WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 5, Number 1, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v5n1.14
Journal homepage: https://wepub.org/index.php/IJCSIT/index

A Myopia Degree Prediction Model for Children and Adolescents Based on Follow-up Data with Different Time Intervals

Pengyuan Ma [1], Shukui Ma [1], Guowei Zhang [1], Fei Ma [2], Guangping Zhuo [1, *]

[1] School of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, China

[2] Department of Computer Science and Technology, Taiyuan University, Taiyuan 030032, China

ABSTRACT

Myopia has become a global public health issue, and existing research primarily focuses on predicting the future spherical equivalent (SE) of children and adolescents based on ocular optical data. However, electronic medical record (EMR) data often exhibit characteristics such as multi-dimensionality, irregular time series, missing values, and non-uniform time intervals, which present significant challenges for prediction. To address these issues, this paper proposes a prediction model named IST-xLSTM, which utilizes a temporal dense prediction module to handle missing data and overcome long-range dependencies and irregular time intervals. Additionally, a hierarchical temporal attention mechanism is introduced to extract both local and global features along the irregular time dimension, enhancing prediction accuracy. Experimental results show that the mean absolute error on the validation set is 0.216 ± 0.176 (D), which is significantly lower than the clinically acceptable prediction threshold of 0.50 (D).

KEYWORDS

Prediction of Myopia in Children and Adolescents; Spherical Equivalent (SE); IST-xLSTM; Hierarchical Temporal Attention Mechanism

1. INTRODUCTION

Myopia has become a global public health issue, severely affecting the vision and mental health of children and adolescents, while also imposing an economic burden on individuals, families, and society [1]. According to the World Health Organization's 2022 report, over 2 billion people worldwide suffer from impaired vision due to uncorrected refractive errors, with at least 1 billion facing preventable vision problems [2]. In China, the prevalence of myopia among adults is growing exponentially, sparking widespread public concern about health. By the age of 15, 78.4% of urban students have myopia, and this proportion rises to 80% by the age of 18 [3, 4]. Studies also show that the incidence of high myopia [SE ≥ -6.0 (D)] is increasing, which is closely associated with lifelong risks such as retinal detachment, myopic macular degeneration, and glaucoma [3]. As there is currently no cure for myopia, preventing its onset and progression has become particularly important. Early intervention and appropriate measures can significantly reduce the risk and consequences of myopia [5]. Therefore, effectively preventing and controlling myopia in children and adolescents has become an urgent task.

With the widespread application of machine learning (ML) algorithms in the healthcare field, artificial intelligence (AI) is expected to bring revolutionary changes to disease diagnosis and prediction, thereby improving medical standards. In-depth analysis of large-scale ophthalmic data

using big data and AI technologies [6] is expected to reveal the patterns and influencing factors of myopia development, enabling the formulation of targeted prevention and management measures to reduce the risk of myopia. Spherical equivalent (SE) is the foundation for screening and diagnosing myopia [7]. The quantitative prediction of SE not only reflects the specific changes in myopia development but also allows for the early implementation of targeted intervention measures. Considering that the vision health status of each child and adolescent may differ, it is particularly important to develop personalized vision protection plans based on individual differences [8]. Through machine learning and AI technologies, personalized vision protection strategies can be created for each child and adolescent based on individual characteristics and environmental conditions, including improving lifestyle habits, scientifically managing eye use, and providing appropriate prescriptions, thereby maximizing vision health protection.

Previous studies have identified various risk factors for the occurrence and development of myopia, such as age, gender, genetics, and outdoor activities [9-11]. In recent years, an increasing number of studies have focused on the prediction of myopia or high myopia in different populations. Most existing studies adopt traditional models, such as linear regression, support vector machines, and decision trees [12-19]. In contrast, deep learning can be trained using complex nonlinear parameters [20] and more effectively learns data features, making it superior to traditional models in various medical prediction tasks [21-23]. However, the application of deep learning in myopia prediction remains relatively scarce. Spadon et al. [24] pointed out that, in addition to observing static symptoms, temporal dynamics also provide important information. However, extracting time features becomes quite complex due to the irregular multivariate time series nature of electronic medical record (EMR) data.

To address the above-mentioned issues, this study proposes a novel network architecture—IST-xLSTM (Irregularly-Sampled-Time xLSTM), aimed at utilizing irregularly sampled electronic medical record (EMR) data for the prediction of myopia in children and adolescents. The IST-xLSTM model consists of two main parts: 1) Missing Data and Irregular Time Series Processing: We introduce a Temporal Dense Prediction Module (Dense Time-xLSTM Model), which is an adjustment and improvement of the original xLSTM [25] network. However, the traditional xLSTM still cannot handle irregular time series, so we integrate a Dynamic Time Decay Module, which effectively resolves the time inconsistency issue in irregular time series. By dynamically adjusting the memory cells of the xLSTM network, this module enables the model to more effectively handle data with irregular time intervals. This mechanism not only improves the accuracy and robustness of time series predictions but also ensures that the model can capture crucial information when faced with varying time intervals. Additionally, considering the possible issue of missing data, this model uses a NanDense layer with Missing Indicators [26], which can adaptively handle missing values, avoiding the drawbacks of data imputation and maintaining competitive performance without adding extra parameters or computational burden. 2) Global and Local Irregularity Handling: This study also introduces a Hierarchical Temporal Attention Mechanism, which effectively extracts features both locally and globally from the data through two attention modules, enhancing the model’s understanding and interaction with data features.

In summary, the contributions of this study are as follows:

(1) Proposed the IST-xLSTM network architecture, specifically designed for predicting myopia in children and adolescents using irregularly sampled electronic medical record data, successfully addressing challenges such as irregular time intervals and missing values.

