# **RESEARCH ARTICLE**

# Federated Learning-Outcome Prediction with Multi-layer Privacy Protection

Yupei ZHANG<sup>1,2</sup>, Yuxin LI<sup>1,2</sup>, Yifei WANG<sup>1,2</sup>, Shuangshuang WEI<sup>1,2</sup>, Yunan XU<sup>1,2</sup>, Xuequn SHANG(🖂)<sup>1,2</sup>

1 School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China

2 MIIT Lab of Big Data Storage and Management, Xi'an 710129, China

© Higher Education Press 2024

Learning-outcome prediction (LOP) is a long-Abstract standing and critical problem in educational routes. Many studies have contributed to developing effective models while often suffering from data shortage and weak generalization to various institutions due to the privacy-protection issue. To this end, this study proposes a distributed grade prediction model, dubbed FecMap, by exploiting the federated learning (FL) framework that preserves the private data of local clients and communicates with others through a global generalized model. FecMap considers local subspace learning (LSL), which explicitly learns the local features against the global features, and multi-layer privacy protection (MPP), which hierarchically protects the private features, including model-shareable features and not-allowably shared features, to achieve client-specific classifiers of high performance on LOP per institution. FecMap is then achieved in an iteration manner with all datasets distributed on clients by training a local neural network composed of a global part, a local part, and a classification head in clients and averaging the global parts from clients on the server. To evaluate the FecMap model, we collected three higher-educational datasets of student academic records from engineering majors. Experiment results manifest that FecMap benefits from the proposed LSL and MPP and achieves steady performance on the task of LOP, compared with the state-of-the-art models. This study makes a fresh attempt at the use of federated learning in the learning-analytical task, potentially paving the way to facilitating personalized education with privacy protection.

**Keywords** federated learning, local subspace learning, hierarchical privacy protection, learning outcome prediction, privacy-protected representation learning

## 1 Introduction

Learning-outcome prediction (LOP) is based on students' learning record data to predict their learning performance in future courses, which helps to achieve personalized instructional resource allocation. Due to its fundamental position in learning analytics and educational data mining, many studies have been developed by using various statistical and machine learning models in wide situations [1]. There are many kinds of noise, such as slipping, guessing, and subjectively grading [2], so LOP is often formulated into grade classification.



**Fig. 1** The formulated data point for a student. Past achievements refer to the student's grades in other courses, where  $g_i$  is the grade in the *i*-th course. Student demographics and course description refers to relevant information about students and courses, where  $s_i$  and  $c_i$  are the features, and  $g_t$  is the target grade to be predicted.

As is shown in Fig. 1, this study takes each grade record corresponding to the event that a student enrolled in a course and achieved the grade. In mathematics, let  $g_i$ ,  $s_i$ ,  $c_i$  and  $g_t$  be the grade obtained at the *i*-th course, the student's *i*-

Received Dec 24, 2023; accepted Jul 6, 2023; issued Sep 28,2023; published Dec 15, 2024. This paper will appear in Front. Comput. Sci.

E-mail: shang@nwpu.edu.cn

th feature, the *i*-th course's feature and the target grade related to this record, respectively. Here, we cast the data feature vector into  $\mathbf{x} = (\mathbf{g}_i; \mathbf{s}_i; \mathbf{c}_i)$  where  $\mathbf{g}_i = [g_1, g_2, \dots, g_k]$ ,  $\mathbf{s}_i = [s_1, s_2, \dots, s_m]$  and  $\mathbf{c}_i = [c_1, c_2, \dots, c_n]$ . And then we denote a grade record by  $(\mathbf{x}, y)$  where  $y = \mathbf{g}_t$ . With these notations, the LOP problem can be defined as follows. Given *n* data points  $\{\mathbf{x}_i, y_i\}_{i=1}^n$ , LOP aims to construct a mapping  $y = \mathcal{F}(\mathbf{x})$  which uses students' past data  $\mathbf{x}_t$  to predict the likelihood grade of the student on the next enrolled course, i.e.,  $\mathcal{F}(\mathbf{x}_t) = y_t$ . Towards LOP, the existing machine learning research methods are mainly divided into matrix factorization-based methods [3, 4], similarity-based methods [5], and mapping-based methods [6, 7]. Refer to the review paper of Zhang et al. [8] for more details.

Nevertheless, data shortage and low generalization caused by data privacy issues often limit current methods. The privacy issue is often ignored, but important to protect personal information and sensitive data in schools, leading to data isolation. Thus, the educational data at hand is always a small sample set with a distribution bias concerning the whole picture of educational behaviors. LOP models on such data often have low generalization accuracy for other schools.

