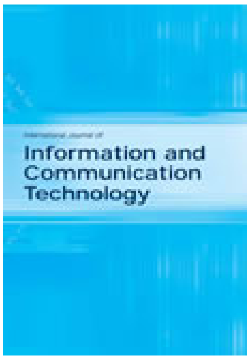


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Mental toughness prediction model of college students based on optimal elastic network regression

Fulian Liu

Article History:

Received:	12 October 2024
Last revised:	28 October 2024
Accepted:	28 October 2024
Published online:	13 December 2024

Mental toughness prediction model of college students based on optimal elastic network regression

Fulian Liu

Mental Health Education Center for College Students,
Wuxi Institute of Technology,
Wuxi, 214121, China
Email: liufl@wxit.edu.cn

Abstract: Predicting the level of mental toughness can help colleges and universities better understand the psychological condition of college students. This paper designs a prediction model of college students' mental toughness based on optimised elastic network regression (ENR) to address the redundant features as well as the overfitting problems of existing studies. Firstly, the ENR is optimised using Bayesian optimisation algorithm (BOENR). Secondly, the important influencing factors are extracted to the maximum extent by using the partial least squares method. Then, linear discriminant analysis (LDA) is used for feature screening of key influencing factors, Pearson's correlation coefficient is used to measure the redundancy relationship among features, and finally, BOENR estimation of regression coefficients is computed based on each feature sample separately. The experimental outcome indicates that the MSE and MAE of the designed model are reduced by 0.0395–0.2264 compared with the other five models.

Keywords: mental toughness prediction; elastic network regression; ENR; Bayesian optimisation; partial least square; linear discriminant analysis; LDA.

Reference to this paper should be made as follows: Liu, F. (2024) 'Mental toughness prediction model of college students based on optimal elastic network regression', *Int. J. Information and Communication Technology*, Vol. 25, No. 10, pp.19–33.

Biographical notes: Fulian Liu graduated from the Henan University. She is currently a Senior Lecturer at the Wuxi Institute of Technology. His research interests are higher vocational education and mental health education.

1 Introduction

The biggest plague in today's society is psychological problems, and college students are at a stage where their values, outlook on life and worldview have not yet been fully formed, making them a high-risk group for the frequent occurrence of mental health problems. Psychological resilience refers to how well an individual adapts in the face of life adversity, trauma, tragedy, threats, or other major life stressors, and it implies the 'ability to bounce back' in the face of life's stresses and setbacks (Gucciardi, 2017). College students' mental toughness has been shown to mediate the relationship between subjective well-being and mental health (Akeman et al., 2020). The prediction of mental toughness can effectively prevent the further occurrence of mental illness and has an important role in promoting the subjective well-being of college students (Li et al., 2020), so how to realise the accurate and effective prediction of college students' mental toughness is of great theoretical significance and applied value.

Luo et al. (2021) manually measured the five core characteristics of mental toughness (meaningful life, perseverance, self-confidence, composure, and sense of presence) through the resilience inventory for mental competence (RISC) and predicted mental toughness through Lasso regression modelling, but the method was vulnerable to outliers. Ye et al. (2023) proposed a psychological resilience prediction method based on grey correlation analysis focusing on the assessment of an individual's adaptability to environmental changes. Babaei (2022) considered personal factors related to stressors, manually selected the main influencing factors, and utilised linear regression methods for mental toughness prediction; however, these factors were highly correlated, resulting in a regression method that was sensitive to small perturbations during matrix inversion. Al Sheeb et al. (2019) proposed the use of ridge regression algorithm for mental toughness prediction, but limited by the characteristics of ridge regression itself, the method could not discard the influence of irrelevant features, resulting in inefficient prediction.

With further research, machine learning (ML) models with black-box models and powerful feature extraction capabilities have attracted the attention of scholars and have been widely used in the field of mental toughness prediction. Hassan (2023) used principal component analysis to select the main influence indicators of mental toughness and used the main indicators as inputs to a decision tree for prediction, but the high computational complexity and complex parameter selection made the prediction unsatisfactory. Sahlan et al. (2021) compared the performance of SVM and decision tree models in predicting mental toughness. The results showed that the SVM performed slightly better than the decision tree in all tests. Yarkoni and Westfall (2017) analysed the main influences affecting mental toughness through collected psychological data, developed a predictive model based on the random forest algorithm, and proposed a method to measure the importance of variables. Wang et al. (2019) used a combination of BPNN and multiple linear regression to predict the psychological still behaviour of college students, which improved the prediction efficiency.

Mental toughness prediction is a high-dimensional, multivariate, and nonlinear problem, and the above ML models are prone to overfitting, while elastic network regression (ENR) can avoid overfitting by adjusting the parameters to balance the penalty terms. Liu et al. (2021) first used high-dimensional variable selection methods such as Lasso to screen redundant variables and construct a class of candidate models, and then applied ENR for mental toughness prediction, but the prediction error was large. Szabo et al. (2022) synthesised personal and contextual factors, decomposed these factors through the VMD method, and utilised resilient network regression for mental toughness prediction with a prediction accuracy of 91.58%.

In summary, existing research on mental toughness prediction models has achieved good results, but feature redundancy and overfitting of the models limit their further development. Aiming at these problems, this paper proposes a college students' mental toughness prediction model based on optimised ENR. The innovative work of the model is mainly reflected in the following four aspects.

1 Optimisation of penalty parameters and regularisation term weights for ENR using Bayesian optimisation algorithm (BOENR), which seeks the parameter combinations that make the objective function optimal through a priori information in order to optimise the prediction performance of ENR.

