

Predicting Cycloplegic Spherical Equivalent in Children and Adolescents: A Comparative Machine Learning Approach with SHAP-Based Model Interpretation

Ouyang Weiyun

The First Affiliated Hospital of Jinan University

Liu Lian

The First Affiliated Hospital of Jinan University

Liang Chunlan

Huizhou Central People's Hospital

Tian Yuan

Jinan University

Zhong Jingxiang

z.jx85221206@126.com

The First Affiliated Hospital of Jinan University

Article

Keywords: cycloplegic refraction, machine learning, myopia, prediction model

Posted Date: September 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10608409/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Predicting Cycloplegic Spherical Equivalent in Children and Adolescents: A Comparative Machine Learning Approach with SHAP-Based Model Interpretation

Ouyang Weiyun, Liu Lian, Liang Chunlan, Tian Yuan, Zhong Jingxiang

First Author: Ouyang Weiyun ; E-mail: 365001850@qq.com; Address: 1.The First Affiliated Hospital of Jinan University, Guangzhou 510630, China, 2.Huizhou Aier Eye Hospital.

Corresponding Author: Zhong Jingxiang; E-mail: zjx85221206@126.com; Address: 1.The First Affiliated Hospital of Jinan University, 2.Shenzhen Bright Star Eye Hospital.

Abstract

Objective: To develop and validate machine learning models for predicting cycloplegic spherical equivalent refraction (SER), myopia status, and the need for cycloplegic dilation using non-cycloplegic clinical parameters in a pediatric population.

Methods: A total of 1,224 eyes from 612 children (aged four to 17 years) were included. Data were collected from two clinical centers with harmonized 27-field standards. Three prediction tasks were established: (1) the need for cycloplegic refraction (based on absolute spherical equivalent shift >0.5 D), (2) cycloplegic myopia status ($SER \leq -0.5$ D), and (3) the exact post-cycloplegic spherical equivalent in children and adolescents. Feature engineering incorporated absolute values, fellow-eye features, interocular differences, and derived biometric indices. Multiple algorithms spanning tree-based ensembles (XGBoost, LightGBM, Random Forest, Extra Trees, Hist Gradient Boosting), neural networks (MLP variants), and classical classifiers were evaluated. Models were trained on 70% of the data and validated on 30% of the data, with stratified sampling. Performance was assessed using accuracy, AUC, RMSE, and R^2 , with bootstrap 95% confidence intervals and SHAP analysis for interpretability.

Results: For predicting the need for cycloplegic refraction, XGBoost achieved the highest accuracy of 0.818 (AUC = 0.780), with precision 0.696 and recall 0.464. For myopia status classification, LightGBM achieved accuracy 0.908 and AUC 0.934, with precision 0.906 and recall 0.974. For post-cycloplegic SER regression, the MLP-linear regressor yielded the lowest RMSE of 0.630 D, MAE 0.424 D, and $R^2 = 0.881$. SHAP analyses showed that the need for cycloplegic refraction was driven by axial length, while myopia status was dominated by non-cycloplegic SER.

Conclusions: Machine learning models using non-cycloplegic parameters can accurately predict cycloplegic refractive outcomes and identify children who need dilation. This approach may offer a scalable, resource-efficient solution for pediatric refractive screening and epidemiological research where cycloplegic examination is impractical.

Keywords: cycloplegic refraction; machine learning; myopia; prediction model

1. Introduction

Myopia has emerged as one of the most critical public visual health concerns worldwide. The World Health Organization (WHO) forecasts that by 2050, the global prevalence of myopia and high myopia will substantially increase, accounting for 49.8% and 9.8% of the world's population, respectively ^[1]. Among them, children and adolescents represent the most vulnerable group ^[2, 3]. In China, the prevalence of myopia among children and adolescents aged 3 to 19 is approximately 84% ^[4], with a tendency toward earlier onset and greater severity. Myopia not only negatively affects children's academic performance and daily quality of life, but—more importantly—high myopia can lead to severe complications such as open-angle glaucoma^[5], cataracts^[6], retinal detachment, and myopic macular degeneration, resulting in irreversible visual impairment or blindness. Consequently, it imposes a significant burden on individuals, families, and society^[7]. Therefore, early detection and timely intervention to slow the progression of myopia in children are not only priorities in ophthalmic clinical practice, but also essential components of global eye health strategies.

Cycloplegic refraction is an internationally recognized gold standard for refractive assessment in children, as it effectively eliminates the influence of accommodation on refractive measurements^[8]. However, in routine clinical practice, cycloplegic testing requires specialized personnel, is time-consuming, and may cause transient side effects, including photophobia and blurred near vision. These limitations hinder the widespread implementation of cycloplegic testing in large-scale school screening programs and epidemiological studies.

In recent years, digital healthcare technologies have leveraged the potential of artificial intelligence (AI) to develop auxiliary solutions characterized by scalability, portability, and reliability^[9]. Such technologies supplement traditional epidemiological approaches by performing correlation analyses and exploiting complex interactions among predictive factors to generate new insights^[10]. In the context of myopia prediction, machine learning (ML) has been used to forecast progression in school-aged children^[11], predict axial length growth in children with myopia^[12], classify myopia status in adolescents^[13], detect pathological myopia^[14], estimate the risk of developing high myopia^[15], and identify risk factors for disease progression^[16].

Although ML shows promise for ophthalmic prediction, existing models often lack interpretability and clinical applicability for myopia management. Most previous studies have focused on single prediction tasks—either SER regression or myopia classification—rather than providing a comprehensive framework that addresses multiple clinical needs simultaneously^[17]. Moreover, the prediction of whether a child actually requires cycloplegic dilation has received limited attention, despite its direct relevance to optimizing clinical workflow.

This study aims to develop and validate a comprehensive ML-based framework for three clinically relevant prediction tasks using non-cycloplegic refractive data and intraocular biometric parameters: (1) classification of whether cycloplegic dilation is needed, (2) classification of myopia status, and (3) regression of post-cycloplegic SER. We leverage a multi-center pediatric dataset with rich feature engineering, including absolute values, fellow-eye features, interocular differences, and derived biometric indices, and we systematically evaluate a broad range of algorithms, including tree-based ensembles, neural network variants, and classical classifiers. SHAP analysis is employed to improve model interpretability and identify key predictors.

