Comparing Classification via Regression and Random Committee for Automatic Sleep Stage Classification in Autism Patients

IN Yulita1,2*, MI Fanany1, AM Arymurthy1
1Faculty of Computer Science, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia
2Departement of Computer Science, Universitas Padjadjaran, Jl. Raya Bandung – Sumedang Km. 21, Sumedang 45363, Indonesia

Email: intanurma@gmail.com*

Abstract. The prevalence of autism children has increased rapidly in the last few periods. There is no cure for autism. But the management and treatment of accompanying medical conditions can be done. One of the effects of his medical condition is a sleep disorder. But children with autism have difficulty communicating the disorders they experience. In medicine, the detection of sleep disorders can be done through a test called polysomnography. One of the purposes of this test is to find the patient's sleep patterns through the sleep stage classification. But the doctors need several days to analyze each test. This study proposes an application that can classify it automatically. The method used was based on machine learning. The two classifiers were classification via regression and random committee. The both performances were compared in sleep stages classification for the autism patients. The result showed that random committees had a higher performance than classification via regression. Its performance got more than 85% for accuracy, precision, recall, and F-measure. This study also implemented resampling to overcome imbalance class problems. It can be seen that this process was useful in improving the performance of both classifiers.

1. Introduction
During the first few months of life, babies enter the normal sleep cycle. They gradually reduce the amount of naps they need and start sleeping for a longer time at night. But some children continue to experience difficulty sleeping. The problems can last a long time. Moreover this sleep disorders are more common in children who suffer from autism (40-80%) [1]. Autism is a severe disorder of the neurobiology development in children. The disorder causes problems to communicate and relate to the social environment [2]. The signs / symptoms apparent before they are 3 years old, and then continues into adulthood if no appropriate intervention is carried out.

People with autism have disorders in social interaction, communication, imagination, and repetitive (repetitive) behavior patterns, and resistance (not easy to follow) to changes in routine [3]. This interference in social interaction causes them to look strange and different from other people / children. Their communication disorders occur in verbal communication (verbal / with words) and non-verbal (do not understand the meaning of gestures, facial expressions, and tone / color / voice intonation). This disturbance in imagination causes children difficulty in terms of activity and play.
Playing and doing activities are different from other people/children. For example, they just imitate and follow things rigidly and repeatedly.

Sleep disorders that commonly occur in children with autism include difficulty in starting to sleep, and often awaken or easily awakened [4]. The lack of sleep in autism children greatly affects various symptoms that often arise in autism children, namely: aggressive behavior, depression, hyperactive behavior, irritability, and low learning and thinking skills. It can also indirectly interfere with sleep for his family. But their weak communication skills make it difficult for their families to overcome their problems. The alternative solution that can be done is through recording the sleep data, which is often called polysomnography.

Polysomnography is a health test used for the diagnosis of sleep disorders. This test is designed to record brain wave activity, blood oxygen levels, heart rate and breathing, also eye and foot movements while the patient asleep [5]. Tests that are able to recognize the disturbances to normal sleep conditions of these patients help doctors diagnose abnormalities so as to facilitate the design of the treatment program needed.

After recording, the doctor analyses the sleep data. One of the target analysis is to identify the patient's sleep patterns. This patient's sleep pattern is obtained through the identification of sleep stages [6]. There are two different levels that alternate in the sleep cycle and reflect different levels of neuron activity. Each level is characterized by the activity of different types of brain waves (electrical activity is recorded with the help of head-placed electrodes). Sleep levels consist of slow eye movement (non-rapid eye movement/NREM) and fast eye movement sleep (rapid eye movement/NREM. NREM consists of a number of stages namely N1, N2, and N3) [7]. NREM and REM stages repeatedly during sleep at night. Stage N1, N2, and N3 are followed by REM. A complete sleep cycle, from N1 to REM, usually takes about one and a half hours.