2 The influencing factors of college students' mental toughness were summarised on the basis of existing studies, standardised, and the important influencing factors were extracted and irrelevant factors were eliminated using partial least squares. Based on this, linear discriminant analysis (LDA) was utilised for feature selection of key influencing factors to remove redundant variables and features.

3 Perform Bayesian bootstrap on the BOENR and calculate the BOENR estimates of the regression coefficients based on each feature sample separately, so that the probability distribution of the BOENR estimates can be obtained as a means of approximating its sampling and asymptotic distributions, and thus making a more accurate inference of the predicted values.

4 The experimental outcome indicates that the MSE and accuracy of the suggested model are 0.0558 and 0.9402, respectively, which are ahead of the comparison model and show the best prediction performance, and can be better applied to the prediction of college students' psychological still behaviour.

2 Relevant theoretical foundations

2.1 Mental toughness framework model

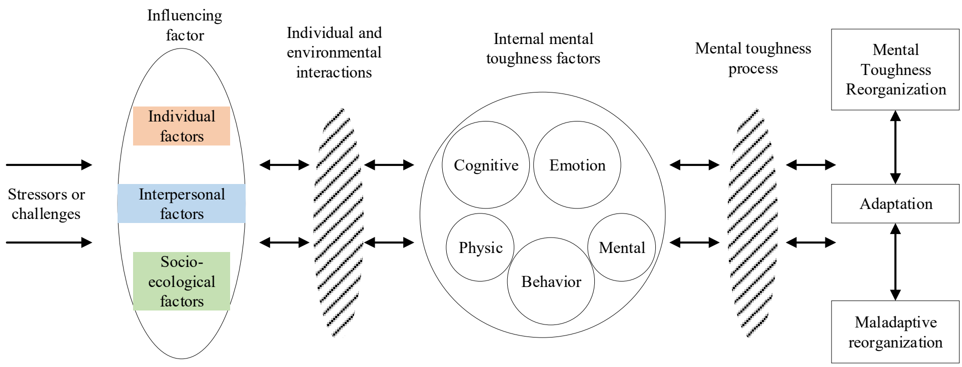
The psychological resilience framework model involves four domains of influence (Farnsworth et al., 2022), namely stressor or challenge, environment/context, individual traits, and adaptive outcomes, as well as two dynamic processes of interactions between the environment and the individual, and between the individual and adaptive outcomes, as shown in Figure 1. The model systematically describes the dynamic process from imbalance to reorganisation of an individual's internal system activated by severe stress or challenge, and provides a more comprehensive and systematic reference of research ideas for conducting research on the prediction of mental toughness (Bédard-Thom and Guay, 2018). The influences on mental toughness within this framework model are categorised into individual, socio-ecological, and interpersonal factors.

1 Individual factors, which consist of traits inherent within the individual that promote resilience, such as the individual physiological indicators, health behaviours (sleep, exercise, etc.), and demographics (gender, age, ethnicity, etc.).

2 Interpersonal factors, highlighting differences between individuals and relationships and personality traits developed or acquired over time, such as family, friends, education, knowledge, skills, and experiences.

3 Socio-ecological factors, which refer to the socio-ecological contexts from which individual situations are linked to environmental contexts that facilitate coping and adaptation, such as informal and formal institutions, geography, and economic income.

Figure 1 Mental toughness framework model (see online version for colours)



2.2 Elastic network regression

The advantages of ENR over other ML models are mainly in its ability to handle complex nonlinear relationships, and its higher tolerance for outliers and missing data (Han and Dawson, 2021). ENR can better address the overfitting issue by adding a regular term to the loss function. For a general linear regression model, assuming that the number of predictor variables is p and the sample size is N, there is equation (1).

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i \\ \varepsilon_i \sim N\left(0, \sigma^2\right), i = 1, 2, \ldots, N \end{cases} \tag{1}$$

where β_0 is the intercept, β_p is the regression coefficient, and ε is the error term, which is usually assumed to be normally distributed with mean 0 and variance σ. The linear regression model is represented by the matrix below. The linear regression model is represented by a matrix as follows.

$$\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon \sim N_N\left(0, \sigma^2 I_N\right) \end{cases} \tag{2}$$

Therefore, the least squares estimate of the regression coefficient is $\hat{\beta}^{LS} = (X^T X)^{(-1)} X^T Y$. For the linear regression model, ENR estimation is defined as follows.

$$\hat{\beta}^{Elastic\text{-}net} = \arg\min \left\{ \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda_1 \sum_{j=1}^{p} \left|\beta_j\right| + \lambda_2 \sum_{j=1}^{p} \beta_j^2 \right\} \tag{3}$$

where λ_1 and λ_2 are penalty parameters. ENR combines the advantages of the linear and ridge regression models described above, and decays between linear and ridge regression through coefficients $0 \leq \lambda \leq 1$. When $\lambda = 1$, the resilient network is equivalent to linear regression; when $\lambda = 0$, the resilient network is equivalent to ridge regression, and when $0 < \lambda < 1$, it is a fold decay between the two methods. β_0 and β_j of ENR are as follows.

$$\min\left(\frac{1}{2T}\sum_{i=1}^{T}\left(y_i-\beta_0-\sum_{j=1}^{r}\beta_j x_{ij}\right)^2+\lambda f_a(\beta)\right) \tag{4}$$

where $f_a(\beta)$ is the penalty term with the following expression.

$$f_a(\beta)=\sum_{i=1}^{r}\left[\frac{(1-\alpha)}{2}\beta_j^2+\alpha\left|\beta_j\right|\right] \tag{5}$$

