Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi
https://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/index
Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru
Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com
e-ISSN: 2685-0893

Machine Learning Classification of Insomnia Using Multidimensional Features and SMOTE

DOI: http://dx.doi.org/10.35889/jutisi.v14i3.3372

Trional Novanza[1*], Mgs. M. Luthfi Ramadhan[2], Femilia Hardina Caryn[3], Rahma Satila Passa[4], Redho Aidil Iqrom[5], Ferlian Seftianto[6], Andre Hardoni[7], Habi Baturohmah[8]

[1,2,3,5]Informatics Engineering, Sriwijaya University, Palembang, Indonesia
[4,6]Computer Systems, Sriwijaya University, Palembang, Indonesia
[7]Computer Engineering, Sriwijaya University, Palembang, Indonesia
[8]Information Systems, Sriwijaya University, Palembang, Indonesia
**Corresponding Author's e-mail:* trionalnovanza@unsri.ac.id

Abstract
Sleep disorders, especially insomnia, are common among adolescents and negatively affect health. Early detection is crucial for appropriate treatment. This study aims to classify insomnia severity in adolescents using a Machine Learning (ML) model and multidimensional features derived from 19 questionnaire instruments. The dataset consists of 95 adolescents aged 16–19 years, categorized into Insomnia, Subclinical Insomnia, and Control classes. The modeling process includes reducing multicollinearity, class balancing with SMOTE, and hyperparameter optimization using GridSearchCV and StratifiedKFold. Feature importance analysis was conducted using decision tree-based methods and permutation importance. The results show that SMOTE improves SVM performance from 0.690 to 0.793 and positively affects Random Forest. Logistic Regression performs best without SMOTE (accuracy 0.759), while XGBoost shows the lowest accuracy (0.614) even with SMOTE. A total of 11 features consistently contribute to all models. In conclusion, ML models, particularly SVM, are effective for classifying insomnia severity in adolescents.
Keywords: *Insomnia; Machine Learning; Multidimensional; Sleep Disorder; SMOTE.*

1. Introduction

Sleep is an essential part of human life. Therefore, obtaining sufficient sleep is vital for everyone [1]. Sleep disorders are significant health issues that can negatively impact an individual [2]. These disorders are not merely trivial issues. They can lead to serious consequences, including an increased risk of neurodegenerative diseases such as dementia and Parkinson's disease, as well as chronic cardiovascular disorders [3]. Biological and social factors contribute to sleep disorders, particularly sleep deprivation. Early detection is crucial for enabling appropriate intervention and effective treatment [4].

As explained by the National Sleep Foundation, approximately 59% of people aged 18 to 29 experience difficulty sleeping due to early morning awakenings and social factors, such as living in a noisy apartment [5]. The detection of sleep disorders traditionally relies on polysomnography (PSG). This method involves manual signal recording and interpretation, which can potentially lead to inconsistent assessment results [6]. The application of PSG for diagnosing insomnia remains a subject of debate because insomnia diagnosis has traditionally relied on subjective patient-reported symptoms [7]. Alternatively, the use of questionnaires to detect symptoms, influencing factors, and the severity of insomnia is commonly conducted using the Insomnia Severity Index (ISI) [8]. The ISI utilizes a questionnaire-based rating scale to detect insomnia symptoms in patients [9]. It consists of seven items designed to evaluate various aspects of sleep disturbance, such as early-morning awakening and sleep maintenance [10]. Sleep quality can also be assessed using standardized instruments, including the Pittsburgh Sleep Quality Index (PSQI). This instrument demonstrates high diagnostic performance in identifying sleep disorder cases [11]. Previous studies have classified insomnia using various methods. For instance, data in the study by Ha et al. were collected using PSG and questionnaires such as the ISI, Epworth Sleepiness Scale (ESS), and PSQI [12]. Schwartz et al. used the Digital Sleep Questionnaire (DSQ) to collect data on general sleep disorders in the community and applied machine learning for classification [13]. Similarly, PSG was used to obtain sleep disorder data in the study by Sharma et al. [14]. However, previous studies often utilized datasets focusing on limited aspects, such as physiological signals via PSG or basic questionnaires. They often did not account for the numerous factors influencing sleep disorders. Adolescents aged 16–19 are considered more vulnerable to sleep disturbances and insomnia symptoms due to academic demands, lifestyle factors, and increased exposure to digital devices before bedtime. Therefore, this age group requires systematic assessment and early screening to prevent long-term impacts.