N1 is a stage of light sleep. It is considered a transition between wakefulness and sleep. During this stage, the muscles begin to relax. It usually accounts for 5-10% of sleep time. An individual can easily wake up at this stage. N2 represents 40-50% of sleep time. During N2, the brain waves slow down with occasional rapid wave bursts. The eye movements stop at this stage. In N3, very slow brain waves called delta waves begin to appear. They are interspersed with smaller and faster waves. This stage represents around 20% of sleep time. This stage is sometimes called deep sleep, where all eye and muscle movements stop. It's hard to wake someone up during these two stages. If someone wakes up while sleeping soundly, he does not immediately adjust and often feels nervous and confused for several minutes after wake up. Some children experience bedwetting, nightmares (sleepwalking) during deep sleep.

REM represents 20% to 25% of sleep time. This stage follows NREM and occurs 4-5 times during the normal sleep period of 8-9 hours. The first REM period during sleepless nights may last less than 10 minutes, while the latter may exceed 60 minutes. In normal night sleep, REM occurs every 90 minutes. If the person is very sleepy, the duration of each REM is very short or maybe even nonexistent. REM sleep is usually associated with dreaming. During REM, the eyeball moves quickly, the heart rate and breathing become fast and irregular, and blood pressure increases. The muscles of the body are almost paralyzed. The brain is very active during REM sleep, and overall brain metabolism can increase by 20%. The electrical activity recorded in the brain during REM sleep is similar to that recorded when awake.

The doctors need a long time to analyze polysomnography. The time efficiency can be done through automatic sleep stage classification. A number of researchers have conducted research related to this. The methods have been carried out including fast convolutional [8], deep belief networks [9], Bayesian approach [10]. Until now, there have been no studies using datasets originating from autism patients. Therefore, this study developed it. The classification method is based on two methods, namely classification via regression and random committee. Furthermore, the two performance of this algorithm are compared.
2. Research method
This study aims to classify the sleep stages in autism patients. The classification is based on classification via regression and random committee. The research process is shown in Figure 1.

![Figure 1. The methodological flow of sleep stage classification.](image)

2.1. Data collecting
The recording was carried out from October to December 2018. Thirteen respondents were involved in this recording. But only seven patients whose data can be used in this study. The rest was due to failure during the recording process. The main cause of failure was tantrum experienced by patients. The recording was done overnight for each patient at the Mitra Keluarga Kemayoran hospital, Jakarta. If the recording had been done, the doctor analysed the recording. The scoring to identify the sleep stages used the AASM rule. Even though many signals were generated during recording. But this study only uses electroencephalography (EEG) data as much as one channel, namely C4 channel.

EEG is a method for recording electrical activity along the scalp. It measures the voltage fluctuations produced by ion currents in brain neurons [11]. The signals are recorded over a period of time that originate from a device called a brain-computer interface (BCI). The electrical signals produced by the brain will be captured by the channel (in the form of electrodes) in BCI. The measurement is done by placing electrodes on the scalp. Each electrode is represented by letters and numbers. The letter indicates the head area on the electrode, such as F and T are frontal and temporal lobe, respectively. Even numbers indicate the right side of the brain and odd numbers are the left sides. Based on the frequency, voltage amplitude, and condition of the object, EEG signals can be divided into 4 waves, namely delta waves (less than 4 Hz), theta (4 - 7 Hz), alpha (8 - 12 Hz), and beta (13 - 49 Hz). The wave is useful for identifying the stages of sleep. But other features are still needed to get more accurate results.

2.2. Preprocessing
This process consists of two main stages, namely segmentation and filtering. Segmentation is a method for dividing data into several windows. In this study, the type of segmentation was constant segmentation, where data were be divided into several windows with the same length. The length in each window was 30 seconds. In many cases, the required signal often mixes with unwanted signals (noise). To get rid of these unwanted signals, this study used several filters, namely band-pass and notch filter. Band-pass filtering aimed to pass one frequency band from Fourier transforms, while notch filter to reject certain frequencies. This study rejected signals with a frequency of 50 Hz through a notch filter. The band-pass filter value in this study was 0.3 half of the sample rate. With the combination of these two filters, the filtered signal was based on the specifications needed in this study.
2.3. Feature extraction
This stage aimed to find features that were able to distinguish characteristics among sleep stages [12].
The input data was the data segmentation. Each data were extracted using a number of methods,
namely fast Fourier transform, dual tree complex wavelet transform with two levels, kurtosis, spectral
mean, fractal exponent, and entropy. From this process, each segment data had 17 real valued features.