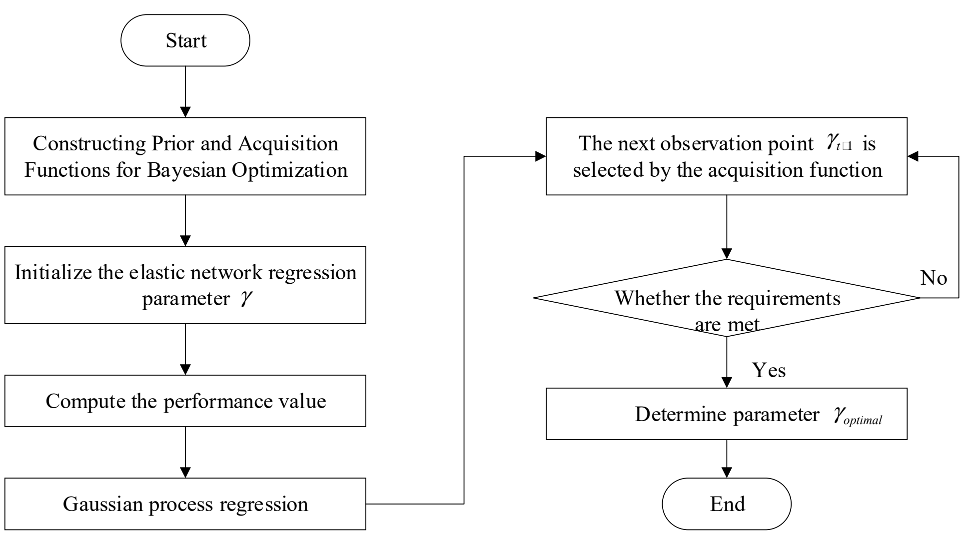
where α is the weight of the regularisation term.

ENR weighs the properties of each regular term by the value of the parameter combination $\gamma = (\alpha, \lambda)$. Therefore, the selection of parameter (α, λ) is particularly critical to the final performance of the model.

3 ENR based on Bayesian optimisation

As can be seen from Subsection 2.2, ENR involves the setting of the penalty parameter λ and the regularisation term weights α. The setting of these parameters has a direct impact on the performance of the model. How to set these parameters reasonably to achieve the best model performance is a problem that needs careful consideration. Therefore, in this paper, Bayesian optimisation (Greenhill et al., 2020) ENR (BOENR) is utilised to seek the parameters that make the objective function optimal through the a priori information, as shown in Figure 2.

Figure 2 Bayesian optimisation ENR



There are theoretically an infinite number of combinations of α and λ. Parameter λ is often selected using a ridge plot, and the parameters are determined by direct observation of the ridge plot. However, ridge maps often contain details that are difficult to interpret by observation alone, and the conclusions drawn are somewhat subjective. To efficiently find the optimal parameter combination γ, this paper adopts the Bayesian optimisation parameterisation, which utilises the existing a priori information to find the parameter γ that makes the performance objective function $\rho(\gamma)$ globally optimal.

$$\rho(\gamma) = \frac{1}{\varepsilon_{RMSE} + \left|1 - R^2\right|} \tag{6}$$

where ε_{RMSE} represents the prediction error and R^2 is the coefficient of determination.

Bayesian optimisation has two main core functions: the prior function (PF) and the acquisition function (ACF). Among them, the PF can be obtained by Gaussian process regression. It is assumed that the initialised t performance values follow a joint Gaussian distribution with expectation 0, as shown below.

$$\rho = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_t \end{bmatrix} \sim N(0, K) = N\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1n} \\ K_{21} & K_{22} & \cdots & K_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ K_{t1} & K_{t2} & \cdots & K_{tn} \end{pmatrix} \right) \tag{7}$$

where K is the covariance matrix.

In this paper, the squared exponential kernel function is used as the K matrix, then we have Kij = $\exp((-(\rho_i - \rho_j)^2/2))$. When the new performance value ρ^* is added, the updated performance distribution still obeys the joint Gaussian distribution as follows.

$$\begin{bmatrix} \rho_1 \\ \vdots \\ \rho_t \\ \rho^* \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K & K_{\rho\rho*}^T \\ K_{\rho\rho*} & K_{\rho*\rho*} \end{bmatrix} \right] \tag{8}$$

where $K_{\rho\rho*} = [k(\rho_1, \rho^*), k(\rho_2, \rho^*), \ldots, k(\rho_1, \rho^*)]$.

Estimating the conditional probability of ρ^*, i.e., finding the updated $N(\mu^*, \Sigma^*)$, given the known data combination (γ_i, ρ_i) and the new parameter γ^*, the new mean μ^* and the new variance Σ^*, computed by Bayes' theorem, are as follows.

$$\begin{cases} \mu^* = K_{\rho\rho*} K_{\rho\rho}^{-1} \rho \\ \Sigma^* = K_{\rho*\rho*} - K_{\rho\rho*} K_{\rho\rho}^{-1} K_{\rho\rho*}^T \end{cases} \tag{9}$$

where μ^* is the expected effect value of the parameter combination, the larger the mean value is, the greater the probability that it will become the optimal solution; Σ^* is the uncertainty of the effect of the parameter combination, the larger Σ^* is, the more possibilities exist at this point. To trade-off exploration and exploitation, the ACF needs to be defined. In this paper, we use the upper confidence bound (UCB) function (Ottens et al., 2017) as the ACF a_{UCB} as follows.

$$a_{UCB}(\gamma) = \mu^*(\gamma) + 1.96\delta^*(\gamma) \tag{10}$$

where $\delta^*(\gamma) = \sqrt{\Sigma^*}$.

