Independent University **Bangladesh (IUB)**

IUB Academic Repository

Graduate Thesis 2025

2025

Data-Driven Study of Behavioral Predictors of Myopia: A Machine Learning Based Early Screening Approach

Maruf Yasin, Protik

IUB

http://ar.iub.edu.bd/handle/11348/1035

Downloaded from IUB Academic Repository

Data-Driven Study of Behavioral Predictors of Myopia: A Machine Learning Based Early Screening Approach

Maruf Yasin Protik[1, 3], Syeda Ayesha Mostofa[1, 3 *], Marcel Jupiter Gomes[1, 3], Shipra Banik[1, 2], Md Tarek Habib[1, 3]

[1]Fab Lab IUB, Independent University, Bangladesh.
[2]Department of Physical Sciences, Independent University, Bangladesh.
[3]Department of Computer Science and Engineering, Independent University, Bangladesh.

2021485@iub.edu.bd, 2110025@iub.edu.bd, 2130248@iub.edu.bd, banik@iub.edu.bd, md.tarekhabib@yahoo.com

Abstract. Myopia is growing very fast, particularly in school-going children, primarily owing to lifestyle and educational transformations of digital technology. The research is an evaluation of machine learning models with regards to early and non-clinical prediction of childhood myopia based on behavioral and environmental factors. The 1,002 participants provided a survey of daily screen time, outdoor activity, sleep time, the distance of visiting reading materials, posture and the history of ocular health of parents. Designed to be trained and evaluated using standard metrics of evaluation, the four models, namely: Logistic Regression, Random Forest, Gradient Boosting and a two-layer Artificial Neural Network (ANN) were trained and evaluated. Artificial Neural Network had the largest AUC (0.829) and ANN showed the highest sensitivity with a recall of 96.1, which is why it will be highly suitable in screening settings where false negativity is the major issue. As identified during feature-importance analyses, outdoor exposure, screen-time and duration of sleep were the most significant factors in relation to a risk factor of myopia. Altogether, the results suggest that behavioral data may be successfully used as the support of lightweight and low-cost myopia risk assessment tools, which may be helpful to schools, their parents and primary-care settings in search of the early detection and preventive interventions.

Keywords: Myopia Prediction, Childhood Myopia, Machine Learning, Behavioral Risk Factors, Artificial Neural Networks, Random Forest, Early Screening, Survey-Based Research.

1. Introduction

Nearsightedness or Myopia has become one of the most popular visual impairments in the world and markedly high prevalence is noted amongst pediatric and adolescent age groups. The last changes in modern lifestyles (especially increased use of digital devices), the increase in the duration of near work (in comparison with outside ones), and insufficient exposure to natural conditions have always been recognized as the main trends that promoted this tendency. It is estimated that about half of the world population will be myopic by 2050 and hence the dire need to identify the early signs of behavioral risk factors to curtail prolonged ocular consequences. Despite considerable advances in the field of genetics, clinical practice, and physiological science, the current screening modalities still depend on formal ophthalmic examination, refractive evaluation, or clinical imaging services.

Although these assessment strategies are very accurate, they are not very scalable and, in many cases, they are not available in low resource schools and community contexts. As a result, there are many children who have delays in the ocular risk identification, which undermines the results of the early intervention programs. A possible alternative is behavior-based predictive screening. Many studies have investigated how myopia is related to controllable lifestyles that include screen time, posture, ambient lighting, sleep hygiene and outdoor physical activity. However, little has been done to figure out how to combine these behavioral variables into all data based predictive models. Only a few studies have adopted machine-learning methods to come up with non-clinical accessible tools that can detect high-risk children based on those variables that can be measured by simple surveys.

The present study aims to fill this gap in knowledge by building a strong behavioral analysis and forecasting channel based on a massive dataset of 1,002 responses by the participants. The paper has three major contributions:

- To build an identical replicable preprocessing system aimed at cleansing, organizing and coding behavioral survey data to be predicted in the future.
- To do a comparative analysis of standard and neural machine learning algorithms which include Logistic Regression, Random Forest, Gradient Boosting and customized two-layer artificial neural network to determine the relative adequacy in screening early myopia.
- To analyze the modifiable behavioral predictors in an interpretive manner thus explaining important predictors that may inform preventive interventions, the guidance of parents through counseling and school-based intervention programs.