Machine Learning (ML), as a modern data analysis tool, can efficiently and accurately detect insomnia symptoms. One of the key advantages of ML lies in its feature selection capability. This capability enables the identification of the most influential variables in detecting sleep-related symptoms [15]. To address the limitations of previous studies, we used a dataset from Kiss et al., which was supported by a grant from the National Heart, Lung, and Blood Institute (NHLBI). This dataset collects sleep disturbance data from general adolescent respondents aged 16 to 19 years using 19 standardized instruments. These instruments represent stress levels, sleep habits and quality, sleep-related cognitions, emotion regulation, coping behaviors, personality, mood, and childhood trauma experiences. Instruments included the ISI, PSQI, MEQ-R, the Adolescent Sleep Hygiene Scale (ASHS), the Cleveland Adolescent Sleepiness Questionnaire (CASQ), the Perceived Stress Scale (PSS), and 13 others [16]. To our knowledge, this dataset has not been previously used in machine-learning-based research on insomnia classification. Thus, this study contributes to the field by accounting for multiple influencing factors.

The main objective of our research is to utilize all 19 questionnaire instruments to classify the severity of insomnia symptoms using machine learning. These instruments represent various conditions or factors causing insomnia in adolescents. The expected benefit of this study is to provide a framework for early screening of insomnia symptoms using multidimensional influencing factors and contribute to the development of a more comprehensive, data-driven sleep assessment approach for adolescents. Classification is performed using Random Forest, Logistic Regression, Support Vector Machine (SVM), and XGBoost. We applied SMOTE for oversampling to balance the data distribution. Furthermore, we analyzed the features contributing most to the model's decisions in classifying adolescent insomnia levels. The results are compared with various other classification methods for evaluation.

2. Related Work

The implementation of machine learning in sleeping disorder classification has been conducted recently. Akın and Polat [17] investigated the risk factors for sleep disorders in patients with multiple sclerosis (MS) using Logistic Regression analysis. This study found a significant and independent association between a sleep disorder diagnosis and fatigue, depression, and generalized body pain. Specifically, the Logistic Regression analysis showed that patients with a sleep disorder were approximately 9.82 times more likely to have depression and 6.54 times more likely to experience fatigue. Meanwhile, other factors evaluated in the model, such as gender, number of MS attacks, anxiety, urinary dysfunction, and use of disease-modifying therapy (DMT), were not found to have an independent association with a sleep disorder diagnosis.

Fang et al. [18] identified Clinically Significant Anxiety (CSA) in individuals with insomnia using several machine learning methods, namely Extreme Gradient Boosting (XGBoost), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Adaptive Boosting (AdaBoost), Multilayer Perceptron (MLP), and Naive Bayes (NB). The study involved 205 participants with short-term insomnia, who were divided into two groups: those with and without Clinically Significant Anxiety (CSA). The collected data included physical activity (PA), sleep patterns (SP), and circadian rhythm (CR), obtained through wrist-worn accelerometers for a minimum of seven days. Among the trained models, XGBoost demonstrated the best performance, achieving an Area Under the Curve (AUC) of 0.777, an F1-Score of 0.545, and an Accuracy of 0.839.

Another study by Simon et al. [19] showed that Random Forest (RF) and Naive Bayes (NB) achieved the best performance in classifying insomnia symptoms based on smartphone usage behavior. The RF model attained an AUC of 0.57, demonstrating better discriminative capacity compared to several other tested models. Similarly, the Naive Bayes (NB) model also performed well, achieving an AUC of 0.58.

In 2025, Nuraeni and Faisal [20] conducted research about sleep disorder classification using an SVM based on the Kaggle Sleep Health and Lifestyle dataset, which consists of 400 data samples. The SVM model was tested across three different data split scenarios (80:20, 50:50, and 60:40) using 4 different kernels: linear, polynomial, RBF, and sigmoid. This research shows an accuracy rate of up to 93% in all kernels, where the data were split into an 80:20 training-testing ratio, and only important features were used. In the same year, Putri et al. [21] also researched a similar topic using the same dataset. However, this research by Putri et al. only used 374 samples from the dataset and focused more on evaluating the influence of the four SVM kernel types using a 90:10 split ratio with 10-fold cross-validation. This paper concluded that the Polynomial kernel produced the best performance with an accuracy of 91.6%, confirming that kernel selection has a significant effect on classification quality.