Federated learning (FL) has the potential to handle the above data-isolated problem. FL enables many clients jointly train a machine-learning model while keeping their local data decentralized. Usually, FL is trained in an iterative manner: In each round of communication, users train client models  $\mathcal{M}_{W^{i}}$  on their respective datasets and then upload them to the server for computation, and the server calculates the global model  $\mathcal{G}_{W^{*}}$  and then distributes it to all client models. More specifically, given multiple datasets for *m* subtasks or clients, the *m* client-model parameters are denoted by

$$W = \left\{ \mathbf{W}^1, \mathbf{W}^2, \cdots, \mathbf{W}^m \right\}$$
(1)

Then, the general problem of federated learning (FL) with m clients aims to

$$\arg\min_{\mathbf{W}^{*}, W} \sum_{t=1}^{m} \mathcal{G}_{\mathbf{W}^{*}} \left( \sum_{i=1}^{n_{t}} \mathcal{L}_{t} \left( \mathcal{M}_{\mathbf{W}^{t}} \left( \mathbf{x}_{i} \right) \right) + \lambda \mathcal{E} \left( \mathbf{W}^{t} \right) \right)$$
(2)

where  $\mathcal{G}_{\mathbf{W}^*}(\cdot)$  indicates the global model with parameter  $\mathbf{W}^*$ ,  $n_t$  is the number of samples in the *t*-th dataset,  $\mathcal{L}_t(\cdot)$  is the loss function of the task indicated by *t*,  $\mathcal{E}$  is a regularization imposed on the model parameters  $\mathbf{W}^t$ ,  $\lambda$  is the trade-off parameter between loss and its regularization and  $\mathcal{M}(\cdot)$  is the local model. Many studies are witnessed in this emerging field, mainly divided into privacy-protection FL framework [9], the federated machine learning models [10, 11], and the convergence of FL algorithm [12]. More FL models can be found in the review paper [13].

However, the existing FL models often focus on acquiring a better global model while failing to maintain the whole local representation in clients. Moreover, there is no FL work to keep private information in a hierarchy that is commonly encountered in real-world applications, such as finance and healthcare. In education, hierarchical privacy can be public, shareable after encryption, and not allowable. Hence, we proposed a novel FL framework with local subspace learning (LSL) and multi-layer privacy protection (MPP), dubbed FecMap, to achieve client-specific classifiers of high performance on LOP in each institution.

#### 1.1 Overview

Figure 2 illustrates the proposed workflow of FecMap running on *m* clients. The workflow is composed of four steps as follows: ① training client model on the privacy-protected data, where FecMap learns a global part, a local part, and a classification head in each client; ② uploading the global part of client model, where the local part and the classification head are kept locally; ③ computing the global model in the server, which is achieved by weighted averaging over all the global parts from clients; ④ returning the global model to update the global part of the client model. As usual, FecMap learns the model parameters in the iteration manner over the four steps. Our contributions are majorly three-fold.

- LSL divides the feature space of the local client into a local subspace and a global subspace. By introducing the maximization subspace separation criterion, we obtained a shared representation for FL communication and a private representation for local classification.
- MPP divides the local client's privacy features into a multi-level privacy hierarchy. By partitioning the model parameters to learn model-shareable features and notallowed features, the FL global communication and the local private protection are achieved, respectively.
- A FL framework, i.e., FecMap, is proposed for LOP, which is carried out online. FecMap trains a local deep network consisting of a global part, a local part, and a classification head. On the server, model averaging is performed on the global parts from the clients.

The remainder of the paper is organized as follows. We review the previous studies on FL in Section 2. The object function and the FecMap framework are presented in Section 3. Section 4 shows the experimental results on three datasets collected from our institution. Section 5 concludes this paper.

# 2 Related work

Current FL studies mainly focus on data distribution, federated machine-learning models, privacy-protected strategy, and personalization [14-16]. i) The FL studies of data distribution are to handle the three data shapes in practice, i.e., horizontal FL where all clients have identical features but different samples [17], vertical FL where all clients have identical samples but different features, and hybrid FL where all clients have small overlaps on both samples and features [15]. ii) The studies on federated machine-learning models are to extend traditional models to train in the FL schema, e.g., federated ridge regression [18], federated K-means [19], federated principal component analysis [20], federated online learning [21], etc. iii) The studies of privacy-protected strategy are to control the parameter communication for holding privacy, such as the differential-privacy strategy [20], the gradientbased aggregation [18], and the homomorphic encryption [14]. iv) The studies on personalization are to deal with the heterogeneity of features and samples [22-25]. Smith et al. proposed a model of federated multi-task learning, i.e., the MOCHA model, by integrating the association between clients [26]. PFedLA trains a dedicated hyper-network per client in the server to implement personalized FL [23]. Li et al. proposed Ditto for personalization by balancing the global model and local models and training a local classification head [9]. The work [25] achieves personalized federated learning through variational Bayesian inference. In addition, many FL applications have already been conducted on training recommendation systems on mobile devices [17,27], predicting treatment outcomes in hospitals [28], and managing health in home [29, 30].

In recent years, extensive FL models have been proposed to pursue better performance. To generate high-quality models for clients, many studies are concatenated on handling the heterogeneity of local datasets, such as the multiple-center FL [31] and clustered FL [32], and enhancing local models, such as the FedRep [11] and FedBN [33]. On the other hand, personalized FL models attract much attention by balancing the server-client dissimilarity [9,25], inheriting private models [24], maximizing correlation with sparse and hierarchical extensions [16] and integrating the similarity between clients [34, 35]. In addition, the system designation of FL is another important direction, such as the scaled FL [36] and the communication-efficient FL [17].