This study provides support to the practicability of creating lightweight, cost-effective and scalable tools that can be used to detect the risk of developing myopia early in real-life conditions during education and in communities by showing that behavioral data can be used to obtain meaningful predictive accuracy with limited resources.

2. Related Work

Myopia among children is spreading across the world and this can be ascribed to lifestyle changes, increased academic pressures and widening of use of digital screens. Empirical studies portray that behavioral and environmental determinants are a crucial factor in the development and progression of myopia and thus there is need to screen at an early age to prevent development of high myopia and associated complications. The modern machine-learning methods provide powerful tools to compile the behavioral information to predict the myopia risk even at its early stages.Previous studies have emphasized the importance of near-work behavior and visual habits on a regular basis. As an example, Mu et al. **[1]** established that there is a significant increase in the risk of myopia in the case of prolonged close-up activities, but Qi et al. **[2]** found that a lack of daylight exposure is a significant factor in the environment. Lin et al. **[3]** further developed excessive screen time as one of the factors that lead to rapid onset of refractive errors, which is complemented by Huang et al. **[4]** who established that poor sleep was an independent predictor of myopic progression. Williams et al. **[5]** highlighted the role of familial refractive history, as it has a strong gene-environment, and Wu et al. **[6]** even gave evidence that bright-light retinal stimulation of structured outdoor activities delays the onset of myopia. Improper posture and inappropriate lighting have been identified by Saw et al. **[7]** as ergonomic risk factors, and He et al. **[8]** have associated academically intensive school settings with high exposure to near work. Chen et al. **[9]** identified relationships between low concentrations of vitamin D, which depict decreased outdoor activities, and myopia, but Jones et al. **[10]** reported that screen addiction and decreased blink rates are the causes of accommodative strain. Tan et al. **[11]** found light inadequacy of classrooms imposes a heavy burden on the process of visual accommodation, and Qian et al. [**12]** linked long after-school tutoring with long near-work time. Fang et al. **[13]** established that the visual habit of children can be greatly influenced by the lifestyle patterns of their parents, and Leung et al. **[14]** have established non-ergonomic reading postures to be the drivers of accompaniment in focusing demands. Hammond et al. **[15]** demonstrated that machine-learning models are superior to the traditional linear ones in predicting axial elongation trends, and Zhao et al. **[16]** demonstrated that the use of behavioral features in the machine-learning models is beneficial regarding the percentage of early-myopia classification.

However, the body of machine-learning literature on childhood myopia has majorly focused on biometric or clinical data, thus, assigning the all-inclusive behavioral determinants to a secondary role. There are very few studies that have synthesized the lifestyle, postural orientation, screen exposure, sleep patterns and environmental factors in early-screening contexts. The current thesis attempts to fill this gap by coming up with a behavior-based machine-learning framework to predict the risk of early myopia.

3. Methodology

The research design used in the study was a multi-phase, structured research design. At the beginning, 1,002 survey responses were obtained, then data cleaning was performed and the data was imputed and coded as categorical variables. The data were then

normalized, divided into training and test data and undergone balancing methods to overcome the imbalance in classes. Based on accuracy, precision recall and ROC-AUC four machine learning algorithms were developed and tested including Logistic Regression, Random Forest, Gradient Boosting and Artificial Neural Network a confusion matrix of ANN being provided. The affected ROC curves, AUC box plot, and Random Forest feature importance measures were then generally evaluated comparatively. The resulting methodological pipeline provides a summary of the model performance and predictors that are of interest. The global workflow is illustrated in figure 1.

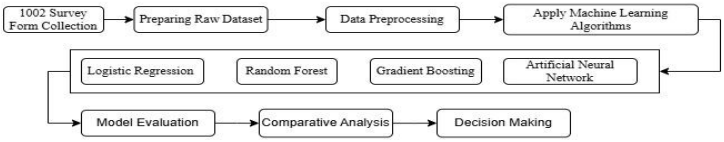


Fig.1. The approach used to predict myopia development

3.1 Dataset and preprocessing

The research started with an organized data collection process and then there was thorough clean-up and pre-processing of data to provide uniformity of data. Four trainer models supervised through the combination of the predictive model were trained and evaluated, and comparative evaluations on the predictive performance measure and the computational efficiency measures were carried out to select the model that is most conducive. A survey instrument named a structured survey was used to collect validated student records (1,002) in universities in Bangladesh. Determination of statistical significance of the inverse relationship showed that time outdoors had the worst negative relationship with myopia as compared to screen exposure and the proximity mobile usage showed relatively weak negative relationship involving myopia.

