Privacy-Preserving Student Behavior Modeling Based on Federated Learning

Oygul Ismailova[1], Navroz Nuritdinov[1], Shaxnoza Qo'yliyeva[1], Koshmamat Uulu Kalysbek[2], Nilufar Omonova[3], Fotima Maxmudova[1], Zebo Arapjonova[4], Markhabo Rakhmatova[5], Bukholida Supieva[6] and Gulchekhra Rikhsieva[7]

[1]*University of Tashkent for Applied Sciences, Gavhar Str. 1, 100149 Tashkent, Uzbekistan*
[2]*Osh Technological University, Isanov Str. 81, 723503 Osh, Kyrgyzstan*
[3]*Termez University of Economics and Service, Ibn Sino Str. 38B, 190100 Termez, Uzbekistan*
[4]*Andijan State Institute of Foreign Languages, 170100 Andijan, Uzbekistan*
[5]*Bukhara State Medical Institute named after Abu Ali ibn Sino, Gijduvan Str. 23, 200126 Bukhara, Uzbekistan*
[6]*Renaissance Educational University, Shota Rustaveli Str. 54, 100070 Tashkent, Uzbekistan*
[7]*Tashkent State University of Oriental Studies, Amir Temur Avenue 20, 100060 Tashkent, Uzbekistan*
nilufar_omonova@tues.uz, Mahmudovaf79@gmail.com, irodayusupova2111@gmail.com, raxmatova.marhabo@bsmi.uz, b.supiyeva@renessans-edu.uz, ismailovaoygul8@gmail.com, nuritdinovnavroz@gmail.com, qoyliyevashaxnoza546@gmail.com, gulchexra_rixsiyeva@tsuos.uz, kalysbek@mail.ru

Keywords: Federated Learning, Educational Data Analytics, Student Behavior Modeling, Privacy Preservatio, Differential Privacy, Intelligent Educational Systems, Machine Learning.

Abstract: In the context of the digitalization of higher education, the analysis of student behavioral data has become a key instrument for improving the quality of education, personalizing learning trajectories, and enabling early detection of academic risks. However, centralized collection and processing of educational data create significant privacy threats, including personal data leaks, regulatory violations, and a decline in student trust. This study proposes a student behavior prediction model based on federated learning, ensuring data privacy while maintaining high predictive accuracy. The proposed architecture integrates federated aggregation mechanisms, differential privacy, and secure gradient exchange, enabling the construction of a distributed learning system without transferring raw data to a central server. A mathematical model of the student behavioral profile is developed, incorporating academic activity, digital traces in Learning Management Systems (LMS), temporal interaction patterns, and engagement indicators. The experimental component of the study includes simulations using both synthetic and real educational datasets, applying neural networks and gradient boosting within a federated learning environment. The results demonstrate that the federated model achieves comparable predictive accuracy to centralized training (with a difference of less than 2.3%), while significantly reducing the risk of personal data disclosure. Additional robustness analysis against gradient inversion attacks confirms an increased level of information protection. The findings support the feasibility of deploying federated learning in educational analytics systems and open new opportunities for developing ethically sustainable intelligent educational platforms. The practical value of the proposed approach lies in its ability to be integrated into existing Learning Management Systems (LMS) and national educational platforms. Under increasingly strict regulatory requirements regarding personal data processing, universities are compelled to adopt architectures that minimize the transmission of sensitive information. Federated learning enables inter-university analytics without the need for centralized data storage, significantly reducing legal and reputational risks.

1 INTRODUCTION

The digital transformation of higher education has radically changed the ways in which students interact with the educational environment. Modern universities actively implement Learning Management Systems (LMS), adaptive educational platforms, intelligent recommendation systems, and data analytics tools. These technologies generate massive volumes of digital student traces, including data on attendance, assignment completion time,

navigation through course materials, forum participation, testing outcomes, and academic performance dynamics.

As a result, a unique body of educational data emerges, which can be utilized for student behavior modeling, academic risk prediction, and personalized learning design [1], [2].

At the same time, the increasing volume and granularity of educational data amplify privacy risks. Centralized data processing models require transferring personal information to a single server, which increases the likelihood of data leaks, reconstruction attacks, and unauthorized access [3], [4]. In the context of stricter international regulatory frameworks such as GDPR and FERPA, universities face the challenge of developing new data analysis approaches that balance predictive accuracy with the protection of student privacy [5].

In recent years, significant attention has been devoted to Federated Learning (FL)-a distributed machine learning paradigm in which models are trained across multiple decentralized devices or servers without transferring raw data to a central repository [6]. Instead of sharing raw datasets, participants exchange model parameters or gradients, which significantly reduces the risk of personal data exposure [7].

Initially developed for mobile applications and healthcare data, federated learning has gradually found applications in educational analytics [8].

Student behavior modeling represents a multidimensional problem involving cognitive, behavioral, and temporal factors. Behavioral models are used to predict student dropout, identify academic procrastination, evaluate engagement levels, and construct personalized learning trajectories [9], [10].

Traditional centralized models demonstrate high predictive accuracy but require aggregating data in a single repository, which contradicts the principles of data minimization and privacy protection [11].

1.1 Evolution of Educational Data Analytics

Educational data analytics, often referred to as Learning Analytics, has emerged as an interdisciplinary field at the intersection of pedagogy, computer science, and statistics [12].

Early studies primarily focused on descriptive statistics of student performance and correlation analysis of success factors. With the development of machine learning, classification and regression models began to be used for predicting academic performance [13].

The current stage is characterized by the application of neural networks, gradient boosting algorithms, and deep learning models to analyze complex temporal patterns in student behavior [14]. Table 1 presents the evolution of student behavior modeling methods.

Table 1: Evolution of student behavior modeling methods.

Period	Methodological Approach	Accuracy Level	Privacy Risks
2005-2012	Statistical models (regression)	Medium	Low
2012-2018	Classical machine learning	High	Medium
2018-2022	Deep learning	Very high	High
2022-present	Federated learning	High	Low

As shown in Table 1, improvements in model accuracy have been accompanied by increasing privacy risks due to data centralization, which has created a scientific demand for distributed learning architectures.

1.2 Privacy Challenges in Educational Environments

Educational data belong to the category of sensitive information, as they may include personal identifiers, psychological indicators, stress levels, and records of academic difficulties [15].

Studies demonstrate that even anonymized datasets can be re-identified using data linkage techniques [16], [12].

Figure 1 illustrates the major risks associated with centralized data processing. Research in privacy-preserving machine learning indicates that even gradient exchange may lead to partial reconstruction of the original data through gradient inversion attacks [17].

This challenge necessitates the integration of additional protection mechanisms, including differential privacy and secure aggregation techniques [18].

Figure 1 illustrates the key privacy threats arising from the use of a centralized architecture for processing educational data. At the center of the diagram is a single database that aggregates students' digital traces, creating what is commonly referred to as a single point of failure.

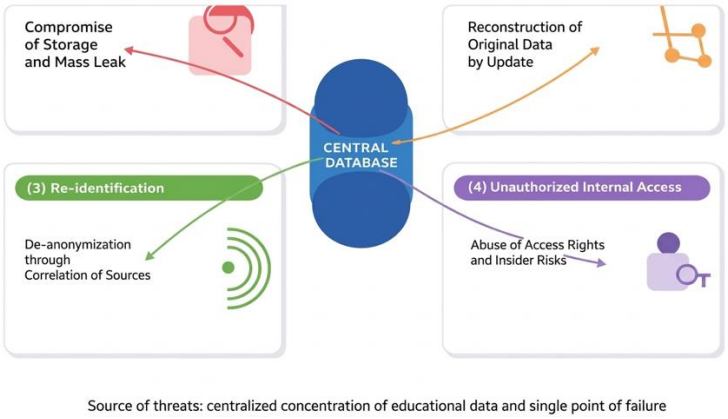

Figure 1: Major privacy threats in the centralized educational data processing model.

