*Nanotechnology Perceptions*
*ISSN 1660-6795*
*www.nano-ntp.com*

# Classifying the Educational Impact of ADHD Using Machine Learning Techniques

**Touhid Bhuiyan[1], Md Azaz Ahamed[2], Abdullah Al Mamun[2], Md Maruf Hassan[3]**

[1]*School of IT, Washington University of Science and Technology, VA, USA*
[2]*Department of CSE, Daffodil International University, Bangladesh*
[3]*Department of CSE, Southeast University, Bangladesh*

This research investigates the utilization of machine learning approach to anticipate the effects of ADHD on education and classification by utilizing a comprehensive dataset. We evaluated the efficacy of numerous models, including SVM, DT, and RF, without incorporating feature selection. The projected accuracies of our experimental results are varied, with SVM achieving an accuracy of 88%, DT at 96%, and RF at 92%. These observations emphasize potential machine learning models to predict performance on education with ADHD student, thereby emphasizing the importance of data-driven approaches in the education performance with neurodevelopmental illnesses student and advocating for automated, scalable solutions.

## 1. Introduction

People with neurodevelopmental disorder such as ADHD may be inattentive, hyperactive, and impulsive, which can greatly affect their general health, schooling, and relationships. Originally believed to solely affect children, subsequent research revealed that ADHD can persist into adolescence and adulthood and that age and gender affect symptom changes. Many students worldwide have ADHD, which makes it difficult for them to control behavior and focus, complete tasks, and manage their time in the classroom. These issues typically lead to poor performance in the classroom, difficulties with friends, and increased risk of mental health issues like depression and anxiety [5].

Identifying ADHD in kids has mostly rested on clinical evaluations and behavioral observations, which can be arbitrary and cannot always reflect the whole complexity of the condition. Though useful, conventional diagnostic techniques are not always exact enough for early intervention. Techniques like machine learning (ML) provide a creative mean to get over this restriction. ML models can find trend predicting academic performance and identify kid at risk of ADHD or other associated disorder by evaluating vast dataset including demographic, academic, and behavioral information.

Recent research using ML methods have categorized pupils according on a range of data points, including academic performance, mental health issues, and social contacts. These analyses imply that in educational environments ML models such SVM, DT, RF, and KNN may be especially helpful in forecasting outcomes. These model have great potential not just in spotting kids who could be at danger but also in revealing important elements impacting academic performance.

Leveraging a dataset comprising factor including age, gender, academic record, sleep pattern, mental health status, and social interaction, this investigate the application of ml approach to predict academic performance of individual with ADHD. This work intend to find important determinants of academic success and failure in ADHD-affected student by using a variety of ML techniques, hence enabling early identification and tailored educational plans. In the end, our study aims to help children with ADHD realize their full academic potential by means of more effective and data-driven solutions developed from this direction.

## 2. Related Work

Muhammad Mahbubur Rahman et al. [1][2024]This study look how hard it to diagnose ADHD, which known for having a confusing cause and a subjective diagnostic method. The author use Fitbit-derived measure of physical exercise look into link with ADHD and test machine learning classifier to see how well they can diagnose. They use 450 people from the ABCD trial as a group to do correlation analysis . Multivariable logistic regression model show that certain Fitbit measure are good at predicting whether someone has ADHD. Random Forest did better than the other that were test. It achieved CV accuracy (0.89), AUC (0.95), precision (0.88), memory (0.90), f1-score (0.89), and test accuracy (0.88), which was better than previou result in classifying ADHD. These results show that wearable data could help doctors make more accurate diagnoses of ADHD and give doctors ideas for future clinical uses.

Sandra Garc´ıa-Ponsoda et al. [2][2024] This paper improve EEG preprocessing and segmentation help diagnose ADHD accurate, which made hard by noise and other problem EEG data. The researcher look at EEG data from kid with ADHD and kid who had not ADHD. To get rid of noise, they use filter, Automatic Subspace Reconstruction (ASR), and Independent Component Analysis (ICA). After splitting the EEG record into segment, they found statistically significant feature and used machine learning model (SVM, KNN, and XGBoost) to find the EEG segments and channel that helped classify ADHD the best. EEG model that use data from later part were more accurate, which could suggest a link between brain fatigue and ADHD. The P3, P4, and C3 channel had the highest classification accuracy (86.1%). Kurtosis, Katz fractal dimension, and Delta, Theta, and Alpha band power spectrum all play a big role.

