(ISSN: 2582-5526) www.thevoiceofcreativeresearch.com

The Voice of Creative Research

Vol. 7 & Issue 2 (April 2025)

https://doi.org/10.53032/tvcr/2025.v7n2.29

Research Article

Analysing the Impact of Irregular Sleep Patterns on Academic Performance Using Machine Learning

Anurag Phalke
Department of Computer Science,
Kirti M. Doongursee College of Arts Science and Commerce
(Autonomous), Mumbai
anuragphalke777@gmail.com

Prabha Kadam
Department of Computer Science,
Kirti M. Doongursee College of Arts Science and Commerce
(Autonomous), Mumbai
prabha.kadam@despune.org

Abstract:
Student academic Grades depend on the sleep quality of the student. The research focused on how the duration of sleep affects students' academic grades by using two machine learning models, namely decision trees and random forest classifiers. The data was collected from students who are currently studying in school and colleges. The attributes like sleep duration and academic scores were considered as a parameter for the study. The data pre-processing and correct split into the train and test set was done under the study. The decision tree and random forest implemented on the records set. The comparative study indicated that random forest returned more accurate results than the decision tree.

***Keywords*:** Sleep Duration, Academic Performance, Machine Learning, Decision Tree, Random Forest, Educational Data Analysis

1. Introduction:

Sleeping duration is really important in our daily lives, and it has an impact on how we feel and think, specifically for students. The one who is sleeping for enough duration is likely to learn better, also remember information more easily, and perform well in studies. In daily life, it is important to identify the sleep patterns of the students who are scoring low grades because

they might be under pressure or may be affected by their learning abilities. The study focused on how sleep affects academic grades by using machine learning techniques. Previous research shows that not getting proper or enough sleep affects memory, focus, and problem-solving skills badly (Curcio, Ferrara, & De Gennaro, 2006). Sleep disturbances are likely to have a negative impact on one's ability to focus in class, resulting in poor academic performance (Dewald et al., 2010). In order to establish this relationship, we use Decision Tree and Random Forest classifiers to understand the correlation between the duration of sleep and academic grades. Machine learning is a useful tool for working with complicated educational data. While traditional statistics provide valuable information, machine learning can capture the most complex and elusive patterns inherent in the data. The random forest algorithm will be compared to the decision tree algorithm, which is known to be easier for finding ways to make decisions and the random forest which relies on multiple prediction trees (Breiman, 2001).

The study was performed to exemplify the extent to which sleep duration impacts academic performance. The study explores how machine learning algorithms can be implemented in the education system for predicting students' academic performance. The study will notify the schools to recognize the sleep health of students as a contributing factor which needs to be addressed for the improvement of academic performance.

2. Literature Review:

The correlation between pupils' sleep patterns and their academic performance is perhaps the most captivating subject for teachers, researchers, and psychologists. Sleep is one of the important things so that we can think clearly, remember things, pay attention, and solve problems. These all are important for studying and scoring good marks in college and school. There is a lot of research that shows that not getting enough sleep can affect our thinking and remembering power in various ways. In 2011, Carskadon explained the "perfect storm" of sleep deprivation among teenagers, stating that irregular sleep directly affects students' grades (Carskadon, 2011). In 2006, Curcio, Ferrara, and De Gennaro found that when students don't get enough sleep, they are not able to focus, take more time to remember information, and experience increased academic stress (Curcio, Ferrara, & De Gennaro, 2006). In 2010, Dewald et al. found a relationship between the quality and amount of sleep students get and their performance grades in school (Dewald et al., 2010). As technology is improving, researchers or students can use advanced techniques like machine learning to analyse educational data better. Tools like decision trees and random forests are famous for predicting the results of students based on multiple aspects like attendance, study habits, and mental health. Decision trees are easier, but they might sometimes overfit, meaning they won't work with newly updated data (Quinlan, 1986). In 2001, to solve this issue, Breiman created the random forest, which uses multiple decision trees for prediction (Breiman, 2001). Previous research shows that the random forest performs better than the single decision tree because a random forest handles data better and is less likely to overfit. Transforming categorical features into numerical values and extracting relevant features is vital in enhancing model performance. The work seeks to analyse the impact of sleep duration on academic grades using a machine learning algorithm

while also comparing a random forest and a decision tree. Our aim is to facilitate the understanding of educational data science and figure out the significance of this approach.