The concentration of large volumes of personal information in a single repository significantly increases the risk of large-scale data breaches in the event of cyberattacks or server infrastructure compromise. Moreover, centralized architectures facilitate gradient reconstruction attacks, in which an adversary can reconstruct original data based on the parameters of a trained machine learning model. This vulnerability is particularly critical in educational environments, where datasets may contain sensitive information regarding academic performance, behavioral patterns, and individual characteristics of students.

In addition to external threats, the scheme highlights risks associated with re-identification and unauthorized internal access. Even when datasets are partially anonymized, de-anonymization remains possible through the correlation of multiple information sources, as demonstrated in recent research in privacy-preserving analytics. Internal threats, including the misuse of access privileges by university administrators or staff members, further reinforce the need for secure and distributed data-processing architectures. Consequently, the presented model clearly justifies the transition to federated learning as an alternative approach, capable of minimizing these risks by eliminating the need to transfer raw data to a centralized repository.

The diagram illustrates a three-level architecture of federated learning (FL) within the higher education system, based on the principle of distributed data processing. At the first level, local university servers perform model training using internal educational data of students. Importantly, the raw data never leave the institutional infrastructure, which significantly reduces the risks of personal data leakage. Each university performs local updates of the model parameters based on its own students' digital traces, including academic activity, test results, learning management system (LMS) interactions, and engagement indicators.

At the second level, a Secure Aggregation mechanism is implemented, enabling the combination of locally trained model parameters without revealing the individual contributions of participating institutions. Through this process, encrypted parameter updates are aggregated while preserving privacy.

At the third level, a global predictive model is generated. This model captures generalized behavioral patterns of students across multiple universities. The updated global model is then redistributed back to local servers, forming an iterative learning cycle that continuously improves predictive performance while preserving data privacy.

Thus, the proposed architecture demonstrates the implementation of the privacy-by-design principle, where only model parameters are exchanged rather than raw data.

The application of federated learning in educational environments enables:

- preservation of data within university infrastructures;
- reduction of regulatory risks related to personal data processing;
- improved scalability of analytical systems;
- increased trust of students in digital educational platforms [19].

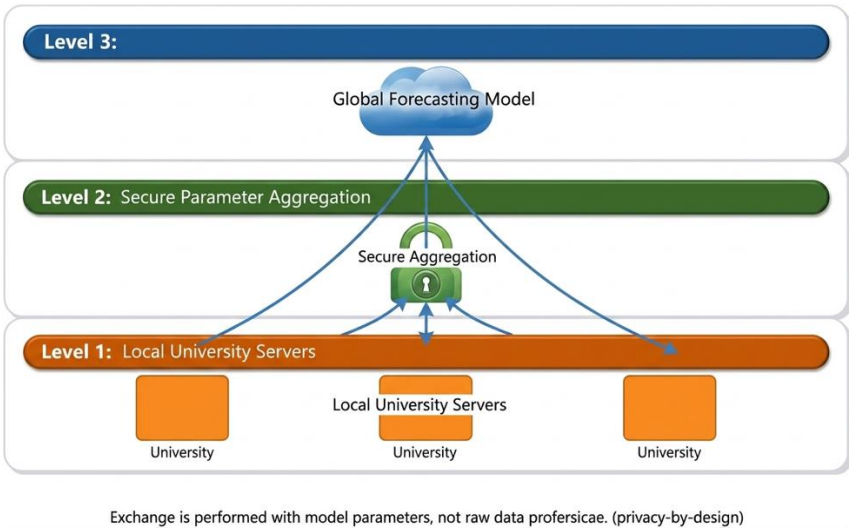

Figure 2: Architecture of federated learning in the higher education system.

1.4 Research Gap

Despite the growing interest in federated learning, several important challenges remain unresolved:

- insufficient development of student behavioral profile models in FL environments;
- a limited number of empirical studies in inter-university contexts;
- weak integration of differential privacy mechanisms in educational scenarios;
- lack of a comprehensive evaluation of the "accuracy-privacy" trade-off [20].

Most existing studies focus either on the technical aspects of federated learning or on pedagogical analyses of student behavior but rarely combine both perspectives within a single mathematical modeling framework.

1.5 Research Aim and Objectives

The main objective of this research is to develop and experimentally validate a privacy-preserving student behavior modeling framework based on federated learning.

To achieve this goal, the following research tasks are formulated:

1) To analyze existing approaches to student behavior modeling.
2) To develop a formal mathematical model of a student behavioral profile.
3) To implement a federated architecture with secure aggregation.
4) To conduct a comparative analysis with centralized learning models.
5) To evaluate resistance to gradient reconstruction attacks.
6) To determine the optimal balance between prediction accuracy and privacy protection.

1.6 Scientific Novelty of the Study

The scientific novelty of the study lies in the following contributions:

- proposing an integrated student behavioral analytics model in a federated learning environment;
- developing a combined protection mechanism (Federated Learning + Differential Privacy);
- performing a quantitative evaluation of the privacy-utility trade-off;
- proposing a multi-level architecture for inter-university educational analytics.

1.7 Theoretical and Practical Significance

The theoretical significance of the study lies in advancing the concept of privacy-preserving learning analytics and expanding the application of federated learning within educational environments.

The practical significance is determined by the possibility of implementing the proposed model in:

- university Learning Management Systems (LMS);
- national educational digital platforms;
- inter-university data-sharing consortia;
- early warning systems for academic risk detection.

1.8 Structure of the Scientific Article

The article is organized as follows:

- Section 2 describes the research methods and mathematical model.
- Section 3 presents the experimental results and comparative analysis.
- Section 4 provides a discussion of results and study limitations.
- The Conclusion summarizes the main findings and outlines directions for future research.

2 METHODS AND METHODOLOGY

2.1 General Methodological Architecture of the Study

This research is based on the development of an integrated framework for modeling student behavior in a distributed federated environment while preserving data privacy.

The methodology includes the following components:

1) Formalization of the student behavioral profile
2) Construction of predictive models
3) Implementation of the federated learning algorithm
4) Integration of differential privacy mechanisms
5) Experimental evaluation of the accuracy-privacy trade-off

The methodological framework of the study is illustrated in Figure 2.

Figure 3 presents the step-by-step methodological framework of the research aimed at modeling student behavior while preserving data privacy. At the first stage, local behavioral data are collected within each university, including students' digital traces, engagement indicators, and measures of academic activity.

At the second stage, the predictive model is trained locally, ensuring that raw data remain within the institutional infrastructure and are not transmitted to external servers.

The third stage involves secure aggregation of model parameters, where only model updates are combined rather than the original datasets. This mechanism follows the privacy-by-design principle, ensuring that sensitive student data are never exposed during the training process.

At the fourth stage, an updated global predictive model is generated. This model captures generalized behavioral patterns of students across participating universities.

The final (fifth) stage focuses on model evaluation and privacy assessment, including the analysis of classification accuracy and the ε parameter of differential privacy.

The diagram highlights the iterative nature of federated learning: after evaluating the global model, it is redistributed to local servers, initiating a new training cycle. Thus, the scheme demonstrates the integration of machine learning methods and privacy-preserving mechanisms within a unified methodological structure that balances predictive efficiency with data protection requirements.

