The Thirty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2025)
February 25–March 4, 2025, Philadelphia, Pennsylvania

Personalized Federated Collaborative Filtering: A Variational AutoEncoder Approach

Zhiwei Li[1], Guodong Long[1], Tianyi Zhou[2], Jing Jiang[1], Chengqi Zhang[3]

[1]Australian Artificial Intelligence Institute, Faculty of Engineering and IT, University of Technology Sydney
[2]Department of Computer Science, University of Maryland, College Park
[3]Department of Data Science and Artificial Intelligence, The Hong Kong Polytechnic University
zhiwei.li@student.uts.edu.au, {guodong.long, jing.jiang}@uts.edu.au, zhou@umiacs.umd.edu, chengqi.zhang@polyu.edu.hk

Abstract

Federated Collaborative Filtering (FedCF) is an emerging field focused on developing a new recommendation framework with preserving privacy in a federated setting. Existing FedCF methods typically combine distributed Collaborative Filtering (CF) algorithms with privacy-preserving mechanisms, and then preserve personalized information into a user embedding vector. However, the user embedding is usually insufficient to preserve the rich information of the fine-grained personalization across heterogeneous clients. This paper proposes a novel personalized FedCF method by preserving users' personalized information into a latent variable and a neural model simultaneously. Specifically, we decompose the modeling of user knowledge into two encoders, each designed to capture shared knowledge and personalized knowledge separately. A personalized gating network is then applied to balance personalization and generalization between the global and local encoders. Moreover, to effectively train the proposed framework, we model the CF problem as a specialized Variational AutoEncoder (VAE) task by integrating user interaction vector reconstruction with missing value prediction. The decoder is trained to reconstruct the implicit feedback from items the user has interacted with, while also predicting items the user might be interested in but has not yet interacted with. Experimental results on benchmark datasets demonstrate that the proposed method outperforms other baseline methods, showcasing superior performance.

Code — https://github.com/mtics/FedDAE

Introduction

In the digital age, Recommendation Systems have become essential tools for filtering online information and helping users discover products, content, and services that match their preferences (Ko et al. 2022). Collaborative Filtering (CF) is widely recognized for its ability to generate personalized recommendations by analyzing the relationships between users and items based on user interaction data (Shen, Zhou, and Chen 2020). However, with the enforcement of data privacy laws like GDPR (Voigt and Von dem Bussche 2017), safeguarding privacy has become increasingly critical.Traditional CF methods typically require centralizing user data on servers for processing, a practice that is no longer viable in today's privacy-conscious environment.

To address this challenge, Federated Collaborative Filtering (FedCF) has emerged, combining the principles of federated learning (FL) and CF (Yang et al. 2020). FedCF enables models to be trained on users' devices, eliminating the need to upload private data to central servers, thus ensuring data privacy while still providing recommendation services (Ammad-Ud-Din et al. 2019). Existing work on FedCF (Chai et al. 2020; Lin et al. 2020b,a) is mostly based on matrix factorization (MF) (Koren, Bell, and Volinsky 2009), which is favored for its computational efficiency and strong interpretability. MF effectively captures the latent relationships between users and items, making it widely used in many practical applications and delivering strong performance in recommendation systems (Mehta and Rana 2017).

However, the fundamental assumption of MF is that the relationships between users and items are linear, which may limit the model's ability to handle complex, non-linear relationships. The success of Neural Collaborative Filtering (NCF) (He et al. 2017) demonstrates the necessarity of building non-linear relationships among users and items. In addition to FedNCF (Perifanis and Efraimidis 2022) which is a federated NCF algorithm, some recent works proposed to preserve the personalization on neural models in federated settings. PFedRec (Zhang et al. 2023a) proposed using personalized item embeddings to capture each user's unique perspective about the relationships among items. FedRAP (Li, Long, and Zhou 2024) decomposes the item embedding to two additive components that preserve shared and personalized knowledge respectively. However, these methods rely on personalized item embedding that may lead to poor performance on model generalization, and maintaining a personalized item embedding is computation consuming.

Main Contribution. To effectively preserve fine-grained personalization and non-linear user-item relationships, we propose a novel **gating dual-encoder** structure that transforms the user-item interaction vector into two separate latent subspaces. Specifically, through collaborative training across clients, the global encoder maps the user profile into a universal latent subspace shared among clients. Meanwhile, the local encoder remains on the client side, where it is trained to map the user profile into a user-specific latent subspace. Consequently, the global encoder retains shared knowledge across clients, while the local encoder preserves personalized knowledge. It is worth noting our proposed

framework preserves personalized information at three levels. First, the user profile includes user-specific interaction records, allowing the global encoder to transform it into a user-specific vector within a universal latent subspace. Second, the personalized local encoder further transforms the user profile into a unique latent subspace. Third, the gating network adaptively adjusts the importance weights of the two encoders, balancing the performance of our proposed model between personalization and generalization.