2. Materials and Methods

2.1 Participants

This retrospective study was conducted from October 2025 to May 2026. Data were primarily collected from 612 children and adolescents (aged range: 4 – 17 years) examined at two clinical centers: Huizhou Aier Eye Hospital (912 eyes, 74.5%) and the First Affiliated Hospital of Jinan University (312 eyes, 25.5%). All patients underwent comprehensive ophthalmic examinations, including non-cycloplegic and cycloplegic autorefractometry, as well as partial ophthalmological examinations.

Inclusion criteria required both pre- and post-cycloplegic SER measurements. Exclusion criteria were as follows: prior myopia treatments such as atropine or orthokeratology lenses; strabismus, amblyopia, or other organic ocular diseases (e.g., congenital cataract, keratoconus, or corneal opacity); a history of ocular surgery or trauma; hypersensitivity to cycloplegic agents; systemic diseases that may interfere with refractive measurements; and incomplete data or failed measurements. This study was conducted in accordance with the Declaration of Helsinki and approved by the Medical Ethics Review Committee of the leading institution (the First Affiliated Hospital of Jinan University).

2.2 Ophthalmic examination and data collection

Demographic information included age and gender. All participants underwent ophthalmic examinations, including uncorrected visual acuity (UCVA) and intraocular pressure (IOP), non-cycloplegic refraction, cycloplegic refraction, and ocular biometry under both conditions. Uncorrected visual acuity was assessed at 5 m using a standard logarithmic visual acuity chart to determine visual clarity without correction. Additionally, automated refraction was performed using an autorefractor (ARK-1, NIDEK, Japan) both before and after cycloplegia, and all measurements were obtained by the same optometrist to ensure consistency and accuracy.

For refractive error measurements, three readings were obtained per eye in both non-cycloplegic and cycloplegic states. If the spherical equivalent refraction (SER) or cylinder values differed by more than 0.5 diopters (D) between two readings for a given eye, the measurement was repeated. Statistical analysis used the mean of the three refractive measurements per eye. Before cycloplegia, all children underwent non-contact intraocular pressure measurement and slit-lamp examination to rule out contraindications for dilation.

Cycloplegia was achieved by instilling 1 drop of 0.5% tropicamide into each eye at 10-minute intervals, for a total of three times. If the pupillary light reflex persisted 30 minutes after the last instillation or the pupil diameter was less than 6.0 mm, an additional instillation of tropicamide was performed. Ocular biometric parameters were measured using an IOL-Master 700 (Carl Zeiss Meditec AG, Jena, Germany) or a NIDEK A-scan (Nidek, Tokyo, Japan). Recorded biometric variables included axial length (AL), lens thickness (LT), anterior chamber depth (ACD), central corneal thickness (CCT), white-to-white distance (WTW), and pupil diameter (PD). The corneal curvature radius (CCR) was derived from the mean keratometry values (in diopters) using the established conversion factor of 337.5 divided by K.

The SER was calculated using the standard formula: $SER = sphere + 1/2 \times cylinder$. Myopia was diagnosed when cycloplegic SER was ≤ -0.50 D. The myopia screening criterion was defined as $UCVA < 5.0$ and non-cycloplegic SER < -0.50 D. To ensure study quality, all personnel received

standardized training before the study commenced, and instrument calibration was performed using standard eyes before formal measurements. Additionally, a 5% subset of ophthalmic examination items was retested to identify and promptly rectify any discrepancies.

2.3 Dataset preprocessing and splitting

To reduce the impact of missing data on model performance, we used median interpolation to impute missing values for continuous variables and mode interpolation to impute missing values for categorical variables. Beyond raw non-cycloplegic measurements, we performed extensive feature engineering: (1) absolute features: all numeric non-cycloplegic variables were transformed to their absolute values; (2) calculated SER: non-cycloplegic SER was computed as $\text{sphere} + \text{cylinder}/2$; (3) SER formula gap: the difference between the recorded SER and the formula-calculated SER; (4) AL/CCR recalculation: axial length divided by corneal curvature radius ($\text{AL} \div \text{CCR}$); (5) fellow-eye features: for each eye, measurements from the contralateral eye were merged; and (6) interocular differences: differences between ipsilateral and contralateral measurements. Numerical features were standardized using StandardScaler, and categorical features were one-hot encoded.

The dataset was split into a 70% training set and a 30% test set using stratified random sampling. For classification tasks, stratification was performed based on the corresponding labels to maintain balanced class distributions between the two subsets. The training set included 856 eyes, and the test set included 368 eyes.

2.4 Model development and validation

In this study, each model was trained as a complete scikit-learn Pipeline to prevent data leakage. For classification, optimal probability thresholds were determined via grid search by maximizing validation accuracy. For regression, models were evaluated directly using continuous predictions.

We evaluated five ML models for the first task (predicting whether cycloplegic dilation is needed): (1) random forest, (2) extreme gradient boosting (XGBoost), (3) multi-layer perceptron classifier with tanh activation (MLP-tanh classifier), (4) histogram-based gradient boosting (Hist Gradient Boosting), and (5) wide multi-layer perceptron classifier (MLP-wide classifier). For the second task (predicting myopia status), we evaluated five ML models: (1) Light Gradient Boosting Machine (LightGBM), (2) XGBoost, (3) Extra Trees, (4) Hist Gradient Boosting, and (5) Gradient Boosting. For the third task (predicting post-cycloplegic SER), we focused on developing multi-layer perceptron regressor models: (1) MLP-linear regressor, (2) MLP-deep regressor, (3) MLP-wide regressor, (4) MLP-regularized regressor, and (5) MLP-tanh regressor.

To assess the predictive performance of the regression models for the targeted SER, we computed the coefficient of determination (R^2), mean absolute error (MAE), median absolute error (MedAE), and root mean square error (RMSE). For classification performance, we used additional evaluation metrics, including area under the curve (AUC), precision, accuracy, F1-score, and recall. For myopia classification, a confidence cutoff was searched under a coverage $\geq 20\%$ constraint to maximize accuracy in high-confidence predictions.