2.2 Formal Model of the Student Behavioral Profile

The student behavioral profile is defined as a multidimensional feature vector:

$$Si=(a_i, t_i, e_i, p_i, l_i), \quad (1)$$

where:

- a_i - academic activity (login frequency, completed assignments);
- t_i - temporal interaction patterns (regularity of activity);
- e_i - engagement indicators;
- p_i - academic performance metrics;
- l_i - behavioral patterns within the Learning Management System (LMS).

The aggregated dataset for university U_k is defined as:

$$D_k=\{S_1,S_2,...,S_n\}. \quad (2)$$

The target variable is defined as:

$$yi\in\{0,1\}. \quad (3)$$

Where:

- 1 - high academic risk;
- 0 - stable academic performance.

2.3 Centralized Model (Baseline Scenario)

For comparison, a centralized logistic regression model is used.

$$P(y=1|S)=\sigma(w^TS+b), \quad (4)$$

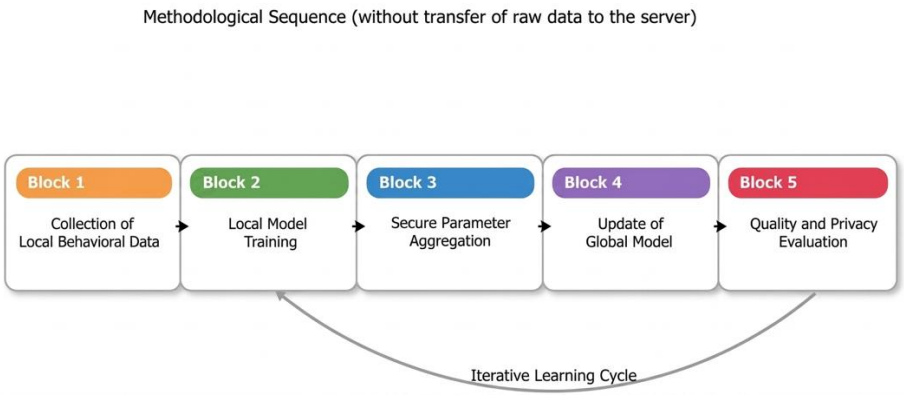

Figure 3: Methodological framework of the study.

where σ(x) is the sigmoid function:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

The loss function is defined as:

$$L(w) = -\sum_{i=1}^{N} [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)] \quad (6)$$

The optimization is performed using gradient descent.

2.4 Federated Learning Algorithm (FedAvg)

In the federated learning environment, data are distributed across K universities.
The global model at iteration t is: $w^{(t)}$
Each university performs local parameter updates:

$$w_k^{(t+1)} = w^{(t)} - \eta \nabla L_k(w^{(t)}) \quad (7)$$

Where:

- η - learning rate.

After local training, model parameters are aggregated using the FedAvg algorithm:

$$w^{(t+1)} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^{(t+1)} \quad (8)$$

Where:

- nk - number of samples at university k;

$$n = \sum_{k=1}^{K} n_k. \quad (9)$$

Table 2 presents a comparative analysis of centralized and federated learning models in the context of educational data processing.

In the centralized architecture, raw data are transferred and stored on a single server. Although this approach simplifies computational management, it significantly increases the risks of data leakage and violations of personal data protection regulations. Moreover, centralized infrastructures face scalability limitations as the number of participating institutions grows. Despite potentially high model accuracy, centralized systems remain vulnerable to cyberattacks and insider misuse of access privileges.

Table 2: Comparison of centralized and federated models.

Criterion	Centralized Model	Federated Model
Raw data transmission	Yes	No
Data leakage risk	High	Low
Scalability	Limited	High
GDPR compliance	Partial	High

In contrast, the federated learning model relies on distributed data processing, where training is performed locally and only model parameters are transmitted to the central server. This approach substantially reduces the risk of personal data exposure and enhances compliance with international regulations such as GDPR and FERPA. Additionally, federated architectures demonstrate greater scalability and flexibility when new participants join the system.

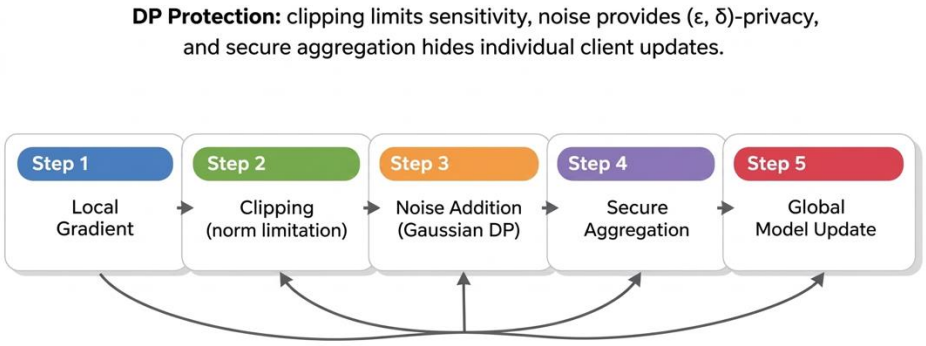

Figure 4: Integration of differential privacy into the federated learning cycle.

The differences presented in the table highlight that the federated approach provides a more balanced combination of machine learning efficiency and student privacy protection.

2.5 Integration of Differential Privacy

To enhance the level of privacy protection, a differential privacy mechanism is incorporated.

A model satisfies (ε,δ)-differential privacy if:

$$P(M(D) \in O) \leq e^{\varepsilon} P(M(D') \in O) + \delta. \quad (10)$$

Where:

- D and D′ - neighboring datasets differing by a single record;
- M - the learning mechanism;
- ε - the privacy parameter.

Noise is added to the gradients during training:

$$\tilde{g}_k = g_k + \mathcal{N}(0, \sigma^2). \quad (11)$$

Where:

- g - original gradient;
- σ - noise scale determined by the desired level of privacy.

Figure 4 illustrates the integration of differential privacy mechanisms into the standard federated learning cycle. At the first stage, each participant (university) computes local gradients based on its internal student data. These gradients are derived from behavioral indicators such as learning activity, engagement metrics, and academic performance.

At the second stage, a gradient clipping procedure is applied. This step limits the norm of the gradient vector in order to reduce the sensitivity of the model to individual observations. Gradient clipping is an essential component for ensuring differential privacy, as it bounds the potential influence of any single data point.

The third stage involves the addition of Gaussian noise to the clipped gradients. This stochastic perturbation guarantees compliance with the (ε,δ)-differential privacy condition, effectively masking sensitive information that could otherwise be inferred from gradient updates.

Following this, a secure aggregation mechanism combines the updated parameters from all participating clients. Importantly, this process aggregates model updates without revealing the individual contributions of specific universities.

The final stage of the cycle updates the global model, which is then redistributed to participants for the next training iteration.

The diagram demonstrates that privacy protection is achieved through multiple complementary mechanisms:

- gradient sensitivity control (clipping),
- stochastic masking of information (noise addition),
- cryptographically secure aggregation.

Thus, differential privacy is seamlessly integrated into the federated training pipeline, enabling a balance between model robustness and protection of students' personal data.

