

A Differential Privacy Method for Publishing Regression Coefficients in Student Performance Analysis

Ketao Zhang, Zhenghui Feng*, Ruofei Zhu

School of Science, Harbin Institute of Technology, Shenzhen 518055, Guangdong, China

*Corresponding Author: Zhenghui Feng, fengzhenghui@hit.edu.cn

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: As educational informatization deepens, data-driven learning analytics improve teaching quality while also raising concerns about the disclosure of sensitive student performance data. Focusing on individual inference risks in regression analysis and coefficient release for international student grades, this paper introduces and optimizes a budget allocation-based differentially private linear regression model (DPBA). Under strict data normalization, the method transforms the ordinary least-squares objective function into a quadratic form, allocates privacy budgets according to the different global sensitivities of its components, and injects Laplace noise. In this way, it balances privacy protection and data utility. An empirical study using real course data from international students shows that, compared with ordinary linear regression, the DPBA model substantially weakens the dominance of a single extreme sample over regression coefficient estimates and prediction outputs, and it effectively resists individual-data inference attacks. At the same time, prediction accuracy remains highly usable, with a coefficient of determination $R^2 > 0.9$. This study provides education administrators with a reliable technical solution for safely releasing teaching statistics under strong privacy protection constraints.

Keywords: Differential privacy; Student performance analysis; Linear regression; Privacy budget allocation; International education; Data security

Online publication: June 26, 2026

1. Research background

1.1. Privacy challenges in educational data mining

As educational informatization continues to advance, data from students' learning processes has become an important basis for teaching evaluation, learning diagnosis, and instructional decision-making. Educational

data mining and learning analytics help teachers identify student performance, detect weak points in course instruction, and provide data support for personalized teaching ^[1]. However, grades and related learning records are highly sensitive. Without appropriate protection during statistical analysis or result release, they may lead to student privacy leakage, harm to student rights and interests, and declining trust in data use ^[2,3].

The development of statistical analysis also expands how complex educational data can be represented and interpreted. Qu et al. ^[4] systematically discuss exploratory functional data analysis and show that statistical descriptions beyond traditional static indicators are valuable when observations involve continuous variation, dynamic trajectories, or high-dimensional structural features. This trend indicates that, as educational data forms and analysis tasks become more complex, preserving structural information and interpretability while controlling sensitive information exposure is becoming a common challenge in applied statistical research.

Existing research on student data privacy in educational settings mainly follows two paths: institutional governance and technical protection. On the governance side, studies emphasize that student data privacy is not merely a technical issue. It also involves data governance in learning analytics, informed consent, risk communication, privacy awareness, and trust among educational actors ^[2,3,5]. On the technical side, common privacy protection strategies in educational data analysis include limiting data access, minimizing data collection, removing direct identifiers, releasing aggregate statistics, and introducing privacy-preserving mechanisms into machine learning workflows ^[6]. These methods help reduce direct re-identification risk, but statistical model outputs such as regression coefficients and predictions can still create indirect inference risks. Such outputs require protection mechanisms with more formal guarantees.

1.2. Grade regression analysis and individual inference risk

Consider the course Probability Theory and Mathematical Statistics offered to international students. A student's final grade is typically composed of regular performance, final examination score, course paper score, and related components. After the course ends, teachers often need to release class averages, average scores for each assessment component, or the distribution of final grades so that students can understand overall course performance and instructors can support later teaching improvement. An average grade appears to be an aggregate statistic, but when the class is small, or an attacker knows other students' grades, directly releasing the average can still reveal an individual grade. For example, if an attacker knows every student's grade except Alice's and the class average is made public, the attacker can use the average and the known grades to infer Alice's grade. Therefore, even aggregate grade statistics require privacy protection mechanisms ^[6,7].

Regression analysis provides more detailed teaching interpretation than averages. In the course setting above, if the final grade y is used as the dependent variable and regular performance x_1 , final examination score x_2 , and course paper score x_3 are used as independent variables, a linear regression model of the form $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ can be constructed. The regression coefficients $\beta_1, \beta_2, \beta_3$ can be interpreted as the relative influence weights of different assessment components on the final grade. Releasing these coefficients helps teachers explain the course assessment structure and helps students understand the relationship between learning components and final performance. However, regression coefficients are fitted from all student records. A single student's grade record, especially an anomalous or boundary case, may have a substantial effect on coefficient estimation. If an attacker has extensive background knowledge, changes in model

parameters or prediction results may be used to infer the grade information of a specific student.

This form of indirect leakage caused by statistical model outputs can be summarized as individual inference risk. Courses for international students often have small class sizes and limited samples, which makes the final regression coefficient estimates especially vulnerable to the influence of individual records. The core problem studied in this paper is therefore not general grade analysis, but the release of regression coefficients under student grade privacy protection.