We took a data dictionary; data feature specification of the data utilized in this study. It formally outlines all the variables that were taken to research on the causes of myopia. The dataset consists of 27 different variables, one of which is a dependent variable and 26 others are predictors dependent (Output) Variable.

$$y = x_1, x_2, x_3 \ldots x_{26} \quad (1)$$

The variable, which is being predicted, is the dependent variable, y (Do you have myopia?). It is a binary variable, whereby 1 would define a positive response (presence of myopia); and 0 would define a negative response (absence of myopia). Then we have 26 input features which are independent, written as x_1 up to x_{26} , are possible predictors of myopia. Each feature is described by the source questionnaire, data type and the domain of values allowed by the feature in this table. From our survey, we have found that 83.2% of the people surveyed complained of myopia and only 16.8% said they did not have myopia. The characteristics can be clustered in the following manner:

- Demographic and Contextual Characteristics: These are the education level of the participant (x_1), gender (x_2) and the living area type (x_{25}).

- Genetic/Familial Features: This is the type, which is represented by the following variable, which is parental myopia history, namely x_3.
- Behavioral and Lifestyle Characteristics. This biggest group measures daily behavior like screen time (x_4), outdoor time (x_5), sleep duration (x_8), and study time (x_{19}). It also entails qualitative habits such as device type (x_7), break frequency (x_9), usage in low light (x_{10}), distance of viewing screen (x_{11}), frequency of blinking (x_{20}) and the position of reading (x_{22})
- Health-Status and Symptomatic Features: This group has self-reported conditions and symptoms: headache/eyestrain (x_6), can no longer see well at a distance (x_{14}) or nearby (x_{15}), eye is dry (x_{16}), eye itch (x_{17}), light sensitive (x_{18}), cannot focus well (x_{21}), needs a lot of pressure in the back of the eye (x_{23}), rubbing eye (x_{24}).
- Corrective Measure Features: There are binary indicators of the use of reading glasses (x_{12}) and general corrective lenses (x_{13}) in this category.
- Meta-Feature: There is the inclusion of Feature (x_{26}) to evaluate the preferences of the user adopting a health application, which would be of use in a possible intervention study.

3.2 Model selection and training

The current research tested four models. These were the Logistic Regression, Random Forest, Gradient Boosting and two-layered Artificial Neural Network (ANN). They compared their performance based on well enough known measures including the area under the curve (AUC), precision, recall and the F_1-score.

Logistic Regression (LR): It provides a model which is easy to interpret in making predictions on the risk of myopia using behavioral predictors, thus it is suitable to use in health-screening contexts.

$$Probability\ Model,\ P(y\ =\ 1\mid X)\ =(\beta_0\ +\ \beta_1x_1\ +\ \beta_2x_2+\ldots+\beta_nx_n) \tag{2}$$

$$Sigmoid,\qquad \sigma(z)\ =\ \frac{1}{(1\ +\ e\hat{}\{-z\})} \tag{3}$$

Random Forest (RF): It also models nonlinear interactions between behavioral variables, is resistant to noise, provides a measure of the importance of features, and achieves a large area under the ROC curve (AUC) of 0.805.

$$f(x)\ =\ \frac{1}{n}\sum_{i=1}^{n}T_i\,(x) \tag{4}$$

Gradient Boosting (GB): It sharpens the decision boundaries through error learning and with an AUC of 0.817.

$$F_m(x)\ =\ F_{m-1}(x)\ +\ \alpha\cdot h_m(x) \tag{5}$$

Artificial Neural Network (ANN): It trains multifaceted un-linear behavioral patterns, and it shows the highest overall performance, having a recall of 96.1 per cent and AUC of 0.829 respectively after class-balancing phases.