Ujunwa Madububambachu et al.[3][2024]This study look at how machine learning use more and more identify mental health disorder in college student,face more problem like stress from schoolwork, change in their social live, money problem, and limited access mental health care. Traditional way of diagnosing, like interview, medical history review, and psychiatric test, take a lot of time and often lead to the wrong diagnosis. The research look into how machine learning model www.ijacsa.thesai.org 1 | P a g e (IJACSA) International Journal of Advanced

Computer Science and Applications, Vol. 15, No. 12,December 2024 like LR,RF,SVM, KNN, and CNN can help doctor be more accurate when they diagnose condition like schizophrenia, bipolar disorder, depression, and the like. The writer compare how well different model worke by first preprocessing a set of mental health data. It stood out because it was 86% accurate, had an F1-score of 0.86, and an AUC of 0.95. CNN, on the other hand, got a perfect memory score of 1.

The work by Nizar Alsharif et al.[4][2024] make it easy to diagnose ADHD, a neurodevelopmental disorder mark by restlessness, impulsivity, and inattention that start in early childhood. Researcher use machine learning model like DT, RT, SVM, and Multilayer Perceptron to look at future data from people with ADHD and a comparison group. At 91% accuracy, the SVM model is best. The MLP model came at 89%, the RF model 87%, and the DT model 78%.

Kuo-Chung Chu et al. [5][2023] It very important to correctly diagnose ADHD early on in order to improve health result and cut down on unnecessary medical cost. This study made a diagnostic support model built on machine learning to better check for ADHD. Three model that were tested: a neural network, a classification and regression tree (CART), and LR.ROC analysis was used to test how well each model worked with 74 people in the ADHD group and 21 people in the control group. The first model, logistic regression, came in second with a ROC value of 0.826, and the third model, neural network, came in third with a value of 0.67. When it came to sensitivity and specificity, the CART model got 78.8% right and 50% wrong. Not only does this model look like it could help with diagnosing ADHD, but it could also help with identifying autism spectrum disorder, Tourette syndrome, and dementia, showing that it has a wider range of useful applications.

Md. Maniruzzaman et al. [6][2022]The goal is use machine learning to find and predict ADHD in kid. Researcher use information from the 2018–2019 National Survey of Children’s Health, which included 45,779 kids age 3 to 17, to try to find important risk factor and tell the difference between kid with ADHD and kid who were healthy.Oversampling and undersampling used to fix the dataset that was not balance (11.4%) of children identified with ADHD. Logistic regression (LR) was used to find important risk factors, and eight machine learning classifiers were tested for classification. These included RF,NB,DT, XGBoost,KNN,MLP,SVM, and 1- dimensional 1D CNN. As well as have an AUC of 0.94, the RF classifier best accuracy (85.5%), sensitivity (84.4%), specificity (86.4%), and accuracy (85.5%). This method worke very well for find ADHD early on, which suggest that an LR-RF model might help with ADHD identification.

Akira Yasumura et al.[7][2020] This study attempt uncover objective biomarker for ADHD using machine learning. The researcher applied a support vector machine (SVM) to predict ADHD severity based on brain function data, including behavioral and physiological indication, age, and reverse Stroop test (RST) performance. The data, acquired from 108 children with ADHD and 108 normally developing (TD) children utilizing the technique of near-infrared spectroscopy (NIRS), assessed changes in prefrontal brain oxygenated hemoglobin during the RST. The SVM model displayed outstanding accuracy with 88.71% sensitivity, 83.78% specificity, and an overall discriminating rate of 86.25%. The results demonstrate that SVM, along with objective RST data, can be a beneficial supplemental biomarker for ADHD diagnosis. Ashley E. Tate et al.

[8][2020] The goal of this project make a model for predicting mental health problems in middleaged teens and to see if machine learning methods can do a better job than standard logistic regression. The researchers used information from 7,638 twins in Sweden, including 474 characteristics from parental reports and registry data. The Strength and Difficultie Questionnaire was use to look at mental health problem. The area under the receiver operating characteristic curve (AUC) was used to test how well a number of model worke. The random forest model had the best AUC (0.739), but there wasn't a big change in performance. Support vector machines came in second (AUC = 0.735).

