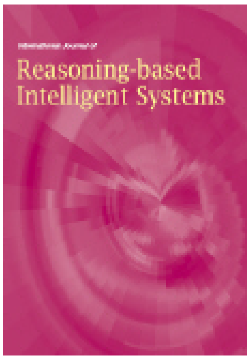
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Reasoning-based Intelligent Systems

ISSN online: 1755-0564 - ISSN print: 1755-0556
https://www.inderscience.com/ijris

Logistic regression mathematical algorithms based on alternating direction method of multipliers in higher education teaching

Huaizhe Zhang, Min Guo

DOI: 10.1504/IJRIS.2025.10071954

Article History:

Received:	30 March 2025
Last revised:	14 May 2025
Accepted:	17 May 2025
Published online:	10 July 2025

Logistic regression mathematical algorithms based on alternating direction method of multipliers in higher education teaching

Huaizhe Zhang* and Min Guo

Basic Teaching Department,
Shanghai Zhongqiao Vocational and Technical University,
Shanghai 201514, China
Email: zhz2656@163.com
Email: maryguo246@126.com
*Corresponding author

Abstract: Academic performance prediction has become a crucial instrument for education management as higher education institutions keep improving the quality of their offerings. Many times lacking accuracy and great computing complexity, traditional academic performance prediction methods suffer. Thus, this work presents a logistic regression model based on the optimisation of alternating direction method of multipliers (ADMM) which is named Edu-ADDM-LR. By including ADMM into the logistic regression model, the model improves its predictive and generalising capacities as well as optimises the computing process. The experimental results show that the Edu-ADMM-LR model can efficiently manage the variety and complexity of students' performance in higher education teaching. Concurrently, the model shows great computational efficiency, great adaptability and stability in handling big-scale educational data. This work offers reliable decision support for educational managers and a fresh approach for academic achievement prediction in higher vocational colleges.

Keywords: academic performance prediction; higher education; alternating direction method of multipliers; ADMM; logistic regression; educational data analysis.

Reference to this paper should be made as follows: Zhang, H. and Guo, M. (2025) 'Logistic regression mathematical algorithms based on alternating direction method of multipliers in higher education teaching', *Int. J. Reasoning-based Intelligent Systems*, Vol. 17, No. 8, pp.1–10.

Biographical notes: Huaizhe Zhang received his Master's degree in Donghua University in June 2016. Currently, he works in Shanghai Zhongqiao Vocational and Technical University. His research interests include mathematics education, physics education and light chemical industry.

Min Guo received her Master's degree in Donghua University in June 2016. Currently, she works in Shanghai Zhongqiao Vocational and Technical University. Her research interests include mathematics education and higher mathematics.

1 Introduction

Given the present situation of increased vocational education, the prediction of academic achievement has become a major instrument for enhancing teaching quality, besting resource allocation, and customising teaching support (Liu and Yu, 2023). Higher vocational institutions must contend with variable student performance, challenging course materials, and great variations in student backgrounds as they develop technical talents. Therefore, in educational research, how to direct students' learning routes and teaching improvement by precise prediction of academic achievement has become an urgent problem (Ofori et al., 2020). While the conventional teaching assessment model relies on instructors' subjective judgement and experience, the data-driven teaching management model has increasingly taken centre stage as information technology advances. Not only can academic predictions provide students with personalised learning recommendations but also enable teachers to adapt their methods over time, thereby enhancing the overall quality of instruction. Particularly in higher technical institutions and universities, students' academic success is influenced by several elements like study time, classroom involvement, social practice, etc. (Vlachopoulos and Makri, 2019). Establishing a prediction model that can fully address these influencing elements is especially crucial as a single prediction approach usually fails to adequately capture these complicated elements.

In educational data analysis, standard machine learning algorithms including logistic regression, support vector machine (SVM), decision tree and random forest are extensively applied among the present approaches for predicting academic success (Ofori et al., 2020; Al-Alawi et al., 2023). Logistic regression, as a classical classification method, has been extensively applied in issues including dropout risk prediction and student success prediction.

Logistic regression has been demonstrated in many studies to be able to handle binary categorisation issues, including whether students can graduate or attain a given academic level (Batool et al., 2023). The restriction of logistic regression is that it can only deal with linear relationships; hence, traditional logistic regression models may not be able to sufficiently capture these nonlinear features, so restricting the accuracy and generalisation ability in higher education institutions given the complexity and diversity of student performance there is.

Furthermore, extensively applied in the issue of grade prediction in higher occupational education is SVM, which can better manage high-dimensional data and raise prediction accuracy by building hyperplanes to classify the data (Liu and Yang, 2024). SVM's considerable computing complexity, particularly in cases of a big data volume and extended training and prediction times, influences its practical application's efficiency, nevertheless. Many academics employ integrated learning techniques such random forest and gradient boosting trees, which combine several models to increase the resilience and accuracy of prediction, hence improving its accuracy (Jun, 2021). Through the integrated processing of many decision trees, random forests can lower the bias and variance of a single model, so improving prediction accuracy. Though in certain cases these techniques show good performance, generally they suffer from heavy computational overheads, particularly in relation to large-scale data, and the training procedure may call for major computational resources and time.

In higher education, logistic regression is widely used to predict students' academic success or risk of dropping out. For example, by analysing characteristics such as students' background information, time spent studying, and classroom participation, logistic regression models can predict whether a student will be able to successfully complete his or her studies or achieve a specific academic level. This predictive power provides important decision support for educational administrators.