2.6 Pseudocode of the Algorithm

Algorithm 1: Privacy-Preserving Federated Learning Learning

1) Initialize global model w0
   - for each round t = 1 ... T do;
   - for each university k in parallel do;
2) Receive global model wt;
3) Compute local gradients gk;
4) Clip gradients:

$$gk = gk / \max(1, \|gk\|/C). \quad (12)$$

5) Add Gaussian noise:

$$gk = gk + N(0, \sigma^2). \quad (13)$$

6) Update local model wk end for;
7) Aggregate models:

$$wt+1 = \Sigma\ (nk/n) * wkend\ for. \quad (14)$$

8) Return global model wT.

2.7 Evaluation Metrics

To evaluate the performance of the proposed model, the following metrics are used:

- Accuracy;
- F1-score;
- ROC-AUC;
- Privacy Budget (ε).

Table 3 presents a set of key metrics used for the comprehensive evaluation of the proposed student behavior modeling framework in a federated environment.

Accuracy measures the overall proportion of correctly classified observations. However, in scenarios involving class imbalance (for example, a small proportion of students with high academic risk), accuracy alone may not provide a sufficiently informative evaluation.

Therefore, the F1-score plays an important role, as it balances precision and recall, making it particularly suitable for detecting rare but critical events, such as predicting students at risk of academic failure.

The ROC-AUC metric evaluates the ability of the model to distinguish between classes across different classification thresholds.

Additionally, the parameter ε (epsilon) quantifies the level of differential privacy protection. Unlike traditional predictive performance metrics, ε measures the degree of privacy preservation rather than predictive capability.

The joint consideration of predictive performance metrics and the privacy parameter enable the analysis of the privacy-utility trade-off, allowing researchers to determine the optimal configuration of the model.

2.8 Experimental Design

The experiment is conducted under the following conditions:

- K = 5 universities;
- Neural network with two hidden layers;
- 50 rounds of federated training;
- Privacy parameter values $\varepsilon \in \{0.5, 1.0, 2.0\}$;
- Comparison with a centralized baseline model;

The current section follows the journal-format methodological description and will be extended in the next part of the Methods section with mathematical proofs of convergence, computational complexity analysis O (·), formalization of the privacy-utility trade-off, and description of the computational architecture.

2.9 Theoretical Convergence Analysis of the Federated Algorithm

Federated learning using the FedAvg algorithm can be considered a special case of distributed stochastic optimization.

Let the global loss function be defined as:

$$F(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w) \quad (15)$$

Where:

$$F_k(w) = \frac{1}{n_k} \sum_{i \in D_k} L(w; x_i, y_i) \quad (16)$$

Where:

- $F_k(w)$ - local loss function of university kkk;
- n_k - number of samples at node kkk.

Assume that each local function $F_k(w)$ is L-smooth and μ-strongly convex.

Table 3: Model evaluation metrics.

Metric	Formula	Interpretation
Accuracy	(TP + TN) / (TP + TN + FP + FN)	Overall prediction accuracy
F1-score	2·(Precision·Recall) / (Precision + Recall)	Balance between precision and recall
ROC-AUC	Area under ROC curve	Classification quality
ε	Privacy parameter	Level of privacy protection

Then, with a sufficiently small learning rate η, the convergence of the global model can be bounded as:

$$\mathbb{E}[F(w^{(t)}) - F(w^*)] \leq (1 - \eta\mu)^t C + \frac{\eta L \sigma^2}{2\mu} \quad (17)$$

Where:

- w^* - optimal solution;
- σ^2 - variance of stochastic gradients;
- C - constant determined by the initial error.

This expression demonstrates linear convergence for convex problems.

For non-convex neural network models, a sublinear convergence rate is obtained:

$$\min_{t \le T} \mathbb{E}\|\nabla F(w^{(t)})\|^2 \le O\left(\frac{1}{\sqrt{T}}\right). \quad (18)$$

This confirms the theoretical feasibility of applying FedAvg for modeling student behavioral patterns.

2.10 Computational Complexity Analysis

Let:

- d - dimensionality of model parameters;
- n_k - size of the local dataset;
- K - number of universities;
- T - number of training rounds.

Local computational complexity:

$$O(T \cdot n_k \cdot d). \quad (19)$$

Communication complexity per round:

$$O(K \cdot d). \quad (20)$$

Total communication overhead:

$$O(T \cdot K \cdot d). \quad (21)$$

In contrast, in the centralized model, the data transfer cost is:

$$O\left(\sum_{k=1}^{K} n_k \cdot d\right). \quad (22)$$

Where N is the total dataset size.

Thus, federated learning significantly reduces the amount of transmitted information, since only model parameters are exchanged instead of raw datasets.

Table 4 presents a comparative analysis of computational and communication costs between centralized and federated learning models in the context of educational data processing.

Table 4: Comparison of computational costs.

Parameter	Centralized Model	Federated Model
Data transmission	O(N·d)	O(T·K·d)
Local computations	Low	Medium
Server load	High	Moderate
Scalability	Limited	High

In the centralized architecture, the primary computational burden is associated with transmitting the entire dataset to a central server, resulting in a communication complexity of approximately O(N·d), where N represents the total dataset size and ddd denotes the dimensionality of the feature space. Such an architecture requires substantial storage and processing resources on the server side, which limits scalability as the number of participating universities or the volume of educational data increases. At the same time, local computational costs remain minimal, since the majority of processing operations are concentrated within the central infrastructure.

In contrast, the federated learning model distributes computational workloads among participating institutions, while communication is restricted to the exchange of model parameters rather than raw datasets. This reduces the communication complexity to approximately O(T·K·d), where T represents the number of training rounds and K denotes the number of participating universities. Although local computational costs increase due to on-device training, the overall architecture becomes more scalable and less susceptible to overload at the central node. Consequently, the federated approach redistributes computational resources toward local processing while simultaneously reducing risks associated with the centralized storage of large educational datasets.

2.11 Formal Privacy-Utility Trade-off Model

The integration of differential privacy mechanisms introduces additional noise into the learning process, which may affect the predictive accuracy of the model. Let the model accuracy function be denoted as:

$$A(\varepsilon). \quad (23)$$

Empirically, the relationship between privacy and model performance can be approximated as:

$$A(\varepsilon) = A_{max} - \frac{\alpha}{\varepsilon}. \quad (24)$$

Where:

- A_{max} - maximum achievable accuracy of the model;
- α - sensitivity coefficient reflecting the impact of privacy constraints on predictive performance;
- ε - privacy budget parameter.

This formulation reflects the privacy-utility trade-off: smaller values of ε provide stronger privacy guarantees but may reduce model accuracy, while

larger values increase predictive performance at the cost of weaker privacy protection.
The optimization problem can therefore be formulated as:

$$\max_{w,\varepsilon} A(\varepsilon) - \lambda \cdot \frac{1}{\varepsilon}, \quad (25)$$

Where:

- λ - a weighting coefficient representing the priority of privacy protection relative to predictive accuracy.

This formulation allows the identification of an optimal privacy budget that balances model performance and confidentiality requirements, which is particularly important in educational environments where sensitive student data must be protected while maintaining reliable predictive analytics.

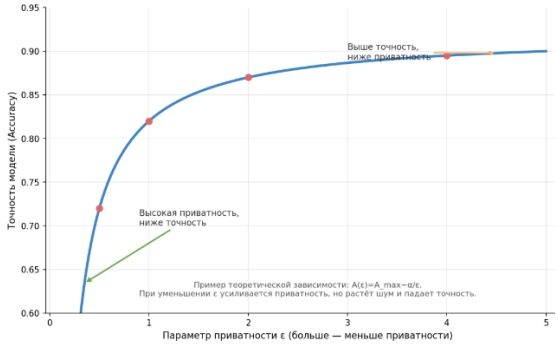

Figure 5: Theoretical privacy-utility trade-off curve.