3. Objective:

- Finding the relationship between the irregular sleep pattern and academic grades using a machine learning model.
- Comparing the machine learning model analysis which are decision tree classifier and random forest classifier.

4. Methodology:

The study will use machine learning models to analyse how sleep duration impacts academic grades. The study includes collecting the data, selecting the models, training the models, and visualizing the results. For this analysis a survey done through Google form as a tool for data collection, in which students provided their sleep duration along with their academic performance. To make the data easier to analyse, authors transformed the categories of sleep duration and academic scores into numerical values. The data was split into two sections: the 80% samples to train the models and 20% samples to test the models. Sleep duration will be the independent variable and academic performance will be the dependent variable. Researchers compared and analysed two machine learning classification models to evaluate the correlation between sleeping hours and academic scores, to determine the most appropriate model describing the relationship between sleep duration and academic performance.

- Decision Tree Classifier: A tree-like model employs a hierarchical arrangement of data organized in accordance to the specific rules of a given issue termed 'Gini.' This model is a decision tree, which is one of the simplest and most intuitive models. Even though they are very simple and easy to follow, there may also be some overfitting problems with deep trees – especially under certain conditions where they become too deep. Tree models have a deep-rooted problem which can be stated as overfitting.
- The Random Forest Classifier: With the Random Forest Classifier, predictions can be made by using multiple decision trees in a particular way. It will create bundles of trees and aggregate their results in decision-making. For the study, the hyper parameters tuned and the model trained using 200 trees while capping the depth of each tree to 5 levels. In this manner, the model will not be too complex for the system to learn effectively.

Both models will undergo training on the data provided and their outcomes analysed through the four metrics of classification, which include the Confusion Matrix, Precision & Recall, Accuracy, and F1-score, shown in the figure1. The multiclass classification performed after class balancing on the data set into Good, Average and Poor academic performance.

```
# Predict and evaluate
y_pred = rf.predict(X_test)
report = classification_report(y_test, y_pred, target_names=le_performance.classes_)
print(report)
```

	precision	recall	f1-score	support
Average	0.00	0.00	0.00	11
Below Average	0.25	0.02	0.03	60
Excellent	0.00	0.00	0.00	1
Good	0.00	0.00	0.00	5
Poor	0.52	0.98	0.68	82
accuracy			0.51	159
macro avg	0.15	0.20	0.14	159
weighted avg	0.36	0.51	0.36	159

Figure1, Classification Report.

Researchers implemented visualisation methods to extract more useful data.

- Confusion matrix heatmaps show graphically how the results have been classified.
- Sleep duration within academic performance groups is shown with a bar plot.

The analysis reveals the advantages and disadvantages of the two models. Decision Trees provide an interpretable structure, while Random Forests offer greater accuracy and generalization through reduced overfitting. The figure 2 indicates the same.