In this study, we used 4 machine learning models, namely XGBoost, Logistic Regression, Random Forest, and SVM, to perform classification. We used the SMOTE oversampling technique and Grid Search users for hyperparameter tuning of the four models.

3. Methodology

3.1 Dataset

We used data collected by Kiss, O. et al., supported and funded by the National Heart, Lung, and Blood Institute (NHLBI), to examine insomnia symptoms as a sleep disorder among adolescents aged 16–19 years[16]. Questionnaires were distributed to 95 adolescents: 48 with healthy sleep, 26 with insomnia, and 21 with insomnia symptoms. These three diagnoses were generated from clinical interviews and according to guidelines in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5).

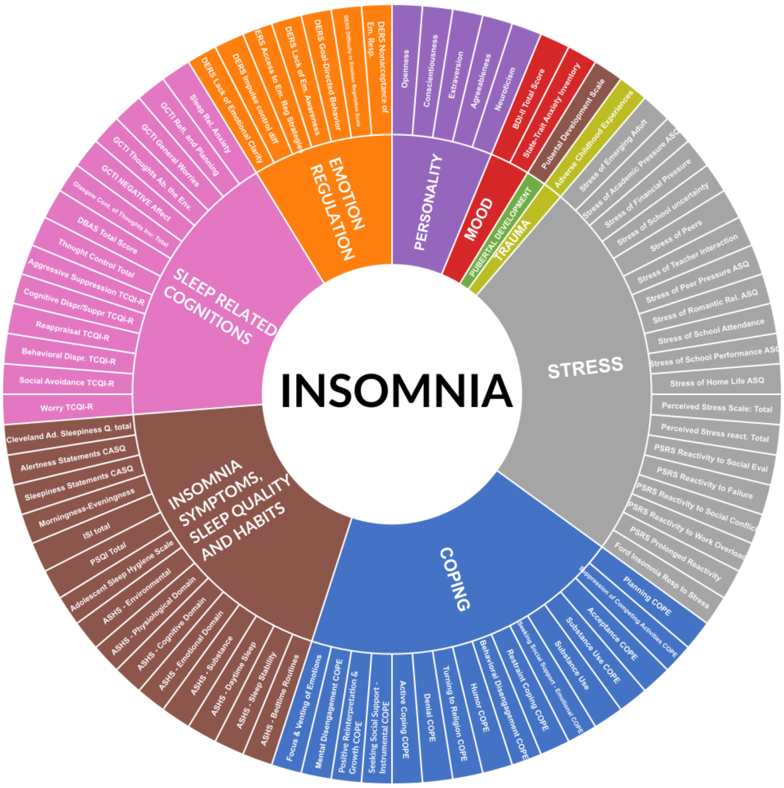


Figure 1. A collection of standardized and validated questionnaire instrument domains. Reillustrated from [16].

The 19 questionnaire instruments used in this dataset include, Insomnia severity index (ISI), Pittsburgh sleep quality index (PSQI), The morningness-eveningness questionnaire (MEQ-R), The sleep hygiene scale (ASHS), Cleveland Adolescent Sleepiness Questionnaire (CASQ), Glasgow content of thoughts inventory (GCTI), Dysfunctional beliefs attitudes and about sleep scale (DBAS-16), Thought control questionnaire insomnia (TCQI-R), Perceived stress scale (PSS), Adolescent stress questionnaire (ASQ), Perceived stress reactivity scale (PSRS), Ford insomnia response to stress test (FIRST), Beck Depression Inventory (BDI-II), The NEO Five-Factory Inventory (NEO-FFI), Adverse Childhood Experiences (ACE), COPE inventory (COPE), Difficulties in Emotion Regulation Scale (DERS), and State-Trait Anxiety Inventory (STAI-Y2).

The questionnaire instruments above represent domains or factors that can cause insomnia symptoms in adolescents, including habits and sleep quality, cognitions associated with sleep, stress and stress responsiveness, emotional regulation and coping behaviors, personality traits, mood states, and experiences of childhood trauma, as in Figure 1.

3.2 Feature Selection

Before using the dataset, we cleaned the data by removing null values for each feature. We checked all features for null values and detected several null values, which we removed to ensure the dataset would produce good accuracy and function correctly.