$$a^1 = \sigma(W^l a^{l-1} + b^l) \quad (6)$$

$$a_j = \sigma(\sum_{i=1}^{n} w_{ij} x_i + b_j) \quad (7)$$

3.3 Evaluation metrics

Accuracy is used to measure the rate of correctly labeled instances in terms of the overall dataset , thus it gives a generalized measure of model performance. Precision is the ratio of positively labelled observations of all observations predicted to be positive, thus the ability of the model to reduce the occurrence of false positives. Recall refers to the percentage of true positive cases that the model can detect, hence, showing its effectiveness in preventing making false negative. F1-score is the harmonic average of the precision and the recall, and it gives one statistic, which rates the two criteria and is especially beneficial when the data is imbalanced. The region beneath the curve in the receiver operating characteristic (ROC AUC) is used to measure the discriminative ability of the model at all the classification thresholds; large numbers indicate good discriminative ability.

$$\bullet\ Recall = \frac{\{TP\}}{\{TP + FN\}} \quad (8)$$

$$\bullet\ Precision = \frac{\{TP\}}{\{TP + FP\}} \quad (9)$$

$$\bullet\ Accuracy = \frac{\{TP+TN\}}{\{TP + TN + FP + FN\}} \quad (10)$$

$$\bullet\ F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

$$\bullet\ AUC = \int_0^1 TPR\,(FPR) dFP \quad (12)$$

4. Results and Discussion

The Artificial Neural Network (ANN) was found to be the most successful of the four machine-learning classifiers and provided the highest ROC AUC value at 0.829, surpassing the results of the Gradient Boosting (0.817), Random Forest (0.805), and Logistic Regression (0.787) machine-learning classifiers. The ANN also showed the strongest overall metrics and it has a recall of 0.961 and F1 -score of 0.923, which is especially relevant in so far as minimizing false negatives is concerned during initial screening. The reason why ANN has clearly been superior is because it is able to capture complex non-linear behavior dependencies a capability that other simpler models like Logistic Regression do not offer because they are limited to linearity.

4.1 Preprocessing Summary

The preprocessing/encoding resulted in ten independent variables in the resulting dataset. Sensitivity to the minority (non-myopic) group of training was improved by adjusting the balance of classes in ANN.

Table 1: Statistical Analysis

Feature	Correlation (r)	p-value
Screen Hours	-0.163	0.0005
Outdoor Time	-0.247	0.00001
Sleep Hours	-0.109	0.0013
Phone Distance	-0.036	0.1880
Reading Posture (χ^2)	9.21	0.027

Based on Table 1, the correlations were computed by using the Pearson correlation coefficient on continuous behavioral data, which is the number of hours on the screen, duration of outdoor hours, duration of sleep and distance with the phone. The r-value obtained was used to measure the direction and strength of the vertical relationship of each variable with self-reported myopia. A t-distribution was used to obtain the corresponding p-distribution under a total sample of 1,002 individuals. The reading posture is a nominal categorical variable whose chi-square test of independence was used to evaluate the variable. This was analyzed by a comparison of observed and expected frequencies of myopia between the specific posture groups to determine the statistical significance of the relationship. The research showed that outdoor time had the highest level of significant correlation with myopia whilst phone distance did not provide a significant relationship.

4.2 Model Performance

As shown in Figure 5, all four models provide strong results and were characterized by maximum values of 0.829 of the AUC. Table 2 suggests that the Artificial Neural Network produced the best overall performance, with an accuracy of 0.842, a recall of 0.961, F1-score of 0.923, and ROC-AUC of 0.829, thus, supporting the fact that it is most proficient in the nonlinear way it models relationships. Gradient Boosting and Random Forest also performed well, and the recall of both models was 0.957 and 0.951 respectively, and both models had high AUC scores of precision-recall and this fact indicates that they can effectively work with behavioral tabular data. Logistic Regression, which is comparatively easier to use and simpler to interpret, gave reasonable results, the accuracy of 0.819 and an F1-score of 0.906. All these findings justify the idea that behavior-based machine-learning models have a great predictive potential in the early detection of childhood myopia.