1.3. Introducing differential privacy into regression coefficient release

Differential privacy provides a provable mathematical protection framework for this problem. Because it focuses on how adding or removing a single record affects released results, it directly corresponds to individual inference risk in the release of regression coefficients for student grades^[7,8]. At the same time, privacy auditing and evaluation for differentially private machine learning continue to develop. This development further shows that a formal definition alone is insufficient for addressing risks in practical applications. Model outputs must also be examined for sensitivity and auditability in the specific application scenario^[9].

In this setting, the intuitive goal of differential privacy is not simply to delete direct identifiers such as names or student numbers. Instead, it aims to ensure that released results no longer depend visibly on whether a particular student appears in the dataset. In other words, if the same release mechanism is run on a dataset containing Alice's grades and on a neighboring dataset that does not contain Alice, the observed outputs should be highly similar. This similarity should make it difficult for an attacker to determine whether Alice's record participated in the computation or to infer her specific grade^[7,8].

For example, whether releasing a class average or regression coefficients fitted from student grade data, an attacker may combine the output with background knowledge to infer individual information if the output is sensitive to one student's record. Differential privacy addresses this problem by injecting appropriate noise into statistical queries, objective functions, or model outputs. The noise limits the influence of one student record on the final released result and thereby reduces individual inference risk caused by grade release.

When differential privacy is used to release regression coefficients for student grades, it must account for the limited sample size of educational data, the interpretability of grade variables, and the high utility requirements of teaching decisions. If the noise is too small, it cannot adequately limit the influence of a single student record on the released result. If the noise is too large, it can weaken the interpretive value of the regression coefficients and reduce prediction performance. This paper therefore adopts the budget allocation-based differentially private linear regression model (DPBA). By allocating the privacy budget to different components of the objective function, the method protects student grade privacy while preserving model utility as much as possible. The empirical results later show that this method substantially reduces parameter changes and prediction changes under leave-one-out sample deletion, and it maintains acceptable grade prediction performance despite a moderate utility loss. It therefore provides technical support for educators who need to release regression analysis results safely.

2. Research method

2.1. Definition of differential privacy and noise calibration

Before discussing privacy risks in student grade release, the attacker's background knowledge assumption must be clarified. Differential privacy typically uses a strong attacker model, assuming that the attacker

can obtain arbitrary information except for the target student, such as Alice. Under this nearly complete background knowledge condition, if a release mechanism can still limit the additional information that an attacker obtains about Alice's grade from the output, the mechanism provides more robust privacy protection.

Under this assumption, the privacy risk in the grade release setting arises because the released result may change observably when one student's record is added or removed. If exact average grades or regression coefficients are directly released, an attacker may use the relationship between the released value and known information to infer the target student's grade. The key to privacy protection is therefore not only removing explicit identifiers such as names and student numbers, but also reducing the identifiable influence of a single student record on the released result.

A direct idea is to add random perturbation to the released result so that the attacker cannot infer individual information from an exact output. Noise, however, cannot be added arbitrarily. If the noise is too small, it cannot mask the change caused by one student record. If it is too large, it destroys the interpretive value and teaching analysis utility of average grades and regression coefficients. A rigorous mathematical condition is therefore needed to describe the requirement that released results should remain sufficiently similar after one student is added or removed, and to determine the appropriate noise scale. Differential privacy provides that formal definition.

Definition 2.1 (ϵ -Differential Privacy^[8]). For any two datasets D and D' that differ in at most one record, let A be a randomized algorithm, and let $A(D)$ and $A(D')$ denote the outputs of the algorithm on the two datasets. The algorithm A satisfies ϵ -differential privacy, if and only if, for any output set S ,

$$\Pr[A(D) \in S] \leq e^\epsilon \cdot \Pr[A(D') \in S] \quad (1)$$

where ϵ is the privacy budget that controls the strength of privacy protection. In general, a smaller ϵ provides stronger privacy protection but lower data utility.

The core idea of differential privacy is that neighboring datasets that contain or omit a specific individual should produce highly similar probability distributions over query outputs. As a result, even after observing the algorithm output, an attacker cannot determine with high confidence whether the database contains that individual's information. In other words, the influence of an individual's presence or absence on the final algorithm output is strictly bounded, which protects each individual in the database. This influence range is controlled and measured by ϵ .

When ϵ is small, meaning privacy protection is strong, the output distributions of two neighboring datasets that differ in only one record tend to overlap. The attacker therefore has difficulty inferring small changes in the dataset from the observed result. In the example above, the goal of differential privacy is to ensure that the distribution of model outputs, such as regression coefficients, remains as similar as possible whether Alice's data is included or excluded.

The definition above specifies the privacy protection goal, but it does not yet answer how much noise should be added. For student grade release, the noise scale must satisfy two constraints. It must be large enough to mask the influence of one student record on an average grade or regression coefficient, but it must not be so large that it destroys the teaching interpretation value of the released result. The key to noise calibration is therefore to first measure the maximum response of the release function to a single-record change.