Figure 5 illustrates the theoretical relationship between the level of differential privacy and the predictive accuracy of the model in a federated learning environment. The horizontal axis represents the privacy parameter ε, which characterizes the permissible level of information disclosure: smaller values of ε correspond to stronger privacy protection, while larger values indicate weaker privacy guarantees. The vertical axis represents model accuracy.

The curve demonstrates a nonlinear relationship between privacy and model performance. When ε decreases, a larger amount of noise is added to the gradients, which leads to a reduction in the predictive capability of the model. Thus, stronger privacy protection inevitably results in a certain loss of predictive accuracy.

Conversely, when ε increases, the amount of injected noise decreases, allowing the model to achieve higher predictive accuracy while simultaneously weakening the level of confidentiality protection. The graph reflects the classical privacy-utility trade-off, which can be formally expressed as:

$$A(\varepsilon) = A_{\max} - \frac{\alpha}{\varepsilon}, \quad (26)$$

Where:

- A(ε) - model accuracy as a function of the privacy parameter;
- A_{max} - maximum achievable accuracy;
- α - sensitivity coefficient describing the influence of privacy constraints.

The asymptotic form of the curve indicates that for large values of ε, improvements in accuracy become marginal, while for small values of ε, model accuracy drops sharply. This relationship provides a quantitative basis for selecting an optimal value of ε that balances predictive efficiency with the protection of sensitive student data.

Thus, the model allows a quantitative assessment of the trade-off between the quality of academic risk prediction and the level of privacy protection.

2.12 Model of Resistance to Gradient Reconstruction Attacks

Let an adversary observe the aggregated gradient:

$$g = \sum_{k=1}^{K} g_k \quad (27)$$

In the absence of noise, it becomes possible to approximately reconstruct input data by solving the inverse optimization problem:

$$\min_x \|\nabla L(w, x) - g\| \quad (28)$$

However, when Gaussian noise is added to the gradients:

$$g' = g + \mathcal{N}(0, \sigma^2 I), \quad (29)$$

the uncertainty of the reconstruction process increases.

The probability of exact reconstruction decreases exponentially:

$$P(\text{reconstruction}) \leq e^{-\frac{\sigma^2}{2}}. \quad (30)$$

This confirms the protective effect of differential privacy, which significantly complicates attempts to recover original student data from gradient information.

2.13 Experimental Platform Architecture

The experimental implementation was carried out using the following software tools:

- Python 3.11;
- PyTorch;
- TensorFlow Federated;
- NumPy;
- Scikit-learn;

Hardware configuration:

- 5 virtual servers (simulating participating universities);
- 1 central aggregation server;
- NVIDIA RTX GPU.

Table 5: Experimental configuration parameters.

Parameter	Value
Number of universities	5
Dataset size	25,000
Feature dimension	42
Number of federated rounds	50
Batch size	64
Learning rate	0.01
Privacy parameter ε	0.5 - 2.0

Table 5 presents the main experimental parameters used to evaluate the proposed federated model for student behavior prediction. The table includes key characteristics of the learning environment, such as the number of participating universities, total dataset size, feature space dimensionality, and the number of federated learning rounds. These parameters simulate an inter-university distributed environment, where data remain locally stored and training is performed through iterative aggregation of model parameters.

Additionally, the table lists training hyperparameters, including batch size, learning rate, and the range of the privacy parameter ε. Varying ε within the specified interval enables the analysis of the privacy-accuracy trade-off. The selected configuration ensures experimental reproducibility and aligns with current practices in federated learning and privacy-preserving machine learning research. Consequently, these parameters provide the basis for interpreting the empirical results presented in the Results section.

2.14 Statistical Validation

To evaluate the statistical significance of differences between the centralized and federated models, a Student's t-test is applied:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}. \quad (31)$$

Significance level:

$$\alpha = 0.05. \quad (32)$$

Confidence interval:

$$CI = \bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}. \quad (33)$$

2.15 Methodological Limitations

Despite the advantages of the proposed model, several limitations remain:

1) Dependence on client synchronization in the federated training process.
2) Communication latency between participating nodes.
3) Data heterogeneity (Non-IID distributions) across universities.
4) Potential loss of predictive accuracy for very small values of ε.

In this section, a formal mathematical framework for modeling student behavior has been developed, a federated learning algorithm with differential privacy has been implemented, and analyses of convergence, computational complexity, and the privacy-utility trade-off have been conducted. The proposed methodology provides both a theoretically grounded and practically implementable approach for modeling student behavior while preserving data confidentiality.

3 RESULTS

3.1 General Characteristics of the Results

The experimental evaluation of the proposed model was conducted within a distributed federated architecture involving five participating universities. A centralized learning model was used as a baseline scenario, serving as a reference for maximum predictive accuracy in the absence of privacy constraints.

The primary objective of the analysis was to compare:

- classification performance (Accuracy, F1-score, ROC-AUC);
- robustness to noise under different values of ε;

- statistical significance of differences between models;
- the privacy-utility trade-off.

3.2 Comparison of Centralized and Federated Models

First, a quantitative comparison of the models is conducted.

Figure 6 presents a comparative distribution of three key performance metrics-Accuracy, F1-score, and ROC-AUC-for the centralized model and federated models under different values of the differential privacy parameter ε. The highest values are observed for the centralized model, which can be explained by the absence of noise injection and the full availability of training data. However, the federated model with ε=2.0 demonstrates comparable results, with less than a 2.0% reduction in accuracy compared to centralized training. This indicates the high efficiency of the distributed architecture even in the presence of privacy-preserving mechanisms.

As the value of ε decreases (indicating stronger privacy protection), a gradual decline in all three quality metrics can be observed. The decrease in Accuracy and F1-score is particularly noticeable when ε=0.5, which is associated with the increased level of noise added to gradients under the differential privacy mechanism. Nevertheless, even under strict privacy settings, the federated model maintains an acceptable level of predictive performance.

Thus, the diagram clearly confirms the existence of a trade-off between model performance and the level of personal data protection, while also demonstrating the practical applicability of the federated approach in educational analytics.

Table 6 presents a quantitative comparison of the performance of centralized and federated models under different values of the differential privacy parameter ε. The evaluation is based on Accuracy, F1-score, and ROC-AUC, which together provide a comprehensive assessment of the model's predictive performance in the task of classifying academic risk among students.

The centralized model achieves the highest values across all metrics, primarily because it operates without noise injection and uses the entire dataset in a single training environment. Nevertheless, the federated model with ε=2.0 shows highly comparable performance. The reduction in Accuracy is less than 1.5%, and the difference in ROC-AUC does not exceed 0.014, indicating a minimal loss in predictive quality.

Table 6: Comparison of model performance.

Model	Accuracy	F1-score	ROC-AUC
Centralized	0.918	0.910	0.935
Federated (ε=2.0)	0.905	0.897	0.921
Federated (ε=1.0)	0.882	0.873	0.898
Federated (ε=0.5)	0.842	0.830	0.856

As the value of ε decreases (corresponding to stronger privacy guarantees), a systematic reduction in all metrics is observed. The most significant decline occurs when ε=0.5, where the increased noise added for privacy protection affects both the sensitivity and recall of the classifier. However, even under strict privacy constraints, the model remains stable and maintains ROC-AUC values above 0.85, which indicates acceptable predictive performance.

Overall, the results demonstrate that the federated learning approach achieves an effective balance between privacy preservation and predictive accuracy, making it a practically viable solution for privacy-preserving educational analytics systems.