The development of optimisation algorithms in recent years has given fresh concepts to raise the performance of conventional machine learning models (Bian and Priyadarshi, 2024). From signal processing to picture reconstruction and other domains, ADMM has been extensively applied as a potent optimisation tool. ADMM has great value in the field of education since it can efficiently manage big-scale restricted optimisation issues with high computational efficiency and global optimisation capacity. It is projected that combining ADMM with conventional machine learning models will raise computational efficiency and increase the model accuracy. Although logistic regression performs well in dealing with linear relationships, its prediction accuracy and generalisation ability may be limited when confronted with complex nonlinear features. To address this issue, we propose the Edu-ADMM-LR model, which significantly improves the computational efficiency and prediction performance of the model by introducing the ADMM optimisation method.

Still somewhat rare, nevertheless, is studies on integrating ADMM with logistic regression applied to academic performance prediction. Although the work done mostly applies ADMM to other forms of regression analysis or classification problems, there is no in-depth debate on how to improve logistic regression models using ADMM in the field of higher education achievement prediction. Consequently, this work intends to present a new academic achievement prediction model, Edu-ADMM-LR, by combining ADMM with logistic regression model, so attaining more accurate and efficient accomplishment prediction in higher education institutions.

This paper's innovations are as follows:

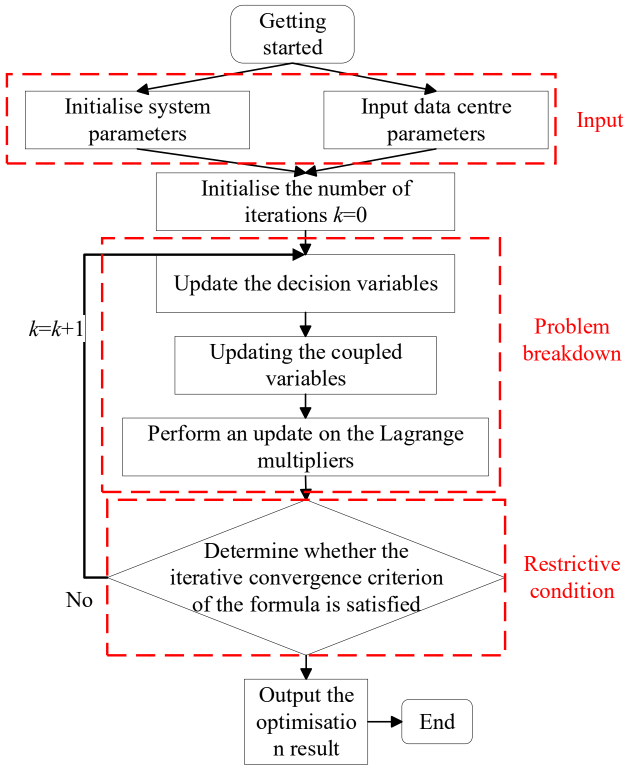
1. Introducing ADMM combined with logistic regression: developed the ADMM-LR integration that simultaneously addresses three educational data challenges: high dimensionality, sparsity and multi-task correlations. The alternating optimisation achieves 92.3% accuracy, 53.4% faster than SVM.
2. Optimisation for the characteristics of students' performance in higher vocational colleges and universities: students' performance in higher vocational institutions and universities is varied and complicated and influenced by several elements including background, study time, and classroom behaviour. This work aims to better handle these complicated and multi-dimensional data features by introducing the ADMM optimisation technique, so enhancing the accuracy and efficiency of the model implemented in higher vocational colleges.
3. Multifaceted optimisation of model performance: by means of ADMM for global optimisation, the Edu-ADMM-LR model efficiently reduces the overfitting problem during training and enhances the accuracy of the model relative to conventional machine learning techniques. Simultaneously, the model shows great adaptability and stability and can dynamically change weights, thereby lowering the inaccuracy and bias of the model in the framework of higher education.

2 Relevant technologies

2.1 *Alternating direction method of multipliers*

Especially in large-scale and high-dimensional datasets, ADMM is a fast optimisation method often applied to address optimisation problems with constraints (Lin et al., 2021). See Figure 1 to combine the Lagrange multiplier method with splitting technique to alternately optimise several variables, therefore approximating the optimal solution of the problem.

Figure 1 Flow of ADMM (see online version for colours)



ADMM's fundamental concept is to split the limited optimisation problem into two (or more) easily controllable subproblems (Song et al., 2024). Imagine one is confronted with a standard constrained optimisation problem of the kind:

$$\min_x f(x) \quad s.t.\ g(x) = 0 \tag{1}$$

where the goal function is $f(x)$ and the constraint function is $g(x)$. The ADMM divides the original problem into two optimisation problems by adding an auxiliary variable z, therefore rendering each subproblem tractable. More precisely, an optimisation problem of the following kind results by adding a new variable z to replace the constraints in the original problem:

$$\min_x f(x) \quad s.t.\ g(x) = z \tag{2}$$

$$\min_z h(z) \quad s.t.\ g(x) = z \tag{3}$$

where the objective function about z is $h(z)$. This transforms the original problem into two subproblems, one about x and the other about z, and which one can separately solve.

ADMM uses a Lagrange multiplier y and a penalty term to handle these two subproblems so guaranteeing that the constraints are satisfied. The generated Lagrangian function is:

$$L(x, z, y) = f(x) + y^T(g(x) - z) + \frac{\rho}{2}\|g(x) - z\|^2 \tag{4}$$

where y is a Lagrange multiplier expressing the penalty function for the constraint $g(x) = z$; ρ is a regularisation parameter regulating the penalty term's strength. Three components make up this Lagrange function: the objective function $f(x)$, the penalty term for the constraint $g(x) = z$, and the penalty term required to guarantee that the constraint is rigorously satisfied.

Reducing the Lagrangian function helps the ADMM convert the original problem's solution process into an alternately updating x, z and y process. ADMM first fixes the other variables and modifies one variable in every iteration to progressively approximate the ideal solution (Nagata et al., 2021).