To effectively train the proposed gating dual encoders, we revisit the FedCF problem from the perspective of Variational Autoencoders (VAEs) (Kingma and Welling 2013) and propose a novel personalized FedCF method, which incorporates a gating dual-encoder VAE, named **FedDAE**. To enhance the model's generalization capability, FedDAE employs a gating network that generates weights based on user interaction data, dynamically combining the outputs of the two encoders to achieve additive personalization. To train our method FedDAE, a global decoder is attached to the dual encoders, and the objective function of FedDAE is optimized through an designed alternating update process.

The paper's main contributions are summarized as below:

- This work reformulates the FedCF problem as a VAE task to capture complex nonlinear relationships and maintain fine-grained personalization;
- We adopt a dual-encoder structure to separate shared and personalized knowledge, coupled with a personalized gating network that dynamically adjusts encoder weights, effectively balancing generalization and personalization to enhance FedCF tasks;
- Comprehensive experimental analyses are conducted to validate the effectiveness of FedDAE.

Related Work

Personalized Federated Learning. Standard federated learning methods, such as FedAvg (McMahan et al. 2017), learn a global model on the server while considering data locality on each device (Li and Long 2024). However, these methods are limited in their effectiveness when dealing with non-IID data. PFL aims to learn personalized models for each device to address this issue (Tan et al. 2022), often necessitating server-based parameter aggregation (Arivazhagan et al. 2019; T Dinh, Tran, and Nguyen 2020; Collins et al. 2021; Li et al. 2024; Zhang et al. 2024b). Several studies (Ammad-Ud-Din et al. 2019; Chen et al. 2023, 2024; Yang et al. 2024) accomplish PFL by introducing various regularization terms between local and global models. Meanwhile, some work focus on personalized model learning by promoting the closeness of local models via variance metrics (Flanagan et al. 2020), or enhancing this by clustering users into groups and selecting representative users for training (Li et al. 2021; Luo, Xiao, and Song 2022; Tan et al. 2023; Yan and Long 2023), instead of random selection. APFL (Deng, Kamani, and Mahdavi 2020) proposes adaptive PFL and derives the generalization bound for the mixture of local and global models, thereby achieving a balance between global collaboration and local personalization.

Federated Collaborative Filtering. As privacy protection becomes increasingly important, many studies (Hegedűs, Danner, and Jelasity 2019; Chai et al. 2020; Zhang and Jiang 2021; Zhang et al. 2024a,d,c) have focused on FedCF. In this paper, we mainly focus on model-based personalized FedCF, which leverages users' historical behavior data, such as ratings, clicks, and purchases, to learn latent factors that represent user and item characteristics (Aggarwal and Aggarwal 2016). To mitigate the impacts caused by client heterogeneity, personalized FedCF has received considerable attention due to its ability to take into account personalized information for each user. Early works (Lin et al. 2020a; Liang, Pan, and Ming 2021; Zhu et al. 2022; Luo et al. 2023; Yuan et al. 2023; Zhang et al. 2023b) primarily focused on modeling user preferences. In contrast, PFedRec (Zhang et al. 2023a) and FedRAP (Li, Long, and Zhou 2024) have implemented dual personalization, which means they personalized both user preferences and item information. However, the nature of most existing work on matrix factorization limits their ability to model complex nonlinear relationships in data.

Variational Autoencoders. VAEs have become increasingly important in recommendation systems due to their ability to model complex data distributions and handle data sparsity issues (Liang et al. 2024). VAEs learn latent representations of user-item interaction data by mapping high-dimensional data to a low-dimensional latent space, which effectively captures complex patterns in user preferences and item characteristics (Kingma and Welling 2013). This latent space enables efficient encoding of both user and item information, facilitating accurate predictions of future interactions (Shenbin et al. 2020). Recent research has focused on enhancing VAE models to improve recommendation performance, including combining VAEs with Generative Adversarial Networks (Lee, Song, and Moon 2017; Yu et al. 2019; Gao et al. 2020), integrating VAEs with more effective priors (Tomczak and Welling 2018; Klushyn et al. 2019; Shenbin et al. 2020; Tran and Lauw 2024), and enhancing VAEs with contrastive learning techniques (Aneja et al. 2021; Xie et al. 2021; Wang et al. 2022). These strategies have significantly boosted the performance of VAEs in recommendation tasks. HI-VAE (Nazabal et al. 2020) has demonstrated the effectiveness of VAE in density estimation and missing data imputation through experiments on heterogeneous data completion tasks. Mult-VAE (Liang et al. 2018) proves that using multinomial distribution as the likelihood function is more suitable for implicit feedback. FedVAE (Polato 2021) represents the first attempt to extend VAE to a federated setting; however, it aggregates gradients for all parameters on the server, which can lead to the loss of personalized information from individual clients.