This response is global sensitivity. Specifically, if a query or model output changes substantially after one student is added or removed, larger noise is required to satisfy differential privacy. Conversely, if the output is not sensitive to a single-record change, the required noise scale is relatively small. The formal

definition of global sensitivity is given below.

Definition 2.2 (Global Sensitivity^[8]). For a query function $f:D \rightarrow R^k$, whose input is a dataset and whose output is a k -dimensional real vector, the global sensitivity Δf of f is defined as

$$\Delta f = \max_{D, D'} \|f(D) - f(D')\|_2 \quad (2)$$

where D and D' are datasets that differ in at most one record.

Thus, global sensitivity describes the maximum change of the query function f over the worst-case neighboring datasets. It is important that this quantity does not depend on the actual grade distribution in the current dataset. Instead, it is mainly determined by the form of the query function, the value range of variables, and the definition of neighboring datasets. Sensitivity therefore characterizes an upper bound on the potential risk of the release task itself, not the fluctuation in a particular computation.

For example, suppose student grades are restricted to $[0,100]$, and the query function is the sum of scores for an assessment component, $f(D) = \sum_i y_i$. Regardless of the other students' grades in the current class, adding or removing one student can change the sum by at most 100. The global sensitivity of this query is therefore 100. This example shows that sensitivity is determined by the summation query and the grade value boundary, not by a specific class dataset.

After global sensitivity is determined, the noise scale can be linked to Δf and the privacy budget ϵ . To balance privacy protection and data utility, differentially private mechanisms usually choose a noise distribution centered at 0, with scale determined by sensitivity and the privacy budget. The Laplace mechanism introduced below is one of the most common implementations.

Theorem 2.3 (Laplace Mechanism^[8]). Given a dataset D , a query function f , and its sensitivity Δf , the randomized algorithm $A(D) = f(D) + \text{Lap}\left(\frac{\Delta f}{\epsilon}\right)$ provides ϵ -differential privacy. Here, $\text{Lap}\left(\frac{\Delta f}{\epsilon}\right)$ denotes noise sampled from a Laplace distribution centered at 0 with scale $\frac{\Delta f}{\epsilon}$, which is then added to $f(D)$.

The probability density function of a Laplace distribution centered at 0 with scale b is

$$\text{Lap}(x | b) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right) \quad (3)$$

The Laplace distribution is a continuous probability distribution centered at 0 and symmetric around 0. Its density is largest near 0 and decays exponentially as the absolute noise value increases. In the Laplace mechanism, the scale parameter $b = \Delta f / \epsilon$ determines noise strength. When query sensitivity increases or the privacy budget decreases, b increases accordingly, and the mechanism is more likely to generate larger random perturbations.

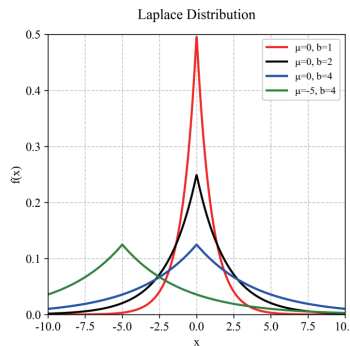


Figure 1. Laplace Distribution.

As Figure 1 shows, most probability mass of the Laplace distribution is concentrated near 0, while both tails retain nonzero probability. This shape indicates that Laplace noise usually does not severely damage the utility of numerical results, but it can still weaken the identifiable influence of a single record on the released result through random perturbation.

The mechanism above provides a basic noise calibration principle for releasing numerical statistics privately. For the release of student grade regression coefficients, perturbing a single statistical query is still insufficient to describe model training and parameter release. The following section first presents the ordinary linear regression model and its objective function, and then introduces the budget allocation-based differentially private linear regression model.

2.2. Budget allocation based differentially private linear regression model

Definition 2.4 (Multiple Linear Regression Model). Let y be the dependent variable, also called the target variable. Let $x = (x_1, x_2, \dots, x_p)^T$ be the independent-variable vector, also called the predictor vector. Let β_0 be the intercept, $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ be the coefficient vector for the independent variables, and ξ be the error term with $\xi \sim N(0, \sigma^2)$. The multiple linear regression model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \xi \quad (4)$$

Given n observations, for the i -th observation, the model in Equation (4) can be written as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \xi_i = x_i^T \beta + \xi_i, \quad i = 1, 2, \dots, n \quad (5)$$

To make the model prediction y_i as close as possible to the observed value, model parameters β are usually estimated by least squares, which constructs the objective function by minimizing the error between observed and predicted values:

$$f(\beta) = \sum_{i=1}^n (y_i - x_i^T \beta)^2 \quad (6)$$

Equation (6) gives the empirical loss function of ordinary linear regression. If this function is minimized directly on raw student grade data and the regression coefficients are released, the model parameters are affected by every student's features and grade record. When an attacker knows information about other students, small changes in the coefficients or objective function may become evidence for inferring the target student's grade. A differential privacy mechanism is therefore needed in model training and parameter release to use random perturbation to limit the identifiable influence of one student record on the final regression result.