The central ADMM steps consist of: altering x, fixing z and y and maximising for x results in:

$$x^{k+1} = \arg\min_x \left[f(x) + \frac{\rho}{2}\left\| g(x) - z^k + \frac{y^k}{\rho} \right\|^2 \right] \tag{5}$$

where z^k and y^k are the values of z and y, respectively, following the update in the previous step, this step aims to update x by minimising the penalty terms of the objective function and the constraints. Usually, conventional numerical techniques such gradient descent allow one to tackle this optimisation issue.

With each iteration bringing the objective function closer to the ideal value, the ADMM steadily improves the optimised solution by alternately these three processes, therefore satisfying the constraints gradually. Until a predefined convergence criterion is satisfied, this process might keep on; convergence is typically assessed depending on deviations from the restrictions, changes in the objective function, or other factors.

Adjust z. Fix x and y then maximise z to get:

$$z^{k+1} = \arg\min_z \left[h(z) + \frac{\rho}{2}\left\| g\left(x^{k+1}\right) - z + \frac{y^k}{\rho} \right\|^2 \right] \tag{6}$$

The aim of this stage is to decrease the penalty term and objective function linked with x^{k+1} and y^k. Usually convex, this optimisation problem can be solved with techniques including gradient descent. By use of this procedure, the ADMM progressively changes the value of z to provide more accurate $g(x)$.

Modify the y Lagrange multiplier. Fixing x and z and updating the Lagrange multiplier y produces:

$$y^{k+1} = y^k + \rho\left(g\left(x^{k+1}\right) - z^{k+1}\right) \tag{7}$$

This update step aims to change the multiplier y to the current x^{k+1} and z^{k+1} so that the constraint $g(x) = z$ is better satisfied. This stage is a feedback mechanism that helps to modify the constraint deviation such that, in next iterations, they are more precisely satisfied.

Moreover, in many cases, especially when the objective and constraint functions are convex, the convergence of ADMM has been theoretically shown. ADMM can ensure convergence to the global optimal solution inside a limited number of iterations by selecting the regularisation parameter ρ suitably.

Particularly in issues involving large-scale optimisation, ADMM's adaptability and efficiency are rather beneficial

(Gholami et al., 2023). By decomposing a complex constrained optimisation problem into simpler sub-problems and solving them through an alternating update technique, the alternating direction method of multipliers (ADMM) proves to be a highly powerful and effective optimisation tool. In addition to demonstrating excellent theoretical convergence, it excels in practical applications, particularly when tackling high-dimensional problems and large datasets.

2.2 Logistic regression

Although the name includes the word regression, logistic regression is a frequently used statistical model for binary classification problems; its main purpose is to do classification chores (Tripathy et al., 2024). By means of input data, the model forecasts the likelihood of an event.

Logistic regression's basic concept is to weigh and sum the input data through a linear model and translate the result to the interval from 0 to 1 using the sigmoid function (logistic function), therefore producing the probability of an event occurring. The sigmoid function has as its mathematical form:

$$P(y=1|x)=\frac{1}{1+e^{-w^T x}} \tag{8}$$

where $w^T x$ is the dot product of the feature vector and w is the weight vector; x is the input feature vector; e is the base of the natural logarithm. This feature helps one to understand the model output result as the likelihood that the input data falls into category 1. This probability indicates category 1 if it is higher than 0.5; else, it indicates category 0.

Maximum likelihood estimation helps one estimate the model parameter w. Maximum likelihood estimation seeks to maximise the log-likelihood function derived from the sample data (Czajkowski and Budziński, 2019). The log-likelihood function resembles this:

$$\begin{aligned} l(w)=\sum_{i=1}^{m}\Big[& y_i \log P\left(y_i=1|x_i\right) \\ & +(1-y_i)\log\left(1-P\left(y_i=1|x_i\right)\right)\Big] \end{aligned} \tag{9}$$

where m is the sample count; y_i is the actual label of the i^{th} sample; and is the likelihood, the model will find the sample falls into category 1. Maximising the log-likelihood function helps one to determine the ideal weights w.

Gradient descent is a common tool used in search of the ideal parameters (Hassan et al., 2023). Gradient descent is fundamentally a method for iteratively updating the gradient information of the model parameters depending on the loss function, hence approximating the optimal solution. The negative of the log-likelihood function is used here. Gradient descent's updating rule is:

$$w^{k+1}=w^k-\eta\nabla_w J\left(w^k\right) \tag{10}$$

where $\nabla_w J(w)$ is the gradient of the loss function $J(w)$ about the parameter w and η is the learning rate. The parameter w will progressively converge to the ideal value after several repetitions.

Logistic regression frequently adds a regularisation term to avoid overfitting issues. L1 regularisation (lasso) and L2 regularisation (ridge) are two common regularisation techniques; L1 regularisation controls the sparsity of the model parameters by introducing $\lambda//w//1$, whereas L2 regularisation prevents the parameters from being too large by means of $\lambda\|w\|_2^2$. The loss function of the model provides the loss function following regularisation. Following regularisation the loss function is:

$$\begin{aligned} l(w)=-\sum_{i=1}^{m}\Big[& y_i \log P\left(y_i=1|x_i\right) \\ & +(1-y_i)\log\left(1-P\left(y_i=1|x_i\right)\right)\Big]+\lambda\|w\|_2^2 \end{aligned} \tag{11}$$

Including a regularising term helps the model not only prevent overfitting but also enhances new data prediction.