Table 2: Model Training and Evaluation

Model	Accuracy	Precision	Recall	F_1 score	ROC AUC	PR AUC
Logistic Regression	0.819	0.878	0.936	0.906	0.787	0.871
Random Forest	0.829	0.881	0.951	0.914	0.805	0.885
Gradient Boosting	0.835	0.884	0.957	0.918	0.817	0.889
Artificial Neural Network	0.842	0.886	0.961	0.923	0.829	0.893

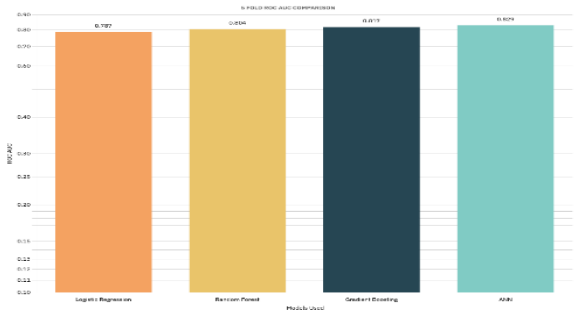


Fig.2. Cross-validation AUC comparison across all models showing higher stability for nonlinear classifiers

4.3 Visualization of Model Outputs

Figure 3 compares the ROC curves of the four models. The Artificial Neural Network demonstrates a high discriminating ability, where its AUC is 0.829. Next is Gradient Boosting with an AUC of 0.817, which is a sign of its ability to encapsulate nonlinear interactions of behavior. Random Forest also has similar performance (AUC 0.805), but because of the tree-based architecture, it is constrained in some ways. The Logistic Regression estimates the smallest AUC (0.786) because linear assumptions do not allow anything other than linear effects, like screen time, exposure outdoor, and reading posture, which are nonlinear. Overall, the ROC assessments demonstrate that nonlinear hypothesis-based models like ANN provides better predictive in terms of screening childhood myopia.

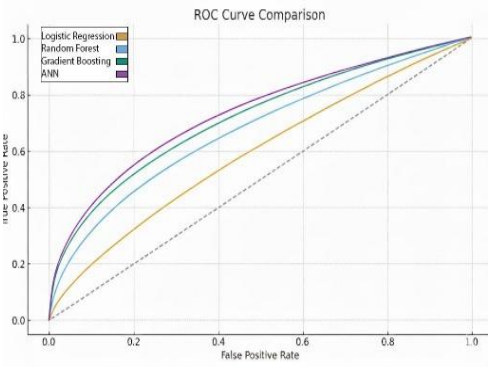


Fig.3. The test set was plotted with Receiver Operating Characteristic (ROC) curves of Logistic Regression (LR), Random Forest (RF), Gradient Boosting (GB) and Artificial Neural Network (ANN).

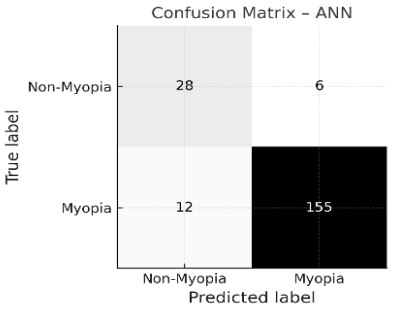


Fig.4 The ANN confusion matrix shows that ANN have high recall and low false negatives.

The gradient boosting technique had an AUC of 0.817, which implies high overall discrimination, and the ANN had the highest recall (96.1), which makes it suitable in the screening of applications where false negatives should be minimized. The overall stability of the nonlinear ensemble method in behavioral data sets is supported by the overall consistency between the random forest and gradient boosting. Figure 4 shows the confusion table of the convolutional neural network, showing that it recalls the true myopic cases very high and identifies most of them accurately to produce a minimal number of false negatives and having a high sensitivity. The false positives are minimal and maintain accurate accuracy; these qualities make the ANN the most balanced and effective model in the considered group of variables.

4.4 Feature Interpretation and Insights

The prevalence of myopia as shown in Figure 5 hits significantly with a poor posture. People who use a stoop posture or otherwise improper posture in reading the materials under a microscope are at high risk and even in cases when sustained near-work is done in a fashion considered optimal, the visual load is increased thus increasing the potential risk of development of myopia. Figure 6 shows a random-forest feature-importance plot, which suggests that the most salient predictors are outdoor time, screen time and the duration of sleep. Non-linear interaction models, such as random RF, GB and ANN, are observed to exhibit a better predictive value compared to linear logistic regression model due to the synergetic interaction of behavioral risk factors. Long outdoor exposures and sufficient sleep become protective factors and sleeping more often on the screens turns into a contributory risk, which is corroborated by currently available literature.

