximetry. After PSG, MSLT was performed the following morning. A total of five 20-minute short naps at 2-hour intervals with a mean sleep latency and the number of sleep-onset rapid eye movement periods were obtained. All participants received HLA typing (DQB1*0602) by blood sampling. The diagnoses of type 1 and type 2 narcolepsy were made by an experienced sleep medicine doctor after completing all examinations, based on the International Classification of Sleep Disorders, third edition, (ICSD-3) diagnostic criteria [3].

Psychiatric evaluation

Schizophrenia and other psychiatric disorders were evaluated and diagnosed through psychiatric diagnostic interviews by experienced psychiatrists, based on the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders, fifth edition, (DSM-5) [14]. The symptoms of schizophrenia, including auditory hallucinations, visual hallucinations, and delusions, were recorded.

Medications

Before any examinations, sleep studies, and PET scans, participants had a drug-free period of at least 7 days, except for antipsychotic medication to prevent the relapse of psychosis.

Positron emission tomography

An ¹⁸F-FDG PET scan was performed on each participant. Initially, participants stayed in a quiet room and were asked to close their

eyes to avoid stimuli after a 5–10 mCi intravenous injection of FDG. They were monitored and supervised by a researcher to stay alert and not talk or move before the PET scan.

The PET image took 10 minutes and was acquired 30 minutes post-injection by using a Siemens mCT PET/CT scanner and a Biograph mMR PET/MR scanner (Siemens Medical Solutions, Malvern, PA, USA). Images were reconstructed from the PET/CT scanner using the 3-D OSEM algorithm (four iterations, 24 subsets, 2-mm Gaussian filter, and zoom 3) with CT-based attenuation correction, and scatter and random corrections as provided by the manufacturer, resulting in a matrix size of $400 \times 400 \times 109$ with a voxel size of $0.68 \times 0.68 \times 2.03 \text{ mm}^3$. PET images were reconstructed from the PET/MR scanner using the 3-D OSEM PET reconstruction algorithm (three iterations, 21 subsets; Gaussian filter: 2 mm; zoom: 3) with MR-based attenuation correction, scatter, and random corrections, resulting in PET images with a matrix size of $344 \times 344 \times 127$ and a voxel size of $0.83 \times 0.83 \times 2.03 \text{ mm}^3$.

Analysis of PET data

The ^{18}F -FDG PET images acquired from both the PET/CT and PET/MR scanners were preprocessed with a pre-optimized scanner-specific filter similarly derived in Joshi et al. [15] so that all PET images from these two scanners resulted in a unified resolution for further analysis [15]. PET data were processed and analyzed using the PMOD image analysis software (version 3.7, PMOD Technologies Ltd, Zurich, Switzerland) and followed the same procedure as in our previous study [16]. For quantitation, a total of 93 regions based on the automated anatomical labeling volume-of-interest templates [17] were selected for future analysis. For calculating the quantitation value, the mean value of regions with an intensity higher than a maximum of 50% in the whole brain was applied as the reference value for calculating the standard uptake value ratio (Figure 1).

Feature selection: filter methods

In machine learning, feature selection is usually applied to reduce the number of features in the model and sometimes to improve the performance of the model. Here, the feature selection was used to select fewer features as biomarkers to improve computation efficiency and accuracy in the differential diagnosis of different disease groups. For this purpose, we used three filter methods for feature selection, including relief-F (RF), term variance (TV), and neighborhood component analysis (NCA), from the filter-feature-selection-toolbox [the tool can be downloaded at <https://github.com/JingweiToo/Filter-Feature-Selection-Toolbox#readme>].

Filter-based feature selection techniques extract relevant features independently of classifier algorithms, mitigating bias, and overfitting. Their key benefit is computational efficiency, making them suitable for high-dimensional data [18]. RF is predominantly engineered for feature selection, rendering it a powerful tool for discerning pertinent features within high-dimensional datasets. However, RF necessitates class labels for feature selection, constraining its utility to supervised learning scenarios. On the other hand, TV was initially customized for text analysis and frequently serves as a method of feature selection. NCA could effectively reduce the dimensionality of the feature space while retaining vital information [19]. Nonetheless, NCA can be computationally demanding, especially in high-dimensional spaces, potentially restricting its application to large datasets.

Classification methods in machine learning

To classify patients of different disease severity, we explored the six most commonly used classification methods of clustering, including the binary decision tree for multiclass classification [17], K-nearest neighbor classification model [20], multiclass Naive Bayes model [21], Neural Network classification model [22], and multiclass models for SVM, with the two kernel