To preserve model utility while controlling the noise scale, this paper adopts the budget allocation-based differentially private linear regression model (DPBA). The method does not simply apply a uniform perturbation to the final regression coefficients. Instead, it first rewrites the ordinary least-squares objective function as a quadratic form in β , and then allocates privacy budgets according to sensitivity differences across components. The objective function in Equation (6) can be transformed as follows:

$$f(\beta) = \beta^T A \beta + \beta^T b + c \quad (7)$$

Equation (7) can be regarded as a quadratic function of β :

$$f(\beta) = g(\beta) + h(\beta) + c \quad (8)$$

where $g(\beta) = \beta^T A \beta$ and $h(\beta) = b^T \beta$.

It is important to note that, for the theoretical upper bounds of later sensitivity calculations to hold, the input feature matrix X and the target variable y must be strictly normalized to the interval $[-1, 1]$ during preprocessing. Under this condition, let $\Delta_g = 2(p+1)^2$ and $\Delta_h = 4(p+1)$. Assign privacy budgets ϵ_g and ϵ_h to g and h respectively, and then apply the Laplace mechanism to obtain a new objective function satisfying differential privacy:

$$f^*(\beta) = \left(A + \text{Lap}\left(\frac{\Delta_g}{\epsilon_g}\right) \right) \beta^2 + \left(b + \text{Lap}\left(\frac{\Delta_h}{\epsilon_h}\right) \right) \beta + c \quad (9)$$

When data dimensionality is high, Δ_g is usually much larger than Δ_h . Because Laplace noise is proportional to sensitivity, a larger privacy budget should be assigned to the more sensitive function to reduce the noise scale. Thus, $\epsilon_g \geq \epsilon_h$ is required to improve the accuracy of model estimation.

Minimizing the objective function in Equation (9) gives the DPBA model parameters and constructs a linear regression model satisfying differential privacy. In addition, the constant term c does not affect the minimization result and can therefore be ignored in the algorithm implementation.

2.3. Algorithm for the budget allocation based differentially private linear regression model

The preceding section has transformed the ordinary least-squares objective function into a quadratic form suitable for noise injection. To further describe the implementation of the DPBA model, suppose the dataset contains n students and d regression input dimensions. Let $X \in \mathbb{R}^{n \times d}$ denote the independent-variable matrix, where each row corresponds to one student's grade features and each column corresponds to an independent variable. An intercept column may be included when needed. Let $y \in \mathbb{R}^n$ denote the dependent-variable vector, representing each student's final grade. Given privacy budget ϵ , the algorithm injects Laplace noise into the quadratic and linear terms of the objective function and estimates model parameters $\hat{\beta}$ under differential privacy.

The algorithm is derived from the budget allocation-based differential privacy (DPBA) model proposed by Fang et al. ^[10]. Considering the characteristics of educational data and practical application requirements, this paper makes adaptive simplifications and adjustments to the model, with the specific steps as follows:

Algorithm 1: Budget Allocation Based Differentially Private Linear Regression Model

Input: Normalized independent-variable matrix $X \in [-1, 1]^{n \times d}$, normalized dependent-variable vector $y \in [-1, 1]^n$, privacy budget ϵ

Output: Estimated model parameters $\hat{\beta}$

1. Compute $a = X^T X$, $b = -2X^T y$, $\Delta_g = 2d^2$, $\Delta_h = 4d$, where d is the number of columns in X .

2. Compute the budget allocation ratio and sub-budgets:

$$\alpha = \frac{\Delta_g}{\Delta_g + \Delta_h}, \quad \epsilon_g = \alpha \epsilon, \quad \epsilon_h = \epsilon - \epsilon_g$$

3. Randomly generate Laplace noise:

$$e_g \sim \text{Lap}\left(\frac{\Delta_g}{\epsilon_g}\right)$$

Input: Normalized independent-variable matrix $X \in [-1,1]^{n \times d}$, normalized dependent-variable vector $y \in [-1,1]^n$, privacy budget ϵ

$$e_h \sim \text{Lap}\left(\frac{\Delta_h}{\epsilon_h}\right)$$

4. Compute the estimated model parameters:

$$\hat{\beta} = \arg \min_{\hat{a}} \{(a + e_g)\beta^2 + (b + e_h)\beta\}$$

3. Empirical data analysis

3.1. Data description and preprocessing

The international student grade dataset used in this study contains 22 variables that cover multiple dimensions of course performance. The variables can be divided into discrete and continuous variables. Discrete variables mainly include department affiliation, class name, grade category for Probability Theory and Mathematical Statistics, course name and code, unique task identifier with a 32-bit ID, training type whose value is always 1, the dataset selected for the course paper, research topic description, group number from 1 to 19, leader indicator as a 0/1 binary variable, and students' countries of origin. Continuous variables are grade indicators, including regular performance measured by weighted average homework and quiz scores, course paper score, final examination score, final grade, four quiz scores from quiz1 to quiz4, four assignment scores from assignment1 to assignment4, final project score, actual attendance count, and attendance percentage. All grade indicators are measured on a 100-point scale. The assignment data contain a few anomalous records and specially marked absence-from-exam cases. After data preprocessing, the final empirical analysis uses course records from 77 international students.