According to equation (11), the next γ_{t+1} that may maximise a_{UCB} is as follows.

$$\gamma_{t+1} = \arg\max \alpha_{UCB}(\gamma) = \arg\max \mu_t(\gamma) + 1.96\delta_t(\gamma) \tag{11}$$

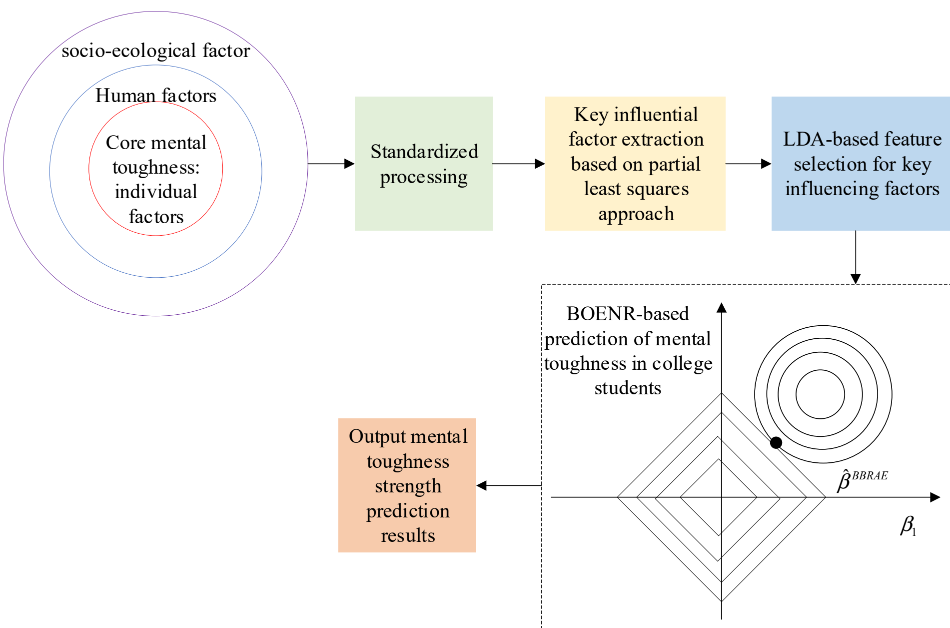
After several rounds of iterations, the optimal combination of parameters that satisfies the performance requirements can be found $\gamma_{optimal}$.

4 Predictive model of college students' mental toughness based on optimised ENR

4.1 Extraction of key influential factors of mental toughness based on partial least squares approach

For the goal of removing redundant variables and features and improve the accuracy of prediction, a prediction model for college students' mental toughness was designed based on Bayesian optimisation ENR, as shown in Figure 3. The influences on mental toughness were first summarised on the basis of existing research, and the key influences were maximally extracted and irrelevant factors were eliminated using partial least squares (Mehmood et al., 2020). The feature selection of the main impact indicators was then carried out using LDA, and the Pearson's correlation coefficient was used to measure the redundancy relationship between the features. Finally the regression coefficient BOENR prediction was calculated based on each feature separately.

Figure 3 ENR-based model for predicting mental toughness in college students (see online version for colours)



Partial least squares (PLS), as an extension of least squares (OLS), is especially suitable for dealing with multivariable forecasting problems with multiple collinearity, and improves the stability and forecasting ability of the model by considering the relationship between variables. In addition, PLS is an optimisation of OLS to maximise the extraction of relevant information, combining the advantages of principal component and linear

Regression, eliminating irrelevant variables, maximising the use of the original data, and quantifying the effect of the independent variable and the dependent variable.

From the mental toughness framework model in the basics section of Subsection 2.1, it can be seen that mental toughness is a process that occurs throughout the life cycle, and its influencing factors include individual, socio-ecological, and interpersonal factors, which are denoted as $\{x_1, x_2, \ldots, x_p\}$, respectively, as the independent variables of the prediction model; and low, medium, and high mental toughness as the dependent variables of the prediction model, which are denoted as $\{y_1, y_2, y_3\}$.

Due to the different scales of the influencing factors, directly using them as sample data will lead to large errors. Standardised processing can transform the influence indicators into values within the range of [0, 1], eliminating the error between the scales of the indicators, thus improving the accuracy of the prediction.

$$X_j = \frac{x_j - \min_{x_j}}{\max_{x_j} - \min_{x_j}} \tag{12}$$

where X_j is the data obtained after standardisation of the original data, x_j is the original data to be standardised; $\min_{xj}$ is the minimum value in the data, $\max_{xj}$ is the maximum value in the data.

The above normalised data matrices are a and b, and the i^{th} pair of components are u_i and v_i. The extracted components are linear combinations of the original variables, so it is possible to set $u_i = a\omega_i$, $v_i = b\tau_i$, then according to the rules for extracting components, the above can be expressed as $\max<a\omega_i, b\tau_i>$ and $\omega_i^T \omega_i = 1$, $\tau_i^T \tau_i = 1$, i.e., take the maximum value of the inner product. Then the regression of a and b on u_1 is implemented as follows.

$$\begin{cases} a = u_1 p_1' + E_1 \\ b = u_1 q_1' + F_1 \end{cases} \tag{13}$$

where $p_1' = (p_{11}, \ldots, p_{1n})$, $q_1' = (q_{11}, \ldots, q_{1n})$ are vectors of regression coefficients, E_1 and F_1 are residual matrices of the regression equation.