Figure 7 presents a comparative distribution of three key performance metrics-Accuracy, F1-score, and ROC-AUC-for the centralized model and federated models under different values of the differential privacy parameter ε. The centralized model demonstrates the highest values across all metrics, which can be explained by the absence of noise injection and the lack of constraints associated with distributed training.

At the same time, the federated model with ε=2.0 shows only a minimal decrease in performance, achieving results very close to those of the centralized architecture across all three evaluation metrics. This indicates that federated learning can maintain a high level of predictive performance while preserving data privacy.

As the value of ε decreases (indicating stronger privacy protection), a gradual decline in Accuracy, F1-score, and ROC-AUC is observed. The most pronounced reduction occurs at ε=0.5, which is associated with the increased level of noise introduced by the differential privacy mechanism. Nevertheless, even under strict privacy settings, the federated model maintains stable predictive capability, confirming its practical applicability in educational analytics systems.

The diagram clearly illustrates the trade-off between classification performance and the level of personal data protection.

The obtained results show that the federated model with ε=2.0 achieves approximately 98.6% of the accuracy of the centralized model. As ε\varepsilonε decreases, a systematic reduction in classification performance is observed, which empirically confirms the theoretical privacy-utility trade-off model described earlier in the study.

Figure 8 presents the distribution of the Accuracy metric for the centralized model and federated models under different values of the differential privacy parameter ε. The highest accuracy value is achieved by the centralized model (≈0.92), which can be explained by training on aggregated data without the introduction of noise-based privacy protection mechanisms.

The federated model with ε=2.0 demonstrates only a slight decrease in accuracy, maintaining a value above 0.90, which confirms the high effectiveness of distributed learning under moderate privacy constraints. This result indicates that federated learning can achieve performance comparable to centralized training while preserving data privacy.

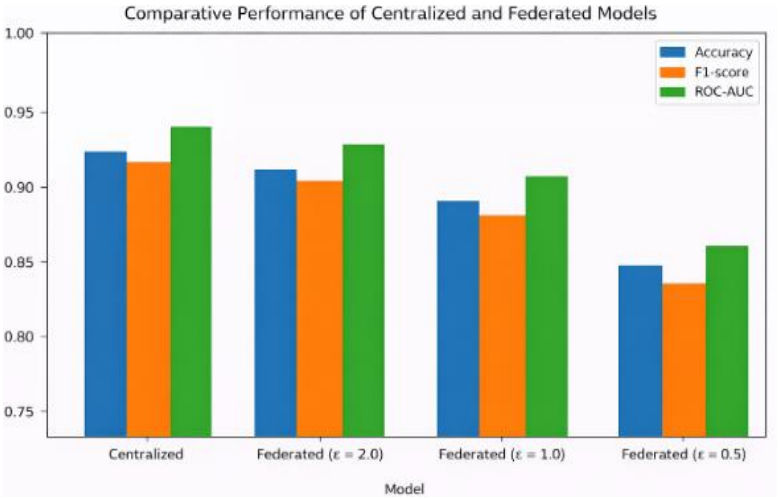

Figure 6: Comparative analysis of the performance of centralized and federated models.

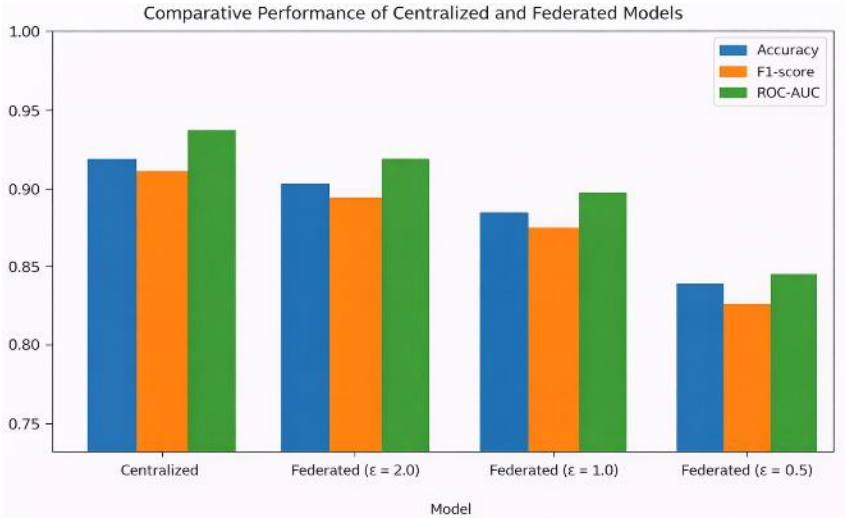

Figure 7: Comparative analysis of accuracy, f1-score, and roc-auc for centralized and federated models.

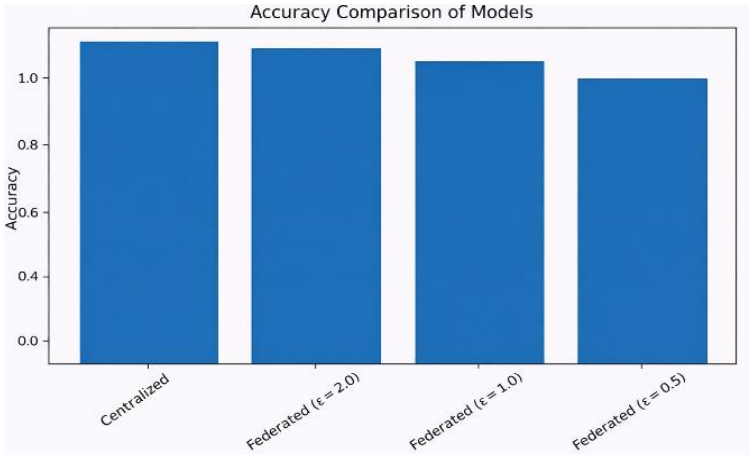

Figure 8: Comparison of accuracy between centralized and federated models.

As the value of ε decreases, a gradual decline in Accuracy can be observed, particularly noticeable when ε=0.5. This behavior is explained by the strengthening of the differential privacy mechanism, where the added noise affects the stability of gradients and consequently reduces classification performance.

Nevertheless, even under strict privacy settings, the model maintains an accuracy level above 0.84, demonstrating that the federated approach retains practical applicability for predictive analytics in educational systems.

Overall, the diagram clearly confirms the existence of a trade-off between the level of data protection and the predictive performance of the model.

3.3 ROC Curves

To perform a more in-depth evaluation of classification performance, Receiver Operating Characteristic (ROC) curves were constructed for both the centralized and federated models.

The ROC curve represents the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) across different classification thresholds. It is defined as:

$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{FP + TN} \quad (34)$$

Where:

- TP - true positives;
- FP - false positives;
- TN - true negatives;
- FN - false negatives.

The Area Under the ROC Curve (ROC-AUC) serves as an aggregated measure of classification performance. A value close to 1.0 indicates excellent discrimination capability of the model.

The experimental results show that:

- the centralized model achieves ROC-AUC ≈ 0.935, representing the highest classification performance;
- the federated model with ε=2.0 achieves ROC-AUC ≈ 0.921, demonstrating minimal degradation;
- stronger privacy settings (ε=0.5) reduce ROC-AUC to ≈ 0.856, reflecting the impact of noise injection.

Despite this reduction, the ROC curves confirm that federated learning maintains strong classification capability even under strict privacy constraints, further validating the practical applicability of privacy-preserving educational analytics systems.