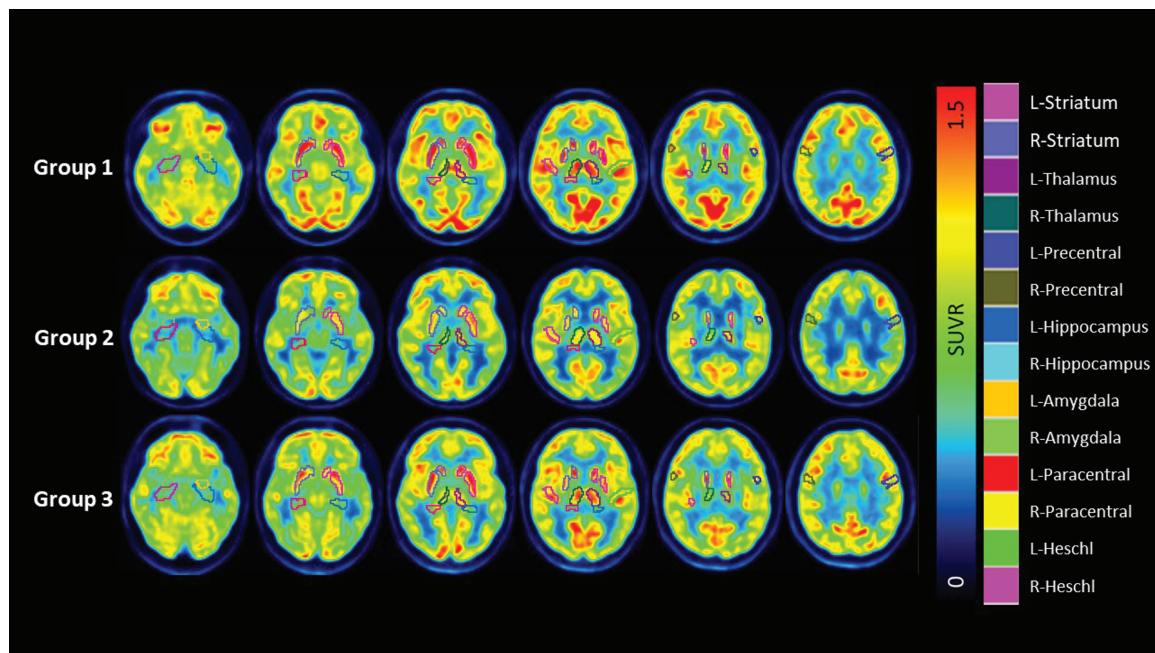


Figure 1. This displays three representative ^{18}F -FDG SUVR images overlaid with VOI masks in groups 1, 2, and 3 patients, respectively. The SUVR color bar is shown with the range [0, 1.5], and some of the subcortical VOIs are indicated with corresponding colors. VOI: volume of interest.

function based on Gaussian and Polynomial functions [23] from the use of the MATLAB 2022a Statistics and Machine Learning Toolbox (Natick, MA, USA: The MathWorks Inc). For the above six machine learning algorithms for classification, a hyperparameter optimization was automatically performed using Bayesian optimization [24, 25], as implemented in the Classification Learner app of the Statistics and Machine Learning Toolbox in MATLAB [26]. The objective of Bayesian optimization is to simultaneously minimize an objective function and reduce model loss by selecting optimal validation settings. Hyperparameters are intrinsic parameters of classifiers, such as the gamma parameter in a support vector machine, have a substantial impact on classifier performance. Nonetheless, optimizing these hyperparameters is often a challenging and time-consuming task. As a result, we employed a default configuration as provided in MATLAB, and that yields optimized results by identifying the optimal value that minimizes cross-validation loss.

For the available optimization options, the default values of 30 iterations and the 300-second training time limit, and the number of hyperparameters 10 for the grid search, were applied in this study. To perform these machine learning tasks, a computer with an Nvidia 1080ti GPU, dual-Xeon E5-2670 Intel CPUs, and 128 Gb RAM was applied along with MATLAB 2022a (Natick, MA, USA: The MathWorks Inc).

The overall procedure of feature selection and machine learning can be summarized as follows (Figure 2):

1. Feature Selection: Three feature selection methods, including Relief-F, TV, and NCA, were applied to evaluate the importance of features from high to low.
2. Model Optimization: A five-fold cross-validation approach was applied in the model optimization. The accuracy was then calculated from the top two to the total of 93 features

obtained from the three feature selection methods for six multiclass classifiers of the decision tree, K-nearest neighbor, Naïve Bayes, Neural Networks, SVM-Gaussian, and SVM-Polynomial. By default, the Classification Learner app in Matlab was utilized to perform a hyperparameter tuning for each classifier using the Bayesian optimization method [25, 26]. The final optimized classification model with associated feature selection was selected as the best validation model for this study by combining both the TV feature selection and the Multiclass Naïve Bayes classifier.

3. Application: Finally, the pretrained model will be applied as the final model for the classification of different disease severity groups using the fixed optimized hyperparameters obtained in the training process.

Statistical analysis

We used SPSS 26.0 to analyze the data. Demographic data were presented as percentages, means, and standard deviations. We used the chi-squared test for group percentage comparisons and the one-way ANOVA and post hoc analysis (Bonferroni test) to analyze the mean differences between the type 1, type 2, and type 1 in schizophrenia groups. A *p*-value of <0.05 was considered significant.