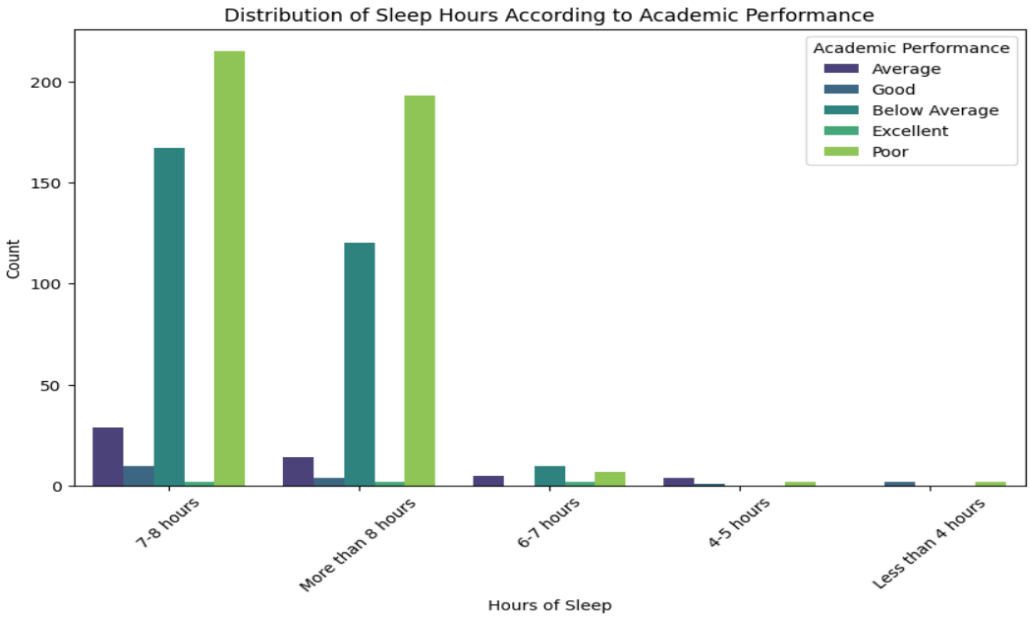

Figure 2, Sleep Duration within Academic Performance Groups.

These results can be used by teachers and researchers interested to identify major contributors to the students' academic grades.

5. Result and Analysis:

The research captures the impact of sleep duration on academic grades. A person can perform well in academics if they sleep for 6-8 hours. A person will perform poorly if they do

not sleep for 6-8 hours. The bar plot Figure 2, shows good academic performance alongside just the right amount of sleep, which is 6-8 hours. The sleep duration affects academic grades, a reason why healthy and enough sleep is important.

A Decision Tree The model lacks consistency and has some overfitting, yet it still provides clarity in a particular decision. Random Forest Its precision and recall is better than that of others, alongside accuracy, enables decreased overfitting, and provides more stable and correct predictions. The Random Forest model exhibits lower error rates. Feature importance analysis indicates that sleep does impact academic performance. The Random Forest model tool is more reliable than the decision tree tool for making predictions, since the accuracy differed by 4%. The Random Forest algorithm returned 60% average accuracy.

6. Conclusion and Future scope:

This research emphasizes the importance of sleep in academic achievement. It shows that students who sleep 6–8 hours each night tend to do better academically, while students who sleep more or less often have more trouble achieving academically. This analysis was undertaken using machine learning models, including Decision Trees and Random Forest algorithms. Decision Trees are simple but can often overfit, making inaccurate predictions with little real-world applicability. On the other hand, Random Forests provide more accurate information because they combine multiple decision trees, which minimizes error and maximizes accuracy. These findings are useful for students as well as teachers. Schools and universities can use these results to create campaigns focused on healthy sleep patterns in order to improve student performance. If educational institutions employ machine learning models in their planning, they can tailor policies aimed at improving academic performance to individual students. For the next analysis, more variables like nutrition, screen exposure, study habits, and anxiety should be added to better explain variation in academic achievement. More sophisticated approaches, such as ensemble learning, can enhance predictive power and give more detail about student achievement. With ongoing research in this area, technology-driven solutions can play a crucial role in optimizing learning outcomes and fostering student well-being.

7. References:

1. Carskadon, M. A. (2011). Sleep in Adolescents: The Perfect Storm. *Pediatric Clinics of North America, 58*(3), 637–647.
2. Curcio, G., Ferrara, M., & De Gennaro, L. (2006). Sleep loss, learning capacity, and the academic performance. *Sleep Medicine Reviews, 10*(5), 323–337.
3. Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32.
4. Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*(1), 81–106.
5. Dewald, J. F., Meijer, A. M., Oort, F. J., Kerkhof, G. A., & Bögels, S. M. (2010).
6. The effect of sleep quality, sleep duration, and sleepiness on the school performance in children and adolescents: A meta-analytic review. *Sleep Medicine Reviews, 14*(3), 179–189.