Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

A Federated Learning Neural Network For Student Dropout Prediction

Anuradha Kumari SINGH[a*] & Karthikeyan S[b]
[a]*Department of Computer Science, Banaras Hindu University, India*
[b]*Department of Computer Science, Banaras Hindu University, India*
*anuradha@bhu.ac.in

Abstract: Federated learning offers a promising solution for training neural networks on distributed datasets while reducing the risk of data leakage, a critical concern in educational data mining where student records are highly sensitive. We propose FedAvg-NNFL, a privacy-preserving neural network for student dropout prediction within a federated framework. Each institution trains a local model on its private data, and a global model is created by aggregating locally trained weights, ensuring privacy and compliance with data-sharing restrictions. Using a benchmark dataset from the Polytechnic Institute of Portalegre, containing demographic, socioeconomic, academic, and macroeconomic features, we compared FedAvg-NNFL against centralized and locally trained models. FedAvg-NNFL achieved accuracy 0.9280, precision 0.9320, recall 0.8972, F1-score 0.9142, and AUC 0.92 under 10-fold cross-validation, outperforming the local model and closely matching centralized performance. These results demonstrate the potential of federated learning for developing accurate, privacy-aware predictive models in educational contexts.

Keywords: Federated Learning, Neural Network, Student Dropout, Predictive model

1. Introduction

Student dropout prediction (SDP) plays a vital role in identifying at-risk students and enabling proactive interventions that enhance retention, institutional performance, and long-term learner success. Leveraging machine learning and optimization techniques, universities can analyze diverse factors including attendance, engagement, socio-economic background, and academic performance to uncover patterns associated with dropout. However, developing a centralized SDP model remains challenging due to the difficulty of data collection and aggregating data across departments and institutions, strict privacy and compliance requirements, variability in data quality and formats, and the scalability demands posed by increasing data volumes.

In the literature review, we found very few studies Fachola et al. (2023); Farooq et al. (2024); Johany´ak, Laufer, and Kov´acs (2024); Kuang et al. (2025); van Haastrecht, Brinkhuis, and Spruit (2024) that employed a federated learning (FL) approach for training a global model for student dropout prediction. To bridge this gap, we developed a federated learning model based on neural networks (FedAvg-NNFL) for student dropout prediction in universities, effectively addressing the challenge of maintaining data privacy.

This study has the following main contributions:

- Collecting student data from the universities is creating data privacy concerns as student records contain sensitive information; to solve this problem, we proposed the use of a federated learning framework which is not explored more in educational data mining.
- Neural network-based federated learning model, FedAvg-NNFL is proposed to predict the student's dropout at university.
- We compared our FedAvg-NNFL model performance with the centralized and locally trained models, and our model outperformed.

The remainder of this article is organized as follows: Section 2 presents a concise overview of recent research on student dropout prediction and the use of federated learning. Section 3 details the proposed methodology and algorithms. Section 4 discusses the experimental results and performance evaluation of the proposed model. Finally, Section 5 highlights the key findings and contributions of the study.

2. Literature Review

Most of the studies have been done on a centralized model, and very few studies have used a decentralized model in educational data mining. The Researcher utilized a federated learning approach to identify students' unique learning patterns and individualized variations across different courses. Another research applied federated learning methods to address two learning analytics challenges: dropout prediction and unsupervised student classification Fachola et al. (2023). The researcher explored the potential of combining federated learning with strategic client selection and sampling techniques to enhance dropout prediction in healthcare applications Nikolaidis and Efraimidis (2025). Researchers introduced a new federated learning approach, DQFed, for learning analytics that consists of aggregating the local models based on their weights computed according to a quality-driven model Bernardi, Cimitile, and Usman (2024).

Another study explored the technical foundations of federated learning, its application in education systems, and its potential benefits in improving learning outcomes and fostering data-driven decision-making in education Surapaneni, Bojjagani, and Sharma (2024). Researchers introduced a new federated learning standard that safeguards dataset confidentiality and enables the prediction of student grades, classified into four categories: low, good, beverage, and drop Chu et al. (2022).