There was an interesting case at this stage, where two features had a large number of null values: PDS_FEMALE (36 null values) and PDS_MALE (59 null values). The Pubertal Development Scale (PDS) is a data set used to measure pubertal development in both boys and girls. Puberty can affect sleep and hormone levels in both boys and girls, making the PDS necessary for measuring insomnia symptoms. In the dataset used, the PDS is divided into two columns: PDS_FEMALE and PDS_MALE. If the respondent is female, the value in the PDS_FEMALE column is filled in, and if the respondent is male, the value is blank or null. Conversely, the PDS_MALE column is filled in if the respondent is male, and the value is blank if the respondent is female. So both columns have many null values, which will affect the data quality and accuracy of the results.

Basically, the PDS_FEMALE and PDS_MALE features are data that are correlated or related to each other but are separated, so that a phenomenon called Feature Codependency, or often called Multicollinearity[22][23], appears. We have measured the correlation between the two columns using the Correlation Coefficient in Figure 2, and it can be seen that there is no overlap and that the data can be combined. We combine these two correlated features into a single feature, PDS, which represents the PDS value for male and female adolescents, and delete the PDS_FEMALE and PDS_MALE features. This removes the null values in PDS_FEMALE and PDS_MALE, as they complement each other in the newly created PDS feature.

We make these efforts to ensure that the datasets used are clean, tidy, and can make the model work well and improve data quality.

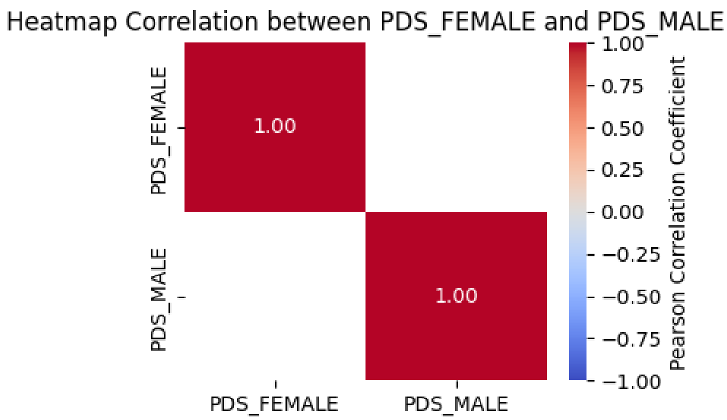


Figure 2. Correlation Heat Map between PDS_FEMALE and PDS_MALE

3.3 Synthetic Minority Over-sampling Technique (SMOTE)

We performed data cleaning, and after these steps, we examined the data distribution across the target classes. We split the dataset before oversampling. The split was performed on the dataset, with 70% of the data used for training and 30% for testing. This shows that the distribution of the three target classes is uneven. The training data for the Clean Insomnia class consisted of 18 data points, the subclinical Insomnia class consisted of 15 data points, and the Control class, which exhibits no insomnia symptoms, had 33 data points.

This imbalanced data distribution raised concerns that the model would be biased toward the majority target class. Therefore, we oversampled samples from training data with the target classes to balance the data distribution across the three target classes. The oversampling technique we used was the Synthetic Minority Oversampling Technique (SMOTE), so that the results of the distribution of target class data in the training data were balanced, namely, Clean Insomnia had 33 data, the Subclinical Insomnia class had 33, and finally the Control class had 33. SMOTE synthesizes data by interpolating between two instances in the minority class. This is achieved through linear interpolation. Let be the feature of the minority class and let be the randomly selected k nearest neighbor. The synthesized instances are computed as follows:

$$Xnew = X_i + \delta(X_j - X_i) \quad (1)$$

Where denotes a random number with a range of 0 to 1. This generates new instances along the line segment connecting and. This research tested the model on both non-SMOTE and SMOTE-enhanced data to assess SMOTE's performance.