(2) Introduced the Temporal Dense Prediction Module and Dynamic Time Decay Mechanism, effectively addressing the time inconsistency in irregular time series, thereby enhancing the model's ability to handle irregular time data and improving prediction accuracy. At the same time, the NanDense layer with missing indicators was employed to adaptively handle missing values, avoiding

the issues caused by imputation and ensuring the model's advantages in both performance and computational efficiency.

(3) Developed the Hierarchical Temporal Attention Mechanism, which enhances the model's understanding and interaction with data features through global and local feature extraction, thereby improving prediction accuracy.

(4) A large number of experimental results demonstrate that, while ensuring data privacy and computational efficiency, the IST-xLSTM network is feasible and effective in handling irregularly sampled electronic medical record data for myopia prediction in children and adolescents.

2. MODEL

To address the issues in myopia prediction for children and adolescents, this paper proposes a novel myopia prediction network, named Irregularly-Sampled-Time xLSTM (IST-xLSTM), the structure of which is shown in Figure 1. To tackle the problems of missing values and irregular time series, the network introduces a Temporal Dense Prediction Module (Dense Time-xLSTM Model).

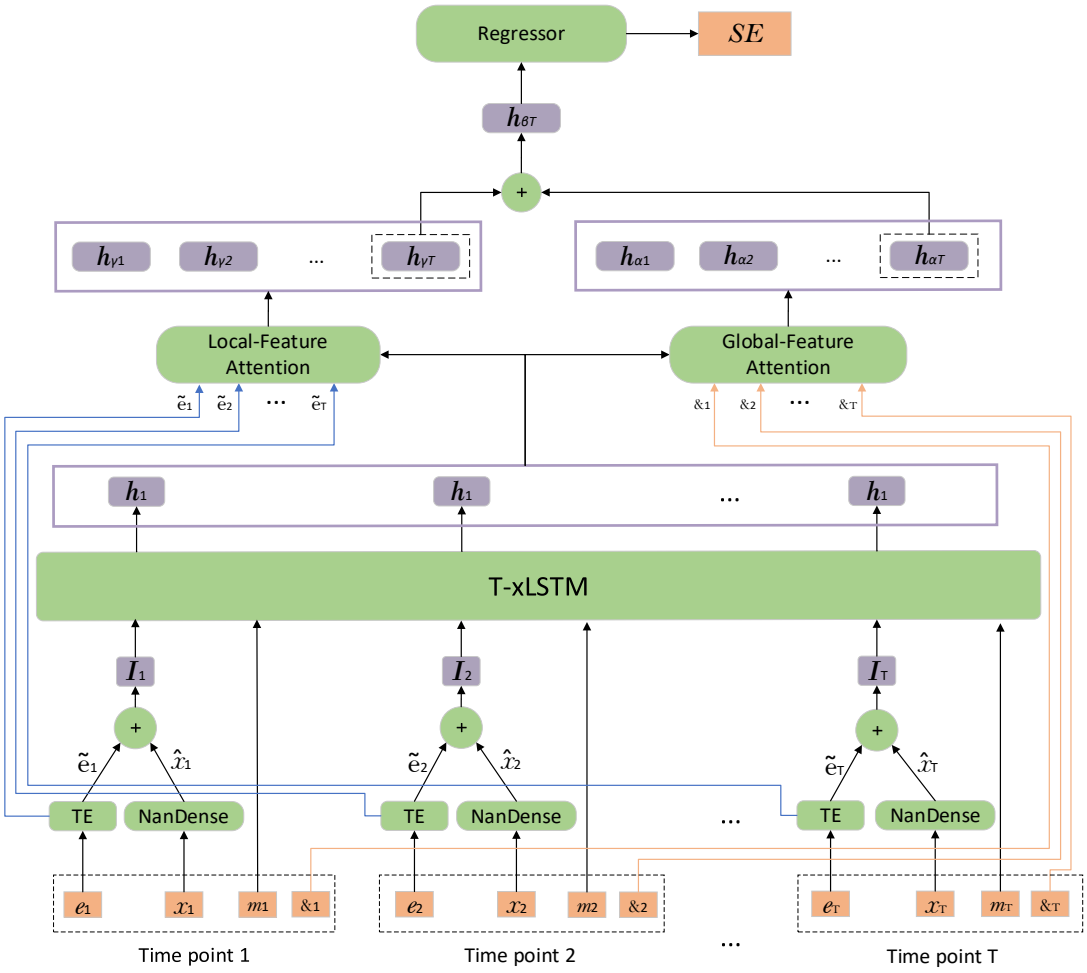

Figure 1. Irregularly-Sampled-Time xLSTM

First, to address the issue of missing values, we employ a non-imputation method. This approach involves using a binary variable for each data point to indicate whether the value is missing, while replacing all missing values with a constant. Although this method is simple, its effectiveness heavily depends on the choice of the constant. Less common methods for handling missing values include likelihood-based methods and sub-model approaches. When dealing with high-dimensional data, the performance of imputation methods is limited, and most alternative methods require manual parameter settings. We adopt NanDense, an adaptive method for handling missing values that requires

no additional parameters and allows neurons to automatically adjust their responses. Additionally, we combine missing indicators to integrate prior knowledge, enabling the model to focus attention on missing values that differ from the observed data. Similarly, the structure of xLSTM is improved, where continuous time dynamics can be applied to weight the hidden states using ordinary differential equations (ODEs), and the model is merged to form T-xLSTM, which improves performance for irregular time series tasks. To handle both global and local irregularities, this study also introduces a Hierarchical Temporal Attention Mechanism. However, existing methods focus only on global irregularities and neglect local irregularities, which could result in the loss of complex details in the predicted data. To address this issue, our research combines two different attention mechanisms: Local-Feature Attention and Global-Feature Attention. These mechanisms respectively handle interactions within a single record, local irregularities, and global irregularities.