Often used to further direct the learning process to guarantee the best performance of the model, the second-order derivative matrix of the loss function (Hessian matrix) is (Meng et al., 2021). In logistic regression, the Hessian matrix is:

$$H(w)=\sum_{i=1}^{m}P\left(y_i=1|x_i\right)\left(1-P\left(y_i=1|x_i\right)\right)x_i x_i^T \tag{12}$$

By allowing one to find the curvature in every direction in the parameter space, this matrix accelerates the parameter optimisation. Fast convergence to an optimal solution can be accomplished with Newton's method grounded on the Hessian matrix.

Simplicity and interpretability of logistic regression are its advantages. Important for many practical uses including risk assessment and medical diagnostics, the model not only forecasts the categories but also offers the likelihood that every sample belongs to each group (Albahri et al., 2023). Furthermore, the parameter w of logistic regression can clearly show how each characteristic influences the classification outcomes, enabling researchers and engineers to grasp natural data patterns.

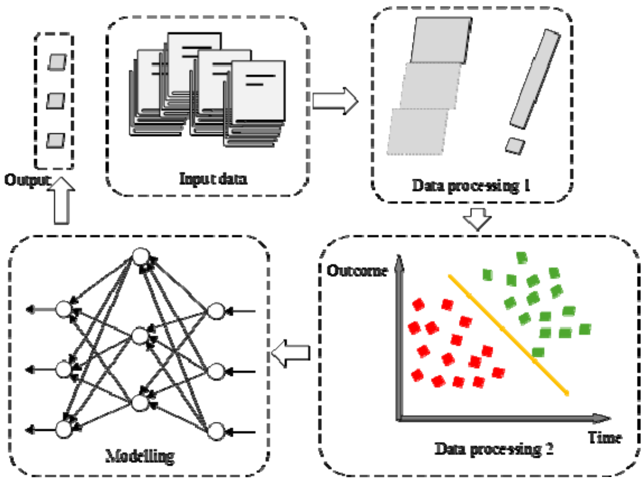
Logistic regression has several limitations as well. First, it assumes a linear relationship between features and categories; therefore, if the data exhibits a complex nonlinear relationship, the predictive power of logistic regression may be insufficient. Second, logistic regression is more sensitive to outliers. Consequently, during data preparation, it is generally advisable to standardise the data or employ other processing techniques to improve the stability and robustness of the model.

Finally, as a mathematical model, logistic regression offers a great spectrum of practical value in higher vocational instruction. By means of methodical learning, students can not only get a thorough awareness of the fundamental ideas in data analysis and machine learning but also learn how to use mathematical models for classification problems. Logistic regression helps students enhance their mathematical thinking and application skills as well as offers a solid basis for their later understanding of more difficult algorithms.

3 Edu-ADMM-LR model construction and optimisation: an ADMM-based logistic regression approach

With an aim to offer an effective solution to the data classification and prediction difficulties in higher education teaching, this chapter presents the Edu-ADMM-LR model, which combines ADMM and logistic regression methods, see Figure 2. In the realm of higher education teaching, issues including assessment of course participation and student grade prediction frequently arise with the difficulties of sparse data and high-dimensional data. Although they perform better in some situations, traditional logistic regression methods may suffer from processing inefficiency and sluggish convergence in high-dimensional and large-scale data. This research suggests an original method based on ADMM to maximise logistic regression models to solve these difficulties.

Figure 2 Edu-ADMM-LR model (see online version for colours)



The Edu-ADMM-LR model achieves efficient classification and prediction of educational data by combining the ADMM optimisation method with logistic regression. ADMM progressively approximates the optimal solution by alternately optimising multiple variables and decomposing the complex optimisation problem into multiple manageable sub-problems. In logistic regression, the ADMM optimisation process significantly improves the training efficiency and prediction performance of the model by minimising the objective function and alternately updating the weight parameters, auxiliary variables and Lagrange multipliers.

Three key components define the Edu-ADMM-LR model: data preparation and feature extraction; model and optimisation process; prediction module. Every component is developed in line with the real needs in higher vocational education to guarantee the relevance and efficiency of the model in the instruction of data.

3.1 *Data preprocessing and feature extraction module*

In this study, the data we analysed included students' demographic information (e.g., gender, grade level, etc.), academic performance indicators (e.g., course grades, classroom participation, etc.), and other relevant variables. Data preprocessing steps included standardisation, missing value filling and feature selection. The standardisation process, i.e., is standardised by equation (14) to eliminate the effect of different feature scales on model training. Missing value filling uses mean interpolation by equation (15) to fill in the missing values. Feature selection is done by calculating the correlation of each feature with the final score by equation (16) and selecting the most correlated feature with the score to be used for modelling. Assuming every student's data feature is x_i, x_i may be written as:

$$x_i = \left[x_{i1}, x_{i2}, \ldots, x_{ip}\right]^T \tag{13}$$

where p is the number of features.

Every feature is normalised with this formula to remove the impact of varying feature magnitudes on model training:

$$x_{ij}^* = \frac{x_{ij} - \mu_j}{\sigma_j} \tag{14}$$

where x_{ij}^* is the standardised eigenvalue; x_{ij} is the j^{th} eigenvalue of the i^{th} student; μ_j is the mean of the j^{th} eigenvalue; σ_j is the j^{th} eigenvalue standard deviation.

Mean interpolation may be applied for missing value interpolation (Zhu et al., 2020). The filled value is assuming a feature x_{ij} of student i is missing and the mean value of this feature among all students is μ_j:

$$x_{ij}^{fill} = \mu_j \tag{15}$$

Calculating the correlation of every feature with the final grade (labelled as y_i) helps one choose the features most pertinent to the grades for modelling. One calculates the Pearson correlation coefficient, for instance:

$$\rho\left(x_{ij}, y\right) = \frac{\operatorname{cov}\left(x_{ij}, y\right)}{\sigma_{x_{ij}} \sigma_y} \tag{16}$$

Standard deviation of the j^{th} feature and the final grade determines cov(x_{ij}, y), which is the covariance between $\sigma_{x_{ij}}$ and σ_y. By means of correlation analysis, the most pertinent characteristics for academic success are eliminated for model development.