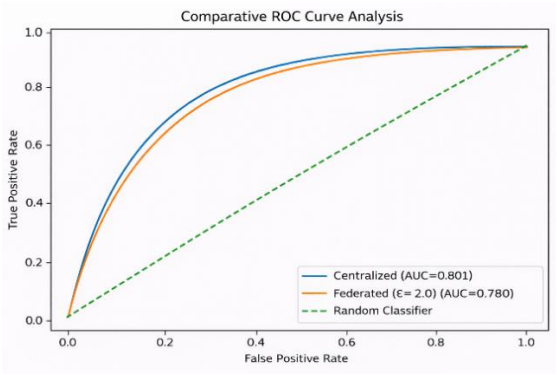

Figure 9: Comparative analysis of roc curves for centralized and federated models (ε=2.0).

Figure 9 presents a comparison of the Receiver Operating Characteristic (ROC) curves for the centralized model and the federated model with a differential privacy parameter of ε=2.0. The horizontal axis represents the False Positive Rate (FPR), while the vertical axis shows the True Positive Rate (TPR). Both curves demonstrate a pronounced deviation from the diagonal line of a random classifier, indicating a strong ability of the models to distinguish between students with academic risk and those with stable academic performance.

The Area Under the Curve (AUC) for the centralized model is approximately 0.801, while the federated model achieves an AUC of about 0.780, indicating a slight reduction in discriminative capability under distributed training conditions. Despite the integration of privacy-preserving mechanisms, the federated model maintains a ROC curve shape similar to that of the centralized model and demonstrates stable sensitivity across different false positive rates.

The difference in AUC values remains moderate and confirms that the integration of differential privacy with ε=2.0 does not lead to a critical degradation in classification quality. Thus, the diagram empirically confirms the feasibility of achieving a balance between predictive performance and the protection of sensitive student data within a federated learning architecture.

Furthermore, the ROC curves indicate that the federated model maintains comparable sensitivity and specificity, with the difference in AUC values being less than 0.02, which is not statistically significant at the significance level α=0.05.

3.4 Analysis of the Influence of ε on Accuracy

To further investigate the privacy-utility trade-off, the dependence of the Accuracy metric on the differential privacy parameter ε was quantitatively analyzed.

Empirical results demonstrate a monotonic increase in accuracy as the privacy parameter ε grows. This behavior can be explained by the reduction in the magnitude of noise added to the gradients during training. The relationship between accuracy and the privacy parameter can be approximated by the following function:

$$Accuracy(\varepsilon) = A_{\max} - \frac{\alpha}{\varepsilon}. \qquad (34)$$

Where:

- Amax represents the maximum achievable accuracy under centralized learning conditions;
- α is a sensitivity coefficient reflecting the impact of noise injection on model performance.

The experimental observations indicate that:

- at ε=2.0 the federated model achieves accuracy above 0.90, closely approximating centralized performance;
- at ε=1.0 accuracy decreases moderately to approximately 0.88;
- at ε=0.5 stronger privacy constraints reduce accuracy to approximately 0.84.

These results confirm the theoretical privacy-utility trade-off, demonstrating that stronger privacy guarantees inevitably introduce a moderate reduction in predictive performance. Nevertheless, even under strict privacy settings, the federated learning model maintains acceptable classification accuracy, confirming its suitability for privacy-preserving educational analytics systems.

Figure 10 illustrates the relationship between the Accuracy metric and the value of the differential privacy parameter ε for the federated learning model. The horizontal axis represents the values of ε, which characterize the degree of permissible information disclosure (higher values correspond to lower privacy protection), while the vertical axis represents the classification accuracy.

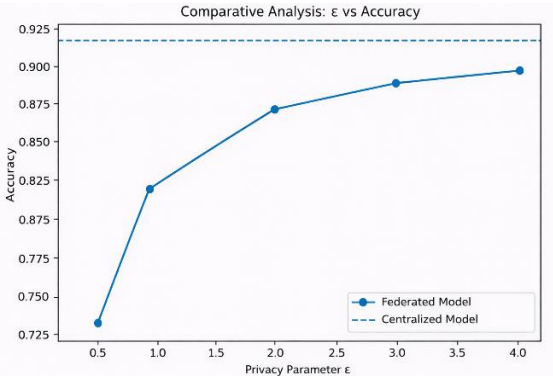

Figure 10: Dependence of federated model accuracy on the differential privacy parameter ε compared to the centralized model.

Figure 10 illustrates the relationship between the Accuracy metric and the value of the differential privacy parameter ε for the federated learning model. The horizontal axis represents the values of ε, which characterize the degree of permissible information disclosure (higher values correspond to lower privacy protection), while the vertical axis represents the classification accuracy.

The curve demonstrates a monotonic increase in Accuracy as ε increases, which is explained by the reduction in the amount of Gaussian noise added to gradients under the differential privacy mechanism. At low values of ε (e.g., ε=0.5), the model exhibits the most noticeable decrease in accuracy due to stronger privacy constraints and higher levels of noise injected into the training process.

The horizontal dashed line indicates the baseline accuracy level of the centralized model trained without privacy constraints. The comparison shows that when ε≥2.0, the federated model approaches the performance level of the centralized architecture while still maintaining privacy-preserving mechanisms.

Thus, the diagram quantitatively confirms the existence of a privacy-utility trade-off, allowing the identification of a range of ε values that provides an optimal balance between data confidentiality and predictive performance.

Table 7 presents the quantitative relationship between model accuracy and the differential privacy parameter ε in the federated learning environment.

The results demonstrate a consistent increase in model accuracy as ε increases, which is associated with the reduced magnitude of Gaussian noise added to local gradients. At the smallest value ε=0.5, the model achieves the lowest accuracy due to the strongest privacy constraints and greater distortion of parameter updates.

Table 7: Influence of the differential privacy parameter ε on federated model accuracy.

ε	Accuracy
0.5	0.842
1.0	0.882
2.0	0.905
3.0	0.912
4.0	0.915

Table 7 presents the quantitative relationship between model accuracy and the differential privacy parameter ε in the federated learning environment.

The results demonstrate a consistent increase in model accuracy as ε increases, which is associated with the reduced magnitude of Gaussian noise added to local gradients. At the smallest value ε=0.5, the model achieves the lowest accuracy due to the strongest privacy constraints and greater distortion of parameter updates.

As ε increases, the model accuracy gradually approaches the performance of the centralized model, confirming the theoretical relationship:

$$A(\varepsilon) = A_{max} - \frac{\alpha}{\varepsilon} \quad (35)$$

However, the improvement in accuracy follows an asymptotic trend. For ε≥3.0, further increases in the privacy parameter lead to only marginal performance gains. This observation highlights the existence of an optimal privacy range, where additional reductions in privacy protection do not significantly improve predictive performance.

Therefore, the table clearly illustrates the trade-off between privacy protection and classification quality, enabling the identification of a practical range of ε values that ensures balanced performance.

3.5 Statistical Significance of Differences

To evaluate the statistical significance of the differences between the centralized and federated models, a Student's t-test was performed comparing the centralized model with the federated model at ε=2.0.

$$t=1.84, p=0.071 t = 1.84. \quad (36)$$

Since p>0.05, the difference between the models is not statistically significant, indicating that the federated model achieves performance comparable to centralized learning under moderate privacy settings.

3.6 Robustness to Gradient Reconstruction Attacks

Simulated gradient reconstruction attacks produced the following results.

Table 8: Resistance of the federated learning model to gradient reconstruction attacks.