We model the irregularly sampled electronic medical record (EMR) data associated with a single patient as a multivariate time series $\mathrm{X} = \{_{\mathrm{X}1}, _{\mathrm{X}2}, \cdots, _{\mathrm{X}T}\}$, which includes up to T patient visit records. The data for each patient visit is modeled as a multidimensional vector $x_t = \{x_t^1, x_t^2, \cdots, x_t^K\}$, containing K diagnostic indicators. The time intervals between two consecutive visits are recorded as a T-dimensional vector$\Delta = \{\Delta_1, \Delta_2, \cdots, \Delta_T\}$, where $\Delta_1 = 0$. Since patient visits depend on clinical needs, the time intervals Δ_t can range from several months to years. Moreover, as different tests may be required at each visit, the multivariate time series X may contain many missing values. This study uses the patient's entire historical record X to predict the ocular spherical equivalent (SE) value at the next time point $(T + 1)$.

2.1. Temporal Dense Prediction Module

It consists of a series of modules, which will be further described later: 1) Input Representation Module; 2) NanDense Layer; 3) Time Decay Module; 4) T-xLSTM Module.

1) Input Representation Module: For each patient, the irregularly sampled electronic medical record (EMR) data is input into the model in chronological order. At time t, four inputs are extracted from the records:

$e_t \in R^{1\times K}$ represents the vector of feature time intervals, capturing local irregularities. If x_{Kt} is observed, then e_{Kt} represents the time difference between t and $t-1$, where $e_{K1} = 0$.

$x_t \in R^{1\times K}$ represents the raw irregularly sampled electronic medical record (EMR) data, which may contain many missing values. These missing values should not be set to 0, as they indicate unmeasured items, not measurements of 0. The NanDense layer handles these missing values, converting them into inputs acceptable by the model while preserving potential information.

$m_t \in R^{1\times K}$ represents a vector used to indicate whether specific test items are missing at time t. Specifically, if a value in x_t is missing, the corresponding value in m_t is assigned as 0; otherwise, it is assigned as 1. The inclusion of missing indicators provides the model with data information, allowing it to distinguish between data values and missing values.

$\delta_t \in R^{1\times K}$ represents the time interval between the current detection time t_k and the last prediction time t_{T+1}, accounting for global irregularities.

2) NanDense Layer: Irregularly sampled electronic medical record (EMR) data is first processed by the NanDense layer. Data containing NaN values is not suitable for direct input into the T-xLSTM model. The NanDense layer is an optimized Dense layer designed to address the missing value problem, especially when there is a large amount of missing data. It mitigates the impact of incomplete datasets by adjusting its internal structure, rather than using imputation. x_t is a vector containing K features. Due to the large number of missing values in x_t, it is split into two parts: the available data x_{at} containing q features, and the missing data x_{mt} containing r missing values

($K = q + m$). When x_t is input into a neural network model with a Dense layer, the activation of the v-th neuron in the Dense layer (where v is determined as needed) can be expressed as:

$$f(v) = \sum_{x_{ti} \in x_{ta}} x_{ti} w(v)_i + \sum_{x_{tj} \in x_{tm}} x_{tj} w(v)_j + b(v) \quad (1)$$

Here, $f(v)$ represents the activation of the neuron, w represents the weight of the neuron, and b represents the bias. When x_t contains multiple missing values, Equation (1) cannot be applied. Therefore, to address this issue, the term involving missing data (the second term in Equation (1)) is removed, and the bias is adjusted to neutralize its effect. The modified activation is expressed as:

$$f(v) = \sum_{x_{ti} \in x_{ta}} x_{ti} w(v)_{ti} + q b(v) + r w(v)_c \frac{1}{K} \quad (2)$$

Here, w_c is a compensation factor used to prevent insufficient neuron output when processing high-dimensional data. w_c is learned during the training process. After processing by the NanDense layer, the vector $x_t \in R^{1 \times K}$ is embedded into the vector $\hat{x}_t \in R^{1 \times K}$. In this way, no additional parameters or models are required, ensuring computational efficiency, especially for high-dimensional data. Moreover, all parameters are derived from the training and learning phases, allowing the data-driven approach to fully optimize the parameters.

3) Time Decay Module: The feature time interval vector $e_t \in R^{1 \times K}$ is modified into another vector $\tilde{e}_t \in R^{1 \times K}$, with the goal of prioritizing more recent test results. The modification process is as follows:

$$f_t = \left(w_{e1} \odot \frac{e_t}{\delta_{max}} + b_{e1} \right) \quad (3)$$

$$\tilde{e}_t = w_{e2} \odot (v - \tanh(f_t \odot f_t)) + b_{e2} \quad (4)$$

Here, $\odot$ denotes element-wise multiplication, and $w_{e1}, b_{e1}, w_{e2}, b_{e2} \in R^{1 \times K}$ are learnable parameters. tanh represents the hyperbolic tangent activation function. The vector $v \in R^{1 \times K}$ has all elements equal to 1, and δ_{max} is the maximum value of δ_t. Next, the output from the NanDense layer $\hat{x}_t$ is added to $\tilde{e}_t$, resulting in a data vector that includes local anomaly awareness: $I_t = \hat{x}_t + \tilde{e}_t$.

4) T-xLSTM Module: The proposed T-xLSTM-based model combines several key components to effectively handle time series prediction tasks. Figure 2 provides an overview of the model architecture.