3.4. Hyperparameter Tuning

Hyperparameter tuning was performed using Grid Search to obtain the best model from the four pooling methods used: Random Forest, XGBoost, SVM, and Logistic Regression. StratifiedKFold was also used as a cross-validation strategy when running the combination using Grid Search to ensure that SMOTE is only applied to the training fold, thus maintaining the original test fold. The following pooling combinations were used:

1) Random Forest
   - n_estimators = 100, 200
   - Max_depth = None, 10
   - Min_samples_split = 2, 5
   - Min_samples_leaf = 1, 2

2) XGBoost
   - n_estimators = 100, 200
   - Max_depth = 3, 5
   - Learning_rate = 0.01, 0.05
   - Subsample = 0.8, 1.0
   - colsample_bytree = 0.7, 0.9

3) SVM
   - C = 0.1, 1
   - Kernel = rbf, linear
   - Gamma = scale, auto

4) Logistic Regression
   - C = 0.01, 0.1, 1
   - Solver = lbfgs
   - Penalty = l2

The combination results performed on the 4 models used include, as many as 32 combinations of Random Forest models, 32 combinations for the XGBoost model, 16 combinations for the SVM model, and 3 combinations for the Logistic Regression model. The use of Grid Search for Hyperparameter tuning of the four models uses training data, where this training data is further divided into training data using k-fold cross-validation for Grid Search and testing data as the final evaluation.

3.5. Feature Important

The dataset used has many features, and all of these features are used to classify the target class. To determine which feature contribute most to the classification results for the target class, we performed feature analysis for four models: Random Forest, XGBoost, SVM, and Logistic Regression. Decision tree-based models like Random Forest and XGBoost use the default attribute feature_importances, while SVM and Logistic Regression use permutation importance. The result is the 15 most contributing features across the four models used.

4. Result and Discussion

4.1 SMOTE Effect

We performed hyperparameter tuning on the four models used before and after using SMOTE with the specified pools to observe the effect of SMOTE on the accuracy and evaluation metrics of each model. The highest accuracy before using SMOTE was 0.759 for the Logistic Regression model, using the best combination of C (0.01), the solver (lbfgs), and the penalty (12). The best accuracy results for each model can be seen in Table 1.

Then, after using SMOTE on the training data and using the same pool combination, several models experienced changes. The highest accuracy was obtained from the SVM model, which previously had a score of 0.690, but increased significantly to 0.793 after using SMOTE. It can be seen that the SVM model performs well when the target class distribution is balanced. Random Forest also saw an increase in accuracy from 0.690 to 0.724 after oversampling. Differences occurred in the XGBoost and Logistic Regression models, where the performance of both models declined after using SMOTE. Logistic Regression decreased from 0.759 to 0.724. Despite this decrease, this model still maintained fairly good accuracy and performed similarly to Random Forest. XGBoost experienced a drastic decrease from 0.655 to 0.614, a model that had low accuracy from the start.

In conclusion, using SMOTE as an oversampling method to balance the distribution of target class data does improve accuracy in certain models, such as SVM and Random Forest. However, SMOTE does not always improve model performance, because the XGBooost and Logistic Regression models experienced a decline. XGBoost already had low performance before oversampling, and after oversampling, it experienced a significant decrease in accuracy, precision, and recall. These results indicate that the effectiveness of using SMOTE is still influenced by the model used and how sensitive the model is to the target class data distribution, as well as the number of data and features used. Thus, with a small amount of data (95) and many features, this can affect the model's performance before or after using SMOTE.

Table 1. Best Combination Results Before SMOTE and After SMOTE

Model	Precision	Recall	F1-Score	Accuracy
Random Forest	0.620	0.608	0.613	0.690
Random Forest (SMOTE)	0.650	0.650	0.650	0.724
XGBoost	0.588	0.567	0.572	0.655
XGBoost (SMOTE)	0.626	0.608	0.614	0.614
SVM	0.566	0.575	0.567	0.690
SVM (SMOTE)	**0.739**	**0.694**	**0.696**	**0.793**
Logistic Regression	0.636	0.658	0.634	0.759
Logistic Regression (SMOTE)	0.603	0.617	0.597	0.724

4.2 Performance Comparison of 4 Machine Learning Models

Based on the evaluation results of the four models in Table 1, the performance of the models varied considerably after data cleaning, hyperparameter tuning, and balancing the target class distribution with SMOTE, resulting in the best classification model for the insomnia data. Model comparisons were conducted using four metrics: F1 score, precision, recall, and accuracy.

The SVM model performed best among the four models, achieving the highest accuracy and F1 score, as well as the highest precision. This indicates that the model can distinguish the three target classes quite well, demonstrating that SVM can handle data with multidimensional features effectively. This strong performance aligns with previous studies by Nuraeni and Faisal [20] and Putri et al. [21], which found that SVM outperformed other models in classifying sleep disorders. Although these two previous studies used simpler sleep disorder datasets with fewer features, the results of this study demonstrate that SVM still performs well even when using datasets with many features (multidimensional), provided that the data is balanced across the target classes. These features consist of 19 instruments representing multiple domains, such as emotion, stress, and sleep quality.