Additionally, the Shapley Additive exPlanations (SHAP) algorithm was used to quantify the importance of each feature in the models^[18]. We implemented the ML models in Python 3.8. The random forest and XGBoost models were developed using scikit-learn and the XGBoost library, respectively.

3. Results

3.1 Participant Characteristics

The study included 912 eyes from 456 participants in the Huizhou Aier Eye Hospital, and 312 eyes from 156 participants in the First Affiliated Hospital of Jinan University. Participants were randomly stratified into training (n = 856, 70%) and test (n = 368, 30%) sets, with stratification by outcome labels to maintain balance across tasks. Demographic and ocular characteristics of participants in the training and test datasets are summarized in Table 1. The mean age was 9.94 ± 2.54 years in the training set and 9.86 ± 2.45 years in the test set. Male participants comprised 53.0% (n = 454) of the training set and 45.1% (n = 166) of the test set. The mean non-cycloplegic spherical equivalent (SER) was -1.57 ± 1.80 D in the training set and -1.74 ± 1.56 D in the test set. Following cycloplegia, the mean SER shifted to -1.24 ± 1.84 D and -1.37 ± 1.62 D, with a mean difference of 0.34 ± 0.69 D in training and 0.37 ± 0.72 D in the test compared with the non-cycloplegic SER, respectively. Uncorrected visual acuity (UCVA) in logMAR was 0.46 ± 0.26 and 0.47 ± 0.27 in the training and test sets, respectively. Intraocular pressure (IOP) was comparable (16.45 ± 2.82 mmHg vs. 16.53 ± 2.72 mmHg). Ocular biometric parameters exhibited similar profiles between the two groups (respectively): AL 23.98 ± 1.05 mm and 24.02 ± 0.98 mm, CCR 7.81 ± 0.25 mm and 7.82 ± 0.27 mm, AL/CCR ratio 3.07 ± 0.12 and 3.07 ± 0.11 , K value 43.29 ± 1.39 D and 43.23 ± 1.42 D, central corneal thickness 557.55 ± 31.25 μ m and 555.38 ± 30.17 μ m, lens thickness 3.48 ± 0.18 mm and 3.48 ± 0.17 mm, anterior chamber depth 3.74 ± 0.24 mm and 3.70 ± 0.24 mm, white-to-white distance 11.98 ± 0.37 mm and 11.95 ± 0.34 mm, and pupil diameter 5.97 ± 0.77 mm and 6.00 ± 0.68 mm.

Post-cycloplegic biometric measurements demonstrated some changes, including an increase in anterior chamber depth (3.83 ± 0.23 mm vs. 3.77 ± 0.24 mm) in the training and test sets, respectively, consistent with the known effects of cycloplegic agents on the anterior segment^[19-21].

Table 1. Baseline characteristics of the participants (n=1,224).

Variable	Training set (n=856) Mean $\pm$ SD / n (%)	Test set (n=368) Mean $\pm$ SD / n (%)
Age (years)	9.94 $\pm$ 2.54	9.86 $\pm$ 2.45
Sex		
Female	402 (47.0)	202 (54.9)
Male	454 (53.0)	166 (45.1)
Non-cycloplegic		
Spherical equivalent (D)	-1.57 $\pm$ 1.80	-1.74 $\pm$ 1.56
Sphere (D)	-1.18 $\pm$ 1.79	-1.36 $\pm$ 1.54
Cylinder (D)	-0.79 $\pm$ 0.92	-0.75 $\pm$ 0.78
Cycloplegic		
Spherical equivalent (D)	-1.24 $\pm$ 1.84	-1.37 $\pm$ 1.62
Sphere (D)	-0.84 $\pm$ 1.85	-1.00 $\pm$ 1.63
Cylinder (D)	-0.80 $\pm$ 0.90	-0.73 $\pm$ 0.77
Uncorrected visual acuity (logMAR)	0.46 $\pm$ 0.26	0.47 $\pm$ 0.27
Intraocular pressure (mmHg)	16.45 $\pm$ 2.82	16.53 $\pm$ 2.72
Ocular biometric values		
Axial length (mm)	23.98 $\pm$ 1.05	24.02 $\pm$ 0.98
Corneal radius (mm)	7.81 $\pm$ 0.25	7.82 $\pm$ 0.27
K value (D)	43.29 $\pm$ 1.39	43.23 $\pm$ 1.42
Central corneal thickness (μ m)	557.55 $\pm$ 31.25	555.38 $\pm$ 30.17
Lens thickness (mm)	3.48 $\pm$ 0.18	3.48 $\pm$ 0.17
Pupil diameter (mm)	5.97 $\pm$ 0.77	6.00 $\pm$ 0.68

Anterior chamber depth (mm)	3.74±0.24	3.70±0.24
White-to-white (mm)	11.98±0.37	11.95±0.34
AL/CCR ratio	3.07±0.12	3.07±0.11
Cycloplegic Ocular Biometric Values		
Axial length (mm)	24.14±1.03	24.14±1.11
Corneal radius (mm)	7.81±0.27	7.83±0.27
K value (D)	43.33±1.38	43.25±1.42
Central corneal thickness (μm)	575.23±32.16	568.78±34.50
Pupil diameter (mm)	7.88±0.80	7.91±0.67
Anterior chamber depth (mm)	3.83±0.23	3.77±0.24
White-to-white (mm)	11.95±0.37	11.93±0.36
AL/CCR ratio	3.07±0.13	3.07±0.11

Abbreviation: D=dioptr; AL=axial length; CCR=corneal curvature radius; K=keratometry reading; logMAR=Logarithm of the Minimum Angle of Resolution.

3.2 Machine Learning Model Performance

We evaluated five ML models for predicting cycloplegic dilation in each eye, as summarized in Table 2. The comparison of receiver operator characteristic (ROC) curves was depicted in Figure 1A. The XGBoost classifier achieved the highest validation accuracy of 0.818 (95% CI: 0.771–0.864), with an AUC of 0.780 (95% CI: 0.713–0.838). The precision was 0.696 (95% CI: 0.538–0.812), recall was 0.464 (95% CI: 0.341–0.578), and F1-score was 0.557 (95% CI: 0.429–0.653).