2.4. Imbalance class handling
Machine learning is designed to generalize data tested as equal and produce the simplest hypothesis.
When this algorithm tests datasets with class imbalanced problem, it tends to focus on the majority
class and ignores the minority class. It causes errors in the minority class classification. The minority
classes are only considered noise. The results in an unbalanced dataset usually have the characteristic
of a miss-classification in the minority class. Therefore, the challenge in overcoming this problem is
how to classify minority classes more accurately. One way to overcome is to resample the original
dataset, both in minority classes (oversampling) or in the majority class (under-sampling) [13]. This
study implemented oversampling so that the amount of minor data will be increased to almost the
same as the amount of major class data. The mechanism was done by taking data randomly for each
stage of sleep so that it reaches the expected threshold. Table 1 shows the proportion of original data
and after resampling. The increase in the amount of data was seen in N1 from 134 to 1207.

| Table 1. The data proportion for each stage in dataset. |
|--------------------------------------------------------|
| Stage | Number of data |
|       | Original | Resampling |
| Awake | 683 | 1207 |
| N1    | 134 | 1207 |
| N2    | 3054 | 1207 |
| N3    | 1429 | 1207 |
| R     | 737 | 1207 |

2.5. Classification
The classification in this study based on classification via regression, and random committee. The
classification via regression is a classification method that can transform problems into regression
functions [14]. This method combines the principles of the decision tree algorithm and linear
regression on several sub-trees (leaves) that are built. This method has two main steps namely [15]:
1. Make an ordinary decision tree, by maximizing the separation of criteria / parameters / attributes
   and their variations according to the target / output values. In making a decision tree,
   it must be calculated the deviation reduction standard.
2. Trimming the decision tree (pruning) on several possible sub-trees, and filling it with the
   regression function (linear model) accordingly, usually on leaves.

The second method was random committee. This method is an ensemble mechanism that is built
from a number of weak hypotheses [16]. The weak hypothesis is a prediction of a base classifier. The
base classifier is generally a tree. Each classifier uses the same data but based on a different random
number seed. Final decisions are calculated from the average predictions generated by each weak
hypotheses [17]. In this study, the weak classifier of classification via regression and random
committee were M5 and random tree, respectively. The model was trained as many as 10 iterations
and the batch size was 100.

2.6. Evaluation
The testing was done using 10-cross validation. Cross validation is a statistical method that can be
used to evaluate the performance of a model or algorithm [18]. The data are separated into two
subsets, namely learning data and validation / evaluation data. The models or algorithms are trained by a subset of learning and validated by a validation subset. The evaluation parameters were based on accuracy, precision, recall, and F-measure.

3. Results and analysis

The analysis was conducted to compare the results between classification via regression and random committee. The performance results are shown in the form of accuracy, precision, recall, F-measure and confusion matrix. Both results are explained in Chapters 3.1 and 3.2.

3.1. Accuracy, precision, recall, F-measure of sleep stage classification

Table 2 and Figure 1 show that the performance of classification via regression was lower than the random committee, without resampling. The accuracy difference was around 2%. If the class imbalance problem was overcome, then there was an increase in both performances. The accuracy of classification via regression rose from 64.6% to 79.6%, as well as the random committee, up from 66.3% to 87.7. The difference in increase from classification via regression was 15.2%, and the random committee was 21.4%. By sampling, random committee outperformed classification via regression with a greater difference. That was 8.1%.

| Table 2. Performance of classification via regression and random committee. |
|---------------------------|-------------------|-------------------|-------------------|-------------------|
|                           | Classification via regression | Random committee | Classification via regression | Random committee |
|                           | Without resampling | With resampling   | Without resampling | With resampling   |
| Accuracy (%)              | 64.4              | 79.6              | 66.3              | 87.7              |
| Precision (%)             | 63.0              | 79.1              | 65.1              | 87.5              |
| Recall (%)                | 64.4              | 79.6              | 66.3              | 87.7              |
| F-measure (%)             | 61.3              | 79.1              | 64.8              | 87.6              |

Figure 2. Differences in results between the two methods with and without resampling
3.2. Confusion matrix of sleep stage classification

Table 3 shows the confusion matrix of classification via regression without resampling. The most difficult class to classify was N1. None of the N1 was successfully classified. Most N1 were classified as N2. Also, most of the N3 and R were classified as N2. The resulting tree model was more likely to predict data to N2. It was due to the existence of the imbalance problem class so that N2 as the major class was more often predicted than the other minor classes. As a major class, the precision of N1 was much greater than N1, N3, and R. 2639 from 3054 data were classified correctly. The second most class easily classified was Awake. This label was dominantly classified as awake, although the remainder was mostly classified as N2.