Results

A total of 314 participants with narcolepsy were enrolled, 204 had NT1 (group 1), 90 had NT2 (group 2), 20 had NT1 + SCZ (group 3), and 26 healthy controls (mean age 19.1 years, 38.5% females). The demographic and clinical data of the participants with narcolepsy are shown in Table 1. The BMI was significantly different between the groups ($p < 0.001$, $3 > 1 > 2$). Significantly fewer participants with NT2 (group 2) had hypnagogic and hypnopompic

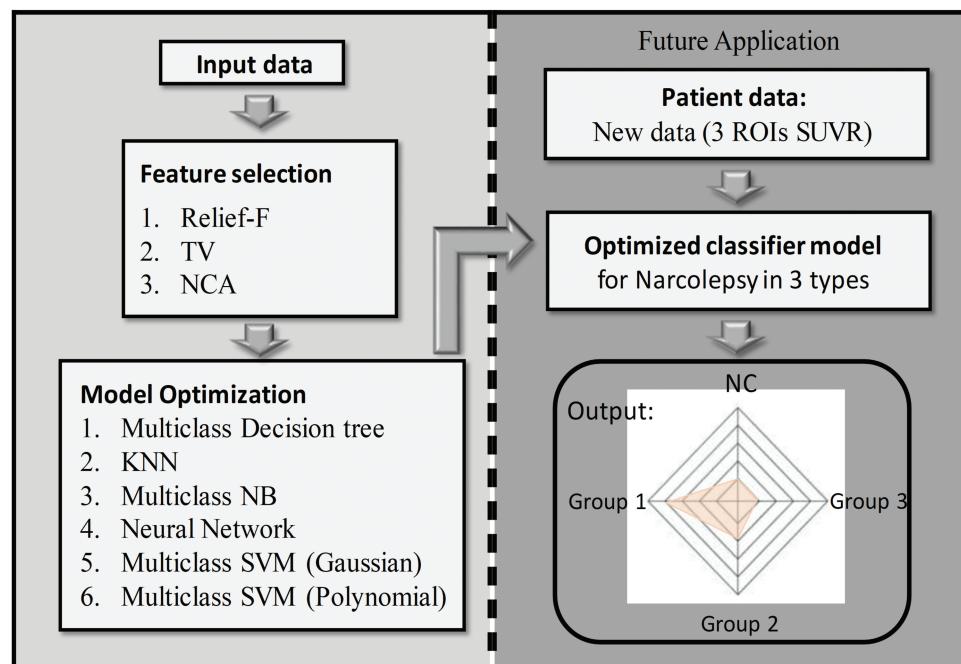


Figure 2. Processing contains four steps: the regional SUVR data is applied as input data in step 1. Three methods of feature selection to rank the importance of features are applied in step 2, while step 3 indicates the automatic optimization of hyperparameters for six classification models. In step 4, an optimized classifier model is chosen as the best model, and a new dataset is applied to the optimized model to predict the subtype of the disease. KNN, K-nearest neighbor; NB, naive bayes; NCA, neighborhood component analysis; ROIs SUVR, region-of-interest's standard uptake value ratio; SVM, support vector machines; TV, term variance.

Table 1. Demographic Data of Patients With Narcolepsy Type 1, Narcolepsy Type 2, and Narcolepsy With Comorbid Schizophrenia

	Group 1: type 1 narcolepsy (n = 204)	Group 2: type 2 narcolepsy (n = 90)	Group 3: type 1 and schizophrenia (n = 20)	P-value	Post hoc (Bonferroni)
Gender (male: %)	52.8%	62.2%	47.4%	0.256	—
Onset age of narcolepsy (y)	13.83 ± 5.77	14.88 ± 7.52	13.25 ± 5.22	0.431	—
Current age(y)	23.23 ± 9.65	24.00 ± 9.52	24.26 ± 6.54	0.771	—
BMI (kg/m ²)	24.00 ± 4.79	21.67 ± 4.26	27.62 ± 6.19	<0.001*	3 > 1 > 2
Symptom/sign					
Hypersomnia (%)	100%	100%	100%	—	—
Cataplexy (%)	75%	—	100%	—	—
Hypnagogic/hypnopompic hallucination (%)	71.2%	34.8%	88.9%	<0.001*	1,3 > 2
Sleep paralysis	72.3%	50.6%	88.9%	<0.001*	1,3 > 2
REM behavior symptom	9.9%	4.5%	5.6%	0.273	—
Sleep disturbance (insomnia)	33.3%	27.0%	27.8%	0.535	—
AH	12.6%	3.4%	77.8%	<0.001*	3 > 1 > 2
VH	8.9%	0.0%	55.6%	<0.001*	3 > 1 > 2
Delusion	3.1%	0.9%	72.2%	<0.001*	3 > 1,2
Psychiatric comorbidity					
MDD	22.4%	22.7%	50.0%	0.030*	3 > 1
ADHD	11.1%	25.8%	11.1%	0.005*	2 > 1
GAD	3.7%	6.7%	5.6%	0.518	—
Insomnia	24.9%	18.0%	27.8%	0.392	—
Tic	2.1%	3.4%	0.0%	0.631	—
Asperger's disorder	4.2%	4.6%	16.7%	0.070	—
Bipolar disorder	1.6%	1.1%	0.0%	0.834	—
Physical comorbidity					
Obesity	34.9%	15.7%	72.2%	<0.001*	3 > 1 > 2
DM	2.1%	3.4%	16.7%	0.005*	3 > 1
HTN	4.2%	2.2%	22.2%	0.001*	3 > 1,2
Thyroid disease	3.2%	4.6%	0.0%	0.598	—
OSA	22.9%	19.1%	38.9%	0.188	—
PLMD	6.4%	10.2%	0.0%	0.245	—
Asthma	6.9%	13.3%	11.1%	0.205	—
Nasal allergy	37.8%	39.3%	33.3%	0.888	—
HLA typing: HLA DQB1*0602	96.8%	31.3%	100%	<0.001*	1,3 > 2
MSLT:					
mean sleep latency	2.81 ± 2.49	5.04 ± 3.14	3.04 ± 2.71	<0.001*	2 > 1,3
SOREMs	3.89 ± 1.21	3.23 ± 1.51	3.83 ± 1.20	0.001*	1 > 2
PSG REM latency	60.46 ± 75.40	76.15 ± 64.73	75.96 ± 88.20	0.310	—