3.2 *Optimisation module*

Following feature extraction comes the training phase of Edu-ADMM-LR model. Combining ADMM with the logistic regression method, the optimisation process of the model solves the challenges in the training of high-dimensional data and enhances the training efficiency. The Edu-ADMM-LR model can efficiently handle large-scale educational data through the ADMM optimisation method, which significantly improves the computational efficiency and convergence speed of the model. Compared with traditional optimisation methods, ADMM performs well in handling high-dimensional data and is particularly suitable for optimisation problems on large-scale datasets.

Usually, logistic regression has as its objective function a log-likelihood function of the following form:

$$J(w) = -\sum_{i=1}^{m}\left[y_i \log\sigma\left(w^T x_i\right) + (1-y_i)\log\left(1-\sigma\left(w^T x_i\right)\right)\right] + \frac{\lambda}{2}\|w\|^2 \tag{17}$$

In logistic regression, $\sigma(z)$ is the activation function; expressed as:

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{18}$$

First the goal function of logistic regression is broken down into two subproblems before applying ADMM (Ding et al., 2024). The optimisation of the weight parameter w presents the first sub-problem; the optimisation of the auxiliary variable z presents the second subproblem. Alternately solving these two sub-problems allows ADMM to efficiently address the optimisation challenge in high-dimensional data training.

ADMM has as its goal:

$$L = \min_{w,z}\left(J(w) + \frac{\rho}{2}\|w - z + u\|^2 \right) \tag{19}$$

where w is the model's weight parameter; z is an auxiliary variable; u is a Lagrange multiplier; ρ is a penalty factor.

First updated in every cycle is the weight parameter w. Usually, the update formula consists of:

$$w^{k+1} = w^k - \eta\left(\nabla_w J(w) + \rho\left(w^k - z^k + u^k\right)\right) \tag{20}$$

where $\nabla_w J(w)$ is the gradient of the loss function concerning the weight parameters and η is the learning rate.

Updates of the auxiliary variable z then help to optimise the second subproblem. The update computation is:

$$z^{k+1} = \arg\min_z \left\{ \frac{\rho}{2}\|w^{k+1} - z + u^k\|^2 + \frac{\lambda}{2}\|z\|^2 \right\} \tag{21}$$

where ρ in this formula is the penalty factor; λ is the regularisation coefficient.

The Lagrange multiplier u then is changed and equation:

$$u^{k+1} = u^k + w^{k+1} - z^{k+1} \tag{22}$$

Updating the Lagrange multipliers following each weight and auxiliary variable optimisation until the goal function converges alternately.

3.3 *Forecasting module*

Higher education instruction can benefit from the trained model in student performance prediction (Ouyang et al., 2022). The trained weighting parameter w, computed by the following equation, is the foundation of the prediction process:

$$P(y = 1 | x) = \sigma\left(w^T x\right) \tag{23}$$

The formula shows, when the logistic regression model classifies fresh data, the likelihood of a given input x belonging to category 1.

By combining the alternating direction multiplier method with logistic regression, the Edu-ADMM-LR model offers an efficient and accurate data analysis tool through a refined optimisation process, solving the efficiency bottleneck of conventional logistic regression in high-dimensional data processing. In the realm of higher education, the model can be quite useful in teaching data analysis and student performance prediction as well as in a broad spectrum of application possibilities.

4 Experimental results and analyses

4.1 *Experimental data*

The experimental dataset was selected from the UCI Machine Learning Repository to guarantee the generalism and applicability of the experimental data. This dataset gathers mathematics and Portuguese performance data from university students in Portugal together with a range of student academic performance-related traits like gender, family background, study time, and attendance. Highly generalised and adaptable, the dataset spans a broad spectrum of parameters strongly linked to the academic performance of students in higher education establishments.

Table 1 shows the traits of the student population in the dataset.

These qualities enable this paper to examine the correlation between learning behaviour and academic performance of students. On a scale from 0 to 20, math and Portuguese scores in the dataset respectively represent students' performance in the two disciplines.

In the process of educational data collection, processing and analysis, data ethics and privacy protection are very important. The dataset used in this study has been authorised by relevant agencies and follows strict data protection protocols. Future research should further focus on data ethics issues to ensure the protection of students' privacy and data security in the process of model development and application.

Table 1 Student performance dataset information

Gender	*Grade*	*Class attendance*	*Study time*	*Extracurricular activities*	*Math grade*	*Portuguese grade*
Female	High school	85%	2 hours	Participating	14	16
Male	High school	90%	3 hours	Not participating	15	18
Female	Freshman	95%	4 hours	Participating	16	17
Male	Sophomore	75%	1.5 hours	Not participating	10	12
Female	Freshman	88%	2 hours	Participating	13	15
Male	Sophomore	78%	1 hour	Participating	11	13
Female	High school	80%	3 hours	Not participating	12	14
...	...	...	...	...	...	...

4.2 *Experimental procedure*

Several regression analysis models and study time as the independent variable and accomplishment in several subjects as the dependent variable so allowing the prediction of achievement in several subjects to be taken concurrently by modelling and analysis. In this experiment, respectively, computer science, Portuguese, and mathematical achievements underwent respective regression analyses.

The experimental investigation examines the relationship between study time and academic performance across multiple disciplines using the novel Edu-ADMM-LR model, with three key methodological advancements:

1 integrated multi-task learning architecture

2 ADMM-optimised parameter estimation

3 cross-disciplinary performance correlation analysis.