For the classification of myopia status, we evaluated five models spanning logistic regression, as detailed in Table 3. The comparison of ROC curves is shown in Figure 1B. LightGBM achieved the highest validation accuracy of 0.908 (95% CI: 0.878–0.936), with an AUC of 0.934 (95% CI: 0.902–0.961). Precision was 0.906 (95% CI: 0.871–0.939), recall was 0.974 (95% CI: 0.955–0.992), and F1-score was 0.939 (95% CI: 0.917–0.958). We further evaluated a high-confidence subset analysis to assess the clinical utility of the LightGBM model for cases where the model exhibited strong predictive certainty. When restricting predictions to samples with model confidence ≥ 0.92 , the coverage was 81.3%, and accuracy improved to 0.968, with an F1-score of 0.981.

For the regression task of predicting continuous post-cycloplegic SER, we evaluated five MLP variants with different architectures and activation functions, as shown in Table 4. The MLP-linear regressor achieved the lowest validation RMSE of 0.630 D (95% CI: 0.541–0.720), with an R^2 of 0.881 (95% CI: 0.826–0.917). The MAE was 0.424 D (95% CI: 0.384–0.469), and the MedAE was 0.311 D (95% CI: 0.273–0.339). The MedAE of 0.311 D is clinically meaningful, as the majority of predictions fall within the ± 0.25 –0.50 D range commonly considered acceptable for refractive prescriptions. Only 6.5% of samples had prediction errors exceeding 1.00 D, suggesting that extreme errors were relatively rare, as shown in Figure 2. The calibration curves of the prediction model across the three tasks are presented in Figure 3.

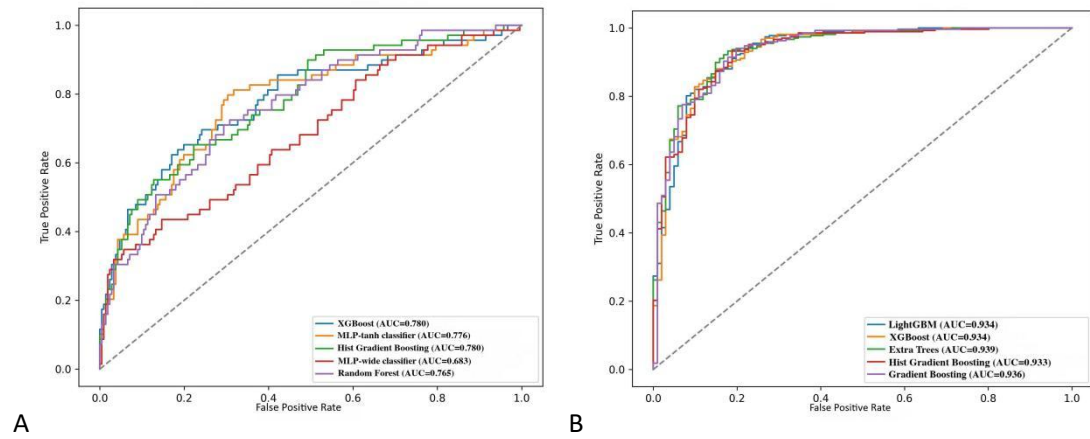


Figure 1. Comparison of ROC curves for predictive performance in cycloplegic dilation (A) and myopia status (B) prediction.

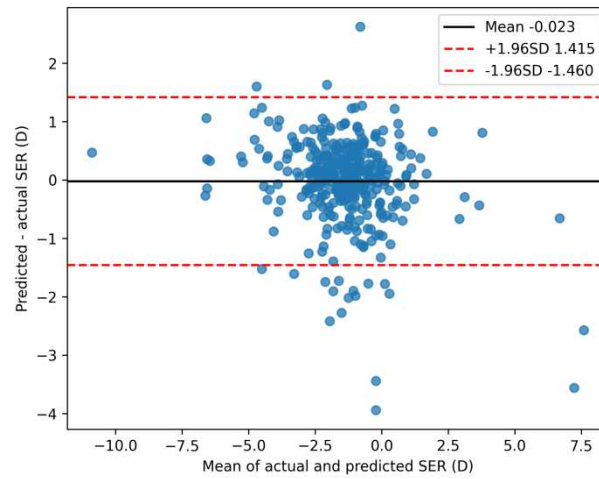


Figure 2. Predictive performance and agreement analysis of predicting post-cycloplegic SER for the overall model.

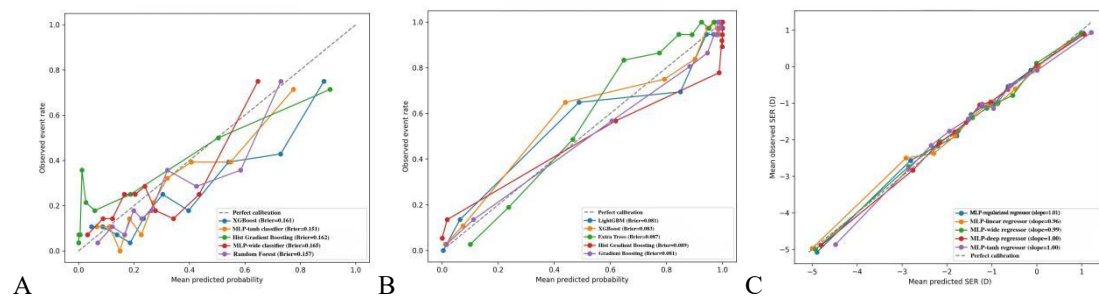


Figure 3. Model calibration curve for cycloplegic dilation prediction (A), myopia status classification (B), continuous post-cycloplegic SER prediction (C).

3.3 SHAP-Based Model Interpretability

To enhance the clinical interpretability of our machine learning models and offer valuable insights for clinical decision-making, we performed SHAP (SHapley Additive exPlanations) analysis for the best-performing model in each task: XGBoost for cycloplegic dilation need, LightGBM for myopia status, and MLP-linear regressor for post-cycloplegic SER. For dilation need prediction (XGBoost),

axial length and pupil diameter were the most influential features, followed by AL/CCR ratio and age. Higher values of these features were associated with increased probability of requiring cycloplegic dilation. For myopia status prediction (LightGBM), non-cycloplegic SER dominated all other predictors, followed by AL/CCR ratio and sphere. More negative SER values, higher AL/CCR ratios, and longer axial length were consistently associated with higher myopia probability. For post-cycloplegic SER regression (MLP-linear regressor), lens thickness, sphere, and cylinder ranked as the top predictors, as shown in Figure 4. Increased lens thickness and more negative cylinder power were associated with more myopic predictions.