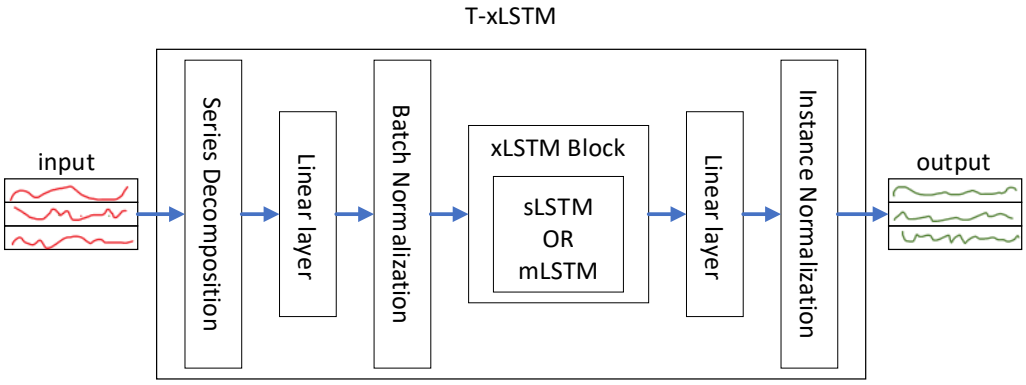

Figure 2. T-xLSTM

We further perform batch normalization to provide stability during the learning process. Batch normalization is a transformation technique in deep learning that normalizes the activations of each

layer, stabilizing the distribution of network inputs. It allows for higher learning rates, accelerates training, and reduces the need for strict initialization and certain forms of regularization, such as Dropout. By addressing internal covariate shift, batch normalization improves the stability and performance of the network across various tasks. It introduces two additional trainable parameters at each layer, enabling deeper networks to train faster and more effectively. The xLSTM module consists of two components: sLSTM and mLSTM. The sLSTM component uses scalar memory and exponential gating to manage long-term dependencies and control the appropriate retention of historical information. The mLSTM component uses matrix memory and covariance update rules to enhance storage capacity and the ability to retrieve relevant information. Depending on the dataset's characteristics, we choose a hybrid component of sLSTM or mLSTM, and dynamically decay each memory cell. The decay is represented as follows:

$$c_{st} = \tanh(w_c \odot c_t) \tag{5}$$

$$adaptive\_decay\_c = \sigma(adaptive\_param\_c) \tag{6}$$

$$c_{st\_dis} = adaptiv\_dacay\_c \odot \tilde{e}_t \odot c_{st} \tag{7}$$

$$c_t = c_t - c_{st} + c_{st\_dis} \tag{8}$$

First, the current memory cell state $c_t \in R^{1\times H}$ is transformed using the tanh activation function to compute the new state $c_{st} \in R^{1\times H}$. Then, the adaptive parameters are normalized using the sigmoid function to obtain a decay value between 0 and 1. Next, this decay value is multiplied by the $\tilde{e}_t$ from the time decay module and the new state c_{st} to obtain the adjusted state $c_{st\_dis} \in R^{1\times H}$. Finally, the cell state c_t is updated by subtracting the original state and adding the decayed state, yielding the final cell state.

2.2. Hierarchical Temporal Attention Mechanism

It includes two attention modules, which will be further described later: 1) Local Attention; 2) Global Attention; 3) Multi-Layer Perceptron Regression Module.

1) Local Attention: The time intervals for each data point may vary, a phenomenon we previously referred to as local irregularity. To model this clinical approach, we developed a local attention module. This module is specifically designed to account for the irregularity of each test item and effectively extract important information from the hidden representations.

First, all hidden representations are transformed into query vectors, represented as:

$$h^* = ReLU(w_h[h_1,\ h_2,\ \cdots h_T]\ + b_h) \tag{9}$$

Here, $w_h \in R^{1\times H}$ and $b_h \in R^{1\times T}$ are learnable parameters. ReLU represents the Rectified Linear Unit activation function. Next, we use the irregular time interval $\tilde{e}_t$ as the key vector to compute attention in $h^* \in R^{1\times T}$, and obtain the new hidden representations, given as follows:

$$\alpha_i = \frac{\tilde{e}_i w_\gamma\, h_*^{\mathrm{T}}}{\sqrt{K}},\ i = 1,\ 2,\ \cdots,\ T-1 \tag{10}$$

$$[\gamma_1,\ \gamma_2,\ \cdots \gamma_{T-1}] = softmax(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_{T-1}) \tag{11}$$

$$h_{\gamma T} = \sum_{t=1}^{T-1} \gamma_t h_{\mathrm{t}} \tag{12}$$

Here, $w_\gamma \in R^{K\times T}$ is the learnable parameter matrix. $h_{\gamma T}$ represents the hidden representation containing local irregularity information.

2) Global Attention: In the overall evaluation, earlier visits are typically given less attention (referred to as global irregularity). To achieve this, we use global attention. The time interval δ_t is transformed into a decay factor and combined with each hidden representation to obtain a globally time-aware hidden representation, expressed as:

$$\tilde{\delta}_t = w_{\delta 2}\left(1 - \tanh\left(\left(w_{\delta 1}\frac{\delta_t}{\delta_{max}} + b_{\delta 1}\right)^2\right)\right) + b_{\delta 2} \tag{13}$$

$$[\beta_1,\ \beta_2,\ \cdots,\ \beta_{T-1}] = softmax\left(\tilde{\delta}_1, \widetilde{\delta}_2,\ \cdots, \widetilde{\delta}_{T-1}\right) \tag{14}$$

$$h_{\beta T} = \sum_{t=1}^{T-1}\beta_t h_t \tag{15}$$

Where $w_{\delta 1}, w_{\delta 2}, b_{\delta 1}, b_{\delta 2} \in R$ are the learnable parameters.

3) Multi-Layer Perceptron Regression Module: So far, we have obtained the final hidden layer data $h_{\gamma T}$ and $h_{\beta T}$ from the local attention and global attention. They are eventually combined as $h_{\tau T} = h_{\gamma T} + h_{\beta T}$, serving as the input to the multi-layer perceptron regression module.

This module predicts the refractive error of the eye. Using $h_{\tau T}$ a three-layer perceptron (MLP) is employed to perform these tasks. The root mean square error (RMSE) and multi-class cross-entropy are chosen as the corresponding loss functions for learning the parameters of each module.