However, these current FL methods fail to maintain the



**Fig. 2** The FecMap model trained in an iterative manner. An FL communication is completed by (1) training the local model, (2) uploading to the server, (3) computing the global model, and (4) updating the local model.

specific features of local clients that can help improve the local model performance. Moreover, there is no FL work to keep private information in a hierarchy. The recent work [37] privatizes some layers in DNN-based federated learning under non-i.i.d. settings to facilitate the learning process. Motivated by aggregation and privatization, we designed a new federated framework by introducing local subspace learning (LSL) and multi-layer privacy protection (MPP) to boost personalization in federated learning.

## **3** The Proposed Method

This section presents our proposed FecMap shown in Figure 2. This paper uses italic lowercase letters for variables (e.g., x), bold lowercase letters for vectors (e.g.,  $\mathbf{x}$ ), bold capital letters for matrices (e.g.,  $\mathbf{X}$ ), italic capital letters for sets (e.g., X), and calligraphic letters for functions (e.g., X).

## 3.1 Motivation

#### 3.1.1 Local Subspace Learning (LSL)

In each round of FL communication, users train client models  $\mathcal{M}_{\mathbf{W}^i}$  on their respective datasets and then upload them to the server to obtain the weight of the global model  $\mathcal{G}_{\mathbf{W}^*}$  by

$$\mathbf{W}^* = \sum_{i=1}^m p_i \mathbf{W}^i \tag{3}$$

where  $p_i$  is the weight assigned to the client model  $\mathcal{M}_{\mathbf{W}^i}$ based on local data. Usually,  $p_i = 1/m$ . However, there is a model bias between  $\mathcal{M}_{\mathbf{W}^i}$  and  $\mathcal{G}_{\mathbf{W}^*}$ , which can be defined as

$$\mathbf{W}^{i} = \operatorname{dist}\left(\mathbf{W}^{*}, \mathbf{W}^{i}\right) \tag{4}$$

where dist  $(\mathbf{A}, \mathbf{B})$  is used to measure the difference between  $\mathbf{A}$  and  $\mathbf{B}$ . This bias makes it often difficult for the global model to be the best model for clients.

Specifically, federated learning's global model primarily captures common subspaces and may ignore individualized features. Existing FL models mostly incentivize the global model  $\mathcal{G}_{\mathbf{W}^*}$  to fit shared features, yet fail to take into account the bias of the global model to local datasets. A strategy is to set another mapping function, i.e., the local submodel  $\mathcal{M}_{\widehat{\mathbf{W}}^i}$ , to learn dist (·) and  $\mathbf{W}^i$ .

#### 3.1.2 Multi-layer Privacy Protection (MPP)

Existing FL studies encrypt all model parameters to upload to the server but fail to consider the levels of privacy security. Inspired by this, an intuitional idea is to protect privacy features by dividing them into shareable features  $\mathbf{F}^*$  and notallowed features  $\mathbf{F}'$ . Wherein, shareable features  $\mathbf{F}^*$  refers to features in the local model that can be shared with the server. This part of the feature does not have strong privacy and can be uploaded to the server in traditional federated communication; not-allowed features  $\mathbf{F}'$  refer to features that are trained only in the local model and do not participate in the communication of the server.

In this study, student grades  $\mathbf{g}_i = [g_1, g_2, \dots, g_k]$  are shareable information that we use to capture the global features by communicating with other clients. Student personal information  $\mathbf{s}_i = [s_1, s_2, \dots, s_m]$  is a sensitive not-allowed feature that we use to complement the local features. As the course description may involve sensitive student information, we also consider  $\mathbf{c}_i = [c_1, c_2, \dots, c_n]$  as the not-allowed feature.

#### 3.2 Model objective

The proposed FecMap combines the strategy for LSL. Let  $\mathcal{M}_{\widetilde{W}^m}$  and  $\mathcal{M}_{\widehat{W}^m}$  be the global sub-model and the local sub-model for the client *m*, respectively. It can be formulated as

$$\mathcal{M}_{\mathbf{W}^m} = \mathcal{F}_m\left(\mathcal{M}_{\widetilde{\mathbf{W}}^m}, \mathcal{M}_{\widehat{\mathbf{W}}^m}\right) \tag{5}$$

where  $\mathcal{F}_m(\cdot)$  is an aggregate function. Replace Eq. (2) with Eq. (5) to get the following rewritten global objective as

$$\arg\min_{\mathbf{W}^{*}, W} \sum_{t=1}^{m} \mathcal{G}_{\mathbf{W}^{*}} \left( \sum_{i=1}^{n_{t}} \mathcal{L}_{t} \left( \mathcal{F}_{t} \left( \mathcal{M}_{\widetilde{\mathbf{W}}^{t}} \left( x_{i} \right), \mathcal{M}_{\widehat{\mathbf{W}}^{t}} \left( x_{i} \right) \right) \right) \right)$$
(6)