There are some recent studies that have applied federated learning techniques to student performance and dropout prediction. Kuang et al. (2025) leveraged federated frameworks to investigate depressive symptoms among students, demonstrating FL's applicability beyond academic metrics. Farooq et al. (2024) proposed a privacy-preserving federated model for predicting student grades while ensuring data confidentiality. Van Haastrecht et al. (2024) conducted a comparative study between local, centralized, and federated learning approaches across multiple educational prediction tasks, including learning outcomes, dropout rates, and question correctness. Meanwhile, Johanyák et al. (2024) introduced a collaborative federated learning initiative between two universities focused on developing a fuzzy logic-based model for predicting student results. Collectively, these studies underscore the growing interest in using federated learning to address privacy concerns while enabling effective predictive modeling in educational environments.

3. Methodology

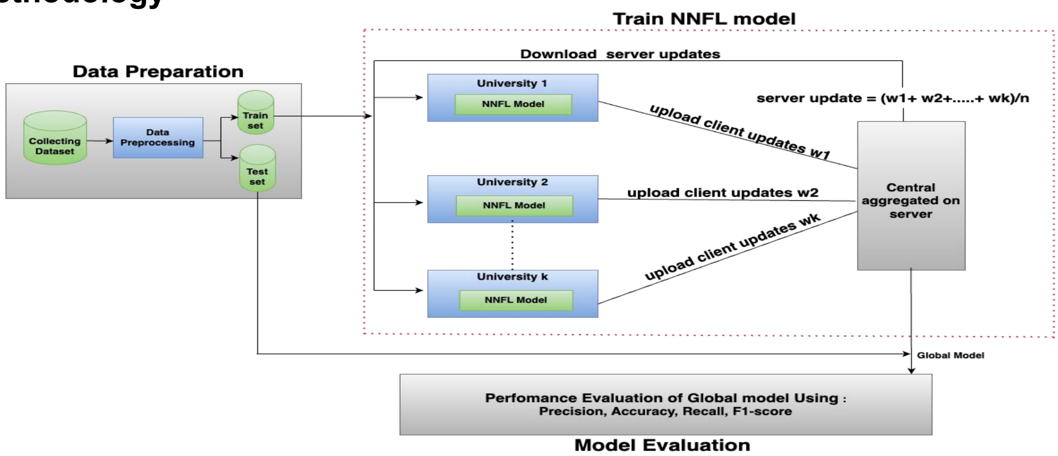

Figure 1. Overview of proposed methodology.

This section contains the discussion of the proposed methodology. Our proposed methodology follows three stages: 1. Data preparation, 2. Training of Federated model, 3. Performance evaluation. **Figure 1** depicts the workflow of our proposed methodology.

3.1 Proposed Algorithm

The Federated Averaging (FedAvg) , **algorithm 1** is a privacy-preserving training strategy that enables collaborative learning of a shared neural network model across multiple clients without centralizing their data. This approach is especially relevant in sensitive domains such as education, where sharing raw student data between institutions may violate privacy regulations.

Algorithm 1 Federated Averaging Neural Network Training (FedAvg-NNFL)

Require:

NUM_CLIENTS ← Number of clients

CLIENT_DATA ← $\{(X_i^{train}, X_i^{test}, y_i^{train}, y_i^{test})_{i=1}^{NUM\_CLIENTS}$

NUM_ROUNDS ← Number of federated training rounds

EPOCHS_PER_CLIENT ← Number of local epochs per round

BATCH_SIZE ← Batch size for local training

Ensure: Final global model weights WG, mean global accuracy and loss

Initialize global model G with weights W_G

for r = 1 to NUM_ROUNDS do

  Initialize empty list client weights ← []

  for i = 1 to NUM_CLIENTS do

    Create local model M_i and set weights $W_i \leftarrow W_G$

    Train Mi on $(X_i^{train}, y_i^{train})$ using EPOCHS_PER_CLIENT and BATCH_SIZE

    Append W_i to client weights

  end for

  Aggregate local weights using element-wise mean: $W_G = \frac{1}{NUM\_CLIENTS} \sum_{i=1}^{NUM\_CLIENTS} W_i$.