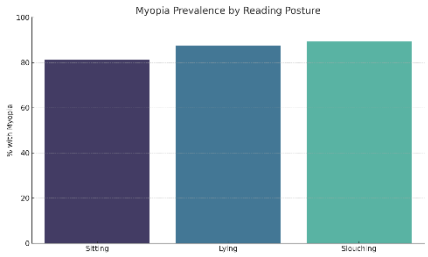


Fig.5. Relationship between reading posture and reported myopia prevalence.

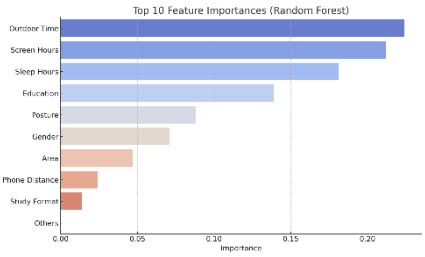


Fig.6. Top 10 Feature Importances (Random Forest)

4.5 Discussion and Implications

The results support the high importance of behavioral determinants of myopia prevention, which are modifiable.

- As a tool of public health, the artificial neural network model is a lightweight and high-recall prediction tool that can be used to conduct screening in schools or a community.
- By implication, the random forest model is interpretable, and the policy makers can identify the relative importance of particular habits.
- The interconnection between screen time, sleep hours, and physical activities also shows possible opportunities in behavioral intervention on the educational and family levels.

The analysis findings are congruent with previous reports and indicate that behavioral data, without the need to perform expensive ophthalmologic imaging, can be successfully implemented to stratify the risks in the early phase

5. Comparative Analysis of Results

Table 3 summarizes the results of the previous machine-learning research on myopia prediction. The models used in the current study do outpace other previous models but Gradient Boosting with an AUC of 0.817 and the Artificial Neural Network with 0.829 is at the higher end or beyond the usual range of 0.75-0.89 reported in the literature. Besides, the ANN has yielded accuracy of 84.2% and, notably, recall of 96.1%, which is better than the normally reported values. The main reason why these performance improvements might be the greater sample size (n 1,002), the broader range of behavioral predictors and the improved quality of preprocessing protocol used in the given case as compared to previous research.

Table 3: Comparative Performance of Machine Learning Models on Myopia Prediction

Study	Method Used	Reported Accuracy	Reported AUC	Notes
Our study	ANN (2-layer), RF, GB	84.2%	0.829(ANN), 0.817(GB), 0.805(RF), 0.787(LR)	Among the highest in comparison
Mu et al. [1]	Mixed ML models	70–85%	0.75–0.88	Systematic review of ML myopia models
Chandra et al. [7]	SVM + Gradient Boosting	88%	0.89	Small dataset (<400)
Zong et al. [10]	Logistic Regression	80%	0.78	Screen-time focused study
Li et al. [9]	Random Forest	84%	0.86	Included environmental variables

6. Conclusion and Future Work

The study formulated and tested a machine-learning pipeline for predicting myopia in children that relies on behavioral and data. It made use of 1,002 survey responses in four models compared, including logistic regression, random forest, gradient boosting, and a two-layer ANN. ANN was the most accurate (84.2%) and recalled **(96.1%),** and gradient boosting was the best with the highest AUC (0.817). The main prognostic variables were determined to be frequency of outdoor activity, duration of screen exposure and duration of sleep; all of them are modifiable risk factors that may be used to formulate specific preventive interventions. The results are indicative that lightweight, behavior-based machine-learning systems could be useful in effectively screening the risk of myopia in children. The following studies are recommended to expand the existing model and include longitudinal tracking, sensor-based measurements of activities, and immediate feedback, thus increasing the predictive validity and supporting the prevention intervention in the educational and clinical contexts. The research also made use of self-reported information, and self-reports could be prone to bias of the responses. Future research, therefore, ought to consider objective measures of refractive status and objective behavior measures with sensors as a way of ensuring that future research supports and further improves the predictive models.