Chi-squared test, one-way ANOVA, and post hoc analysis; *p-value < 0.05. BMI, body mass index; REM, rapid eye movement; AH, auditory hallucination; VH, visual hallucination; ADHD, attention-deficit/hyperactivity disorder; DM, diabetes mellitus; GAD, generalized anxiety disorder; HTN, hypertension; PLMD, periodic limb movement disorder; OSA, obstructive sleep apnea; MDD, major depressive disorder; MSLT, multiple sleep latency test; SOREM, sleep-onset rapid eye movement; PSG, polysomnography.

hallucinations and sleep paralysis, ($p < 0.001$, 1, 3 > 2). Significantly more participants with NT1 + SCZ (group 3) had auditory/visual hallucinations ($p < 0.001$, 3 > 1 > 2) and delusions ($p < 0.001$, 3 > 1, 2). Significantly more participants with NT1 + SCZ (group 3) had major depressive disorders ($p = 0.03$, 3 > 1), obesity ($p < 0.001$, 3 > 1 > 2), diabetes mellitus ($p = 0.005$, 3 > 1, 2) and hypertension (HTN) ($p = 0.001$, 3 > 1, 2). Significantly more participants

with NT2 (group 2) had attention deficit hyperactivity disorders ($p = 0.005$, 2 > 1).

Significantly more participants with a positive HLA DQB1*0602 were noted in the NT1 and NT1 + SCZ groups (groups 1 and 3) ($p < 0.001$, 1, 3 > 2). The sleep studies showed that the mean sleep latency by MSLT of participants with NT1 and NT1 + SCZ (groups 1 and 3) was significantly higher ($p < 0.001$, 2 > 1, 3), and the NT1

Table 2. Performance Measures for all Classification and Feature Selection Methods. The “No.ROI (Min)/Acc. % (Max)” Indicates the Maximum Accuracy of Acc. % (Max) for Each Classification Model (of Six Models) Using Corresponding Minimal Number of Features by No.ROI (Min), Where the Features Were Selected From the Three Associated Feature Selection Approaches

No.ROI (Min)/Acc. % (Max)	Feature selection methods		
	RF	TV	NCA
Classification methods	Dtree	3/65.92	89/63.70
	KNN	23/99.65	17/99.66
	NB	6/99.60	3/99.60
	NN	25/99.68	27/99.64
	SVM-G	11/99.62	13/99.61
	SVM-P	23/99.61	25/99.61

Acc., accuracy; Dtree, binary decision tree for multiclass classification; KNN, k-nearest neighbor classification model; NB, multiclass naive Bayes model; NCA, neighborhood component analysis; NN, neural network classification model; RF, relief-F; ROI, region of interest; SVM-G, support vector machine-Gaussian; SVM-P, support vector machine-Polynomial; TV, term variance.

Table 3. The Best Selected Features by RF, TV, and NCA

Multiclass NB	Feature selection methods		
	RF	TV	NCA
No.ROI/Acc.	6/99.60	3/99.60	6/99.60
Choose the best features	1st ROI	Left basal ganglia	Left basal ganglia
	2nd ROI	Left thalamus	Left Heschl
	3rd ROI	Left precentral	Left striatum
	4th ROI	Left Cb	Right Heschl
	5th ROI	Left paracentral	Left hippocampus
	6th ROI	Right thalamus	Right amygdala

Acc., accuracy; Cb, cerebellum; NB, naive bayes; NCA, neighborhood component analysis; RF, relief-F; ROI, region of interest; TV, term variance. Bold values refer to the ROI selected to achieve the goal of high classification accuracy.

group (group 1) had significantly more SOREMs than the NT2 group (group 2) ($p = 0.001$, $1 > 2$).

In this study, a comparative analysis of the abovementioned six classifiers was performed on the dataset of 314 participants and 93 features. Table 2 displays the classification results from the six classifiers using the narcolepsy dataset from three different feature selection methods, each showing the results of the minimum number of regions of interest (ROIs) and the maximum accuracy value. There are five classification methods with high accuracy values (>99%) of only a small number of ROIs, except the binary decision tree.

Based on the criteria of the minimal number of ROIs and maximal accuracy, the naïve Bayesian classifier generates the best results with fewer features selected from the RF, TV approach, and NCA with featured ROI numbers of 6, 3, and 6, respectively, where accuracy can reach 99.6%.

For the final optimized multiclass Naïve Bayes classifier, the results for the optimized features and accuracy of the three featured selection methods (RF, TV, and NCA) are shown in Table 3. The top six most important ROIs are listed for each feature selection method. To apply a small number of ROIs and high classification accuracy, the Naïve Bayes classifier with the feature selection TV achieves the goal with only three ROIs (left basal ganglia, left Heschl, and left striatum) and an accuracy of more than 99%.