3. EXPERIMENTAL DATA

3.1. Dataset

The dataset used in this study was collected by the Shanxi Provincial Ophthalmic Hospital between March 2019 and September 2023 through regular and irregular ophthalmic examinations, aiming to explore the eye health status of adolescents and to study and predict the development trends of myopia. The dataset includes 941 adolescent patients aged between 3 and 20 years, with a total of 6,918 independent examination records. The number of follow-up visits for each patient ranges from 2 to 6, with follow-up intervals varying from 1 month to 4 years. The dataset records detailed ophthalmic physiological parameters and refractive status information, including axial length, anterior chamber depth, lens thickness, aqueous depth, pupil diameter, central corneal thickness, corneal curvature, spherical equivalent, and cylindrical power. Additionally, the demographic information of the patients, such as age, gender, date of birth, and the specific dates of each examination, are also thoroughly recorded. The inclusion criteria for participants were: aged between 6 and 18 years, at a critical stage of myopia onset, with no other serious eye diseases (such as retinal lesions, cataracts, glaucoma, etc.), and having at least two complete ophthalmic examination records covering ocular physiological parameters and refractive error data. All patients or their legal guardians have signed an informed consent form and agreed to participate in the study. The exclusion criteria included: having other eye diseases that might interfere with vision assessment, undergoing any form of eye surgery (such as refractive surgery) within one year prior to the start of the study, missing key ophthalmic physiological parameters or refractive error data, having systemic diseases that severely affect vision development or prevent participation in follow-up, or being unable to meet the age requirements or provide informed consent. In-depth analysis of this dataset will provide strong support for the research and prediction of myopia in adolescents.

The relevant myopia diagnostic standards are based on the "2021 Asian Myopia Management Consensus" issued by the Asian Optometric Management Association (AOMA) and the Asian

Optometric Conference (AOC). Based on the spherical equivalent (SE) of the eye when relaxed, the criteria for myopia are as follows: $\mathrm{SE} \leq -0.5\mathrm{D}$. The severity of myopia is classified as: (1) Mild myopia: $-3.0\mathrm{D} < \mathrm{SE} \leq -0.5\ \mathrm{D}$; (2) Moderate myopia: $-6.0\mathrm{D} < \mathrm{SE} \leq -3.0\mathrm{D}$; (3) High myopia: $\mathrm{SE} \leq -6.0\mathrm{D}$.

3.2. Data Cleaning and Data Transformation

The cleaned dataset includes 890 children and adolescents with 1,780 eyes (samples). Each sample is associated with 2 to 6 records. The sample distribution is as follows: 706 samples with 2 records, 514 samples with 3 records, 276 samples with 4 records, 154 samples with 5 records, and 130 samples with 6 records. The time interval between the first and last record of any sample ranges from 1 quarter (<1 quarter) to 16 quarters (≥15 quarters, <16 quarters). Each record contains 13 features, as detailed in Table 1.

Table 1. Non-standard Time Electronic Medical Record Data Distribution Statistics

Features	Statistical Data
Id	
Gender	0 (55.7%): Male; 1 (44.3%): Female
Age	9.17, [6, 18]
School Age Stage	0 (59.6%): Primar; 1 (25.7%): Middl; 2 (14.7%): High
OAL	-0.36, [-14.25, 10.50]
CYL	-0.75, [-6.75, 4.00]
SE	-0.74, [-14.88, 10.12]
AL	23.39, [19.01, 28.36]
ACD	3.06, [1.88, 3.84]
LT	3.52, [2.67, 4.77]
CCT	548.15, [449, 688]
K1 Value	42.54, [37.47, 47.95]
K2 Value	44.26, [39.62, 51.47]

In statistical analysis, appropriate methods are used to detect and handle outliers. For the detected outliers, corresponding handling strategies are adopted based on specific circumstances: obvious data entry errors are corrected, implausible outliers are deleted, or reasonable outliers are retained with the support of medical professional knowledge. This decision process is comprehensively judged based on medical background and data analysis results. For categorical variables (e.g., gender, school stage), one-hot encoding is used to convert non-numeric features into a numerical format that can be processed by the model. For numeric data, Z-score standardization is performed, adjusting the data to have a mean of 0 and a standard deviation of 1. This method not only retains the distribution characteristics of the data but also ensures that all numeric features are on the same scale, improving the model's performance. Based on medical knowledge and exploratory data analysis results, features closely related to myopia development are selected to reduce model complexity and improve prediction accuracy. For numeric data, Z-score standardization is applied as follows:

$$Z = \frac{X-\mu}{\sigma} \tag{16}$$

Where μ represents the mean of the dataset and σ represents the standard deviation of the dataset.

3.3. Data Augmentation

To increase the sample size, this study employed a data truncation technique by selecting specific segments from the time series to enhance the diversity of the dataset. The main advantage of the data truncation technique is its ability to systematically expand the variety of samples within the dataset. By extracting data from different time periods, multiple data subsets can be created that represent different characteristics at various time points. This method is particularly suitable for situations where the sample size is limited but data records are frequent, effectively improving the model's generalization ability and its adaptability to time variations.

Specifically, the historical records of children or adolescents are divided into multiple samples, ensuring that all input data used for training and prediction are recorded before the label (i.e., SE value). For example, the four records of a child or adolescent (a, b, c, d) can be split into 11 samples, as shown in Table 2.