where  $W = \{\{\widetilde{\mathbf{W}}^1, \widehat{\mathbf{W}}^1\}, \{\widetilde{\mathbf{W}}^2, \widehat{\mathbf{W}}^2\}, \cdots, \{\widetilde{\mathbf{W}}^m, \widehat{\mathbf{W}}^m\}\}$ . To widen the distance between two submodels of clients, a regulariza-

tion term is introduced into Eq. (6) as

$$\mathcal{E}\left(\mathbf{W}^{t}\right) = \operatorname{dist}\left(\mathcal{M}_{\widetilde{\mathbf{W}}^{t}}, \mathcal{M}_{\widetilde{\mathbf{W}}^{t}}\right) = \exp\left(-\frac{\left\|\widetilde{\mathbf{W}}^{t} - \widehat{\mathbf{W}}^{t}\right\|_{F}}{\left\|\widetilde{\mathbf{W}}^{t}\right\|_{F} + \left\|\widehat{\mathbf{W}}^{t}\right\|_{F}}\right) \quad (7)$$

where dist (A, B) measures the difference between A and B;  $\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$  calculates the Fibonacci norm of the matrix **A** which aims to collect the difference between all parameters between two identical models, correspondingly; exp(·) is the exponential function.

Thus, the objective function is implemented by combining Eqs. (6) and (7), where  $\alpha \ge 0$  is a balance parameter.

$$\arg\min_{\mathbf{W}^{*}, W} \sum_{t=1}^{m} \mathcal{G}_{\mathbf{W}^{*}} \left( \sum_{i=1}^{n_{t}} \mathcal{L}_{t} \left( \mathcal{F}_{t} \left( \mathcal{M}_{\widetilde{\mathbf{W}}^{\prime}} \left( x_{i} \right), \mathcal{M}_{\widetilde{\mathbf{W}}^{\prime}} \left( x_{i} \right) \right) \right) + \alpha \operatorname{dist} \left( \mathcal{M}_{\widetilde{\mathbf{W}}^{\prime}}, \mathcal{M}_{\widetilde{\mathbf{W}}^{\dagger}} \right)$$
(8)

The proposed FecMap combines the strategy of MPP. Denote by  $\mathbf{F}^*$  and  $\mathbf{F}'$  the shareable features and the privacy features, respectively. By MPP, it can be cast as

$$\mathbf{F}^* = \mathbf{g}_i, \mathbf{F}' = \mathbf{s}_i \oplus \mathbf{c}_i \tag{9}$$

where student grades  $\mathbf{g}_i = [g_1, g_2, \dots, g_k]$  are employed as shareable features, while student personal information  $\mathbf{s}_i = [s_1, s_2, \dots, s_m]$  and course description  $\mathbf{c}_i = [c_1, c_2, \dots, c_n]$ are used as not-allowed features in this study.

Finally, substitute Eq. (9) into Eq. (8) and then obtain the following global objective problem.

$$\arg\min_{\mathbf{W}^{*}, \mathbf{W}} \sum_{t=1}^{m} \mathcal{G}_{\mathbf{W}^{*}} \left( \sum_{i=1}^{n_{t}} \mathcal{L}_{t} \left( \mathcal{F}_{t} \left( \mathcal{M}_{\widetilde{\mathbf{W}}^{t}} \left( \mathbf{F}_{i}^{*} \right), \mathcal{M}_{\widehat{\mathbf{W}}^{t}} \left( \mathbf{F}_{i}^{*} \right), \mathcal{H}^{t} \left( \mathbf{F}_{i}^{'} \right) \right) \right) + \alpha \operatorname{dist} \left( \mathcal{M}_{\widetilde{\mathbf{W}}^{t}}, \mathcal{M}_{\widehat{\mathbf{W}}^{t}} \right)$$
(10)

where  $\mathcal{M}_{\widetilde{W}^{t}}$ ,  $\mathcal{M}_{\widehat{W}^{t}}$ , and  $\mathcal{H}^{t}$  are fulfilled by the shared sub-model, the not-shared sub-model, and the local head.

## 3.3 The FecMap Algorithm

The proposed FecMap algorithm is implemented in the commonly used iteration manner, as shown in Algorithm 1.

**Client Update.** The task for each client is to train the LOP classifier on their respective datasets. In the *k*-th round of FL communication, the network weights are updated by

$$\left(\widetilde{\mathbf{W}}_{(k)}^{t}, \widehat{\mathbf{W}}_{(k)}^{t}\right) \leftarrow \text{SGD}\left(\mathbf{W}_{(k-1)}^{*}, \widehat{\mathbf{W}}_{(k-1)}^{t}, \mathcal{H}_{(k-1)}^{t}\right)$$
 (11)

where SGD (·) is the stochastic gradient descent and  $W^*_{(k-1)}$  is the global model from the previous round.