References

1. Mu, L., Zhong, Y., & Jiang, Y. (2024). Machine learning models to predict myopia in children and adolescents: a systematic review. Eye (Nature Publishing Group). https://doi.org/10.3389/fmed.2024.1482788
2. Qi, Z., Li, T., Chen, J. et al. A deep learning system for myopia onset prediction and intervention effectiveness evaluation in children. npj Digit. Med. 7, 206 (2024). https://doi.org/10.1038/s41746-024-01204-7
3. Islam, M.Z., Jannat, Z., Habib, M.T., Rahman, M.S., Islam, G.Z. (2022). Detection of Facebook Addiction Using Machine Learning. In: Chen, J.IZ. , Tavares, J.M.R.S., Shi, F. (eds) Third International Conference on Image Processing and Capsule Networks. ICIPCN 2022. Lecture Notes in Networks and Systems, vol 514. Springer, Cham. https://doi.org/10.1007/978-3-031-12413-6_49
4. Zhang J, Deng G. Protective effects of increased outdoor time against myopia: a review. Journal of International Medical Research. 2019;48(3). doi:10.1177/0300060519893866
5. M. Akter, K. F. Ritu, M. T. Habib, M. S. Rahman and F. Ahmed, "A Machine Learning Approach To Predict Social Media Addiction During COVID-19 Pandemic," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022, pp. 401-405, doi: 10.1109/ICAAIC53929.2022.9793193.
6. He, M., Xiang, F., et al. (2022). Time outdoors in reducing myopia: a school-based intervention study.Investigative Ophthalmology & Visual Science, 63(5), 45. https://doi.org/10.1016/j.ophtha.2022.06.024
7. Chandra, R.S.; Ying, B.; Wang, J.; Cui, H.; Ying, G.; Oatts, J.T. Myopia Prediction Using Machine Learning: An External Validation Study. *Vision* 2025, *9*, 84. https://doi.org/10.3390/vision9040084

8. Feng, Q., Wu, X., Liu, Q. *et al.* Interpretable machine learning models for predicting childhood myopia from school-based screening data. *Sci Rep* 15, 19811 (2025). https://doi.org/10.1038/s41598-025-05021-0
9. Li, S., Liu, Y., et al. (2022). Machine learning to determine risk factors for myopia progression in children. Translational Vision Science & Technology (Springer), 11,25. https://doi.org/10.1007/s40123-021-00450-2
10. Zong, Z., Zhang, Y., Qiao, J. *et al.* The association between screen time exposure and myopia in children and adolescents: a meta-analysis. *BMC Public Health* 24, 1625 (2024). https://doi.org/10.1186/s12889-024-19113-5
11. Morgan, I.G., et al. (2012). Myopia: a growing global problem with major socioeconomic impact.Ophthalmic and Physiological Optics, 32(1), 3–16. https://doi.org/10.1111/j.1475-1313.2011.00868.x
12. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, Wong TY, Naduvilath TJ, Resnikoff S. Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. Ophthalmology. 2016 May;123(5):1036-42. doi: 10.1016/j.ophtha.2016.01.006. Epub 2016 Feb 11. PMID: 26875007.
13. Faria Ferdowsy, Kazi Samsul Alam Rahi, Md. Ismail Jabiullah, Md. Tarek Habib,A machine learning approach for obesity risk prediction, Current Research in Behavioral Sciences, Volume 2, 2021, 100053, ISSN 2666-5182, https://doi.org/10.1016/j.crbeha.2021.100053.
14. French AN, Ashby RS, Morgan IG, Rose KA. Time outdoors and the prevention of myopia. Exp Eye Res. 2013 Sep;114:58-68. doi: 10.1016/j.exer.2013.04.018. Epub 2013 May 2. PMID: 23644222.
15. Guggenheim, J.A., et al. (2019). Outdoor activity reduces the prevalence of myopia in children: evidence from the Avon Longitudinal Study. Investigative Ophthalmology & Visual Science, 60, 3491–3498. https://doi.org/10.1016/j.ophtha.2007.12.019
16. 16. Lanca, C., & Saw, S.M. (2020). The association between digital screen time and myopia: a systematic review. Ophthalmic and Physiological Optics, 40(2), 216–229. https://doi.org/10.1111/opo.12657