Table 2. Augmented Sample Table Extracted from 4 Records (a, b, c, d) of Children or Adolescents

Input	Labels	Time Interval
[a]	[b]	[ab]
[a]	[c]	[ac]
[a]	[d]	[ad]
[b]	[c]	[bc]
[b]	[d]	[bd]
[c]	[d]	[cd]
[a, b]	[c]	[ab, bc]
[a, b]	[d]	[ab, bd]
[a, c]	[d]	[ac, cd]
[b, c]	[d]	[bc, cd]
[a, b, c]	[d]	[ab, bc, cd]

After data preprocessing, the sample size of the dataset increased to 17,212. Among them, the number of samples with sequence lengths of 2, 3, 4, 5, and 6 were 7,394, 5,758, 2,996, 934, and 130, respectively. Subsequently, the dataset was stratified by sequence length, and each layer was further divided into training (80%), validation (10%), and test (10%) sets.

4. EXPERIMENTS AND DISCUSSION

4.1. Experimental Environment and Parameter Settings

The experiment is based on the Linux operating system, and the training hardware environment primarily includes 8 GeForce RTX 3090 GPUs to ensure the efficiency of model training and prediction. The software environment uses Python 3.9.13 and the PyTorch 1.13 framework. The model employs an improved Long Short-Term Memory network (xLSTM) as the base network, incorporating a hierarchical temporal attention mechanism to capture both local and global features of irregular data.

The main parameter settings for the model are as follows: the initial learning rate is set to 1e-4, and the Adam optimizer is used. The sLSTM network uses 4 attention heads, each with a size of 128 and a dropout rate of 0.15. In the mLSTM network, 4 attention heads are also used, each with a size of 128, but the dropout rate is set to 0.25. Each experiment runs for a maximum of 800 epochs, and the learning rate decreases by one order of magnitude every 50 epochs. If the loss on the test set does not decrease for 10 consecutive epochs, training is terminated early.

Table 3. MAE Table of SE Under Different Sequence Lengths Corresponding to Different Stratification Conditions

Strat. Cond.		The MAE of SE at different sequence lengths (M ± SD)				
		2	3	4	5	Total
Quarter	1-2	0.163±0.122	0.159±0.096	0.107±0.083	0.083±0.058	0.116±0.102
	3-4	0.208±0.153	0.198±0.103	0.153±0.086	0.118±0.052	0.163±0.122
	5-6	0.219±0.106	0.208±0.112	0.172±0.126	0.136±0.062	0.192±0.146
	7-8	0.253±0.164		0.202±0.122		0.226±0.152
	9-10	0.303±0.219				0.303±0.219
MD	SE > -0.5	0.254±0.150	0.204±0.132	0.196±0.129	0.123±0.097	0.232±0.124
	-3.0<SE≤-	0.242±0.148	0.201±0.108	0.158±0.103	0.135±0.083	0.217±0.131
	-5.0<SE≤-	0.243±0.129	0.216±0.123	0.172±0.095	0.118±0.069	0.192±0.103
	SE≤-5.0	0.226±0.124	0.210±0.112	0.185±0.108		0.216±0.127
Age	6-9 yrs	0.256±0.146	0.217±0.131	0.193±0.126		0.206±0.129
	9-12 yrs	0.233±0.140	0.201±0.108	0.178±0.102	0.112±0.091	0.192±0.118
	12-18 yrs	0.209±0.113	0.196±0.092	0.153±0.95	0.106±0.126	0.173±0.106

4.2. Experimental Evaluation Metrics

In this experiment, the loss function of the model is Mean Squared Error (MSE), which has a range of [0, +∞). MSE calculates the average of the squared differences between the predicted and actual values, and it is commonly used for regression problems. The calculation formula is as follows:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \tag{17}$$

Where y_i is the i-th true value, $\hat{y}_i$ is the i-th predicted value, and m is the number of samples. MSE is a smooth, continuous, and differentiable function, making it well-suited for gradient descent algorithms. The smaller the MSE, the better the model's predictive performance.

At the same time, the model's predictive performance is also evaluated using the Mean Absolute Error (MAE). MAE measures the average absolute difference between the predicted values and the actual values, with a range of [0, +∞). The formula is as follows:

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \tag{18}$$

MAE is less sensitive to outliers, making it a robust error metric. The smaller the MAE, the better the model's predictive performance. When the MAE is 0, it indicates that the model's predictions are perfectly accurate.

4.3. Experimental Results Analysis

After 780 training iterations, the model converged, and the loss (i.e., MSE) during the training process is shown in Figure 3. The MAE of future SE on the validation set is 0.216±0.176 (D). The MAE results under different stratification conditions are detailed in Table 3. When the sequence lengths are 2, 3, 4, and 5, the corresponding MAE ranges are observed for different interval quarters, different degrees of myopia, and different ages.

The table presents the changes in the mean and standard deviation of MAE (Mean Absolute Error) for SE (Spherical Equivalent) under different stratification conditions (quarters, myopia degree, and age) for various sequence lengths. In the stratification by quarters, sequences with shorter time

intervals show smaller prediction errors, with the overall best performance occurring in the 1-2 quarters group, achieving 0.116±0.102 (D). Age stratification shows that the prediction errors are generally higher for younger groups (6-9 years old) compared to older groups, with this difference being more pronounced in shorter sequences. Overall, as the sequence length increases, the MAE value decreases, demonstrating better predictive accuracy. It is worth noting that an MAE within 0.50 (D) is considered clinically acceptable prediction error. Based on the model's high accuracy and robustness, it holds significant clinical application value in predicting short-term vision outcomes for children and adolescents.