**Server Update.** After client updating, all clients upload the global part  $\widetilde{W}_{(k)}^{t}$  to the server. The server then performs



**Fig. 3** Schematic diagram of the FecMap client neural network model. Circle points are network nodes; green box is sharable; red boxes are non-sharable.  $\mathbf{g}_i$  is the grade, and *dist* is the distance between the two networks. The learned representations are composed of student features  $\mathbf{s}_i$  and course features  $\mathbf{c}_i$ . Outputs are the levels of the course scores.

an average of all clients and returns the new global model to clients. The server model is then computed by

$$\mathbf{W}_{(k+1)}^{*} = \frac{1}{m} \sum_{t=1}^{m} \widetilde{\mathbf{W}}_{(k)}^{t}$$
(12)

## **4** Model Implementation and Evaluation

#### 4.1 An Online FecMap Implementation

To use the proposed FecMap for LOP in multiple institutions, we have developed an online FL platform by employing the open-source technology of Streamlit. The online workflow is shown in Fig. 4, including the interfaces of data loading, parameter setting, model training, results visualization, and result evaluation. Our experiments are conducted on multiple computers and a server via this online version.

## 4.2 The Used Dataset

To evaluate the FecMap model, we collected our university's undergraduate student performance records in the years 2015–2020. Each student has two semesters a year, each free for course selection. A final examination is usually administered to assess the student's learning outcomes at the end of a course. Most results are given in (0 - 100) or grades (A - E), and other forms are removed, such as "pass" and "fail".

In this paper, records from three majors, i.e., Computer Science and Technology (CST), Software Engineering (SE), and Electronic Information Engineering (EIE), are selected for the experiment. Detailed information, including sample sizes and basic classification statistics for each dataset,

Algorithm 1 FecMap Algorithm
<b>Require:</b> Participation rate $\gamma$ ; step size $\alpha$ ; the number of
local updates $\tau$ ; the number of communication rounds T
<b>Initialization</b> $\mathbf{W}_0^*; \mathcal{H}_0^1, \cdots, \mathcal{H}_0^n; \widehat{\mathbf{W}}_0^0, \cdots, \widehat{\mathbf{W}}_0^n$
1: <b>for</b> $t = 1, 2,, T$ <b>do</b>
2: Server receives a batch of clients $L^t$ of size $\gamma n$
Server sends current global model $\mathbf{W}_{t-1}^*$ to clients
3: <b>for</b> each client $i$ in $L^t$ <b>do</b>
4: Client <i>i</i> initializes
$\widetilde{\mathbf{W}}_{t}^{\mathrm{i}} \leftarrow \mathbf{W}_{t-1}^{*}, \widehat{\mathbf{W}}_{t}^{\mathrm{i}} \leftarrow \widehat{\mathbf{W}}_{t-1}^{i}, \mathcal{H}_{t}^{i} \leftarrow \mathcal{H}_{t-1}^{i}$
Updates the local sub-model $\widehat{\mathbf{W}}_t^i$ and local head $\mathcal{H}_t^i$
5: <b>for</b> $s = 1$ to $\tau$ <b>do</b>
6: $\left(\widetilde{\mathbf{W}}_{t+1,s}^{i}, \widetilde{\mathbf{W}}_{t+1,s}^{i}\right) \leftarrow \operatorname{SGD}\left(\mathbf{W}_{t,s}^{*}, \widetilde{\mathbf{W}}_{t,s}^{i}, \mathcal{H}_{t,s}^{i}, \alpha\right)$
7: end for
Client <i>i</i> sends updated global model $\mathbf{W}_{t,\tau}^i$ to server
8: end for
9: <b>for</b> each client <i>i</i> not in $L^t$ <b>do</b>
10: Set $\widehat{\mathbf{W}}_{t,\tau}^{i} \leftarrow \widehat{\mathbf{W}}_{t-1,\tau}^{i}, \mathcal{H}_{t,\tau}^{i} \leftarrow \mathcal{H}_{t-1,\tau}^{i}$
11: end for
Server computes the new global model as:
$\mathbf{W}^*_{t+1} = rac{1}{m}\sum_{i=1}^m \widetilde{\mathbf{W}}^i_t$
12: end for

12: end for is shown in Table 1. In addition, student characteristics include age, class, gender, country, and ethnicity. Course characteristics include total course hours and weekly course hours, credits, category, textbook used, course examination method, and course offering unit. All attributes are converted to numerical values and then normalized between 0 and 1.

#### 4.3 Experimental Settings

We conducted the experiments via the following settings.

**Data partitioning.** The dataset in hand is equally divided for all clients with the same class distribution. The dataset allocated to each client is then partitioned into a training set and a test set by a ratio of 7:3.

**Model parameters.** FecMap in this experiment employs MLP for  $\mathcal{H}$ , ReLU for the activation function, the softmax function for classification, and the cross-entropy for loss computation. There are in total 8 fully connected layers in the overall network. Concretely, the global part has  $dim_{in}$ -22-11-6 units; the local part has  $dim_{in}$ -22-11-6 units; the classification head has 27-20-13-5 units. The feature dimension determines the model input  $dim_{in}$ .