Jae-Won Kim et al. [9][2015] This study look at how machine learning can be used to guess how methylphenidate treatment will work for kid with ADHD, using a variety of factor such as genetic, cognitive data, neuroimaging data, and demographic data. 83 people took part in the study and were given a lot of test and measurement before they started an 8- week methylphenidate treatment. It was 84.6 percent accurate (AUC = 0.84) for the support vector machine (SVM) model to guess how a treatment would work. Some of the most important thing that were used to make prediction were age, weight, genetic polymorphism, lead level, Stroop test result, and oppositional symptom. The result show that SVM could be a useful method for guessing how ADHD treatment will work, but more study is needed to make it more accurate.

Hudson F. Golino et al. [10][2014] This study use both psychometric and machine learning to guess how well high school kid will do in school. The group was made up of 135 student who took three psychological test. Based on their grade, Random Forest, a machine learning model, put the student into two group: high achievement and low success. Random Forest was 75% accurate in the training set, with a precision of 73.69% and a sensitivity of 68%. The accuracy dropped to 68.18% in the testing set.TDRI that was the best indicator. The study show that machine learning could be used to predict how well student will do in school in the field of educational psychology.

## 3. Methodology

This research study the use of machine learning (ML) model to categorize academic performance and determine the effect of ADHD on pupil. The technique consists of multiple stage: data collection, preprocessing, model selection, assessment, and hyperparameter adjustment.

**A.** Data Description: The dataset used in this study was taken from MIT (ADHD dataset) and contains numerous student records with 10 variables. These include demographic factors (age, gender), academic performance, behavioral aspects (e.g., note-taking, liking presentations, challenges with tasks), and mental health indicators (e.g., depression status, sleep patterns, number of friends). This dataset serves as the basis for forecasting students' academic performance.

**B.** Preprocessing: To verify data quality and compatibility with ML models, the following preparatory activities were performed: Missing values were addressed using proper imputation methods. For instance, the field with missing values was provided via mean replacement.Numerical columns, such as SleepPerDayHours and NumberOfFriend, were analyzed for outliers using statistical approaches, and extreme values were adjusted or eliminatedCategorical variables, including Gender, DepressionStatus, and AcademicPerformance, were encoded into numerical representations using label encoding to guarantee model compatibility.Numeric data were normalized to a standard range to boost model performance.