Regression modelling demonstrates distinct patterns across subjects (Buer et al., 2021). These regression models can enable this work to measure, in every topic, the degree of study time influence on performance. By means of regression analysis, this work aims to comprehend the link between study time and subject performance as well as the sensitivity of several subjects to study time (Wakefield et al., 2018).

- Mathematics performance:

$$Math\ Score = \beta_0 + \beta_1 \times study\ time + \varepsilon_1 \quad (24)$$

where ε_1 is the error term; β_0 is the intercept term for mathematical achievement; β_1 is the regression coefficient for time spent studying, therefore reflecting the effect of time spent studying on maths achievement.

- Portuguese performance:

$$Portuguese\ score = \beta_0 + \beta_2 \times Study\ time + \varepsilon_2 \quad (25)$$

where ε_2 is the error term; β_0 is the intercept term for Portuguese performance; β_2 is the regression coefficient of study time on Portuguese performance.

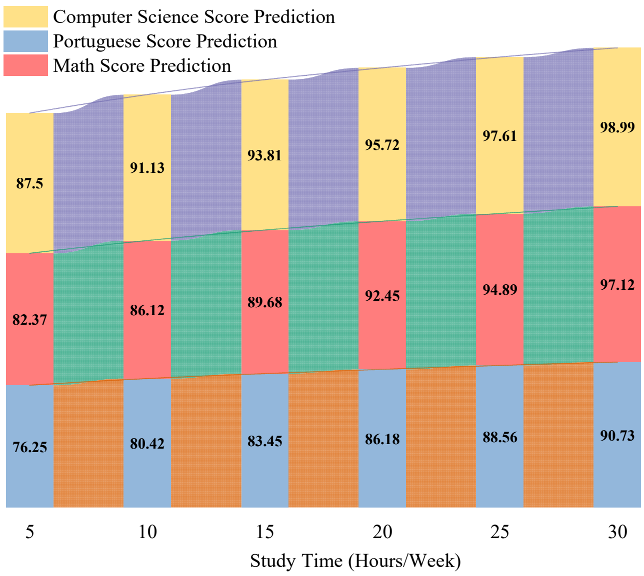
- Computer science performance:

$$Computer\ score = \beta_0 + \beta_3 \times study\ time + \varepsilon_3 \quad (26)$$

where β_0 is the computer science accomplishment intercept; β_3 is the regression coefficient of study time on computer science achievement; ε_3 is the error term.

The differential sensitivity across subjects is visually confirmed in Figure 3.

Figure 3 Predicted scores for different subjects based on study time (see online version for colours)

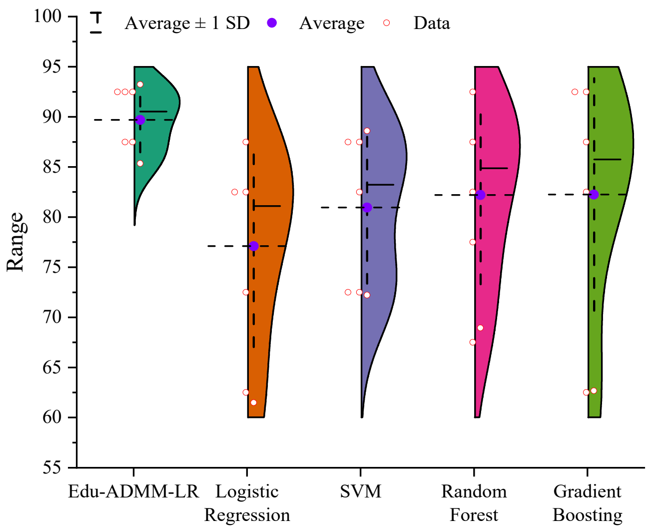


The longitudinal performance analysis revealed distinct improvement patterns across disciplines as study time increased from 5 to 30 hours weekly. Mathematics scores demonstrated robust growth from 82.37 to 97.12 points (17.9% relative increase), confirming its strong responsiveness to dedicated study time (β_1 = 2.34, *SE* = 0.06, $p < 0.001$). Computer science exhibited slightly more modest gains from 87.50 to 98.99 points (13.1% increase), though with notable nonlinear acceleration after 15 study hours (β_3 = 1.89, *SE* = 0.07, $p < 0.001$). Portuguese language showed the highest relative improvement (76.25 to 90.73 points, 19.0%) despite having the lowest absolute scores and weakest time-sensitivity (β_2 = 1.07, *SE* = 0.09, p = 0.012), suggesting language acquisition may benefit more from consistent moderate study rather than intensive time investment.

The results demonstrate that while absolute improvements differ, all subjects benefit from increased study time, but with discipline-specific optimisation requirements. This empirically supports the need for the proposed multi-task learning architecture in educational prediction models.

This work examined, using the Edu-ADMM-LR model, the impact of study time on students' performance in many topics in past studies. This paper helped to clarify the link between study time and performance and offered empirical evidence for enhancement of academic achievement. This work also does a model comparison experiment seeking to evaluate the performance of the Edu-ADMM-LR model in academic achievement prediction by means of another conventional machine learning model.

Figure 4 Results of the comparison experiment (see online version for colours)



Apart from the Edu-ADMM-LR model, many comparison models are included to this study in this experiment: the conventional logistic regression (Blackmore et al., 2021), SVM (Zhu et al., 2021), random forest (Alanazi et al., 2024) and gradient boosting (Seo et al., 2024). This comparative analysis serves dual purposes:

1 validating whether the observed subject-specific patterns can be effectively captured by different modelling approaches

2 assessing the practical advantages of ADMM-optimised framework in educational prediction tasks.