Table 2. Performance of Machine Learning Models for Predicting Cycloplegic Dilation

Models	Accuracy(95%CI)	Precision(95%CI)	Recall(95%CI)	F1-score(95%CI)	AUC(95%CI)
XGBoost	0.818(0.771,0.864)	0.696(0.538,0.812)	0.464(0.341,0.578)	0.557(0.429,0.653)	0.780(0.713,0.838)
MLP-tanh classifier	0.814(0.768,0.859)	0.743(0.583,0.886)	0.377(0.263,0.497)	0.500(0.371,0.604)	0.776(0.707,0.841)
Hist Gradient Boosting	0.811(0.768,0.850)	0.667(0.537,0.807)	0.464(0.344,0.573)	0.547(0.432,0.645)	0.780(0.717,0.838)
MLP-wide classifier	0.807(0.761,0.854)	0.759(0.587,0.909)	0.319(0.215,0.428)	0.449(0.320,0.561)	0.683(0.602,0.747)
Random Forest	0.804(0.754,0.852)	0.750(0.585,0.909)	0.304(0.197,0.419)	0.433(0.304,0.555)	0.765(0.697,0.820)

95% CI=95% confidence interval; AUC=area under the receiver operating characteristic curve; XGBoost=eXtreme gradient boosting; MLP-tanh classifier=multi-layer perceptron classifier with tanh activation function; Hist Gradient Boosting=histogram-based gradient boosting; MLP-wide classifier=wide multi-layer perceptron classifier

Table 3. Performance of Machine Learning Models for Predicting Classification of Myopia Status

Models	Accuracy(95%CI)	Precision(95%CI)	Recall(95%CI)	F1-score(95%CI)	AUC(95%CI)
LightGBM	0.908(0.878,0.936)	0.906(0.871,0.939)	0.974(0.955,0.992)	0.939(0.917,0.958)	0.934(0.902,0.961)
XGBoost	0.905(0.872,0.932)	0.906(0.871,0.939)	0.970(0.948,0.989)	0.937(0.914,0.956)	0.934(0.900,0.960)
Extra Trees	0.902(0.868,0.935)	0.929(0.899,0.958)	0.936(0.907,0.964)	0.933(0.909,0.956)	0.939(0.908,0.964)
Hist Gradient Boosting	0.902(0.871,0.932)	0.917(0.883,0.948)	0.951(0.925,0.974)	0.934(0.912,0.955)	0.933(0.902,0.961)
Gradient Boosting	0.902(0.870,0.932)	0.920(0.887,0.950)	0.948(0.923,0.973)	0.934(0.911,0.955)	0.936(0.903,0.962)

LightGBM=light gradient boosting machine

Table 4. Performance of Machine Learning Models for Predicting Post-Cycloplegic SER

Models	MAE(95%CI)	RMSE(95%CI)	MedAE(95%CI)	R2(95%CI)
MLP-linear regressor	0.424(0.384,0.469)	0.630(0.541,0.720)	0.311(0.273,0.339)	0.881(0.826,0.917)
MLP-deep regressor	0.502(0.453,0.554)	0.733(0.642,0.828)	0.355(0.302,0.403)	0.839(0.764,0.886)
MLP-wide regressor	0.516(0.461,0.573)	0.742(0.646,0.842)	0.365(0.345,0.409)	0.835(0.757,0.884)
MLP-regularized regressor	0.556(0.507,0.605)	0.770(0.686,0.857)	0.423(0.367,0.471)	0.823(0.738,0.868)
MLP-tanh regressor	0.511(0.454,0.573)	0.801(0.662,0.939)	0.344(0.306,0.390)	0.808(0.753,0.846)