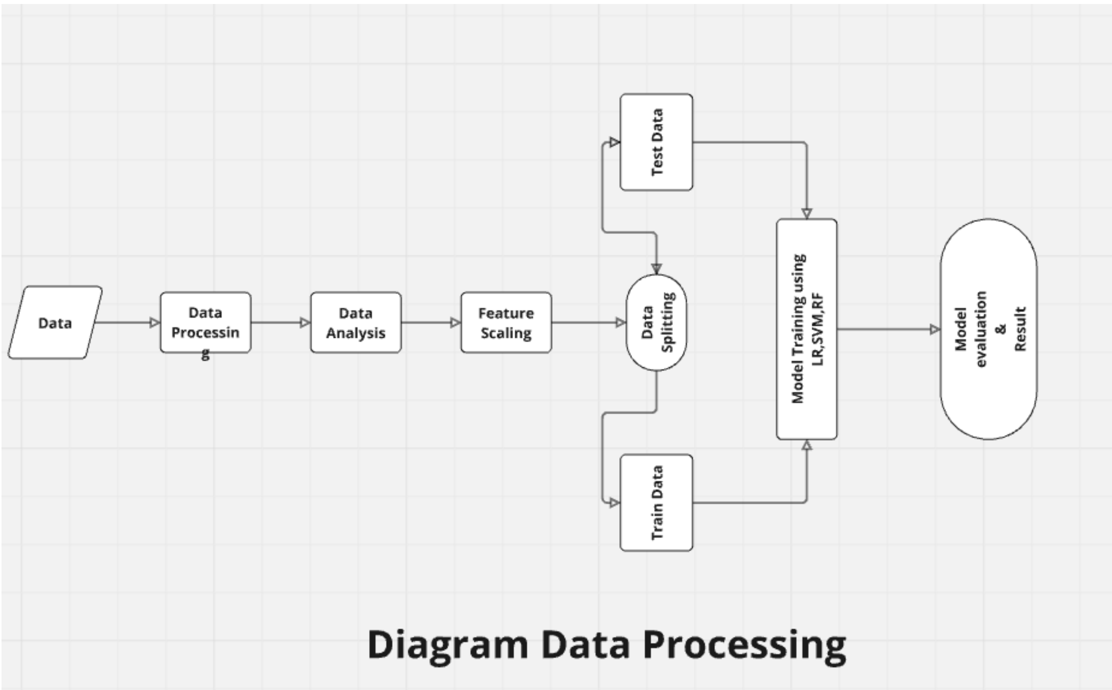


Figure 1. Data Proccessing Diagram.

**C.** Feature Selection:

To determine the most relevant characteristic influencing academic success,exploratory data analysis (EDA) undertaken.Correlation heatmap and statistical test used to discover relationship between variable. Key element,including Sleep-PerDayHours, DepressionStatus, and FaceChallengesToCom-pleteAcademicTask, recognized as significant factor.

**D.** Machine Learning Model Implementation:

Several supervised ML technique developed to classify student academic performance.The dataset divide into two set: a training set and a test set.80% of the dataset used for training, while the remaining 20% served as the test set.ML-based classifier were fitted, ncluding RF, DT, and SVM.For the top-performing classifier (RF, DT, and SVM),hyperparameter optimization performed using a grid search algorithm.In this study, we employed 5-fold CV to identify the optimal hyperparameter values that maximized classification accuracy.Once the ideal hyperparameter values were identified, the ML-based classifiers were refitted, and predictions were made for the test set. Performance scores for each classifier were then calculated, highlighting their ability to predict academic outcomes in children with ADHD.

A. Support Vector Machine (SVM)

SVM, a classifier that determines categorically different data by an optimal hyperplane margin between them.

$$(x) = \sum_{i=1}^{n} \alpha i \ yi \ K(xi, x) + b$$

Where:

- x: Input data (features) for the sample that we want to classify.
- xi : Support vectors (training samples that closest to decision boundary).
- yi : True labels of support vectors (+1 for one class, −1 for other class).
- αi : Weights (Lagrange multipliers) assigned to support vectors.
- K(xi , x): Kernel function that measures how similar xi is to x.
- b: Bias term (helps shift decision boundary).

The predicted class (yˆ) is determined by the sign of f(x):

$$y^{\hat{}} = \begin{cases} +1 \text{ if } f(x) > 0 \\ -1 \text{ if } f(x) < 0 \end{cases}$$

When probability estimation is enabled, the model applies a sigmoid function to f(x) to estimate class probabilities:

$$P(y = 1|x) = \frac{1}{1 + \exp(-Af(x) - B)}$$

zxHere:

- P(y = 1|x): The likelihood that sample belongs to class +1.
- A and B: These are values learned during training to fit sigmoid function.
- f(x): The result ofdecision function.

B. Random Forests (RF)

Random forests, or random decision forests, are a type of ensemble method and specifically a type of decision tree used for classification, regression, and other tasks. They aggregate many trees that have been trained on the same dataset, with the final output representing either the majority vote or average prediction. Random Forests are very robust; they manage large datasets efficiently and also constrain overfitting. Equations are given below:

Classification

The prediction for single decision tree is denoted as ht(x), where t represents t-th tree in forest and x is input. The final predicted class (yˆ) is attained by predominant voting over T trees:

$$\hat{y} = \arg\max c \sum_{t=1}^{T} I(ht(x) = c)$$

Where:

• c: Label or category of class.

• I: Function that gives 1 if ht(x) = c, and 0 if it is not.

Regression

Each decision tree gives a numerical prediction(x). The final prediction ($\hat{y}$)obtained by averaging the predictions of all the trees:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} ht(x)$$

Where:

• T: Total trees in forest

• ht(x): Prediction made by t-th tree.

C. Decision Tree (DT)

Decision Tree, such as the C4.5 model, is an analytical tool that classifies data into different categories based on a set of conditions. It has found applications in many disciplines due to its explicit and reasonable approach to making predictions.