Discussion

The overall results of this study are promising, and the predictive models can serve as a good tool for sleep medicine specialists to

facilitate narcolepsy diagnosis. Machine learning has, in recent years, been widely applied to different medical fields and previous studies have used different machine learning methods in the diagnosis of narcolepsy. A neural network analysis is proven to efficiently diagnose narcolepsy from non-narcoleptics, by analyzing the PSG sleep stages of patients with narcolepsy [27]. Another study that analyzed the clinical features of narcolepsy from the European Narcolepsy Network database using the Stochastic Gradient Boosting model also reported fair results [28], and found that cataplexy and MSLT findings can contribute to the subtype classification of narcolepsy. Our findings identified features with our PET data and proved that PET scans can be useful in the diagnosis of narcolepsy when facilitated by machine learning methods, further expanding the application of machine learning in diagnosing narcolepsy.

The major clinical challenge of narcolepsy diagnosis is differentiating between NT1 and NT2. The severity and presentation of cataplexy vary individually and can be absent or trivial in the early stages of the disease [29]. Patients with narcolepsy may not notice cataplexy themselves, and younger children may have difficulty describing their symptoms due to their limited speech ability. Studies have also found that the severity and frequency of cataplexy declines in older ages [30], demonstrating the potential challenges of differential diagnoses in adulthood by clinical symptoms. Although low CSF hypocretin levels can accurately diagnose NT1 from NT2, it is not widely used due to low availability and concerns about lumbar punctures. Our model achieved an accuracy of more than 99%. By using this model, clinicians can make a differential diagnosis of narcolepsy subtypes more efficiently, and the model

can be invaluable when diagnosing NT1 in those with mild cataplexy, of a younger age, or with limited speech ability.

It is crucial to diagnose schizophrenia among narcoleptic patients, as patients with a dual diagnosis are more obese and depressed [31], and have poorer neurocognitive functions than those with only schizophrenia or narcolepsy [11, 13]. They require more medical attention, so an early diagnosis and intervention is necessary. Although the comorbidity rate of schizophrenia in narcoleptic patients is still uncertain, it is been reported to be as high as 14.1% in young narcoleptic patients [12]. It can be unrecognized, and diagnosis can then be delayed. Sleepiness could be a sedative effect of antipsychotics and other commonly prescribed psychotropics for patients with schizophrenia [32], and some anti-dopaminergic medications have been reported to mask cataplexy in narcoleptic patients [33]. Furthermore, there are overlapping symptoms between the two diagnoses, such as hallucinations and hypnagogic/hypnopompic hallucinations. Clinically, close monitoring in a psychiatric or neurology ward is sometimes warranted for diagnoses in difficult cases [13], but this is time and resource-consuming and can be unacceptable for patients and their families. Previous studies found significant differences in the several brain regions between NT1 and NT2, but the only significant difference between NT1 and NT1 + SCZ was hypometabolism in the bilateral lingual [13]. Narcolepsy and schizophrenia may also share some physiopathological factors, supported by similar PET findings of hypermetabolism of the striatum, basal ganglia, and thalamus in patients with NT1 and patients with SCZ [13]. Dopamine dysfunction relates to the pathophysiology of schizophrenia, and its function also regulates sleep and wakefulness [34]. Our model, with a fair accuracy of the differential diagnosis of the dual diagnosis group, can support clinicians in the diagnosis of comorbid schizophrenia among patients with narcolepsy.

Our analysis by feature selection found specific brain regions for different classification methods, including basal ganglia, thalamus, precentral gyrus, Heschl, striatum, and hippocampus. The mechanism of the sleep and wake cycle is currently not fully understood, but there's no doubt that the thalamus can regulate waking and sleeping states by acting as an integrator of subcortical sleep-wake inputs [35]. Although basal ganglia and striatum are known for regulating motor functions, habit formations, emotions, and addictive behaviors, they also maintain wakefulness and suppress sleep to achieve their fundamental functions [36]. Besides, a significant hypometabolism of the striatum, hippocampus, thalamus, and basal ganglia has been reported to a greater extent in patients with NT2 than those with NT1 [11]. However, the hypothalamus is regarded as an important brain region related to narcolepsy but is not included as a feature in our model. It is clear that NT1 is caused by an early loss of the hypothalamus neurons that produce hypocretin [37]. An autopsy study also revealed the destruction of the hypocretin neurons in the lateral hypothalamus [38]. The low ranking of the hypothalamus in all three feature selection methods can be explained as all the participants have narcolepsy and the hypothalamus can be influenced in all the participants. This important brain region can relate to the common pathophysiology of hypersomnia of the three groups. Thus, it does not assist with differential diagnosis, while the impacts on the other brain regions by different subtypes of narcolepsy differ and lead to different clinical presentations. Our previous PET studies also never found any differences in the hypothalamus between the three groups of narcoleptic patients [11, 13]. We should note that feature selection reveals the contribution of single brain regions in the diagnosis of narcolepsy, but the analysis by machine learning is often not interpretable.

Although the cost-effectiveness of PET in diagnosing narcolepsy is not clear, it is more available in different areas than hypocretin tests, less invasive, and more acceptable for patients, compared to lumbar punctures. Besides, it can also help to rule out other brain-related conditions which may contribute to narcolepsy-like symptoms. The effectiveness of the combined PET and machine learning can be invaluable, and the analysis of brain images by machine learning can also be applied to other sleep disorders and neurological diseases [39, 40]. Future implications of this model are of interest. Patients who are initially diagnosed with NT2 may progress to have cataplexy, and, in our experience, some patients can develop comorbid schizophrenia years after their narcolepsy diagnosis. Prospective studies applying PET scans and developing predictive models on patients with narcolepsy may have the potential to diagnose these changes earlier; early interventions and treatments could be provided which may be able to prevent disease progression.