Feature selection was conducted before linear regression modeling. Regular performance and final examination score were ultimately included as independent variables, and final grade was used as the dependent variable. Specifically, the independent-variable matrix in the subsequent models is denoted by X , which consists of regular performance and final examination score. Each row of X corresponds to one student's grade features, and each column corresponds to an independent variable. The dependent-variable vector is denoted by Y , also written as y in the algorithm, and represents each student's final grade. This selection reduces the cumulative noise effect caused by high-dimensional features in the differential privacy mechanism, and it is also supported by empirical feature selection results showing that the coefficients of other variables are not significant. The limited privacy budget can therefore be concentrated on key variables to improve overall model utility.

3.2. Empirical analysis design

The empirical analysis uses ordinary linear regression as the baseline and compares it with the DPBA-based differentially private linear regression model as the privacy-preserving method. The privacy budget is set to $\epsilon=5$. This value is chosen mainly because the empirical dataset is small, containing only 77 students. Under a small-sample condition, a much smaller ϵ would excessively amplify the injected Laplace noise and make the regression model lose its teaching interpretation value and practical utility. This paper therefore makes a moderate practical compromise to preserve model utility under limited sample size. As data scale

increases in future work, the baseline value of epsilon can be further reduces, and sensitivity analysis over different ϵ values can be conducted.

Table 1. Model performance results

Metric	Ordinary Linear Regression	DPBA Model
MSE	0.0038	0.0070
R^2	0.9689	0.9151

To reduce the contingency caused by a single random split and to reduce the influence of randomness in the differential privacy noise injection process, this paper performs 100 repeated random splits on the 77 student samples. Repeated analysis more objectively demonstrates the combined advantages of the DPBA model over ordinary linear regression in utility preservation and privacy protection. Each split uses an 80% training set and a 20% test set, with a reproducible random splitting process. Subsequent analyses of model utility and privacy protection ability are both based on this repeated split framework. Utility is evaluated on the test set, while privacy protection ability is evaluated through leave-one-out sample deletion analysis on the training set obtained in each split.

3.3. Model utility and privacy protection analysis

In educational data applications, the practical value of a privacy protection method first depends on whether it can protect student grade privacy while maintaining acceptable model utility. This section first compares the prediction performance of the DPBA model with ordinary linear regression and quantifies the utility loss introduced by differential privacy. It then analyzes model parameter and output changes from the perspective of removing student samples one at a time, thereby examining the protection provided for individual student grade records.

3.3.1. Model utility analysis

This study evaluates model utility using mean squared error (MSE) and the coefficient of determination (R^2), with ordinary linear regression as the baseline model. The table reports averages over 100 repeated splits, and the results are shown in **Table 1**.

The empirical results show that the MSE of the DPBA model increases from 0.0038 for ordinary linear regression to 0.0070, an absolute increase of about 0.0032. The coefficient of determination decreases from 0.9689 to 0.9151, an absolute decrease of about 0.0538. This indicates that differential privacy noise does introduce a utility loss, but the loss is limited. The DPBA model still has an MSE below 0.01 and an R^2 above 0.9, showing that it retains strong explanatory ability for grade variation. In student performance analysis, the DPBA model therefore provides privacy protection with only a small loss in prediction accuracy, which is an important advantage over directly releasing ordinary regression results.

Figure 2 further shows the prediction performance of the two models across repeated analyses. The left panel shows that the prediction deviation introduced by the DPBA model is about 0.5 points on a 100-point scale, which is within an acceptable range for teaching management. The middle panel shows that privacy processing does not change the basic relationship among regular performance, final examination score, and final grade. The right panel shows that DPBA predictions are slightly more dispersed, but they remain concentrated near the diagonal overall. In the high-score range, where the final grade is >80, DPBA prediction

accuracy is basically consistent with the ordinary model. The advantage of the DPBA model is therefore not that it eliminates utility loss, but that it controls the loss within a small range while providing privacy protection for subsequent regression coefficient release.