The residual matrices E_1 and F_1 are subsequently extracted, and letting the rank of a be $r(r \leq p)$, there exist r components such that the following equation holds.

$$\begin{cases} a = u_1 p_1' + \ldots + u_r p_r' + E_1 \\ b = u_1 q_1' + \ldots + u_r q_r' + F_1 \end{cases} \tag{14}$$

The regression equation of y_k on the independent variable $x_1, x_2, \ldots, x_p$ can be obtained by bringing $u_i = a\omega_i$, $v_i = b\tau_i$ into the above equation: $y_k = \beta_{j1}\mathrm{x}_1 + \ldots + \beta_{jp}x_p$. Finally when the components are extracted and the precision reaches a satisfactory value the algorithm can be stopped, the precision is calculated similarly to least squares so that the sum of squares of errors is minimised, and when the sum of squares of errors almost no longer varies, the number of components at this point is the final principal component variables $x_1, x_2, \ldots, x_q$, $q < p$.

4.2 Feature selection of key influencing factors based on LDA

Due to the strong redundancy relationship between the features of different principal component variables, this paper selects the features of principal component variables based on LDA. The non-diagonal elements of the intra-class scatter matrix of the traditional LDA algorithm are the covariance between the features, which can easily lead to the model being biased towards retaining the noise features with low redundancy (Ji and Ye, 2008), so in this paper, we adopt the squared Pearson correlation coefficient to measure the feature redundancy relationship, and the specific steps are as follows.

Suppose $X = \{\chi_1, \chi_2, \ldots, \chi_N\}$ is a vector set $\chi_i \in R^n$ consisting of principal component variables, characterised by $\{f_1, f_2, \ldots, f_m\}$, $\Phi_1, \Phi_2, \ldots, \Phi_c$ is C different pattern classes, the interclass scattering matrix S_b is $\sum_{i=1}^{C}(m_i - m)(m_i - m)^T$, the intraclass scattering matrix S_w is $\sum_{i=1}^{C}\sum_{x_i \in c}\{(x_i - m_i)(x_i - m_i)^T\}$, and the overall scattering matrix S_t is $S_b + S_w$, where m is the overall mean of the samples, ${}_cm_i$ is the mean of the samples of the i^{th} class, and N_i is the number of samples of the i^{th} class.

The non-diagonal elements of the intraclass scatter matrix are computed by replacing the squared Pearson's correlation coefficient as follows, controlling the weights of the diagonal and non-diagonal elements by the parameter α. The weights of the diagonal and non-diagonal elements are then calculated as implied in equation (16). Since the Pearson's correlation coefficient takes values between (–1, 1), the value obtained according to equation (15) is equivalent to penalising feature pairs with high redundancy relationships

$$P_{ij} = \left(\frac{Cov(f_i, f_j)}{\sigma_i \cdot \sigma_j}\right)^2 \tag{15}$$

$$S_{sp} = \alpha \times P + (1 - \alpha) \times diag(S_w) \tag{16}$$

The objective function of the LDA algorithm based on Pearson's correlation coefficient can then be obtained as shown below.

$$w^* = \arg\min_{w} - \frac{w^T S_b w}{w^T S_{sp} w} \tag{17}$$

where w is an n-dimensional non-zero column vector. Convert the above objective function to a Lagrangian function.

$$L(w, \lambda) = tr\left(\left(w^T S_t w\right) - \lambda\left(w^T w - I\right)\right) \tag{18}$$

Solving equation (18) with respect to the partial derivatives of the column vectors of w and making the partial derivatives 0 yields $S_b w = \lambda S_w w$, which is decomposed to obtain $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_m]$ and is sorted in descending order, where $w = [w_1, w_2, \ldots, w_d]$ consists of the eigenvectors corresponding to the first d largest nonzero eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_d$.

4.3 Prediction of mental toughness in college students based on optimised ENR

After obtaining the main characteristics of the key influences on mental toughness, this paper utilises the BOENR to predict mental toughness in college students. By performing Bayesian bootstrapping on the BOENR, the probability distributions of multiple features are obtained, and the BOENR estimates of the regression coefficients are computed based on each feature sample separately, the probability distributions of the BOENR estimates can be obtained as a way of approximating their sampling distributions and asymptotic distributions, which allows for more robust statistical inference of the predicted values.

First, based on the basics of Subsection 2.2, the BOENR estimates of the regression coefficients $\hat{\beta}^{OLS} = (X^T X)^{-1} X^T y$ are computed to obtain the predicted values of the dependent variable, $\hat{y}^{OLS} = (\hat{y}_1^{OLS}, \ldots, \hat{y}_n^{OLS})^T$, and the sample estimates of the random noise, $\hat{\varepsilon}^{OLS} = (\hat{\varepsilon}_1^{OLS}, \ldots, \hat{\varepsilon}_n^{OLS})^T$, that satisfy $E(\hat{\varepsilon}^{OLS}) = 0_{n\times 1}, D(\hat{\varepsilon}^{OLS}) = \sigma^2 M$, where $M = I - H = I - X(X^T X)^{-1} X^T$, and hence $E(\hat{\varepsilon}_i^{OLS}) = 0, D(\hat{\varepsilon}_i^{OLS}) = \sigma^2 h_{ii}$, where h_{ii} is the i_{th} element on the diagonal of the identity matrix, H, and σ^2 is the variance.