MedAE=median absolute error

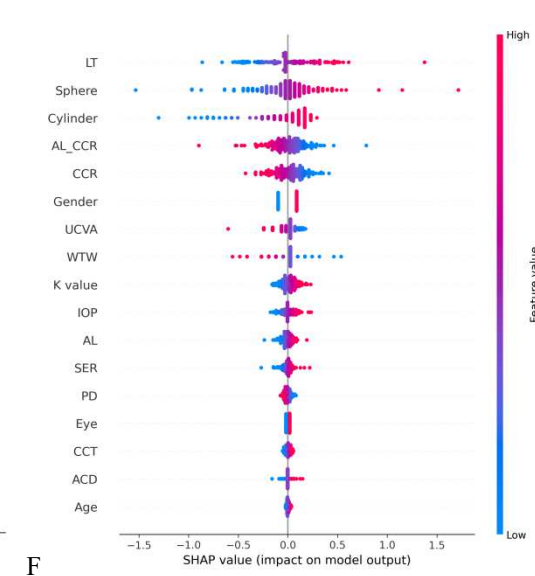
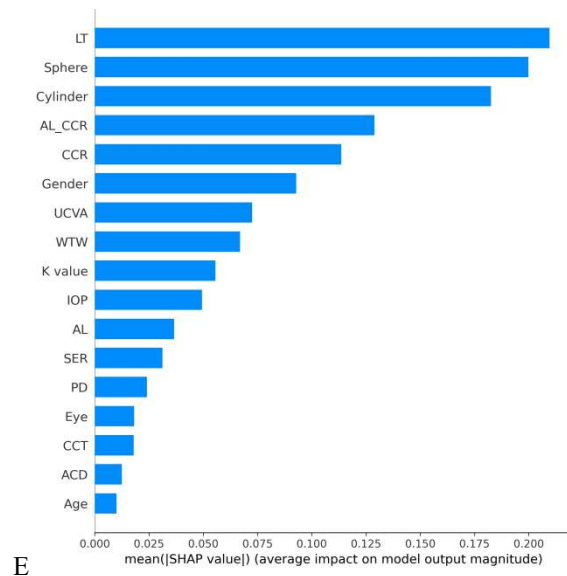
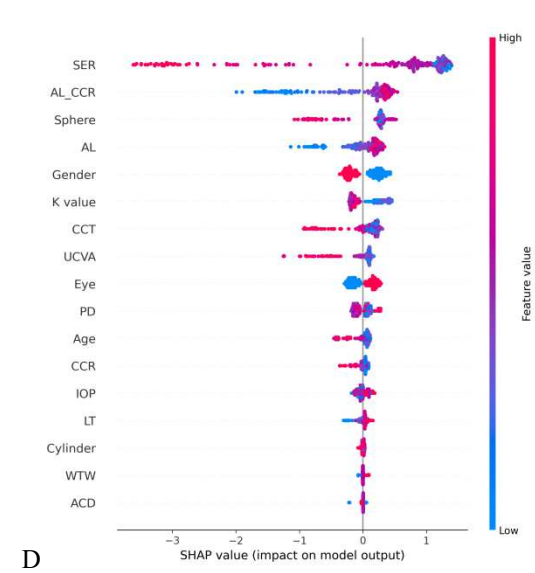
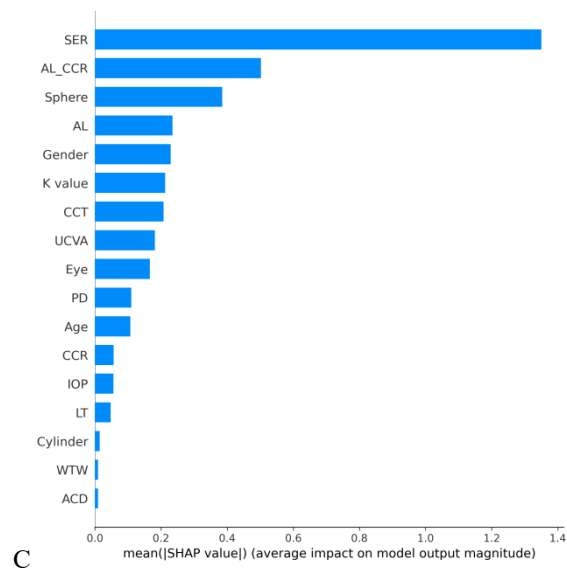
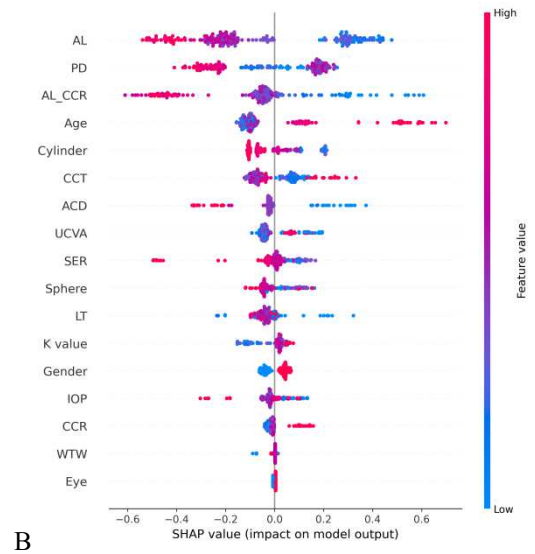
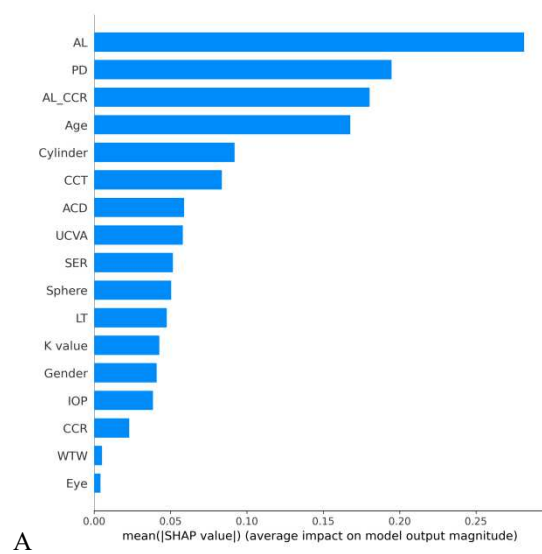


Figure 4. SHAP (SHapley Additive exPlanations) of independent variable importance analysis across three models. AB: Feature importance for cycloplegic dilation using XGBoost; CD: Feature importance for myopia status using LightGBM; EF: Feature importance for post-cycloplegic SER using MLP-linear regressor.

4. Discussion

In this study, we developed and validated a comprehensive machine learning framework that simultaneously addresses three clinically relevant prediction tasks—post-cycloplegic SER regression, myopia classification, and dilation need prediction—using non-cycloplegic parameters and ocular biometric values in a multi-center pediatric cohort. Most prior studies have focused on isolated tasks^[22]. Our results demonstrate that machine learning models can accurately estimate cycloplegic refractive status with clinically acceptable precision, offering a practical alternative when cycloplegic refractive error cannot be obtained.

It is clinically valuable to determine when cycloplegia may be omitted and when it is required for accurate refractive measurement, thereby avoiding unnecessary cycloplegia. We derived 27 parameters from routine non-cycloplegic optometric examinations and machine learning analyses to predict the need for cycloplegia in the first prediction task. Our XGBoost model achieved an accuracy of 0.818 and an AUC of 0.780. Du et al.^[23] reported comparable performance, with AUC values of 0.762–0.833 and accuracy ranging from 68.5% to 77.0% across different thresholds.

For myopia classification, our LightGBM model achieved an accuracy of 0.908 and an AUC of 0.934, with outstanding sensitivity of 0.974. This compares favorably with the results of Ying et al.^[24] (AUC 0.987, sensitivity 91.1%, specificity 97.2%) and Chandra et al.^[25] (AUC 0.99, sensitivity 93.7%, specificity 96.4%). The high-confidence subset analysis for myopia classification (accuracy 0.968 at confidence ≥ 0.92 , covering 81.3% of cases) is particularly promising. In clinical practice, this means that for over 80% of children, this model can make accurate predictions with a high degree of reliability, thereby improving screening efficiency.