Random Forest also performed quite well after using SMOTE, in line with the results of research conducted by Sharma et al. [15], where oversampling an imbalanced class distribution can help improve model performance, albeit not significantly. The performance increase, although not significant, could be due to the small dataset used.

Logistic Regression experienced a decrease in accuracy after implementing SMOTE, but still showed good and consistent performance, as evidenced by its relatively good accuracy before oversampling. This result was also observed in the study by Simon et al. [19], explaining that the Logistic Regression model performed well when using datasets where the data patterns between the target classes were already quite distinct and easily separated. The introduction of synthetic data disrupted the patterns and decreased model performance. The XGBoost model's performance was already quite poor before oversampling, and after using SMOTE, its accuracy declined further.

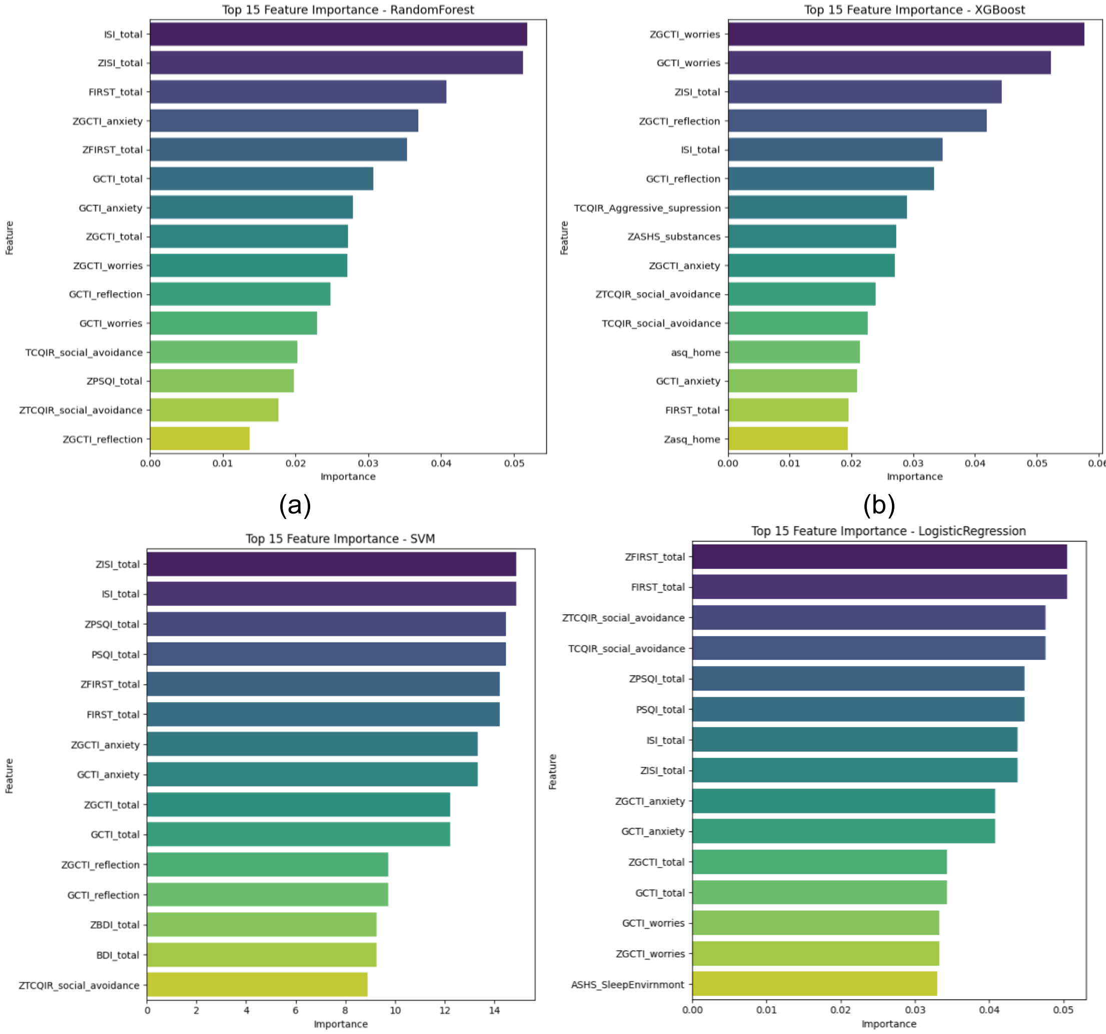
These results indicate that the SVM model is the most effective for classifying insomnia levels in adolescents, especially with small data sets and a large number of features, and the use of SMOTE to balance the target class distribution. However, when talking about the target class data, which is initially unbalanced, the Logistic Regression model is the most effective.