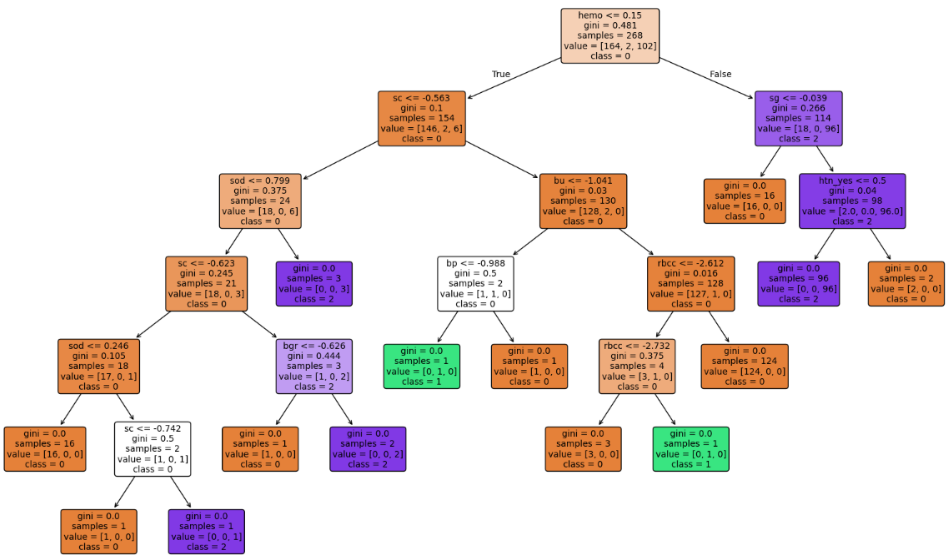


Fig. 2. Decision Tree Visualization (Train Data)

D. sParameter Optimization

Parameter optimization is an important process for improving a model’s performance by determining the best values for its settings. Hyperparameter optimization can be done using a variety of techniques, such as Grid Search, Random Search, and Bayesian Optimization. In our research, we used GridSearchCV from the scikit-learn library to fine-tune the hyperparameters of all of our predictive models. This approach investigates various parameter combinations to determine which one produces the best results. As a result, these optimization techniques significantly improved the accuracy of our predictive analysis.

Performance Evaluation Metrics: Several assessment metrics have been used on this have a look at to gauge how well the device studying fashions achieved. These metrics offer facts approximately how well the version predicts Chronic Kidney Disease (CKD). The metrics indexed underneath were employed:

Precision (P): Tells us the percentage of fine predictions that have been really accurate. The model has fewer false alarms (fake positives) whilst its precision is high.

$$P = \frac{TP}{TP + FP} \qquad (1)$$

Recall (R): The model's take into account gauges how nicely it recognizes real CKD instances. It shows the share of CKD patients that the model efficaciously identified. In scientific studies, in which failing to stumble on a wonderful case will have foremost repercussions, that is mainly crucial.

$$R = \frac{TP}{TP + FN} \qquad (2)$$

F1-Score: Precision and Recall are blended into a single price referred to as the F1-Score. It gives a truthful angle, particularly in cases where in the distribution of CKD and nonCKD cases within the dataset is uneven.

$$\text{F1-score} = \frac{2 \times P \times R}{P+R} \qquad (3)$$

Accuracy (A): The version's typical accuracy is determined by how many predictions it efficaciously made. It is determined via dividing the overall range of predictions by the percentage of correct predictions (CKD and non-CKD).

$$A = \frac{TP + TN}{TP + TN + FN + FP} \qquad (4)$$

## 5. Experimental Results & Discussions

A. Experimental Result: Analysis The Decision Tree (DT) and Random Forest (RF) models performed admirably with the numerical and categorical factor in the dataset, as they both achieve 100% accuracy during cross-validation. As an example of it capacity to discover complicated pattern in massive amounts of data, these models correctly grouped the students' academic performance. Conversely, SVM accuracy was 94%, which is slightly lower. While the SVM model outperformed the other model generally, it lagged somewhat behind due to its inability to distinguish between the various student groups. These findings suggest that SVM may have limitations when it comes to capturing feature connections, particularly in cases where the data is not homogeneous.