Table 3. Confusion matrix of classification via regression without resampling

| Stage | Classified as | Awake | N1 | N2 | N3 | R |
|-------|---------------|-------|----|----|----|---|
| Awake |               | 481   | 1  | 151| 46 | 4 |
| N1    |               | 14    | 0  | 104| 15 | 1 |
| N2    |               | 84    | 3  | 2639| 241| 87|
| N3    |               | 42    | 1  | 771| 588| 27|
| R     |               | 30    | 0  | 499| 26 | 182|

If this method balanced the proportion of data through resampling, then there was an increase in the performance, as shown in Table 4. N1 was the most easily classified label. Almost all data were classified as N1. The improvements also occurred in N3 and R. The data in both labels was successfully classified higher than before. All three labels showed a high increase in precision. Although not a minor class, awake also experienced an increase in precision. 1067 of 1207 data were classified correctly. It means that the precision reached 88.4%. But this precision increase had an impact on the precision of the major class, N2. The N2 precision was decreased. Only 662 of 1207 data were successfully classified as N2. The rest was classified as N3 and R.

Table 4. Confusion matrix of classification via regression with resampling

| Stage | Classified as | Awake | N1 | N2 | N3 | R |
|-------|---------------|-------|----|----|----|---|
| Awake |               | 1067  | 36 | 42 | 36 | 26|
| N1    |               | 3     | 1187| 6  | 7  | 4 |
| N2    |               | 47    | 86 | 662| 229| 183|
| N3    |               | 36    | 32 | 172| 889| 78|
| R     |               | 25    | 26 | 100| 55 | 1001|

Table 5 shows the random committee performance without resampling in classifying the sleep stages. Similar to classification via regression, the most difficult class to be classified was N1. Only 8 of 134 data were successfully classified as N1. But the precision was better than classification via regression. The most easily class to be classified was N2. Unlike the case with N1 in classification via regression, this method had lower precision. Overall, this method was better than classification via regression because the precision of all labels except N2 was higher.
Table 5. Confusion matrix of random committee without resampling

| Stage | Classified as | Awake | N1 | N2 | N3 | R |
|-------|---------------|-------|----|----|----|---|
| Awake |               | 509   | 8  | 124| 31 | 11|
| N1    |               | 19    | 8  | 91 | 12 | 4 |
| N2    |               | 81    | 15 | 2484|322 |152|
| N3    |               | 38    | 4  | 624|734 |29 |
| R     |               | 32    | 1  | 388|46  |270|

After the resampling has been implemented, the random committees had the highest success in classifying N1. The precision was 100%, as shown in Table 6. Through the resampling, it changed the performance of N1 as a minor class. Without resampling, it was the most difficult label to be classified. But with resampling, it became the label which was to be easily classified. The other second minor class, R, could also be better classified after resampling. Unlike the case with N2 as a major class, the resampling decreases its precision. The precision was the lowest compared with other labels. 872 of 1207 N2 data N2 were classified as N2. It means that precision was still above 70%. It was higher than classification via regression. Also, the second easily class to be classified was awake. Most awake data had been classified as its self.

Table 6. Confusion matrix of random committee with resampling

| Stage | Classified as | Awake | N1 | N2 | N3 | R |
|-------|---------------|-------|----|----|----|---|
| Awake |               | 1156  | 5  | 29 | 15 | 2 |
| N1    |               | 0     | 1207| 0  | 0  | 0 |
| N2    |               | 49    | 35 | 872|160 |91 |
| N3    |               | 23    | 6  | 176|966 |36 |
| R     |               | 18    | 3  | 74 |19  |1093|

4. Conclusion
Based on the research that has been done, it can be seen that:
1. Resampling can improve the performance of classification via regression and random committees, as shown in Table 2.
2. Resampling was more suitable to be applied in random committees than classification via regression. The difference in increase was higher.
3. The sleep stage classification was more suitable to be applied in random committees because it provided the highest performance, especially after resampling.
4. Before the clas imbalanced problem has beeed handled, the major class was the highest precision. But it got the lowest precision after resampling. The resampling had influenced the data distribution which has caused changes in tree models that tend to be of minor class.