The performance of our SER regression model (RMSE 0.630 D, MAE 0.424 D, R^2 0.881) is comparable to that reported in several recent studies. Chen et al.^[26] reported R^2 of 0.920–0.934 and MAE of 0.385–0.413 D in their validation cohort. Ying et al.^[24] achieved R^2 of 0.935 and an MAE of 0.393 D using XGBoost. Sankaridurg et al.^[27] developed a linear regression model for children and adolescents aged 4–15 years and predicted cycloplegic SE using age, uncorrected visual acuity, and non-cycloplegic SE. The model achieved an R^2 of 0.91, and 77% of participants were correctly classified in terms of refractive status; however, no independent validation set was used to evaluate the model's effectiveness. The slightly higher RMSE in our study may reflect differences in cohort composition and the use of a simpler MLP architecture compared to more complex ensembles. Moreover, the RMSE reported for a previous multiple linear regression model was 1.213 D^[28], indicating lower accuracy than that of machine learning models. Notably, our median absolute error of 0.311 D is clinically acceptable and falls within the 0.25–0.50 D range considered tolerable for clinical decision-making.

However, the relatively low recall indicated that the model identified fewer than half of the true positive cases, reflecting the inherent difficulty of predicting which children would show clinically meaningful shifts after cycloplegia in our study. The moderate performance in this task is expected given the inherent complexity of predicting individual accommodative responses; the 0.5D threshold represents a clinically meaningful shift but is influenced by multiple factors including age, baseline refraction, and individual variability in accommodation force. These findings suggested that ensemble tree methods, particularly gradient-boosted trees, were better suited for this task than neural network architectures, likely due to their ability to handle the nonlinear interactions among biometric predictors while maintaining robustness to the class imbalance.

Our SHAP analyses found marked differences in feature importance across tasks, providing actionable insights for future model development. Non-cycloplegic SER dominated myopia classification, but showed minimal importance for dilation need prediction. This suggests that predicting the magnitude of cycloplegic-induced change depends more heavily on biometric and demographic factors than on baseline refractive error alone. Furthermore, we systematically evaluated feature contributions and their directional effects across all three tasks with SHAP analysis, an approach that remains uncommon in this field. Feng et al.^[29] recently demonstrated the value of SHAP for interpretable myopia prediction, showing that AL/CCR ratio, non-cycloplegic SE, and age were key predictors. While most prior studies have reported feature importance rankings alone, our SHAP summary plots explicitly show that higher AL/CCR ratios, longer AL, and more negative non-cycloplegic SER are consistently associated with myopic outcomes across all tasks. This transparency enables clinicians to understand the model's decision-making process and builds trust in its predictions, which is a critical prerequisite for clinical adoption.

Most notably, lens thickness emerged as the most influential predictor of SER regression, surpassing non-cycloplegic SER. This finding is highly consistent with Su et al.^[30], who showed that incorporating lens-related features significantly improved predictive performance, achieving an R^2 of 0.964 and an RMSE of 0.472 D. Taken together, the findings of their study and ours indicate that the lens—a factor frequently overlooked in earlier machine learning studies that primarily focused on axial length and corneal curvature—plays a crucial role in determining refractive outcomes^[31]. Therefore, we suggest that lens parameters should not be overlooked in future refractive prediction models.

We found that the MLP-linear regressor outperformed deeper neural network variants for SER regression, which is noteworthy. Although deep networks offer greater capacity to model complex interactions, they also carry a higher risk of overfitting, particularly with modest sample sizes. The linear MLP architecture may have benefited from regularization and simplicity, effectively capturing the dominant linear relationships while avoiding overfitting to noise. This suggests that, for SER prediction, the underlying relationships may be sufficiently captured by relatively simple models, and that the primary challenge lies in accurate feature measurement rather than complex nonlinear modeling.

Currently, most large-scale screenings rely on non-cycloplegic autorefraction, which systematically overestimates myopia, particularly in younger children with higher diopter values^[32]. The findings of our present study indicate that, in school-based vision screening and epidemiological investigations in which cycloplegia may be impractical, machine learning models can serve as an effective approach for

estimating cycloplegic refractive error and refractive status at the individual level using non-cycloplegic data. Notably, improved assessment of refractive status and more accurate measurement of refractive diopters enable better characterization of the population-level distributions of both refractive status and refractive error, thereby providing substantial value for epidemiological studies.

5. Limitations and Future Directions

There were several limitations in the present study. First, the relatively small sample size compared to other studies may result in more predictive error for cases with larger or smaller refractive errors, reduced generalizability of the findings, or potential biases in the machine learning methodology. Second, we did not incorporate environmental or behavioral risk factors—such as outdoor time, near-work hours, or parental myopia—that are known to influence myopia development and progression. Third, considering that the machine learning process largely depends on observed data, the model needs external validation despite being confirmed using a split test set independent of the training set. Fourth, in this cross-sectional study, we were only able to evaluate the ML models for predicting the cycloplegic refractive error at a single time point. Future longitudinal studies are needed to evaluate how these ML models perform for monitoring the progression of refractive error over time or other cycloplegic medications. Finally, despite achieving balanced and robust performance, concerns about the quality of real-world clinical data and the potential for overfitting issues should be approached with caution.

6. Conclusions

We developed a machine learning framework that accurately predicts post-cycloplegic SER, myopia status, and dilation need using non-cycloplegic parameters and ocular biometric values. SHAP analysis confirmed the biological plausibility of key predictors and revealed the critical role of lens thickness in SER regression—a finding that challenges the field's conventional emphasis on axial length and corneal curvature alone. Our study provides a promising strategy for screening and monitoring myopia status in children and adolescents, and for providing myopia interventions in a precise manner.

Clinical trial number: Not applicable.

Funding: The author(s) declare that no financial support was received for the research and/or publication of this article.

Ethics statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Medical Ethics Review Committee of the leading institution that the First Affiliated Hospital of Jinan University (Approval No. KY-2026-127). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Conflict of interest: No conflicts of interest.

Data Availability Statement:

The datasets used and analysed during the current study available from the corresponding author on reasonable request.