This study has several limitations. First, the participants of this study are relatively young and are in the earlier stages of their respective diseases. Hence, we need to be cautious when applying our results to different age groups. Second, most participants received a PET study before initiating medication treatment, and some received medication before enrollment. Furthermore, due to the risk of relapse, antipsychotics were not discontinued before the examination. Medication usage can have an impact on the results of the PET scans, although there was an adequate drug-free interval before the examinations of most medications, except for antipsychotics. However, long-term medication effects on the brain can be minimized by selecting younger participants. Third, we did not include other potential impacts of other factors that are significantly different between groups, such as age, BMI, and obesity on the predictive models. Besides, comorbidities are not excluded and can also impact PET scan results, such as major depressive disorders, attention deficit hyperactivity disorders, Asperger's disorder, and tic disorder. A previous work has applied these clinical factors for differentiation among the narcolepsy subgroups but the classification accuracy is relatively lower [41]. Fourth is the imbalanced sample size of the three groups since type 1 narcolepsy is more prevalent than type 2 narcolepsy and narcolepsy with comorbid schizophrenia. Fifth, there is no external validation of the predictive models. Despite these limitations, our study is the first to analyze the PET findings of patients with NT1, NT2, and NT1 + SCZ by feature selection and machine learning, which can be helpful in future differential diagnoses of narcolepsy. We have published several studies of PET findings among different subtypes of narcolepsy before, and with accumulating data in recent years, the novelty of the current study is the potential for faster and more convenient differential diagnosis by predictive models. It is also important to note that the optimal combination of selected features and classification methods may vary depending on specific data and requirements, so multiple iterations are necessary to achieve the best solution. Further studies with larger sample sizes and more diverse data sources may be necessary to validate and refine the predictive models.

Note that due to small sample size, the results reported in this study were obtained using the original samples based on the five-fold cross-validation in the model optimization step. Due to rare cases of these diseases, thus the same size in each group is imbalanced, and there is no extra cohort available for independent validation now. Nevertheless, to mitigate the imbalance dataset in the three groups, as shown in the supplementary, we have also applied resampling approach based on standard uptake value ratio permutations to augment the initial dataset while

retaining key statistical properties of the original dataset [42] in the second step for model optimization. Note that the permutation augmentation can be repeatedly applied to the dataset without duplicating entries, thereby preventing overfitting [43]. We expanded the original data for each group into 500 pieces of data through permutation, so that each group has the same amount of data. Then the optimized classification models (TV + NB) and the selected features (left basal ganglia, left Heschl, left striatum) were applied to the augmented dataset, and the resulted accuracy, sensitivity, specificity, or recall rate are like those in the original dataset as shown in Supplementary Table S1 and Table 2. Ultimately, extra data is needed for external validation of the optimal model to have confidence in the performance of the proposed optimal model and the reported results in the future.

In conclusion, the accuracy of our predictive model using PET data are promising and can facilitate clinicians to diagnose narcolepsy subtypes. Brain regions selected by machine learning methods can be investigated to clarify their roles in the pathophysiology of different narcolepsy subtypes. Based on the findings of this study, future research with a larger sample size could further refine the predictive model of narcolepsy.

Supplementary Material

Supplementary material is available at SLEEP online.

Acknowledgments

This study was partially supported by Taiwan's Ministry of Science and Technology grant # MOST 107-2314-B-182A-128-MY2 and 109-2314-B-182A-112-MY3 and Chang Gung Memorial Hospital Research Grants #CMRPG 3M0051 to YS Huang and #CMRPG 3L0291 and 3L0292 to Wei-Chih Chin.

Author Contributions

Prepared the initial draft of the manuscript: W.C. Chin, S.Y. Huang. Conduct the statistical analyses: S.Y. Huang, I.Tang, I.T. Hsiao. Conceive the study design: Y.S. Huang, I.T. Hsiao, F.Y. Liu, C.H. Wang. Provide the interpretation of the data: W.C. Chin, S.Y. Huang. Collect study data: Y.S. Huang, F.Y. Liu, I.T. Hsiao. Edit the final manuscript: W.C. Chin, Y.S. Huang.

Disclosure Statement

Financial disclosure: The authors declare no competing financial interests. Nonfinancial disclosure: The authors declare no competing nonfinancial interests.

Data Availability

Anonymized data that support the findings of this study are available on reasonable request from the corresponding author, YS Huang.