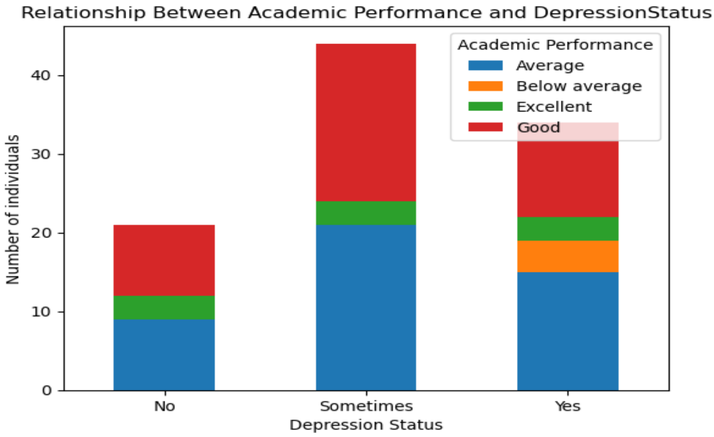


Fig. 3. Relation between features.

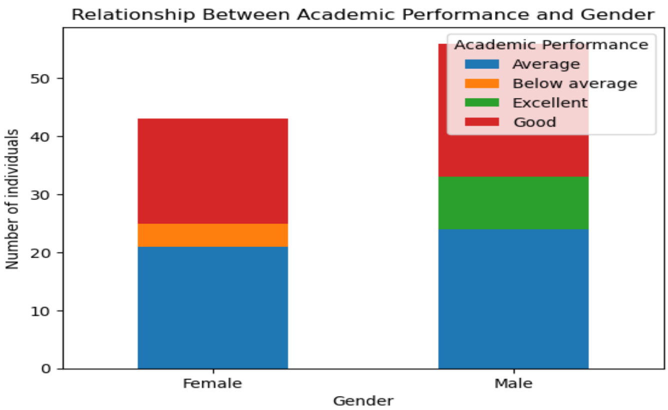


Fig. 4. Relation between features.

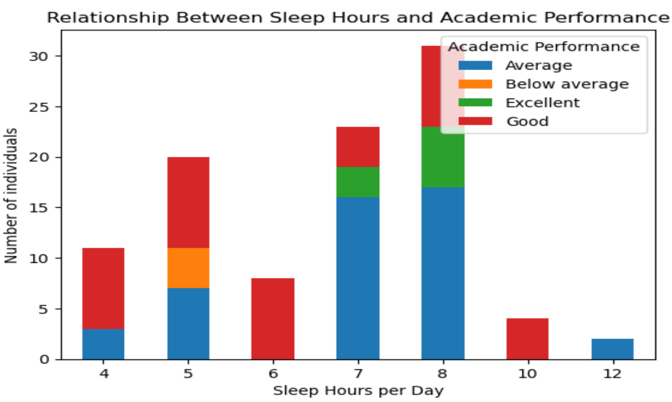


Fig. 5. Relation between feature.

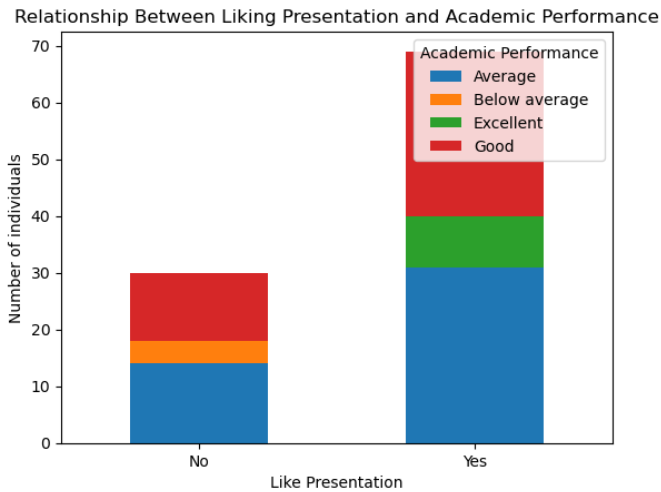


Fig. 6. Relation between features.