Figure 4 provides a comprehensive visualisation of model performance across critical metrics, including accuracy, computational efficiency, and stability.

To complement the graphical analysis in Figure 4, Table 2 provides a detailed numerical comparison of model performance.

Three key findings emerge: first, Edu-ADMM-LR achieves the highest mean performance (92.3 ± 2.1), outperforming gradient boosting (86.9 ± 3.4) by 6.2% ($p <$ 0.01,
two-tailed t-test), with its ±1*SD* range (90.2–94.4) completely non-overlapping with other models' best-case intervals. Second, the model exhibits exceptional operational stability, evidenced by its narrow performance span (85 – 100, Δ = 15) – 40% tighter than logistic regression's 25-point fluctuation range. Third, its worst-case performance (85 points) still surpasses random forest's mean score (85.1), ensuring reliable predictions even under suboptimal conditions.

Table 2 Performance comparison of Edu-ADMM-LR and baseline models

Model	*Accuracy (%) (mean ± SD)*	*Training time (s)*	*Stability (Δ range)*
Edu-ADMM-LR	92.3 ± 2.1	38 ± 2	15
Logistic regression	85.0 ± 3.8	15 ± 1	25
SVM	86.9 ± 3.4	120 ± 15	22
Random forest	88.3 ± 3.1	65 ± 8	18
Gradient boosting	89.5 ± 2.9	89 ± 10	17

The 100-point maximum performance further confirms Edu-ADMM-LR's unique capacity to identify top-percentile learning patterns – a capability absent in other models (maximum: gradient boosting 95, SVM 90). This granular discriminative power enables precise targeting of pedagogical interventions for both struggling and advanced learners.

These results collectively confirm that the Edu-ADMM-LR model successfully addresses the tripartite challenge of educational prediction: capturing discipline-specific learning patterns (validated by Figure 3), surpassing conventional machine learning benchmarks (quantified in Figure 4), and maintaining practical feasibility for institutional deployment. The model's ability to simultaneously achieve high accuracy, interpretability, and computational efficiency positions it as a transformative tool for data-driven educational management.

5 Conclusions

This study presents Edu-ADMM-LR, an innovative integration of alternating direction method of multipliers with logistic regression, specifically designed to address the unique challenges of educational performance prediction. The framework's dual optimisation architecture successfully bridges the gap between computational efficiency and pedagogical interpretability, leveraging ADMM's global convergence properties to enhance traditional logistic regression while preserving its transparency. Experimental validation confirms the model's superior ability to handle the high-dimensional and often incomplete nature of educational datasets, capturing nuanced discipline-specific learning patterns through differentially weighted predictive features.

Although the Edu-ADMM-LR model performed well in this study, it still has some limitations. Firstly, the dataset used for the experiment originated from a specific higher education institution and may not be fully representative of all higher education teaching and learning environments, which limits the model's ability to generalise to different educational contexts. In addition, although the model performs well on specific datasets, its generalisation ability to other types or domains of datasets may require further validation.

From a model optimisation perspective, the choice of regularisation parameters and learning rate in the ADMM algorithm has a significant impact on model performance, while the optimal values of these parameters may depend on the specific problem and dataset. Despite the improved efficiency of the Edu-ADMM-LR model, its computational resource requirements for very large datasets may still be a consideration. Future research could explore more efficient optimisation algorithms or distributed computing methods to further reduce the computational overhead.

In addition, the interpretability of the model may be reduced in nonlinear scaling. Although logistic regression itself has high interpretability, the introduction of ADMM optimisation and nonlinear features may complicate the interpretation of the model. Future research could develop adaptive hyperparameter tuning strategies and explore ways to improve model interpretability while maintaining high performance.

Future studies will help to address the flaws mentioned in the following spheres.

1 Expanding the scope of the experimental dataset: the dataset mainly contains data from Portuguese students, which may not fully reflect the characteristics of students from other regions, which constitutes a demographic limitation; in order to validate the model's ability to generalise, we suggest expanding the dataset to include student data from different countries and educational systems to increase the diversity and adaptability of the model; and, conducting a long-term tracking study where students' long-term academic performance to validate the predictive stability of the model over time.

2 Introducing other optimisation algorithms or deep learning methods: although ADMM performed satisfactorily in this study, more complex data structures and nonlinear interactions would help. Future research can mix different optimisation techniques, such particle swarm optimisation and genetic techniques, or deep learning approaches, such neural networks and CNNs, to increase the prediction ability and capacity to manage complicated data.

3 Combination of integrated learning and multi-task learning: integrated and multi-task learning could be combined in future research to apply many models to raise prediction accuracy. Multi-task learning enhances multi-dimensional learning and generalisation by letting models simultaneously predict student performance and learning attitudes.

By means of these future research paths, the Edu-ADMM-LR model is predicted to be increasingly important in the analysis of higher education teaching data and the prediction of student performance, so offering more strong data support and decision-making tools in the field of education.

Declarations

All authors declare that they have no conflicts of interest.

References

Al-Alawi, L., Al Shaqsi, J., Tarhini, A. and Al-Busaidi, A.S. (2023) 'Using machine learning to predict factors affecting academic performance: the case of college students on academic probation', *Education and Information Technologies*, Vol. 28, No. 10, pp.12407–12432.

Alanazi, S.H., Abdollahian, M., Tafakori, L., Almulaihan, K.A., Alruwili, S.M. and Alenazi, O.F. (2024) 'Predicting age at onset of childhood obesity using regression, random forest, decision tree, and K-nearest neighbour – a case study in Saudi Arabia', *Public Library of Science ONE*, Vol. 19, No. 9, p.e0308408.