**Model training.** FecMap randomly initializes all models for T = 500 communication rounds of training. Clients are randomly selected at a rate of  $\gamma = 0.1$  to participate in federated machine learning. The local model in clients performs 15 SGD in each local update, where the learning rate and the moment are set to 0.01 and 0.5, respectively.

Major	Student	Course	Record	0≤grade<60	60≤grade<70	70≤grade<80	80≤grade<90	90≤grade<100
CST	1463	44	46512	1874(4.0%)	7434(15.9%)	10755(23.1%)	15720(33.7%)	10729(23.0%)
SE	1621	46	37952	1445(3.8%)	7050(18.5%)	9437(24.8%)	13048(34.3%)	6972(18.3%)
EIE	1937	43	49546	1917(3.8%)	9128(18.4%)	10655(21.5%)	16715(33.7%)	11131(22.4%)

 Table 1
 Data set information, where the values in brackets are the ratios.



Fig. 4 The Online FecMap flowchart.

**Model comparisons.** Five state-of-the-art FL methods, FedAvg [17], FedProx [10], LG-Fed [38], FedPer [39] and FedRep [11] are meanwhile evaluated for comparison.

#### 4.4 Experiment Results

#### 4.4.1 Data Representation Visualization

In order to observe the discriminative power of the model, we visualized the representations by the last layer of the local and global parts in Fig. 3 for two classes. The resulting 8D subspace features are projected into 2D representations by using TSNE [40] for visualization. Fig. 5 shows the global representation, the local representation, and the combined representation from the 1st to the 3rd column, respectively. Besides, the discriminative representation from the layer before the softmax layer is also displayed in the 4th column.

As shown in Fig. 5, the results of three clients are shown, where different colors represent the categories labeled with numbers. From the results, the global representations are much mixed, while the local representations potentially have

#### Table 2 Accuracy comparison of various methods.

Methods	CST	SE	EIE
FedAvg ([17])	80.46	76.82	77.86
FedProx ([10])	79.23	78.23	78.19
LG-Fed ([38])	76.50	75.85	78.51
FedPer ( [39])	81.58	78.29	77.52
FedRep ( [11])	81.87	83.95	82.99
FecMap(Ours)	83.10	84.92	85.41

improvements. Finally, the combination of both local and global representations is more discriminative for distinguishing the two classes clearly.

## 4.4.2 Prediction Results

In order to test the prediction accuracy of the model for student performance, the proposed FecMap and these comparison methods are tested on each of the three datasets. The client model is evaluated on the local test set of all clients, and then the results are averaged to obtain the final accuracy. Table 2 lists all results.

As shown in Table 2, FecMap obtains higher prediction performance than the state-of-the-art FL methods. Specifically, on the CST dataset, the accuracy outperformed the FedAvg, FedProx, LG-Fed, FedPer and FedRep methods by 2.64%, 3.87%, 6.6%, 1.52% and 1.23%, respectively. Similar results can be observed on the SE and EIE datasets. It is demonstrated that the proposed model outperforms existing FL methods overall and performs well in the task of LOP. The results imply that FecMap can benefit from the proposed LSL and MPP.

Fig. 6 compares the confusion matrixes achieved from FecMap and FedRep on the test sets. In comparison to the FedRep framework, the proposed method achieves higher prediction accuracy in all five categories. In particular, our FecMap achieves an accuracy of 93% in detecting students with course grades below 60 (grade 1), while FedRep was only 85%. This result shows that this method is practical and feasible in an academic early warning system. For the prediction of grades 3, 4 and 5, FecMap also delivers higher accuracy than FedRep, although there are more errors in grade 2. In addition, FecMap achieves excellent performance in distinguishing between grade 1 and grade 5.



Fig. 5 Visualization of data representation. The first column is the global representations, the second is the local representations, the third is the combined representations, and the fourth is the discriminative representations. Each row represents a client, and each color represents a category.



Fig. 6 Confusion matrix for FecMap (left) and FedRep (right).



To evaluate the impacts of the introduced LSL and MPP, we evaluated the FecMap model with only MPP (FecMap-MPP) and the FecMap model with only LSL (FecMap-LSL) on the three datasets, respectively. Besides, we compared with FedRep as it is our base model. Fig. 7 shows the prediction performance of FecMap-LSL vs. FedRep and FecMap-MPP vs. FedRep. The results show that FecMap can benefit from the introduced LSL and MPP. The final accuracy is listed in Table 3, showing that both LSL and MPP can improve the performance of FedRep, where LSL seems more powerful than MPP. Overall, FecMap benefits from MPP and LSL to achieve the best results.

#### 4.4.4 Parameter Discussion

We adjust the number of clients and the size of the dataset



Fig. 7 The accuracy of FecMap-LSL and FecMap-MPP against the number of communications, compared to FedRep.