References
[1] Mannion A, Leader G and Healy O 2013 An Investigation of Comorbid Psychological Disorders, Sleep Problems, Gastrointestinal Symptoms And Epilepsy in Children And Adolescents With Autism Spectrum Disorder Research in Autism Spectrum Disorders 7(1) 35-42
[2] Levin A, Fox N, Zeannah Jr C and Nelson C 2015 Social Communication Difficulties and Autism in Previously Institutionalized Children Journal of the American Academy of Child
(3) Bishop S, Hus V, Duncan A, Huerta M, Gotham K, Pickles A 2013 Subcategories of Restricted and Repetitive Behaviors In Children With Autism Spectrum Disorders Journal of autism and developmental disorders 43(6) 1287-1297
(4) Johnson C, Turner K, Foldes E, Brooks M, Kronk R and Wiggs L 2013 Behavioral Parent Training to Address Sleep Disturbances in Young Children With Autism Spectrum Disorder: A Pilot Trial Sleep medicine 14(10) 995-1004
(5) Elliott R, McKinley S, Cistulli P and Fien M 2013 Characterisation of Sleep In Intensive Care Using 24-Hour Polysomnography: An observational Study Critical Care 17(2)
(6) Fadhullah MU, Resahya A, Nugraha DF and Yulita IN 2018 Sleep Stages Identification in Patients With Sleep Disorder Using K-Means Clustering Journal of Physics: Conference Series 1013(1) 012162
(7) Berry R, Gamaldo C, Harding S, Brooks R, Lloyd R, Vaughn B and Marcus C 2015 AASM Scoring Manual Version 2.2 Updates: New Chapters For Scoring Infant Sleep Staging and Home Sleep Apnea Testing Journal of Clinical Sleep Medicine 11(11) 1253-125
(8) Yulita IN, Fanany MI and Arymurthy AM 2018 Fast Convolutional Method for Automatic Sleep Stage Classification Healthcare informatics research, 24(3) 170-178
(9) Längkvist M, Karlsson L and Loutfi A 2012 Sleep Stage Classification Using Unsupervised Feature Learning Advances in Artificial Neural Systems
(10) Fu C, Zhang P, Jiang J, Yang K and Lv Z 2017 A Bayesian Approach For Sleep And Wake Classification Based on Dynamic Time Warping Method Multimedia Tools and Applications 76(17) 17765-17784
(11) Grønli J, Rempe MJ, Clegern WC, Schmidt M and Wisor JP 2016 Beta EEG Reflects Sensory Processing in Active Wakefulness and Homeostatic Sleep Drive in Quiet Wakefulness Journal of sleep research 25(3) 257-268
(12) Herrera LJ 2011. Symbolic Representation of The EEG For Sleep Stage Classification 2011 11th International Conference on Intelligent Systems Design and Applications (ISDA) (pp. 253-258) IEEE
(13) Yap B, Rani K, Rahman H, Fong S, Khairudin Z and Abdullah N 2014 An Application of Oversampling, Undersampling, Bagging and Boosting in Handling Imbalanced Datasets In Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013) (pp. 13-22). Singapore: Springer
(14) Ruan Y, Lin H and Tsai M 2014 Improving Ranking Performance With Cost-Sensitive Ordinal Classification Via Regression Information retrieval 17(1) 1-20
(15) Arora T and Dhir R 2017 Correlation-Based Feature Selection and Classification Via Regression of Segmented Chromosomes Using Geometric Features Medical & biological engineering & computing 55(5) 733-745

Acknowledgments
This work supported by Center of Excellence for Higher Education Research Grant funded by Indonesian Ministry of Research and Higher Education. This paper also backed by GPU grant from NVIDIA.