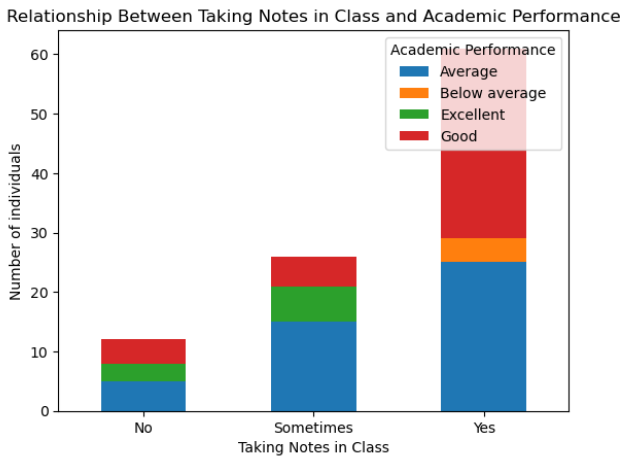


Fig. 7. Relation between features.

## 6. Conclusion

With an eye toward ADHD and associated problems, this study predicted student performance using machine learning (ML) model. Crucially variable determining academic performance was found using a dataset comprising demographic, mental health, social characteristic, and academic activity. Changing hyperparameter raised model accuracy. While the Random Forest model attained 92% on the test set, the Decision Tree model attained 96% with optimal setting. These result indicate that, even for challenging dataset like this one, hyperparameter tuning influences model performance. By incorporating multidimensional elements including sleep patterns, social interactions, and mental health condition, we discovered that these ML models—namely Decision Tree and Random Forest—can detect adolescents at risk of academic issues, especially ADHD. Future research could include cognitive tests or classroom behavior to raise projection accuracy. Deep learning and reinforcement learning among other ML techniques help to enhance models. For children with ADHD and similar disorders, integration of longitudinal data to monitor academic performance and real-time prediction algorithms for focused therapies could be of use. These gains would target academically challenged pupils and enhance educational initiatives.

Acknowledgment

This research work is supported by School of IT, Washington University of Science and Technology, VA, USA

## References

1. Rahman, M. M. (2024). Enhancing ADHD prediction in adolescents through Fitbit-derived wearable data. medRxiv.
2. García-Ponsoda, S., Mate, A., & Trujillo, J. (2024). Refining ADHD diagnosis with EEG: The impact of preprocessing and temporal segmentation on classification accuracy. Computers in Biology and Medicine, 183, 109305.
3. Madububambachu, U., Ihezue, U., Ukpebor, A., & Ayo, S. (2024). Predictive machine learning approaches for mental health diagnoses in college students. International Journal of Engineering Technology & Information, 5(3), 1-9.
4. Alsharif, N., Al-Adhaileh, M. H., & Al-Yaari, M. (2024). Accurate identification of attention-deficit/hyperactivity disorder using machine learning approaches. Journal of Disability Research, 3(1), 20230053.
5. Chu, K. C., Huang, H. J., & Huang, Y. S. (2023). Validity of diagnostic support model for attention deficit hyperactivity disorder: A machine learning approach. Journal of Personalized Medicine, 13(11), 1525.
6. Maniruzzaman, M., Shin, J., & Hasan, M. A. M. (2022). Predicting children with ADHD using behavioral activity: A machine learning analysis. Applied Sciences, 12(5), 2737.
7. Yasumura, A., Omori, M., Fukuda, A., Takahashi, J., Yasumura, Y., & Nakagawa, E. (2020). Applied machine learning method to predict children with ADHD using prefrontal cortex activity: A multicenter study in Japan. Journal of Attention Disorders, 24(14), 2012-2020.
8. Tate, A. E., McCabe, R. C., Larsson, H., Lundström, S., Lichtenstein, P., & Kuja-Halkola, R. (2020). Predicting mental health problems in adolescence using machine learning techniques. PLOS One, 15(4), e0230389.
9. Kim, J. W., Sharma, V., & Ryan, N. D. (2015). Predicting methylphenidate response in ADHD using machine learning approaches. International Journal of Neuropsychopharmacology, 18(11), pyv052.
10. Golino, H. F., Gomes, C. M. A., & Andrade, D. (2014). Predicting academic achievement of high-school students using machine learning. Psychology, 5(18), 2046-2057.