  Update global model G with weights W_G

  for i = 1 to NUM_CLIENTS do

    Evaluate G on $(X_i^{train}, y_i^{train})$ to obtain accuracy acc_i

  end for

  Compute and print mean training accuracy: $\text{Accuracy}_r = \frac{1}{NUM\_CLIENTS} \sum_{i=1}^{NUM\_CLIENTS} acc_i$

end for

Final Evaluation:

for i = 1 to NUM_CLIENTS do

  Evaluate G on (X_i^{test}, y_i^{test}) to get loss l_i and accuracy acc_i

end for

Compute:

Mean Accuracy: $\bar{A} = \frac{1}{NUM\_CLIENTS} \sum_{i=1}^{\text{NUM\_CLIENTS}} acc_i$

Mean Loss: $\bar{L} = \frac{1}{NUM\_CLIENTS} \sum_{i=1}^{\text{NUM\_CLIENTS}} l_i$

return Final weights W_G , $\bar{A}$, and $\bar{L}$

The Federated Averaging Neural Network Training (FedAvg-NNFL) algorithm begins with the initialization of a global neural network model G with random weights. This global model is iteratively refined over multiple federated training rounds based on contributions from distributed clients. During each round, every client receives a copy of the global model and initializes its local model with the same weights. The local model is then trained on the client's private dataset $(X_i^{train}, y_i^{train})$ using a fixed number of epochs and batch size. After training, only the updated local model weights W_i are sent back to the central server. The server then performs federated averaging by computing the element-wise mean of the collected weights:

$$W_G = \frac{1}{NUM\_CLIENTS} \sum_{i=1}^{NUM\_CLIENTS} W_i.$$

This aggregated weight set becomes the updated global model, effectively integrating knowledge from all clients while preserving their data privacy. Optionally, the global model may be evaluated on each client's training data to monitor intermediate performance, such as mean training accuracy. After completing all training rounds, the final global model is evaluated on the test datasets of all clients to measure its performance. For each client, the test accuracy and loss are recorded and averaged to compute the overall global performance:

$$\text{Mean Accuracy } \bar{A} = \frac{1}{NUM\_CLIENTS}\sum_{i=1}^{NUM\_CLIENTS} acc_i \text{ ;}$$
$$\text{Mean Loss } \bar{L} = \frac{1}{NUM\_CLIENTS}\sum_{i=1}^{NUM\_CLIENTS} l_i .$$

The algorithm ultimately returns the final global model weights W_G, which are suitable for inference or further fine-tuning, along with the average accuracy and loss across all clients, reflecting the overall effectiveness of the federated learning process.

3.2 Dataset and Data Preprocessing

We used the publicly available dataset from the Polytechnic Institute of Portalegre (Realinho et al., 2022), previously employed in our earlier study. The dataset (743 KB) contains 4,424 records and 37 attributes covering demographic, socioeconomic, macroeconomic, enrollment, and first- and second-semester academic data. Originally, it had three target classes dropout, graduate, and enrolled. For this study, we excluded the enrolled category (students whose final status was unknown), leaving 3,630 records with two definitive outcomes: dropout and graduate. Preprocessing involved loading the data from Excel, removing rows with missing values, and separating features (X) from the target (y). Numerical features were standardized using StandardScaler. To address class imbalance, we applied the SMOTE–Tomek method, which oversamples the minority class and removes borderline samples. Finally, the data was split into 80% training (2,904 records) and 20% testing (858 records).

4. Results and Discussion

Experiments were conducted in Python using scikit-learn (v1.3.2) for preprocessing and TensorFlow for neural network implementation, training, and testing. All runs were performed on Google Colab with a T4 GPU. The proposed FedAvg-NNFL model combines neural networks with federated learning for student dropout prediction. The neural network comprises three dense layers: 16 ReLU-activated neurons (input: 36 features), 8 ReLU-activated neurons, and a final sigmoid-activated neuron for binary classification.