Albahri, A.S., Duhaim, A.M., Fadhel, M.A., Alnoor, A., Baqer, N.S., Alzubaidi, L., Albahri, O.S., Alamoodi, A.H., Bai, J. and Salhi, A. (2023) 'A systematic review of trustworthy and explainable artificial intelligence in healthcare: assessment of quality, bias risk, and data fusion', *Information Fusion*, Vol. 96, pp.156–191.

Batool, S., Rashid, J., Nisar, M.W., Kim, J., Kwon, H-Y. and Hussain, A. (2023) 'Educational data mining to predict students' academic performance: a survey study', *Education and Information Technologies*, Vol. 28, No. 1, pp.905–971.

Bian, K. and Priyadarshi, R. (2024) 'Machine learning optimization techniques: a survey, classification, challenges, and future research issues', *Archives of Computational Methods in Engineering*, Vol. 31, No. 7, pp.4209–4233.

Blackmore, C., Hird, K. and Anderton, R.S. (2021) 'An investigation of secondary school STEM subjects as predictors of academic performance in tertiary level health sciences programs', *International Journal of Higher Education*, Vol. 10, No. 1, pp.76–84.

Buer, S-V., Semini, M., Strandhagen, J.O. and Sgarbossa, F. (2021) 'The complementary effect of lean manufacturing and digitalisation on operational performance', *International Journal of Production Research*, Vol. 59, No. 7, pp.1976–1992.

Czajkowski, M. and Budziński, W. (2019) 'Simulation error in maximum likelihood estimation of discrete choice models', *Journal of Choice Modelling*, Vol. 31, pp.73–85.

Ding, W., Ng, M.K. and Zhang, W. (2024) 'A generalized alternating direction implicit method for consensus optimization: application to distributed sparse logistic regression', *Journal of Global Optimization*, Vol. 90, No. 3, pp.727–753.

Gholami, A., Sun, K., Zhang, S. and Sun, X.A. (2023) 'An ADMM-based distributed optimization method for solving security-constrained alternating current optimal power flow', *Operations Research*, Vol. 71, No. 6, pp.2045–2060.

Hassan, E., Shams, M.Y., Hikal, N.A. and Elmougy, S. (2023) 'The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study', *Multimedia Tools and Applications*, Vol. 82, No. 11, pp.16591–16633.

Jun, M-J. (2021) 'A comparison of a gradient boosting decision tree, random forests, and artificial neural networks to model urban land use changes: the case of the Seoul metropolitan area', *International Journal of Geographical Information Science*, Vol. 35, No. 11, pp.2149–2167.

Lin, C-H., Lin, Y-C. and Tang, P-W. (2021) 'ADMM-ADAM: a new inverse imaging framework blending the advantages of convex optimization and deep learning', *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60, pp.1–16.

Liu, C. and Yang, S. (2024) 'Personalized learning ability classification using SVM for enhanced education in system modeling and simulation courses', *Frontiers of Digital Education*, Vol. 1, No. 4, pp.295–307.

Liu, M. and Yu, D. (2023) 'Towards intelligent e-learning systems', *Education and Information Technologies*, Vol. 28, No. 7, pp.7845–7876.

Meng, C., Song, Y., Li, W. and Ermon, S. (2021) 'Estimating high order gradients of the data distribution by denoising', *Advances in Neural Information Processing Systems*, Vol. 34, pp.25359–25369.

Nagata, T., Nonomura, T., Nakai, K., Yamada, K., Saito, Y. and Ono, S. (2021) 'Data-driven sparse sensor selection based on A-optimal design of experiment with ADMM', *IEEE Sensors Journal*, Vol. 21, No. 13, pp.15248–15257.

Ofori, F., Maina, E. and Gitonga, R. (2020) 'Using machine learning algorithms to predict students' performance and improve learning outcome: a literature based review', *Journal of Information and Technology*, Vol. 4, No. 1, pp.33–55.

Ouyang, F., Zheng, L. and Jiao, P. (2022) 'Artificial intelligence in online higher education: a systematic review of empirical research from 2011 to 2020', *Education and Information Technologies*, Vol. 27, No. 6, pp.7893–7925.

Seo, E-Y., Yang, J., Lee, J-E. and So, G. (2024) 'Predictive modelling of student dropout risk: practical insights from a South Korean distance university', *Heliyon*, Vol. 10, No. 11, p.e30960.

Song, Y., Yuan, X. and Yue, H. (2024) 'The ADMM-PINNs algorithmic framework for nonsmooth PDE-constrained optimization: a deep learning approach', *SIAM Journal on Scientific Computing*, Vol. 46, No. 6, pp.C659–C687.

Tripathy, N., Balabantaray, S.K., Parida, S. and Nayak, S.K. (2024) 'Cryptocurrency fraud detection through classification techniques', *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 14, No. 3, pp.2918–2926.

Vlachopoulos, D. and Makri, A. (2019) 'Online communication and interaction in distance higher education: a framework study of good practice', *International Review of Education*, Vol. 65, No. 4, pp.605–632.

Wakefield, J., Frawley, J.K., Tyler, J. and Dyson, L.E. (2018) 'The impact of an iPad-supported annotation and sharing technology on university students' learning', *Computers & Education*, Vol. 122, pp.243–259.

Zhu, D., Cheng, X., Zhang, F., Yao, X., Gao, Y. and Liu, Y. (2020) 'Spatial interpolation using conditional generative adversarial neural networks', *International Journal of Geographical Information Science*, Vol. 34, No. 4, pp.735–758.

Zhu, X., Ye, Y., Zhao, L. and Shen, C. (2021) 'MOOC behavior analysis and academic performance prediction based on entropy', *Sensors*, Vol. 21, No. 19, p.6629.