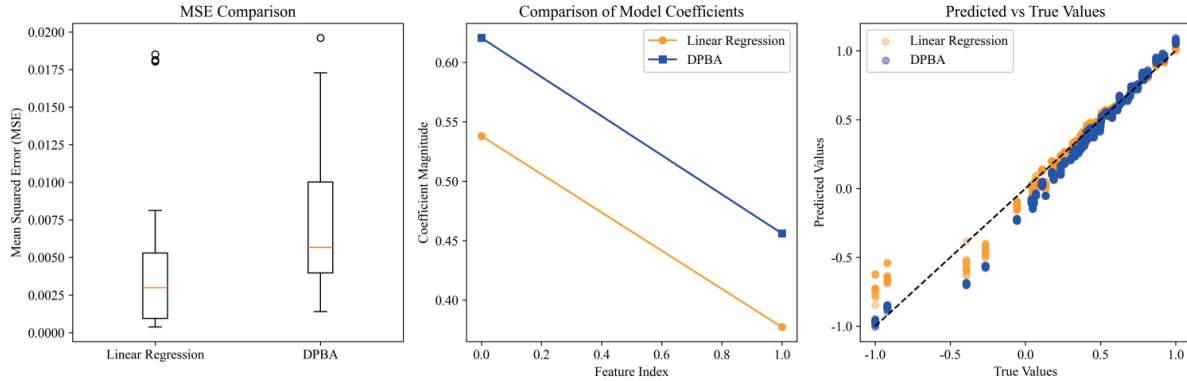


Figure 2. Comparison between the DPBA model and the linear regression model.

3.3.2. Privacy protection analysis

After confirming that model utility remains acceptable, this paper further examines the privacy protection ability of the DPBA model. According to the definition of differential privacy, the attacker is assumed to know almost all background information in the database except for the target individual. The key to privacy protection is therefore to limit the influence of a single student sample on model parameters and prediction outputs. For this purpose, this paper conducts leave-one-out sample deletion analysis on the training set obtained in each random split. In each round, one student sample is removed from the training set while all remaining training samples are retained. The model is then retrained, and changes relative to the model trained on the complete training set are computed. In the next round, the previously removed sample is restored, and another student sample is removed. This process is repeated in each of the 100 random splits, and the results are averaged over the same sample-order positions within the training set. The horizontal coordinate, Index, denotes the order number of the removed sample within the training set, allowing the average influence associated with different training sample positions to be observed.

As **Figure 3a** shows, compared with ordinary linear regression, the DPBA model has clearly smaller parameter changes after any student sample is removed, and the curve is smoother overall. The larger fluctuations in the ordinary linear regression curve indicate that some student samples exert stronger influence on regression parameters. By introducing differential privacy noise into the objective function, the DPBA model effectively weakens the dominance of a single sample over parameter estimation. This comparison shows that the DPBA model can reduce the risk that an attacker identifies influential samples or infers student grade information through parameter changes.

As **Figure 3b** shows, changes in prediction outputs follow the same pattern as changes in parameters. Ordinary linear regression produces more visible prediction fluctuations after some student samples are removed, whereas the DPBA model produces smaller and more stable output changes overall. This indicates that differential privacy processing reduces not only the sensitivity of regression coefficients to individual student samples, but also the dependence of prediction results on individual records. For educational

scenarios that require releasing model coefficients or prediction analysis results, this stability can further reduce indirect inference risk caused by model outputs.

The protective effect of differential privacy can also be observed from the distributional perspective. As **Figure 4** shows, **Figure 4a** and **Figure 4b** present the frequency distributions of changes in model parameters and prediction results during leave-one-out sample deletion under one repeated split. The distribution for ordinary linear regression is more likely to contain large change values, indicating that some student samples may strongly influence parameter estimation and prediction results. The DPBA model has a more concentrated distribution, and extreme changes are suppressed. The advantage of the DPBA model is therefore reflected not only in improved average stability, but also in compression of tail risk. It reduces the excessive perturbation that a small number of high-influence samples can impose on released results, thereby protecting student grade privacy more effectively.