Federated learning enables collaboration among multiple parties to develop a model without directly sharing their data. This approach involves training the model on a decentralized network, where the data stays on local client devices or servers at various locations. In this study, we introduced a FedAvg-NNFL model. NUM_ROUNDS, EPOCHS_PER_CLIENTS, and BATCH_SIZE are the key hyperparameters affecting the local training process of each client. We trained our FedAvg-NNFL model with 50 clients, NUM_ROUNDS=10, EPOCHS_PER_CLIENTS=20, and BATCH_SIZE=64.

Figure 2 depicts the comparison between the local model and federated model FedAvg-NNFL training over different clients. The performance of local and federated training across epochs reveals distinct trends. In local training (represented by the blue line with circles), the accuracy starts relatively high at epoch 5 (0.923) but then drops to its lowest point by epoch 10 (0.911). It slightly improves at epoch 20 (0.912), reaches a peak again at epoch 30 (0.924), and then dips slightly by epoch 50 (0.916). This fluctuation indicates instability, likely caused by the model's sensitivity to local data distributions, suggesting possible overfitting or limited generalization.

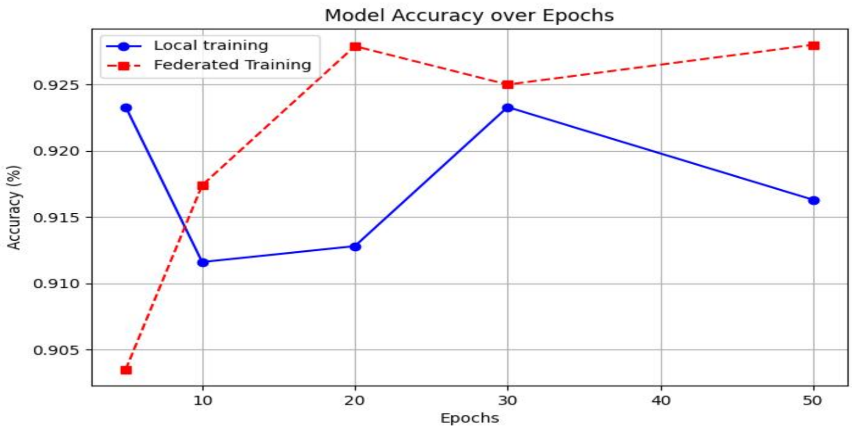

Figure 2. Comparison between Local model and federated model(FedAvg-NNFL) training over different clients at 10-fold cross-validation.

In contrast, federated training (depicted by the red line with squares) starts at a lower accuracy 0.904 at epoch 5 but improves steadily, reaching 0.917 at epoch 10 and peaking at 0.927 by epoch 20. Although there's a minor dip at epoch 30 (0.925), it regains peak accuracy 0.927 by epoch 50. This consistent improvement and stabilization demonstrate the robustness and generalization capabilities of federated learning, likely due to the aggregation of knowledge from diverse client data. Notably, federated training consistently outperforms local training from epoch 10 onward.

Table 1. *Comparison between centralized, local and federated model (FedAvg-NNFL) at 10-fold cross-validation*

Training Mode	Accuracy	Precision	Recall	F1-score	AUC
Centralized Training	0.9628	0.9083	0.9036	0.9026	0.96
Local Training	0.9175	0.9384	0.8909	0.9140	0.92
FedAvg-NNFL (federated)	0.9280	0.9320	0.8972	0.9142	0.92

Table 1 compares centralized, local, and federated (FedAvg-NNFL) training. Centralized training achieved the highest accuracy (0.9628) and AUC (0.96) under 10-fold cross-validation, as expected from full data access. The local model attained the highest precision (0.9384) but lower accuracy (0.9175) and recall (0.8909) due to client-specific training. FedAvg-NNFL offered a balanced performance, with accuracy 0.9280, precision 0.9320, recall 0.8972, F1-score 0.9142, and AUC 0.92, outperforming the local model in most metrics while preserving privacy and closely approximating centralized results.