Then the Bayesian bootstrap of $\hat{\varepsilon}^{OLS}$, from the uniform distribution $G(0, 1)$ to generate $(n - 1)$ random samples and in accordance with the order from small to large, to get $g(1), \ldots, g(n - 1)$, and then make $g(0) = 0$, $g(n) = 1$, calculate their difference $\theta_i = g(i) - g(i - 1)$, $i = 1, 2, \ldots, n$, can be obtained from the Bayesian weight $\theta = (\theta_1, \ldots, \theta_n)^T$, at this time, $(\theta_1, \ldots, \theta_n) \sim Dir(1, \ldots, 1)$, $Dir(1, \ldots, 1)$ for the Dirichlet distribution, by the nature of the Dirichlet distribution can be obtained from the equation (19).

$$\begin{cases} E[\theta_i] = \dfrac{1}{n} \\ D[\theta_i] = \dfrac{n-1}{n^2(n+1)} \end{cases} \tag{19}$$

Obviously sampling of θ_i and Bayesian fitting are two separate processes, so θ_i and $\hat{\varepsilon}_i$ are independent of each other, and multiplying the residuals using Bayesian weights yields the set of residuals $\varepsilon^{BBR} = (\varepsilon_1^{BBR}, \ldots, \varepsilon_n^{BBR})^T$, where $\varepsilon_i^{BBR} = n\theta_i \hat{\varepsilon}_i^{OLS}$, and hence $E(\varepsilon_i^{BBR}) = 0, D(\varepsilon_i^{BBR}) = \sigma^2 h_{ii}((n-1)/(n+1))$. Based on this the sample dependent variable y^{BBR} can be obtained as follows.

$$y^{BBR} = X\hat{\beta}^{OLS} + \varepsilon^{BBR} \tag{20}$$

This results in a new sample (y^{BBR}, X). A BOENR fit on (y^{BBR}, X) yields equation (21) as well as the new predicted value of the dependent variable, where again $\hat{\beta}^{BBRAE}$ is noted as the estimated value of the BOENR coefficients on the sample of characteristics obtained using BOENR.

$$\hat{\beta}^{BBRAE} = \left(1 + \frac{\lambda_2}{n}\right)\left\{\arg\min_{\beta} y^{BBR} - X\beta_2^2 + \lambda_2 \|\beta\|_2^2 + \lambda_1^* \sum_{j=1}^{p} \hat{w}_j |\beta_j|\right\} \tag{21}$$

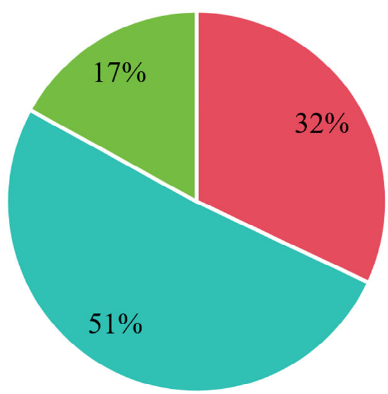
Repeating the process k times, allowing the calculation of the Monte Carlo simulation $(\hat{\beta}^{BBRAE} - \beta_p) \sim G(F_n)$ of the BOENR predictions, an approximation of

$(\hat{\beta}^{AE} - \beta^{*}) \sim G(F)$, which can be used to analyse the properties of β_p based on $G(F_n)$, to calculate confidence intervals for the parameters, or to test the predictions.

5 Experimental results and analyses

In this paper, the basic personal information of 12,076 college students in the class of 2023 in a university in China, and the mental toughness scale data of college students are used as the experimental dataset, and the total score of the scale is the average of the scores of each item, of which the score in the range of 25–49 points indicates low toughness, in the range of 50–74 points is medium toughness, and in the range of 75–125 points is high toughness, and Figure 4 is the number of students with different strengths of mental toughness in the percentage of the situation. Twenty factors were selected as independent variables from the basic personal information. The experiments were conducted using MATLAB program for dataset segmentation, selection of independent variables and modelling, running on MATLAB version 2021. The computing platform was DELL PowerEdge T430 with Intel Xeon E5-2678 v3 @ 2.5 GHz processor and 128 GHz RAM. The ratio of the minimum and maximum values of the regression coefficients β was defaulted to 0.0001, and the model was trained and tested once for each value of β. 100 sets of regression coefficients, mean square error (MSE), and MSE + 1SE were obtained.

Figure 4 Percentage of the number of students with different levels of mental toughness (see online version for colours)



The blue dashed vertical line in Figure 5 indicates that at the minimum MSE + 1SE, there are 12 independent variables with non-zero regression coefficients and MSE = 0.0219; the green dashed vertical line indicates that at the minimum MSE, there are 8 independent variables with non-zero regression coefficients and MSE = 0.0215. The difference between the two MSEs is very small, but the difference in the number of independent variables is large, and the minimum MSE + 1SE is usually taken. This is because the more independent variables there are, the more complicated the model computation is, the more difficult it is to collect the data, and the more it involves the privacy of the individuals, so the fewer the independent variables are, the better, under the premise of guaranteeing the performance of the model.