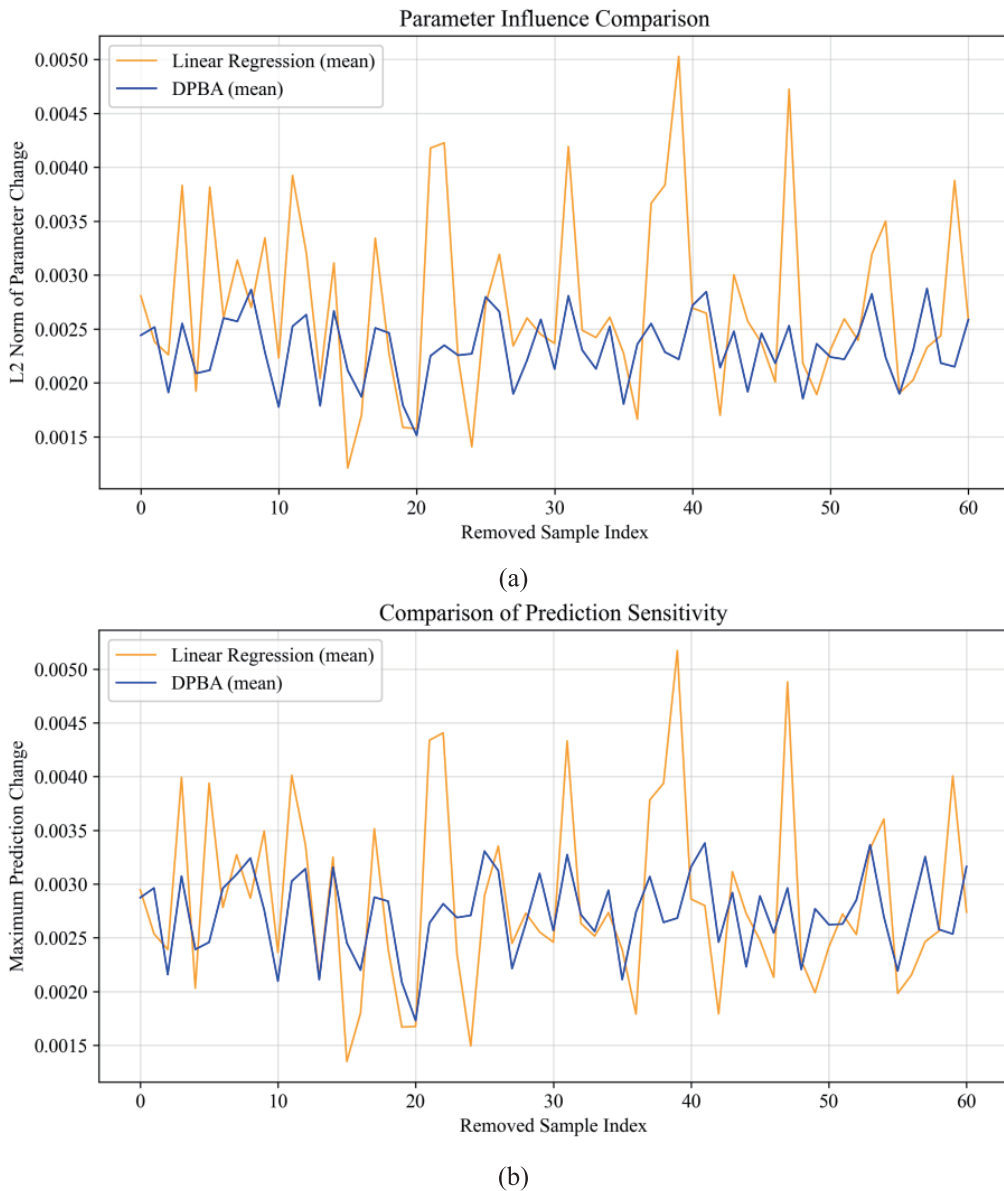


Figure 3. Changes in model parameters and outputs after leave-one-out sample removal.

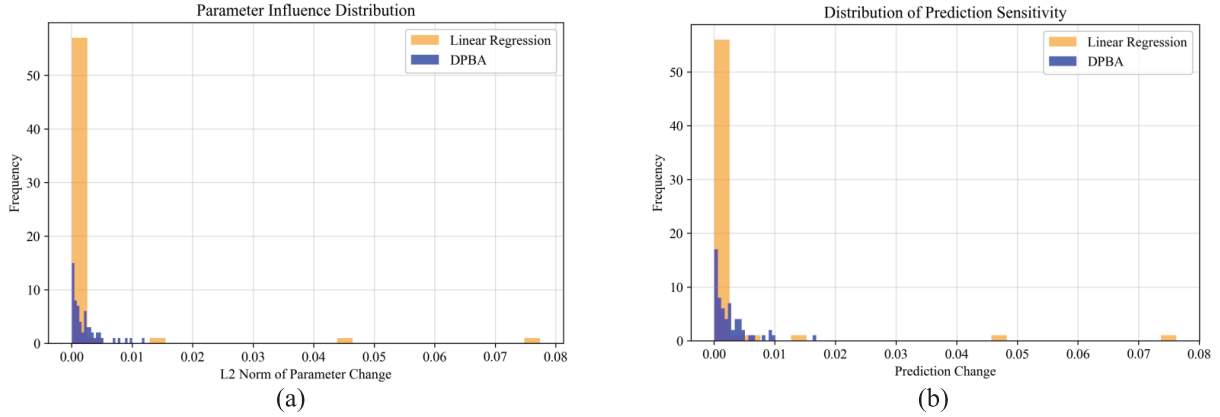


Figure 4. Frequency histograms of parameter and prediction changes after leave-one-out sample removal.

3.4. Summary

The empirical analysis shows that the DPBA model retains good model utility after differential privacy protection is introduced. Compared with ordinary linear regression, its MSE increases and its R^2 decreases slightly, but the overall prediction error remains low and the explanatory ability stays within an acceptable range.