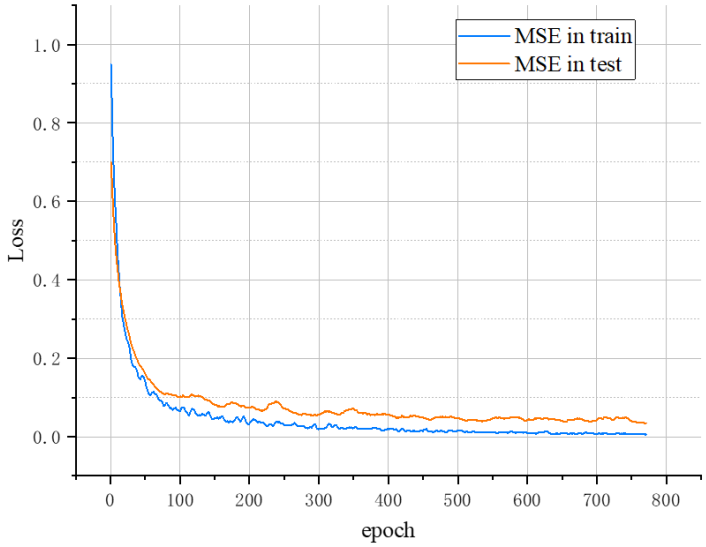

Figure 3. Training loss and test loss values of the model

4.4. Ablation

The main purpose of the ablation experiment is to evaluate the impact of each feature variable and model component on overall prediction performance. By gradually removing or adjusting certain parts of the model, this method helps identify the features or components that are crucial for the model's prediction accuracy, thus providing a basis for optimizing the model structure and improving overall performance.

4.4.1. Feature Ablation Experiment

We performed an ablation study by gradually removing input features and observing the changes in model performance. Specifically, we sequentially removed key features such as age and gender, spherical and cylindrical refractive errors, K1 and K2 values, axial length, lens thickness, anterior chamber depth, central corneal thickness, and all eye parameters, recording the model performance after each ablation. To ensure stability and reproducibility, each experiment was repeated ten times, and the average value was taken. The results are shown in Table 4.

Table 4. MAE Table of SE Under Different Feature Removal Conditions

Remove features	MAE of SE (Mean ± SD)
Baseline	0.216±0.176 (D)
Age and Gender	0.353±0.289 (D)
OAL & CYL	0.416±0.354 (D)
k1, k2	0.275±0.193 (D)
AL	0.264±0.206 (D)
LT, ACD, CCT	0.278±0.195 (D)
AOP	0.304±0.235 (D)

Through the ablation experiments, we found that features such as age, gender, and spherical equivalent (SE) significantly impacted the model's prediction performance. In particular, after removing the spherical equivalent and cylindrical power features, the model's MAE increased significantly, indicating that these features are crucial for the model's predictive accuracy. Although removing the ocular parameters caused a decrease in model performance, there was no significant deterioration. This phenomenon may be attributed to the fact that spherical equivalent and cylindrical power, as composite features, can largely reflect ocular.

To further explore the specific impact of ocular parameters on myopia prediction, we will first remove the spherical equivalent and cylindrical power as the baseline, and then progressively remove other ocular parameters to analyze their contribution to the prediction performance in more detail. The results are shown in Table 5.

Table 5. MAE Table of SE After Removing Spherical and Cylindrical Lenses with Different Feature Removals

Remove features	MAE of SE (Mean ± SD)
Baseline (Removing OAL and CYL)	0.416±0.354 (D)
k1, k2	0.575±0.539 (D)
AL	0.654±0.593 (D)
LT, ACD, CCT	0.559±0.487 (D)
AOP	0.948±1.145 (D)

Based on the results of the comprehensive ablation experiment, the impact of removing different features on the model's prediction performance varies significantly. Removing the ocular parameters such as K1, K2, or lens thickness, aqueous depth, and central corneal thickness has a relatively small effect on the model's performance. However, when the axial length is removed, the MAE increases significantly, indicating that the axial length is a key feature in the prediction model. When all ocular parameters are removed, the MAE increases dramatically, highlighting the overall importance of ocular parameters for the model's prediction accuracy.

4.4.2. Model component ablation experiment

We conducted an ablation experiment on the model components and recorded the performance metrics after each ablation. The results are shown in Table 6.

Based on this, we found that the xLSTM network had a particularly significant impact on the performance of the IST-xLSTM model. The xLSTM module consists of two parts: sLSTM and mLSTM, each responsible for different tasks. The sLSTM manages long-term dependencies through scalar memory and exponential gating mechanisms, effectively controlling the storage and use of historical information. The mLSTM, on the other hand, enhances storage capacity and information retrieval ability through matrix memory and covariance update rules. It has a higher memory capacity, allowing it to capture complex patterns in the time series more effectively. Additionally, the hierarchical time attention mechanism is also an important factor influencing the model's performance. The two attention mechanisms we designed address both global and local irregularities in the irregular time series data, highlighting the importance of time interval features in processing non-regular time series data. This indicates that the dynamic changes in time intervals are crucial for improving the model's prediction performance.

Table 6. MAE Table of SE Under Different Model Component Removals

Ablation of Model Components	MAE of SE (Mean ± SD)
Baseline	0.216±0.176 (D)
NanDense Module	0.296±0.318 (D)
sLSTM Module	0.527±0.396 (D)
mLSTM Module	0.419±0.327 (D)
xLSTM Network	0.948±0.942 (D)
Hierarchical Time Attention Mechanism	0.472±0.336 (D)

In summary, the xLSTM network and the hierarchical time attention mechanism have the most significant improvement effects on the model, especially when dealing with complex and irregular time series data. These components significantly enhance the model's memory capacity and its ability to model global temporal dependencies, thereby improving overall prediction accuracy.

5. CONCLUSION

This study achieves higher prediction accuracy on a relatively small dataset. Traditional methods struggle to fully utilize time-related information in electronic health records with irregular time intervals and inconsistent record lengths. The proposed IST-xLSTM network model is capable of handling data with uncertain sequence lengths and can effectively capture temporal trends, even when time intervals are irregular, by independently processing time features. Compared to traditional prediction models, the xLSTM we use has stronger memory capabilities, making it suitable for predicting longer sequence data. By incorporating a hierarchical time attention mechanism, the model's ability to handle irregular time intervals is further enhanced, enabling it to fully extract data features and significantly improve prediction accuracy. This research holds significant value for statistical analysis in healthcare institutions and provides strong evidence for parents to better understand their children's vision decline. This can guide caregivers to take timely corrective and early intervention measures, which are more crucial for proactive prevention and control of myopia, compared to the reactive measures taken by healthcare institutions. This helps to effectively prevent and control early myopia in children and adolescents.