References

- 1 Holden BA, Fricke TR, Wilson DA, et al. Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. *Ophthalmology*, 2016; 123(5): 1036-1042.
- 2 Baird PN, Saw SM, Lanca C, et al. Myopia. *Nat Rev Dis Primers*, 2020; 6(1): 99.
- 3 Grzybowski A, Kanclerz P, Tsubota K, et al. A review on the epidemiology of myopia in school children worldwide. *BMC Ophthalmol*, 2020; 20(1): 27.
- 4 Dong L, Kang YK, Li Y, et al. Prevalence and time trends of myopia in children and adolescents in China: a systemic review and meta-analysis. *Retina*, 2020; 40(3): 399-411.
- 5 Marcus MW, De Vries MM, Junoy Montolio FG, et al. Myopia as a risk factor for open-angle glaucoma: a systematic review and meta-analysis. *Ophthalmology*, 2011; 118(10): 1989-1994.e1982.
- 6 Pan CW, Cheng CY, Saw SM, et al. Myopia and age-related cataract: a systematic review and meta-analysis. *Am J Ophthalmol*, 2013; 156(5): 1021-1033.e1021.
- 7 Haarman AEG, Enthoven CA, Tideman JWL, et al. The Complications of Myopia: A Review and Meta-Analysis. *Invest Ophthalmol Vis Sci*, 2020; 61(4): 49.
- 8 Morgan IG, Iribarren R, Fotouhi A, et al. Cycloplegic refraction is the gold standard for epidemiological studies. *Acta Ophthalmol.*, 2015; 93(6): 581-585.
- 9 Reid JE, Eaton E. Artificial intelligence for pediatric ophthalmology. *Current opinion in ophthalmology*, 2019; 30(5): 337-346.
- 10 Foo LL, Ng WY, San Lim GY, et al. Artificial intelligence in myopia: current and future trends. *Current Opinion in Ophthalmology*, 2021; 32(5): 413-424.
- 11 Lin H, Long E, Ding X, et al. Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical records: a retrospective, multicentre machine learning study. *PLoS Medicine*, 2018; 15(11): e1002674.
- 12 Tang T, Yu Z, Xu Q, et al. A machine learning-based algorithm used to estimate the physiological elongation of ocular axial length in myopic children. *Eye and Vision*, 2020; 7(1): 50.
- 13 Yang X, Chen G, Qian Y, et al. Prediction of myopia in adolescents through machine learning methods. *International Journal of Environmental Research and Public Health*, 2020; 17(2): 463.
- 14 Lu L, Zhou E, Yu W, et al. Development of deep learning-based detecting systems for pathologic myopia using retinal fundus images. *Communications biology*, 2021; 4(1): 1225.
- 15 Foo LL, San Lim GY, Lanca C, et al. Deep learning system to predict the 5-year risk of high myopia using fundus imaging in children. *NPJ Digital Medicine*, 2023; 6(1): 10.
- 16 Li S-M, Ren M-Y, Gan J, et al. Machine learning to determine risk factors for myopia progression in primary school children: the Anyang childhood eye study. *Ophthalmol. Ther.*, 2022; 11(2): 573-585.
- 17 Chandra R, Ying B, Wang J, et al. External validation of machine learning models to predict cycloplegic refractive error and myopia status using non-cycloplegic data. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 2025; 29(4): 104557.
- 18 Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 2017; 30.
- 19 Momeni-Moghaddam H, Maddah N, Wolffsohn JS, et al. The Effect of Cycloplegia on the Ocular Biometric and Anterior Segment Parameters: A Cross-Sectional Study. *Ophthalmol Ther*, 2019; 8(3): 387-395.

- 20 Li L, Fu J, Chen W, et al. Difference of refractive status before and after cycloplegic refraction: the Lhasa Childhood Eye Study. *Jpn J Ophthalmol*, 2021; 65(4): 526-536.
- 21 Pei R, Liu Z, Rong H, et al. A randomized clinical trial using cyclopentolate and tropicamide to compare cycloplegic refraction in Chinese young adults with dark irises. *BMC Ophthalmol*, 2021; 21(1): 256.
- 22 Hao Y, Wang X, Sun B, et al. Machine-learning random forest algorithms predict post-cycloplegic myopic corrections from noncycloplegic clinical data. *Optom Vis Sci*, 2025; 102(3): 138-146.
- 23 Du B, Wang Q, Luo Y, et al. Prediction of spherical equivalent difference before and after cycloplegia in school-age children with machine learning algorithms. *Front Public Health*, 2023; 11: 1096330.
- 24 Ying B, Chandra RS, Wang J, et al. Machine Learning Models for Predicting Cycloplegic Refractive Error and Myopia Status Based on Non-Cycloplegic Data in Chinese Students. *Transl Vis Sci Technol*, 2024; 13(8): 16.
- 25 Chandra RS, Ying BL, Wang JY, et al. Myopia Prediction Using Machine Learning: An External Validation Study. *Vision*, 2025; 9(4): 14.
- 26 Chen B, Tian L, Tian F, et al. Machine learning-driven prediction of cycloplegic refractive error in Chinese children. *Front Cell Dev Biol*, 2025; 13: 1608494.
- 27 Sankaridurg P, He X, Naduvilath T, et al. Comparison of noncycloplegic and cycloplegic autorefraction in categorizing refractive error data in children. *Acta Ophthalmol.*, 2017; 95(7): e633-e640.
- 28 Lee. Correlations among Refractive Error, Axial Length, and Corneal Power in Childhood. *JOURNAL OF THE KOREAN OPHTHALMOLOGICAL SOCIETY*, 2018; 59(5).
- 29 Feng Q, Wu X, Liu QW, et al. Interpretable machine learning models for predicting childhood myopia from school-based screening data. *Scientific Reports*, 2025; 15(1): 11.
- 30 Su Q, Du B, Li B, et al. Predictive Modeling of Cycloplegic Refraction Using Non-Cycloplegia Ocular Parameters With Emphasis on Lens-Related Features. *Transl Vis Sci Technol*, 2025; 14(10): 3.
- 31 Kim SR, Kang DH, Choe GS, et al. Ensemble machine learning prediction model for clinical refraction using partial interferometry measurements in childhood. *PLoS One*, 2025; 20(7): e0328213.
- 32 Hyewon M, Jing Y, Yang S, et al. [Changes in refraction before and after cycloplegia among children and adolescents aged from 2 to 15 years]. *Chinese Journal of Ophthalmology and Otorhinolaryngology*, 2018; 18(3): 196-200.