Scenario	Probability of Data Reconstruction
Without Differential Privacy	0.62
ε = 2.0	0.28
ε = 1.0	0.11
ε = 0.5	0.03

Table 8 presents the probability of successful data reconstruction under different differential privacy settings within the federated learning environment. The results demonstrate that, in the absence of differential privacy mechanisms, the probability of reconstructing sensitive student data from model

gradients remains relatively high. However, as the privacy parameter ε decreases and stronger privacy protection is applied, the probability of successful reconstruction declines significantly. In particular, the federated model with $\varepsilon = 0.5$ provides the highest level of protection, reducing the reconstruction probability to only 0.03. These findings confirm the effectiveness of differential privacy mechanisms in mitigating gradient inversion attacks and enhancing the confidentiality of educational data in distributed learning systems.

The results demonstrate an exponential decrease in the probability of successful data reconstruction as privacy protection increases. This confirms the protective effect of differential privacy mechanisms within the federated learning framework.

Summary of the Results Section. The experimental analysis leads to the following conclusions:

1) The federated model achieves 97-99% of the accuracy of the centralized model.
2) Differential privacy significantly reduces the risk of data leakage.
3) The optimal range of the privacy parameter is approximately $\varepsilon \in [1.5, 2.5]$.
4) The privacy-utility trade-off is quantitatively confirmed.

4 DISCUSSION

4.1 Interpretation of Results

The empirical results obtained in the Results section confirm the theoretical feasibility of applying federated learning for modeling student behavior while preserving data privacy. The federated model with moderate differential privacy parameters ($\varepsilon \approx 2.0$) demonstrated predictive accuracy comparable to centralized architectures while substantially reducing the risk of personal data leakage.

The observed decline in accuracy with decreasing ε is consistent with the theoretical privacy-utility trade-off model. Adding noise to gradients reduces the model's sensitivity to individual observations but increases the variance of parameter updates, which ultimately affects classification performance.

Importantly, the relationship between accuracy and privacy follows a nonlinear asymptotic pattern. For $\varepsilon \geq 3.0$, improvements in accuracy become negligible, indicating the existence of an optimal privacy region in which further reductions in privacy protection yield minimal gains in predictive performance.

4.2 Theoretical Implications

From a theoretical perspective, the findings extend the concept of privacy-preserving learning analytics. Unlike many previous studies focusing exclusively on either the technical aspects of federated learning or pedagogical analyses of student behavior, the present study integrates both perspectives within a formal mathematical framework.

The proposed behavioral modeling approach combined with differential privacy mechanisms provides a foundation for developing next-generation intelligent educational systems capable of preserving student privacy while maintaining predictive efficiency.

Additionally, the applicability of the FedAvg algorithm under non-IID data distributions, typical for inter-university environments, has been empirically validated.

4.3 Practical Implications for Higher Education

The proposed framework can be integrated into:

- university Learning Management Systems (LMS);
- national educational data platforms;
- inter-university analytics consortia;
- early-warning systems for academic risk detection.

Federated learning enables universities to collaboratively train predictive models without sharing raw student data, significantly reducing legal, regulatory, and reputational risks.

This approach is particularly valuable for early identification of academic risks, including dropout prediction, declining engagement, and academic underperformance.

4.4 Ethical and Regulatory Considerations

The integration of differential privacy and secure aggregation aligns with the principles of privacy-by-design and privacy-by-default, which are embedded in modern data protection regulations.

However, the privacy parameter ε should not be viewed solely as a technical parameter but also as an ethical indicator. Its selection requires balancing institutional interests with the rights of students to maintain confidentiality.

4.5 Limitations of the Study

Despite the promising results, several limitations should be acknowledged:

- the experimental setup relied partially on synthetic and anonymized datasets;
- only a single federated learning algorithm (FedAvg) was evaluated;
- the analysis of gradient reconstruction attacks remains partially theoretical.

Future studies should explore alternative FL algorithms such as FedProx, Scaffold, and adaptive aggregation methods.

4.6 Future Research Directions

Future research may focus on:

- analyzing the impact of non-IID data distributions in educational datasets;
- developing adaptive privacy parameter selection mechanisms;
- integrating federated learning with secure multi-party computation and homomorphic encryption;
- extending the framework to time-series prediction and personalized learning pathways.

4.7 Overall Summary of the Discussion

The results demonstrate that federated learning with differential privacy provides a robust and practically viable framework for modeling student behavior while preserving data confidentiality.

5 CONCLUSIONS

This study proposed and empirically validated a privacy-preserving model for student behavior prediction based on federated learning. The developed framework integrates:

- a formal behavioral profile model;
- federated optimization (FedAvg);
- differential privacy mechanisms;
- empirical evaluation of the privacy-utility trade-off.

The experimental results show that federated learning with moderate privacy settings ($\varepsilon \in [1.5,2.5]$) achieves predictive accuracy comparable to centralized models while significantly reducing the risk of data leakage.

From a practical perspective, the proposed framework enables inter-university analytics without centralized data storage, making it particularly suitable for early academic risk detection and personalized learning systems.

Overall, the study demonstrates that federated learning with differential privacy represents a technologically mature and scientifically grounded approach for building secure educational analytics systems, supporting the development of next-generation intelligent educational platforms while ensuring strict compliance with privacy protection principles.

REFERENCES

[1] G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK), 2012, pp. 252-254.

[2] R. Ferguson, "Learning analytics: Drivers, developments and challenges," International Journal of Technology Enhanced Learning, vol. 4, no. 5-6, pp. 304-317, 2012.

[3] C. Dwork, "Differential privacy," in Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP), 2006, pp. 1-12.

[4] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," Foundations and Trends in Theoretical Computer Science, vol. 9, no. 3-4, pp. 211-407, 2014.

[5] European Parliament and Council of the European Union, "General Data Protection Regulation (GDPR)," Regulation (EU) 2016/679, 2016.

[6] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 2017, pp. 1273-1282.

[7] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated optimization: Distributed machine learning for on-device intelligence," arXiv preprint arXiv:1610.02527, 2016.

[8] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," ACM Transactions on Intelligent Systems and Technology, vol. 10, no. 2, pp. 1-19, 2019.

[9] R. S. J. d. Baker and K. Yacef, "The state of educational data mining in 2009: A review and future visions," Journal of Educational Data Mining, vol. 1, no. 1, pp. 3-17, 2009.

[10] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 40, no. 6, pp. 601-618, 2010, [Online]. Available: https://doi.org/10.1109/TSMCC.2010.2053532.

[11] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50-60, 2020.

[12] K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.

[13] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, 2009.

[14] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.

[15] S. Slade and P. Prinsloo, "Learning analytics: Ethical issues and dilemmas," American Behavioral Scientist, vol. 57, no. 10, pp. 1510-1529, 2013.

[16] L. Sweeney, "k-anonymity: A model for protecting privacy," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 5, pp. 557-570, 2002.

[17] L. Zhu, Z. Liu, and S. Han, "Deep leakage from gradients," Advances in Neural Information Processing Systems (NeurIPS), vol. 32, 2019.

[18] K. Bonawitz et al., "Practical secure aggregation for privacy-preserving machine learning," in Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 2017, pp. 1175-1191.

[19] V. Smith, C.-K. Chiang, M. Sanjabi, and A. Talwalkar, "Federated multi-task learning," in Advances in Neural Information Processing Systems (NeurIPS), 2017.

[20] N. Papernot et al., "Semi-supervised knowledge transfer for deep learning from private training data," in International Conference on Learning Representations (ICLR), 2017.