Figure 5 Curve of MSE with parameter beta (see online version for colours)

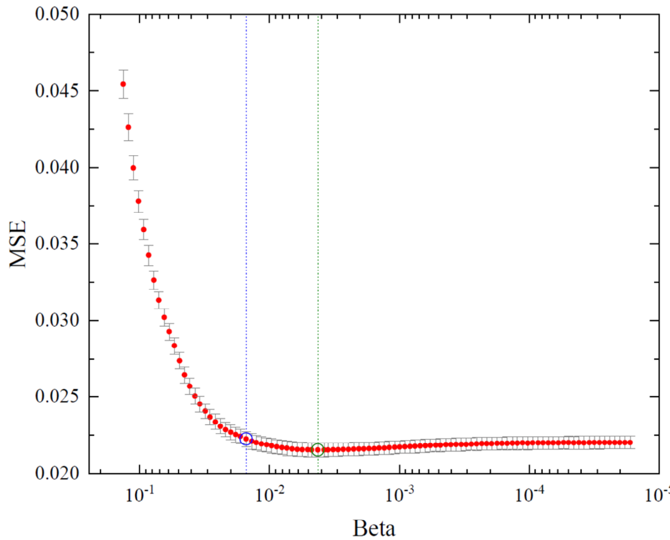
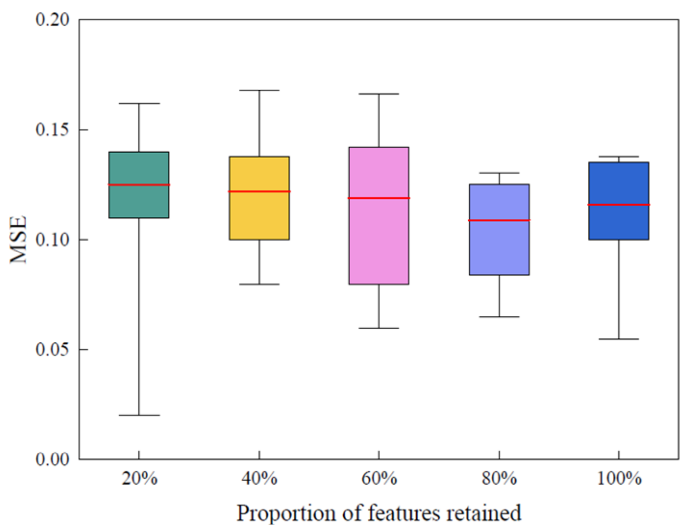


Figure 6 Comparison of MSE with different proportions of selection characteristics (see online version for colours)



When the independent variables contain features of 20%, 40%, 60%, 80% and 100% ratios, their MSE counterparts are shown in Figure 6. The model was able to get to the lowest prediction error when 80% of the features were included in the model, with an MSE of 0.109 and a standard deviation of 1.12, and the next lowest prediction error when 100% of the ratio features were included, with an MSE of 0.116 and a standard deviation of 2.33, followed by 60%, 40% and 20%. The best predictive performance is achieved when 80% of the features are retained using the LDA algorithm, so the feature redundancy relationship is measured by the squared Pearson's correlation coefficient instead of the original covariance, and the removal of low-correlation invalid features reduces the training cost of the model and improves the prediction efficiency.

To further validate the prediction effectiveness of the proposed model, this paper integrates the prediction performance of the proposed model BOENR as well as the comparison models PCADT (Hassan, 2023), CSIRF (Yarkoni and Westfall, 2017), BPLAS (Wang et al., 2019), LAENR (Liu et al., 2021), and VMDNR (Szabo et al., 2022) for comparative tests of prediction performance. MSE, MAE and R^2 are commonly used indicators to measure the prediction error, the smaller the value of MSE and MAE, the higher the prediction accuracy. R^2 is an important indicator to assess the fit between the predicted and actual values in the model, the closer the value is to 1, the better the fit is. Accuracy, F1 are the key indicators for assessing the accuracy of prediction, where F1 is the reconciled average of precision and recall, which can fully reflect the accuracy of prediction. A comparison of the performance metrics of different models is shown in Table 1.

Table 1 The predictive performance of different mental toughness prediction models

Model	*MSE*	*MAE*	R^2	*Accuracy*	*F1*
PCADT	0.2822	0.3027	0.8361	0.7826	0.7791
CSIRF	0.2563	0.2951	0.8537	0.8065	0.8235
BPLAS	0.1335	0.1806	0.9079	0.8541	0.8533
LAENR	0.1852	0.2193	0.9242	0.8811	0.8639
VMDNR	0.1061	0.1464	0.9518	0.9158	0.8982
BOENR	0.0558	0.1099	0.9815	0.9402	0.9517

As can be seen from Table 1, BOENR has the lowest MSE and MAE, which are reduced by 0.0395–0.2264 compared to PCADT, CSIRF, BPLAS, LAENR and VMDNR. The R^2 of BOENR is 0.9815, which is closest to 1. The predicted value is closest to the actual value, which has good fitting effect and high prediction accuracy. In addition, the Accuracy and F1 of BOENR are 0.9402 and 0.9517, respectively, which are higher than the comparison model and have high prediction accuracy.

PCADT and CSIRF have comparable prediction performance. Decision trees and random forests are both ML algorithms based on tree structures with high computational complexity and less than ideal prediction. Both BPLAS and LAENR are based on linear regression algorithms for prediction, with the difference that BPLAS may affect the fit of the model when the independent variable and the dependent variable may be in a nonlinear relationship or when encountering high-dimensional data. While LAENR removes the redundant variables, although it achieves a good prediction performance, it does not screen the redundant features of the variables. VMDNR predicts by dimensionality reduction of the impact indicator variables, and then uses ENR to predict, and the prediction accuracy reaches 0.9158, but it does not optimise the traditional ENR, so the prediction effect is still to be improved. BOENR is able to mine the key important features; through the dimensionality reduction of high-dimensional data, the constructed model is more concise; at the same time, through the Bayesian optimisation algorithm continuously adjusts and optimises the parameters, it can improve the prediction performance of the model. Therefore, BOENR can improve the prediction accuracy and generalisation ability of the model.