References

1. Silber MH, Krahn LE, Olson EJ, Pankratz VS. The epidemiology of narcolepsy in Olmsted County, Minnesota: a population-based study. *Sleep*. 2002;25(2):197–202. doi: [10.1093/sleep/25.2.197](https://doi.org/10.1093/sleep/25.2.197)
2. American Academy of Sleep Medicine. *The International Classification of Sleep Disorders: Diagnostic and Coding Manual*. 2nd Ed. Westchester, IL: AASM; 2005.
3. American Academy of Sleep Medicine. *International Classification of Sleep Disorders*. 3rd ed. Darien, IL: AASM; 2014.
4. Nishino S, Okuro M, Kotorii N, et al. Hypocretin/orexin and narcolepsy: New basic and clinical insights. *Acta Physiol (Oxf)*. 2010;198(3):209–222. doi: [10.1111/j.1748-1716.2009.02012.x](https://doi.org/10.1111/j.1748-1716.2009.02012.x)
5. Dye TJ, Gurbani N, Simakajornboon N. Epidemiology and pathophysiology of childhood narcolepsy. *Paediatr Respir Rev*. 2018;25:14–18. doi: [10.1016/j.prrv.2016.12.005](https://doi.org/10.1016/j.prrv.2016.12.005)
6. Juji T, Satake M, Honda Y, Doi Y. HLA antigens in Japanese patients with narcolepsy All the patients were DR2 positive. *Tissue Antigens*. 1984;24:316–319. doi: [10.1111/j.1399-0039.1984.tb02144.x](https://doi.org/10.1111/j.1399-0039.1984.tb02144.x)
7. Thannickal TC, Moore RY, Nienhuis R, et al. Reduced number of hypocretin neurons in human narcolepsy. *Neuron*. 2000;27:469–474. doi: [10.1016/s0896-6273\(00\)00058-1](https://doi.org/10.1016/s0896-6273(00)00058-1)
8. Tandon R, Keshavan MS, Nasrallah HAS. “just the facts” what weknow in 2008 2 Epidemiology and etiology. *Schizophr Res*. 2008;102(1-3):1–18.
9. Ruoff CM, Reaven NL, Funk SE, et al. High rates of psychiatric comorbidity in narcolepsy: Findings from the Burden of Narcolepsy Disease (BOND) study of 9,312 patients in the United States. *J Clin Psychiatry*. 2017;78(2):171–176. doi: [10.4088/JCP.15m10262](https://doi.org/10.4088/JCP.15m10262)
10. Douglass AB, Hays P, Pazderka F, Russell JM. Florid refractory schizophrenias that turn out to be treatable variants of HLA-associated narcolepsy. *J Nerv Ment Dis*. 1991;179(1):12–7; discussion 18. doi: [10.1097/00000503-199101000-00003](https://doi.org/10.1097/00000503-199101000-00003)
11. Huang YS, Hsiao T, Liu FY, et al. Neurocognition, sleep, and PET findings in type 2 vs type 1 narcolepsy. *Neurology*. 2018;90(17):e1478–e1487.
12. Huang YS, Liu FY, Lin CY, Hsiao IT, Guilleminault C. Brain imaging and cognition in young narcoleptic patients. *Sleep Med*. 2016;24:137–144. doi: [10.1016/j.sleep.2015.11.023](https://doi.org/10.1016/j.sleep.2015.11.023)
13. Chin WC, Liu FY, Huang YS, Hsiao IT, Wang CH, Chen YC. Different positron emission tomography findings in schizophrenia and narcolepsy type 1 in adolescents and young adults: A preliminary study. *J Clin Sleep Med*. 2021;17(4):739–748. doi: [10.5664/jcsm.9032](https://doi.org/10.5664/jcsm.9032)
14. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, 5th ed (DSM-5). Washington, DC: Author. 2013.
15. Joshi A, Koeppe RA, Fessler JA. Reducing between scanner differences in multi-center PET studies. *Neuroimage*. 2009;46(1):154–159. doi: [10.1016/j.neuroimage.2009.01.057](https://doi.org/10.1016/j.neuroimage.2009.01.057)
16. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15(1):273–289. doi: [10.1006/nimg.2001.0978](https://doi.org/10.1006/nimg.2001.0978)
17. Quinlan JR. Induction of decision trees. *Mach Learn*. 1986;1(1):81–106. doi: [10.1007/bf00116251](https://doi.org/10.1007/bf00116251)
18. Pudjihartono N, Fadason T, Kempa-Liehr AW, O'Sullivan JM. A review of feature selection methods for machine learning-based disease risk prediction. *Front Bioinform*. 2022;2:927312. doi: [10.3389/fbinf.2022.927312](https://doi.org/10.3389/fbinf.2022.927312)
19. Malan NS, Sharma S. Feature selection using regularized neighbourhood component analysis to enhance the classification performance of motor imagery signals. *Comput Biol Med*. 2019;107:118–126. doi: [10.1016/j.combiomed.2019.02.009](https://doi.org/10.1016/j.combiomed.2019.02.009)
20. Altman NS. An introduction to kernel and nearest-neighbor non-parametric regression. *The American Statistician*. 1992;46(3):175–185. doi: [10.2307/2685209](https://doi.org/10.2307/2685209)
21. Možina M, Demšar J, Kattan M, Zupan B. Nomograms for visualization of naive Bayesian classifier, in Proceedings of the 8th European