4.3 Feature Importance Analysis

After examining the performance of the four models used in this study that is Random Forest, XGBoost, SVM, and Logistic Regression, we performed feature importance analysis on all four models after using SMOTE, as the best-performing model, SVM, was achieved after SMOTE. The goal was to determine which features contributed most to or influenced the model's decision to classify the three target classes across the four models.

The analysis showed that some feature contributions were similar across all models, while others differed. This is normal, as each model uses a different mechanism or method. Random Forest and XGBoost, which are based on decision trees, use the built-in feature importances in scikit-learn to measure how frequently a feature is used in tree splits [24]. Models that lack this feature importance use alternative methods, SVM, and Logistic Regression, which measure feature contribution linearly using permutation importance [25].

There are at least 11 features that consistently appear in the 15 most contributing features in each model used in this study, namely ISI_total, ZISI_total, FIRST_total, ZGCTI_anxiety, ZFIRST_total, GCTI_total, GCTI_anxiety, ZGCTI_total, ZGCTI_reflection, GCTI_reflection, ZTCQIR_social_avoidance. Based on Figure 3, these features have the highest contribution to the four models in classifying the target class.



(c) (d)

Figure 3. 15 Most important features contributing to the four models: (a) Random Forest, (b) XGBoost, (c) SVM, (d) Logistic Regression.

5. Conclusion

In the study, Machine Learning was effective in classifying insomnia levels in adolescents using a variety of multidimensional features derived from 19 questionnaires related to insomnia symptoms. SVM was the most effective model for classifying the dataset, with an accuracy of 0.793. Hyperparameter tuning using grid search and validation using StratifiedKFold were very useful because they identified the best pool combination for the four models used: Random Forest, XGBoost, SVM, and Logistic Regression. The use of SMOTE to balance the distribution of target class data also improved the performance of several models, particularly the SVM and Random Forest models. Conversely, Logistic Regression achieved its best performance when the target class data was unbalanced, and performance decreased after SMOTE. XGBoost appeared to be the worst-performing model despite using SMOTE. These results indicate that the effectiveness of SMOTE depends on the characteristics of the model used. Feature importance analysis was conducted to see which features had the most influence on classifying the level of insomnia, where many features were used, so that this analysis showed information from which features contributed the most to the model's decision in classifying the level of insomnia.

Future research research can try more machine learning models that still perform well even though the dataset used is small and has many features. It can also be tried using Deep Learning models in the future to see the results.