At the same time, in the leave-one-out analysis of training samples, the parameter changes and prediction changes of the DPBA model are both smaller than those of ordinary linear regression. This indicates that the DPBA model weakens the influence of a single student sample on model outputs and reduces grade privacy inference risk caused by the release of regression coefficients or prediction results.

4. Discussion

This paper positions the DPBA model as a contextualized application of an existing method. It applies the DPBA model to the release of regression coefficients for student grades and focuses on its applicability in educational data analysis. The empirical results show that the model achieves relatively stable privacy protection under a moderate utility loss. On the one hand, the prediction error of the DPBA model remains low and R^2 stays above 0.9, indicating that the results can still support basic grade analysis and teaching interpretation. On the other hand, when training samples are removed one at a time, the parameter changes and prediction changes of the DPBA model are both smaller than those of ordinary linear regression, indicating that it reduces the influence of a single student sample on model outputs.

In terms of application value, the DPBA model provides educators with a relatively operable privacy protection path. For scenarios that require releasing regression coefficients, model prediction results, or related teaching analysis conclusions, the model can preserve the main relationship structure among regular performance, final examination score, and final grade without directly exposing student grade records. Teachers and school administrators can therefore continue to use regression models for course evaluation and learning analysis while reducing grade privacy inference risk.

Several limitations remain. First, the empirical data come from a single course setting, and the sample size is relatively limited. Future work can examine model stability using more courses, different majors, and larger student datasets. Second, because of the small current sample size, this paper uses a relatively large fixed privacy budget, $\epsilon=5$, to preserve model utility. It has not yet systematically compared changes in

model utility and privacy protection strength under different privacy budgets. As data scale increases, ϵ can be reduced further and a comprehensive budget sensitivity analysis can be conducted. Finally, this paper mainly compares the DPBA model with ordinary linear regression. Future work can introduce other differentially private regression methods or more fine-grained attack evaluation metrics to assess the applicability boundaries of different methods in educational data analysis more comprehensively.

Overall, the DPBA model achieves a favorable balance on the student grade data used in this paper. Compared with ordinary linear regression, its prediction performance declines but remains usable, while it substantially reduces the influence of a single student sample on regression parameters and prediction outputs. The model can therefore serve as a feasible privacy protection solution for educational data release and grade analysis. Future work can further improve the study through multi-dataset validation, privacy budget selection, and method comparison.

Funding

Guangdong Provincial Education Science Planning Project (Higher Education Special Project) (Project No.: 2025GXJK0601); the Humanities and Social Sciences Research Project of the Ministry of Education of China (Project No.: 25YJA910002); the Basic Research Fund in Shenzhen Natural Science Foundation (Project No.: JCYJ20240813104924033); the Startup Research Funding for Newly Recruited High-Caliber Talents of Harbin Institute of Technology (Shenzhen), China (Project No.: HA11409087)

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Romero C, Ventura S, 2020, Educational Data Mining and Learning Analytics: An Updated Survey. *WIREs Data Mining and Knowledge Discovery*, 10(3): e1355.
- [2] Prinsloo P, Slade S, Khalil M, 2022, The Answer Is (Not Only) Technological: Considering Student Data Privacy in Learning Analytics. *British Journal of Educational Technology*, 53(4): 876–893.
- [3] Mutimukwe C, Viberg O, Oberg LM, et al., 2022, Students' Privacy Concerns in Learning Analytics: Model Development. *British Journal of Educational Technology*, 53(4): 932–951.
- [4] Qu Z, Dai W, Euan C, et al., 2025, Exploratory Functional Data Analysis. *Test*, 34(2): 459–482.
- [5] Jones KML, VanScoy A, Bright K, et al., 2022, A Measurement of Faculty Views on the Meaning and Value of Student Privacy. *Journal of Computing in Higher Education*, 34(3): 769–789.
- [6] Schleiss J, Gunther K, Stober S, 2022, Protecting Student Data in ML Pipelines: An Overview of Privacy-Preserving ML. In Rodrigo MM, Matsuda N, Cristea AI, et al. (eds.), *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium*. Cham: Springer International Publishing, 532–536.
- [7] Yang Y, Zhang Z, Miklau G, et al., 2012, Differential Privacy in Data Publication and Analysis. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*. Scottsdale, AZ: ACM, 601–606.
- [8] Dwork C, McSherry F, Nissim K, et al., 2006, Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory*

of Cryptography Conference. Berlin: Springer, 265–284.

- [9] Namatevs I, Sudars K, Nikulins A, et al., 2025, Privacy Auditing in Differential Private Machine Learning: The Current Trends. *Applied Sciences*, 15(2): 647.
- [10] Fang X, Yu F, Yang G, et al., 2019, Regression Analysis with Differential Privacy Preserving. *IEEE Access*, 7: 129353–129361.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.