- Conference on Principles and Practice of Knowledge Discovery in Databases. Springer-Verlag: Pisa, Italy. 2004:337–348.
22. Zhang GP. Neural networks for classification: A survey. *IEEE Trans Syst Man Cybern C Appl Rev*. 2000;30(4):451–462. doi: [10.1109/5326.897072](https://doi.org/10.1109/5326.897072)
 23. Daqi G. and Tao Z. Support vector machine classifiers using RBF kernels with clustering-based centers and widths. in 2007 International Joint Conference on Neural Networks. Orlando, Florida, USA, 2007.
 24. Snoek J, Larochelle H, Adams RP. Practical bayesian optimization of machine learning algorithms. *Adv Neural Info Process*. 2012;25:2960–2968.
 25. Passos D, Mishra P. A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks. *Chemometr Intell Lab Syst*. 2022;223:104520.
 26. Isabona J, Imoize AL, Kim Y. Machine learning-based boosted regression ensemble combined with hyperparameter tuning for optimal adaptive learning. *Sensors (Basel)*. 2022;22(10):3776. doi: [10.3390/s22103776](https://doi.org/10.3390/s22103776)
 27. Stephansen JB, Olesen AN, Olsen M, et al. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. *Nat Commun*. 2018;9(1):1–15.
 28. Zhang Z, Mayer G, Dauvilliers Y, et al. Exploring the clinical features of narcolepsy type 1 versus narcolepsy type 2 from European Narcolepsy Network database with machine learning. *Sci Rep*. 2018;8(1):1–11.
 29. Dauvilliers Y, Siegel JM, Lopez R, Torontali ZA, Peever JH. Cataplexy—clinical aspects, pathophysiology and management strategy. *Nat Rev Neurol*. 2014;10(7):386–395. doi: [10.1038/nrneurol.2014.97](https://doi.org/10.1038/nrneurol.2014.97)
 30. Kanayashi T, Yano T, Ishiguro H, et al. Hypocretin (orexin) levels in human lumbar CSF in different age groups: Infants to elderly persons. *Sleep*. 2002;25:337–339.
 31. Huang YS, Guillemainault C, Chen CH, Lai PC, Hwang FM. Narcolepsycataplexy and schizophrenia in adolescents. *Sleep Med*. 2014;15(1):15–22. doi: [10.1016/j.sleep.2013.09.018](https://doi.org/10.1016/j.sleep.2013.09.018)
 32. Fang F, Sun H, Wang Z, Ren M, Calabrese JR, Gao K. Antipsychotic drug induced somnolence: Incidence, mechanisms, and management. *CNS Drugs*. 2016;30(9):845–867. doi: [10.1007/s40263-016-0352-5](https://doi.org/10.1007/s40263-016-0352-5)
 33. Okura M, Riehl J, Mignot E, Nishino S. Sulpiride, a D2/D3 blocker, reduces cataplexy but not REM sleep in canine narcolepsy. *Neuropsychopharmacology*. 2000;23(5):528–538. doi: [10.1016/S0893-133X\(00\)00140-8](https://doi.org/10.1016/S0893-133X(00)00140-8)
 34. Dauvilliers Y, Tafti M, Landolt HP. Catechol-O-methyltransferase, dopamine, and sleep-wake regulation. *Sleep Med Rev*. 2015;22:47–53. doi: [10.1016/j.smrv.2014.10.006](https://doi.org/10.1016/j.smrv.2014.10.006)
 35. Gent TC, Bassetti CL, Adamantidis AR. Sleep-wake control and the thalamus. *Curr Opin Neurobiol*. 2018;52:188–197. doi: [10.1016/j.conb.2018.08.002](https://doi.org/10.1016/j.conb.2018.08.002)
 36. Lazarus M, Chen JF, Urade Y, Huang ZL. Role of the basal ganglia in the control of sleep and wakefulness. *Curr Opin Neurobiol*. 2013;23(5):780–785. doi: [10.1016/j.conb.2013.02.001](https://doi.org/10.1016/j.conb.2013.02.001)
 37. Nishino S, Ripley B, Overeem S, Lammers GJ, Mignot E. Hypocretin (orexin) efficiency in human narcolepsy. *Lancet*. 2000;355(9197):39–40. doi: [10.1016/S0140-6736\(99\)05582-8](https://doi.org/10.1016/S0140-6736(99)05582-8)
 38. Blouin AM, Thannickal TC, Worley PF, Baraban JM, Reti IM, Siegel JM. Nar immunostaining of human hypocretin (orexin) neurons: Loss in narcolepsy. *Neurology*. 2005;65(8):1189–1192. doi: [10.1212/01.wnl.0000175219.01544.c8](https://doi.org/10.1212/01.wnl.0000175219.01544.c8)
 39. Gautam R, Sharma M. Prevalence and diagnosis of neurological disorders using different deep learning techniques: A meta-analysis. *J Med Syst*. 2020;44(2):49. doi: [10.1007/s10916-019-1519-7](https://doi.org/10.1007/s10916-019-1519-7)
 40. Vieira S, Pinaya WH, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neurosci Biobehav Rev*. 2017;74:58–75. doi: [10.1016/j.neubiorev.2017.01.002](https://doi.org/10.1016/j.neubiorev.2017.01.002)
 41. Syu HY (2018). Prediction Model of Narcolepsy Based on Ensemble Learning Approach [Master thesis]. National Taipei University of Nursing and Health Sciences, Taipei]. Taipei, Taiwan: National Digital Library of Theses and Dissertations in Taiwan. <https://hdl.handle.net/11296/2d9ud3>
 42. Iglesias G, Talavera E, González-Prieto A, Mozo A, Gómez-Canaval S. Data augmentation techniques in time series domain: A survey and taxonomy. *Neural Comput & Applic*. 2023;35:10123–10145. doi: [10.1007/s00521-023-08459-3](https://doi.org/10.1007/s00521-023-08459-3)
 43. Wei J, Zou K, Eda: Easy data augmentation techniques for boosting performance on text classification tasks. *arXiv preprint arXiv* 2019;1901:11196.