6 Conclusions

College students are a high-risk group for frequent mental health problems, and predicting their mental toughness can help colleges and universities provide targeted mental health education and support. Existing mental toughness prediction models have redundant features as well as overfitting, resulting in low prediction performance. Intending to the above issues, this paper designs a college students' mental toughness prediction model based on optimised ENR. Firstly, the Bayesian optimisation algorithm is used to tune the ENR, and the parameter combinations that make the objective function reach the optimal are sought through the a priori information, so as to achieve the purpose of optimising the prediction performance of the ENR. Then the analysis summarised the influencing factors of college students' heart resilience, used LDA for feature selection of important influencing indicators, and used Pearson's correlation coefficient to measure the redundancy relationship between features. Finally, the BOENR estimates of the regression coefficients are computed based on each feature sample separately, and the probability distribution of the BOENR estimates can be obtained, which allows for more accurate statistical inference of the predicted values. The experimental outcome indicates that the designed model has low prediction error and high prediction accuracy, and can efficiently realise the accurate prediction of college students' mental toughness.

Acknowledgements

This work is supported by the 2020 Student Work Research Project of Wuxi Institute of Technology in China (No. XGZ202001).

References

Akeman, E., Kirlic, N., Clausen, A.N. et al. (2020) 'A pragmatic clinical trial examining the impact of a resilience program on college student mental health', *Depression and Anxiety*, Vol. 37, No. 3, pp.202–213.

Al Sheeb, B., Abdella, G.M., Hamouda, A.M. et al. (2019) 'Predictive modeling of first-year student performance in engineering education using sequential penalization-based regression', *Journal of Statistics and Management Systems*, Vol. 22, No. 1, pp.31–50.

Babaei, T. (2022) 'Psychological burnout in elite athletes with disabilities: the predictive role of the mood states and mental toughness', *Journal of Humanistic Approach to Sport and Exercise Studies*, Vol. 2, No. 2, pp.188–200.

Bédard-Thom, C. and Guay, F. (2018) 'Mental toughness among high school students: a test of its multidimensionality and nomological validity with academic achievement and preference for difficult tasks', *Social Psychology of Education*, Vol. 21, No. 4, pp.827–848.

Farnsworth, J.L., Marshal, A. and Myers, N.L. (2022) 'Mental toughness measures: a systematic review of measurement properties for practitioners', *Journal of Applied Sport Psychology*, Vol. 34, No. 3, pp.479–494.

Greenhill, S., Rana, S., Gupta, S. et al. (2020) 'Bayesian optimization for adaptive experimental design: a review', *IEEE Access*, Vol. 8, pp.13937–13948.

Gucciardi, D.F. (2017) 'Mental toughness: progress and prospects', *Current Opinion in Psychology*, Vol. 16, pp.17–23.

Han, H. and Dawson, K.J. (2021) 'Applying elastic-net regression to identify the best models predicting changes in civic purpose during the emerging adulthood', *Journal of Adolescence*, Vol. 93, pp.20–27.

Hassan, A.A.S. (2023) 'The components of the spiritual intelligence predicting the mental toughness and emotional creativity for the university students', *Education Research International*, Vol. 2023, No. 1, p.1631978.

Ji, S. and Ye, J. (2008) 'Generalized linear discriminant analysis: a unified framework and efficient model selection', *IEEE Transactions on Neural Networks*, Vol. 19, No. 10, pp.1768–1782.

Li, C., Zhang, Y., Randhawa, A.K. et al. (2020) 'Emotional exhaustion and sleep problems in university students: does mental toughness matter?', *Personality and Individual Differences*, Vol. 163, p.110046.

Liu, Y., Pan, H., Yang, R. et al. (2021) 'The relationship between test anxiety and emotion regulation: the mediating effect of psychological resilience', *Annals of General Psychiatry*, Vol. 20, pp.1–9.

Luo, J., Ren, S., Li, Y. et al. (2021) 'The effect of college students' adaptability on nomophobia: based on lasso regression', *Frontiers in Psychiatry*, Vol. 12, p.641417.

Mehmood, T., Sæbø, S. and Liland, K.H. (2020) 'Comparison of variable selection methods in partial least squares regression', *Journal of Chemometrics*, Vol. 34, No. 6, p.e3226.

Ottens, B., Dimitrakakis, C. and Faltings, B. (2017) 'DUCT: an upper confidence bound approach to distributed constraint optimization problems', *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol. 8, No. 5, pp.1–27.

Sahlan, F., Hamidi, F., Misrat, M.Z. et al. (2021) 'Prediction of mental health among university students', *International Journal on Perceptive and Cognitive Computing*, Vol. 7, No. 1, pp.85–91.

Szabo, Z.P., Kun, A., Balogh, B.E. et al. (2022) 'Dark and strong?! The associations between dark personality traits, mental toughness and resilience in Hungarian student, employee, leader, and military samples', *Personality and Individual Differences*, Vol. 186, p.111339.

Wang, J., Zhang, Z., Luo, H. et al. (2019) 'Research on early warning model of college students' psychological crisis based on genetic BP neural network', *American Journal of Applied Psychology*, Vol. 8, No. 6, pp.112–120.

Yarkoni, T. and Westfall, J. (2017) 'Choosing prediction over explanation in psychology: lessons from machine learning', *Perspectives on Psychological Science*, Vol. 12, No. 6, pp.1100–1122.

Ye, S., Wang, R., Pan, H. et al. (2023) 'Development and validation of a prediction nomogram for academic burnout among Chinese adolescents: a cross-sectional study', *BMJ Open*, Vol. 13, No. 5, p.e068370.