References

[1] P. Liu et al., “Automatic sleep stage classification using deep learning: signals, data representation, and neural networks,” Artificial Intelligence Review, vol. 57, no. 11, p. 301, 2024.
[2] M. Zhao, H. Tuo, S. Wang, and L. Zhao, “The effects of dietary nutrition on sleep and sleep disorders,” Mediators of Inflammation, vol. 2020, no. 1, p. 3142874, 2020.
[3] Y. J. Lee, J. Y. Lee, J. H. Cho, and J. H. Choi, “Interrater reliability of sleep stage scoring: a meta-analysis,” Journal of Clinical Sleep Medicine, vol. 18, no. 1, pp. 193–202, 2022.
[4] S. Khanmohmmadi et al., “Revolutionizing sleep disorder diagnosis: a multi-task learning approach optimized with genetic and Q-learning techniques,” Scientific Reports, vol. 15, no. 1, p. 16603, 2025.
[5] J. F. Gaultney, “The prevalence of sleep disorders in college students: impact on academic performance,” Journal of American College Health, vol. 59, no. 2, pp. 91–97, 2010.
[6] T. S. Alshammari, “Applying machine learning algorithms for the classification of sleep disorders,” IEEE Access, vol. 12, pp. 36110–36121, 2024.
[7] L. Frase, C. Nissen, K. Spiegelhalder, and B. Feige, “The importance and limitations of polysomnography in insomnia disorder—a critical appraisal,” Journal of Sleep Research, vol. 32, no. 6, p. e14036, 2023.
[8] W. Lee et al., “The simplification of the Insomnia Severity Index and Epworth Sleepiness Scale using machine learning models,” Scientific Reports, vol. 13, no. 1, p. 6214, 2023.
[9] S. Chung et al., “Psychometric properties of the Insomnia Severity Index and its comparison with the shortened versions among the general population,” Psychiatry Investigation, vol. 21, no. 1, p. 9, 2024.
[10] K. Sadeghniiat-Haghighi, Z. Yazdi, and M. Firoozeh, “Comparison of two assessment tools that measure insomnia: the Insomnia Severity Index and polysomnography,” Indian Journal of Psychological Medicine, vol. 36, no. 1, pp. 54–57, 2014.
[11] D. J. Buysse et al., “Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical/polysomnographic measures in a community sample,” Journal of Clinical Sleep Medicine, vol. 4, no. 6, pp. 563–571, 2008.
[12] S. Ha et al., “Predicting the risk of sleep disorders using a machine learning–based simple questionnaire: development and validation study,” Journal of Medical Internet Research, vol. 25, p. e46520, 2023.
[13] A. R. Schwartz et al., “Brief digital sleep questionnaire powered by machine learning prediction models identifies common sleep disorders,” Sleep Medicine, vol. 71, pp. 66–76, 2020.
[14] M. Sharma, J. Tiwari, V. Patel, and U. R. Acharya, “Automated identification of sleep disorder types using triplet half-band filter and ensemble machine learning techniques with EEG signals,” Electronics, vol. 10, no. 13, p. 1531, 2021.
[15] H. Yeo et al., “Exploring predictors of insomnia severity in shift workers using machine learning model,” Frontiers in Public Health, vol. 13, p. 1494583, 2025.
[16] O. Kiss, D. Yuksel, D. E. Prouty, F. C. Baker, and M. de Zambotti, “A dataset reflecting the multidimensionality of insomnia symptomatology in adolescence using standardized questionnaires,” Data in Brief, vol. 44, p. 108523, 2022.
[17] Y. A. Akın and B. S. Arıca Polat, “Sleep disorders in multiple sclerosis,” Gulhane Medical Journal, vol. 63, no. 2, pp. 141–146, 2021.
[18] L. Fang et al., “Machine learning models to identify clinically significant anxiety in short-term insomnia using accelerometers,” Depression and Anxiety, vol. 2025, no. 1, p. 3082856, 2025.
[19] L. Simon, Y. Terhorst, C. Cohrdes, R. Pryss, L. Steinmetz, J. D. Elhai, and H. Baumeister, “The predictive value of supervised machine learning models for insomnia symptoms through smartphone usage behavior,” Sleep Medicine: X, vol. 7, p. 100114, 2024, doi: 10.1016/j.sleepx.2024.100114.
[20] N. Nuraeni and M. Faisal, “Classification of sleep disorders using support vector machine,” Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics and Control, 2025. doi: 10.22219/kinetik.v10i1.2054.
[21] T. R. Putri, P. N. Utamy, M. Wahyudi, Sumanto, and A. S. Budiman, “Application of support vector machine algorithm for classification of sleep disorders,” Journal of Artificial Intelligence and Engineering Applications, vol. 5, no. 1, pp. 83–89, 2025.

[22] J. H. Kim, “Multicollinearity and misleading statistical results,” Korean Journal of Anesthesiology, vol. 72, no. 6, pp. 558–569, 2019.
[23] J. Y.-L. Chan et al., “Mitigating the multicollinearity problem and its machine learning approach: a review,” Mathematics, vol. 10, no. 8, p. 1283, 2022.
[24] X. Li, Y. Wang, S. Basu, K. Kumbier, and B. Yu, “A debiased MDI feature importance measure for random forests,” Advances in Neural Information Processing Systems, vol. 32, 2019.
[25] B. D. Topçuoğlu, N. A. Lesniak, M. T. Ruffin IV, J. Wiens, and P. D. Schloss, “A framework for effective application of machine learning to microbiome-based classification problems,” mBio, vol. 11, no. 3, pp. 10–1128, 2020.