Limitations and Future Work

While this study demonstrates the feasibility of federated neural networks for student dropout prediction, several limitations remain. First, our implementation uses a standard FedAvg algorithm without customization for handling non-IID data or client-specific behavior. Future work will explore client personalization, asynchronous updates, and adaptive aggregation to enhance model robustness. Second the use of signal dataset. We plan to incorporate Dirichlet-based non-IID partitioning to better assess generalizability. Additionally, fairness analysis across demographic subgroups and performance under class imbalance conditions will be introduced to ensure equitable model behavior. These enhancements will strengthen both the technical contribution and the practical utility of our approach in privacy-sensitive educational settings.

5. Conclusion

Ensuring student data privacy is a major challenge in educational data mining, as university records often contain highly sensitive information. To address this, we propose FedAvg-

NNFL, a privacy-preserving federated learning framework for predicting student dropout without sharing raw data. Each institution trains a local neural network on its private dataset, and a global model is constructed by aggregating locally trained weights, ensuring compliance with data-sharing restrictions. We evaluated FedAvg-NNFL on the benchmark dataset from the Polytechnic Institute of Portalegre, which includes macroeconomic, socioeconomic, demographic, admission, and academic performance attributes. Compared with a centralized neural network and a locally trained model, FedAvg-NNFL achieved balanced results: accuracy 0.9280, precision 0.9320, recall 0.8972, the highest F1-score of 0.9142, and an AUC of 0.92 under 10-fold cross-validation. The findings highlight federated learning's potential to deliver accurate, privacy-aware predictive models, enabling institutions to identify at-risk students and design timely interventions to improve retention and success.

Data Availability Statement

The data presented in this study are publicly available at
https://doi.org/10.5281/zenodo.5777339.

Acknowledgements

We would like to thank all the people who prepared and revised previous versions of this document.

References

Bernardi, M. L., Cimitile, M., & Usman, M. (2024). An Entropy-based Federated Learning approach for the student's dropout prediction. In *Book of Abstracts* (p. 48).

Chu, Y. W., Hosseinalipour, S., Tenorio, E., Cruz, L., Douglas, K., Lan, A., & Brinton, C. (2022, October). Mitigating biases in student performance prediction via attention-based personalized federated learning. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 3033-3042).

Fachola, C., Tornaría, A., Bermolen, P., Capdehourat, G., Etcheverry, L., & Fariello, M. I. (2023). Federated learning for data analytics in education. *Data*, *8*(2), 43.

Farooq, U., Naseem, S., Mahmood, T., Li, J., Rehman, A., Saba, T., & Mustafa, L. (2024). Transforming educational insights: strategic integration of federated learning for enhanced prediction of student learning outcomes. *The Journal of Supercomputing*, *80*(11), 16334-16367.

Johanyák, Z. C., Laufer, E., & Kovács, L. (2024, October). Fuzzy Logic-Based Preliminary Risk Assessment for a Supervised Federated Learning Solution in Predicting Student Results. In *2024 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)* (pp. 1-8). IEEE.

Kuang, Y., Liao, X., Jiang, Z., Gu, Y., Liu, B., Tan, C., ... & Li, K. (2025). Federated learning-based prediction of depression among adolescents across multiple districts in China. *Journal of Affective Disorders*, *369*, 625-632.

Nikolaidis, C. C., & Efraimidis, P. S. (2025). Advancing elderly social care dropout prediction with federated learning: client selection and imbalanced data management. *Cluster Computing*, *28*(2), 114.

Realinho, V., Machado, J., Baptista, L., & Martins, M. V. (2022). Predicting student dropout and academic success. *Data*, *7*(11), 146.

Surapaneni, P., Bojjagani, S., & Sharma, N. K. (2024, November). Federated Learning-based Big Data Analytics For The Education System. In *2024 International Conference on Intelligent Computing and Emerging Communication Technologies (ICEC)* (pp. 1-6). IEEE.

van Haastrecht, M., Brinkhuis, M., & Spruit, M. (2024, July). Federated learning analytics: Investigating the privacy-performance trade-off in machine learning for educational analytics. In *International Conference on Artificial Intelligence in Education* (pp. 62-74). Cham: Springer Nature Switzerland.