There are still some limitations in this study. The current dataset is hard to obtain, and future research will gradually expand the scale of the dataset. Additionally, the dimensions of the longitudinal data still need improvement, as the onset and progression of myopia (future refractive errors) are influenced by various factors, including environmental factors (such as eye usage habits, learning habits, lighting conditions, sunlight exposure, outdoor activities, and parental education level) and genetic factors (such as whether parents are myopic, whether there is high myopia, and genetic testing). Relying solely on ocular biometric parameters is insufficient for accurately predicting future refractive errors. Although retinal images could serve as features, our dataset only includes visual screening records, and retinal images were not available in this study. Combining retinal images and screening records for multimodal learning is expected to further improve prediction accuracy..

ACKNOWLEDGEMENTS

Shanxi Province Natural Science Foundation: Research on Core Scalable and Assurance Intelligent Service Model for Cross-Network Virtual Reality (201801D121147)

REFERENCES

[1] World Health Organization. Vision impairment and blindness. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (2023).

[2] W. H. Organization, World report on vision (2019).

[3] Pärssinen O, Lyyra AL. Myopia and myopic progression among schoolchildren: a three-year follow-up study. Investigative ophthalmology & visual science. 1993 Aug 1; 34(9):2794-802.

[4] Jonas, Jost B., et al. Myopia in China: a population-based cross-sectional, histological, and experimental study. The Lancet 388 (2016): S20.

[5] Sankaridurg, Padmaja, et al. Imi impact of myopia, Investigative Ophthalmology & Visual Science 62 (2021) 2. doi:10.1167/iovs.62.5.2.

[6] Liao, J., Tang, Z., & Liu, X. (2020). Discussion on building hospital information operation and maintenance platform using big data and artificial intelligence technologies. Information and Computer, 32(24), 124-126..

[7] Jong, M., Resnikoff, S., & Tan, K. O. (2022). Asia Myopia Management Consensus. Chinese Journal of Optometry and Visual Science, 24(3), 161-169.

[8] World Health Organization. Child and adolescent health in humanitarian settings: operational guide: a holistic approach for programme managers [J]. 2021.

[9] French AN, Morgan IG, Mitchell P, Rose KA. Risk factors for incident myopia in Australian schoolchildren: the Sydney adolescent vascular and eye study. Ophthalmology. 2013 Oct 1; 120(10):2100-8.

[10] Fan Q, Guo X, Tideman JW, et al. Childhood gene-environment interactions and age-dependent effects of genetic variants associated with refractive error and myopia: The CREAM Consortium. Scientific reports. 2016 May 13;6(1):25853.

[11] Tedja MS, Haarman AE, Meester-Smoor MA, et al. IMI–myopia genetics report. Investigative ophthalmology & visual science. 2019 Feb 28; 60(3):M89-105.

[12] Jones LA, Sinnott LT, Mutti DO, et al. Parental history of myopia, sports and outdoor activities, and future myopia. Investigative ophthalmology & visual science. 2007 Aug 1; 48(8):3524-32.

[13] Chua SY, Sabanayagam C, Cheung YB, et al. Age of onset of myopia predicts risk of high myopia in later childhood in myopic Singapore children. Ophthalmic and Physiological Optics. 2016 Jul; 36(4):388-94.

[14] Jones LA, Sinnott LT, Mutti DO, et al. Parental history of myopia, sports and outdoor activities, and future myopia. Investigative ophthalmology & visual science. 2007 Aug 1; 48(8):3524-32.

[15] 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 Nov 6; 15(11): e1002674.

[16] Ma Y, Zou H, Lin S, et al. Cohort study with 4-year follow-up of myopia and refractive parameters in primary schoolchildren in Baoshan district, Shanghai. Clinical & Experimental Ophthalmology. 2018 Nov; 46(8): 861-72.

[17] Chen Y, Han X, Guo X, et al. Contribution of genome-wide significant single nucleotide polymorphisms in myopia prediction: findings from a 10-year cohort of Chinese twin children. Ophthalmology. 2019 Dec 1; 126(12):1607-14.

[18] 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 Jan; 17(2):463.

[19] Huang J, Zhang Q, Zhao N, et al. Myopia contributing factors and myopia prediction based on vision examination data. J. Univ. Electron. Sci. Technol. China. 2021; 50:256-60.

[20] Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN computer science. 2021 Nov; 2(6):420.

[21] Adam G, Rampášek L, Safikhani Z, et al. Machine learning approaches to drug response prediction: challenges and recent progress. NPJ precision oncology. 2020 Jun 15; 4(1):19.

[22] Jalali A, Lonsdale H, Do N, et al. Deep learning for improved risk prediction in surgical outcomes. Scientific reports. 2020 Jun 9; 10(1):9289.

[23] Sun C, Hong S, Song M, et al. Predicting COVID-19 disease progression and patient outcomes based on temporal deep learning. BMC Medical Informatics and Decision Making. 2021 Dec; 21:1-6.

[24] Spadon G, Hong S, Brandoli B, et al. Pay attention to evolution: Time series forecasting with deep graph-evolution learning. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021 Apr 27; 44(9):5368-84.

[25] Beck M, Pöppel K, Spanring M, et al. xLSTM: Extended Long Short-Term Memory. arxiv preprint arxiv: 2405.04517. 2024 May 7.

[26] Kia SM, Rad NM, van Opstal D, et al. PROMISSING: Pruning Missing Values in Neural Networks. arxiv preprint arxiv: 2206.01640. 2022 Jun 3.