Table 3         Ablation study							
Methods	CST	SE	EIE				
FecMap - LSL	82.58	84.56	85.30				
FecMap - MPP	82.01	84.36	83.18				
FedRep	81.87	83.95	82.99				
FecMap	83.10	84.92	85.41				

to investigate their effects. As can be seen in Fig 8(left), the accuracy of FecMap has 2%-5% higher than other compared methods when varying the number of clients. The results show that the performance gets better as the number of clients increases. A similar observation could be reached in Fig. 8(right) in which the size of the dataset in clients increases. On the other hand, the proposed FecMap delivers the highest accuracy among all the compared state-of-the-art FL models consistently in all experiment cases. Overall, FecMap achieves the best performance on the task of LOP and benefits from more clients and data.



Fig. 8 Parameter discussion on the number of clients *n* (left) and the data set size *d* (the number of samples per client, right) in the proposed model and other comparison methods.



**Fig. 9** Loss function (left) and accuracy (right) for FecMap in the case study with 20 communication rounds.

## 4.4.5 A Case Study of Online FecMap

In this section, we present a case study on the SE dataset that was collected from the major of software engineering in our institution. The main steps are as follows.

**STEP 1**: Uploading the input data following the provided example on the homepage of our online system.

**STEP 2**: Setting up the training parameters. The parameters include the number of clients, the number of communication rounds, the number of training sessions, the proportion of participants in the training, the learning rate, and the federated learning algorithm. We here set parameters to 5, 20, 10, 0.1, 0.01, and FecMap, respectively.

**STEP 3**: Training FecMap while monitoring local models' performance. The client models are trained in parallel based on the set number of clients and participation rate. After all participating clients are trained, all local models are uploaded to the server. The server averages all model parameters and distributes the resulting model to each client. The local training loss and accuracy are computed per epoch.

**STEP 4**: Calculating and showing the results in terms of various metrics, including prediction accuracy, loss function,

and confusion matrix, displayed on the visualization page.

Fig. 9 shows the loss-function curves and accuracy curves with 20 communication rounds. The FecMap training of this case is completed in 115.557 seconds. Finally, we achieved the following results: the training loss converges in the 17th round; the global model achieves the test accuracy of 66.48%, while the average accuracy is 68.353% on the five clients. The obtained test accuracy is lower than the results in Table 2 because this case study used a smaller number of clients.

# 5 Conclusion

In this paper, we proposed a new FL framework with local subspace learning (LSL) and multi-layer privacy protection (MPP) to deal with data privacy for LOP. LSL is to learn and retain the local features against the global features that are sent into the federated average. MPP is to hierarchically perform and protect the private features, including model-shareable and not-allowed features. Experiment results on our three datasets validate that FecMap has a stable and consistent prediction performance in comparison with other FL models. This study paves the way for data analysis between education institutions with privacy protection.

**Acknowledgements** This study was supported in part by the National Natural Science Foundation of China (62272392, U1811262, 61802313), the Key Research and Development Program of China (2020AAA0108500), the Key Research and Development Program of Shaanxi Province (2023-YBGY-405), the fundamental research funds for the central university (D5000230088), and the Higher Research Funding on International Talent Cultivation at NPU (GJGZZD202202).

# References

- Zhang Y, An R, Liu S, Cui J, Shang X. Predicting and understanding student learning performance using multi-source sparse attention convolutional neural networks. IEEE Transactions on Big Data, 2021
- Zhang Y, Dai H, Yun Y, Liu S, Lan A, Shang X. Meta-knowledge dictionary learning on 1-bit response data for student knowledge diagnosis. Knowledge-Based Systems, 2020, 205: 106290
- Symeonidis P, Malakoudis D. Multi-modal matrix factorization with side information for recommending massive open online courses. Expert Systems with Applications, 2019, 118: 261–271
- Zhang Y, Yun Y, Dai H, Cui J, Shang X. Graphs regularized robust matrix factorization and its application on student grade prediction. Applied Sciences, 2020, 10(5): 1755
- Bydžovská H. Student performance prediction using collaborative filtering methods. In: International Conference on AI in Education. 2015, 550–553
- Al-Shehri H, Al-Qarni A, Al-Saati eL. Student performance prediction using support vector machine and k-nearest neighbor. In: IEEE 30th Canadian conference on electrical and computer engineering. 2017, 1–4
- Polyzou A, Karypis G. Feature extraction for next-term prediction of poor student performance. IEEE Transactions on Learning Technologies, 2019, 12(2): 237–248
- Zhang Y, Yun Y, An R, Cui J, Dai H, Shang X. Educational data mining techniques for student performance prediction: Method review and comparison analysis. Frontiers in psychology, 2021, 12
- Li T, Hu S, Beirami A, Smith V. Ditto: Fair and robust federated learning through personalization. In: International Conference on Machine Learning. 2021, 6357–6368
- Li T, Sahu A K, Zaheer M, Sanjabi M, Talwalkar A, Smith V. Federated optimization in heterogeneous networks. Proceedings of Machine Learning and Systems, 2020, 2: 429–450
- Collins L, Hassani H, Mokhtari A, Shakkottai S. Exploiting shared representations for personalized federated learning. In: International Conference on Machine Learning. 2021, 2089–2099
- Haddadpour F, Mahdavi M. On the convergence of local descent methods in federated learning. arXiv:1910.14425, 2019
- Li T, Sahu A K, Talwalkar A, Smith V. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 2020, 37(3): 50–60
- Li Q, Wen Z, Wu Z, Hu S, Wang N, Li Y, Liu X, He B. A survey on federated learning systems: vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering, 2021
- Tan A Z, Yu H, Cui L, Yang Q. Towards personalized federated learning. IEEE Transactions on Neural Networks and Learning Systems, 2022
- 16. Li Y, Liu X, Zhang X, Shao Y, Wang Q, Geng Y. Personalized federated learning via maximizing correlation with sparse and hierarchical

extensions. arXiv preprint arXiv:2107.05330, 2021

- McMahan B, Moore E, Ramage D, Hampson S, Arcas y B A. Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics. 2017, 1273–1282
- Chen Y R, Rezapour A, Tzeng W G. Privacy-preserving ridge regression on distributed data. Information Sciences, 2018, 451-452: 34–49
- Dennis D K, Li T, Smith V. Heterogeneity for the win: One-shot federated clustering. In: International Conference on Machine Learning. 2021, 2611–2620
- Ribero M, Henderson J, Williamson S, Vikalo H. Federating recommendations using differentially private prototypes. Pattern Recognition, 2022, 129: 108746
- Zhou P, Wang K, Guo L, Gong S, Zheng B. A privacy-preserving distributed contextual federated online learning framework with big data support in social recommender systems. IEEE Transactions on Knowledge and Data Engineering, 2019, 33(3): 824–838
- Zhang Y, Xu Y, Wei S, Wang Y, Li Y, Shang X. Doubly contrastive representation learning for federated image recognition. Pattern Recognition, 2023, 139: 109507
- Ma X, Zhang J, Guo S, Xu W. Layer-wised model aggregation for personalized federated learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022, 10092– 10101
- Li X C, Zhan D C, Shao Y, Li B, Song S. Fedphp: Federated personalization with inherited private models. In: Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part I 21. 2021, 587–602
- Zhang X, Li Y, Li W, Guo K, Shao Y. Personalized federated learning via variational bayesian inference. In: International Conference on Machine Learning. 2022, 26293–26310
- Smith V, Chiang C K, Sanjabi M, Talwalkar A S. Federated multi-task learning. Advances in neural information processing systems, 2017, 30
- Duan M, Liu D, Chen X, Liu R, Tan Y, Liang L. Self-balancing federated learning with global imbalanced data in mobile systems. IEEE Transactions on Parallel and Distributed Systems, 2020, 32(1): 59–71
- Bercea C I, Wiestler B, Rueckert D, Albarqouni S. Federated disentangled representation learning for unsupervised brain anomaly detection. Nature Machine Intelligence, 2022, 4(8): 685–695
- Wu Q, Chen X, Zhou Z, Zhang J. Fedhome: Cloud-edge based personalized federated learning for in-home health monitoring. IEEE Transactions on Mobile Computing, 2020, 21(8): 2818–2832
- Wang N, Chen Y, Hu Y, Lou W, Hou Y T. Feco: Boosting intrusion detection capability in iot networks via contrastive learning. In: IEEE INFOCOM 2022-IEEE Conference on Computer Communications. 2022, 1409–1418
- Long G, Xie M, Shen T, Zhou T, Wang X, Jiang J. Multi-center federated learning: clients clustering for better personalization. World Wide Web, 2023, 26(1): 481–500
- Sattler F, Müller K R, Samek W. Clustered federated learning: Modelagnostic distributed multitask optimization under privacy constraints.

IEEE transactions on neural networks and learning systems, 2020, 32(8): 3710-3722

- Li X, Jiang M, Zhang X, Kamp M, Dou Q. Fedbn: Federated learning on non-iid features via local batch normalization. arXiv preprint arXiv:2102.07623, 2021
- Zhang Y, Wei S, Liu S, Wang Y, Xu Y, Li Y, Shang X. Graph-regularized federated learning with shareable side information. Knowledge-Based Systems, 2022, 257: 109960
- Yang L, Huang J, Lin W, Cao J. Personalized federated learning on non-iid data via group-based meta-learning. ACM Trans. Knowl. Discov. Data, 2023, 17(4)
- Bonawitz K, Eichner H, Grieskamp eW. Towards federated learning at scale: System design. Proceedings of Machine Learning and Systems,

2019, 1: 374-388

- Li X C, Gan L, Zhan D C, Shao Y, Li B, Song S. Aggregate or not? exploring where to privatize in dnn based federated learning under different non-iid scenes. arXiv preprint arXiv:2107.11954, 2021
- Liang P P, Liu T, Ziyin L, Allen N B, Auerbach R P, Brent D, Salakhutdinov R, Morency L P. Think locally, act globally: Federated learning with local and global representations. arXiv preprint arXiv:2001.01523, 2020
- Arivazhagan M G, Aggarwal V, Singh A K, Choudhary S. Federated learning with personalization layers. arXiv preprint arXiv:1912.00818, 2019
- Maaten V. d L, Hinton G. Visualizing data using t-sne. Journal of machine learning research, 2008, 9(11)