Problem Formulation

Given n users and m items, let $\mathcal{U}$ and $\mathcal{I}$ be the sets of users and items respectively. We assume of knowing the users' implicit feedback, i.e., $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n]^T \in \{0,1\}^{n \times m}$, where $r_{ui} = 1$ if the u-th user interacted with the i-th item, and $r_{ui} = 0$ otherwise. For the u-th user, we use $\mathbf{r}_u \in \{0,1\}^m$ to indicate the interaction history, $\mathcal{I}_u$ to de-

note the set of all interacted items, and its length is m_u.

Federated Collaborative Filtering (FedCF)

In recommendation tasks with implicit feedback, Collaborative Filtering (CF) methods typically rely solely on users' interaction patterns $\mathbf{R}$ to capture the relationships between users $\mathcal{U}$ and items $\mathcal{I}$, allowing for generating item recommendations without the need for exogenous information about the users or items (Koren, Rendle, and Bell 2021). However, in the context of FL, the task of FedCF is to model the latent relationships between each user u and items $\mathcal{I}$ at their client using the observed portions of their interaction data $\mathbf{r}_u$ locally. The goal is to generate predicted ratings $\hat{\mathbf{r}}_u$ while simultaneously ensuring the protection of user privacy. Thus, we have the following for recommendation:

$$\min \sum_{u=1}^{n} \mathcal{L}_{recon}(\hat{\mathbf{r}}_u, \mathbf{r}_u), \tag{1}$$

where $\mathcal{L}_{recon}$ measures the reconstruction loss between $\hat{\mathbf{r}}_u$ and $\mathbf{r}_u$. By optimizing Eq. 1, we ensure that $\hat{\mathbf{r}}_u$ follows the same distribution as $\mathbf{r}_u$, i.e., $\hat{\mathbf{r}}_u \sim p(\mathbf{r}_u)$ $(u = 1, 2, \ldots, n)$.

A VAE Perspective For FedCF

From the perspective of VAEs, we first sample a latent variable $\mathbf{z}_u \sim \mathcal{N}(0, \mathbf{I}_k)$, which is assumed to model the u-th user's decision logic of the preference in recommendation system. It is worth noting that the work (Liang et al. 2018) suggested that multi-nominal distribution might be an appropriate option for recommendation with implicit feedback.

Given the sampled variable $\mathbf{z}_u$, the user's preference probabilities for items can be generated by a nonlinear function $f_\theta : \mathbb{R}^k \to \mathbb{R}^m$, parameterized by θ, applied to $\mathbf{z}_u$:

$$\pi(\mathbf{z}_u) = f_\theta(\mathbf{z}_u), \tag{2}$$

s.t. $\mathbf{r}_u$ is drawn from $\mathbf{r}_u \sim \mathrm{Mult}(m_u, \pi(\mathbf{z}_u))$. To measure the correctness of the predicted user preference, we use log-likelihood of $\mathbf{r}_u$ conditioned on $\mathbf{z}_u$ defined as follows:

$$\log p_\theta(\mathbf{r}_u|\mathbf{z}_u) = \sum_{i=1}^{m} \mathbf{r}_{ui} \log \pi_i(\mathbf{z}_u). \tag{3}$$

We apply variational inference (Jordan et al. 1999) to approximate the intractable posterior distribution $p(\mathbf{z}_u|\mathbf{r}_u)$ using variational distribution to estimate θ for the function f_θ:

$$q_{\mathbf{\Phi}_u}(\mathbf{z}_u|\mathbf{r}_u) = \mathcal{N}(\mu_{\mathbf{\Phi}_u}(\mathbf{r}_u), \mathrm{diag}\{\sigma^2_{\mathbf{\Phi}_u}(\mathbf{r}_u)\}). \tag{4}$$

Given $k << \{n, m\}$, both of the vectors $\mu_{\mathbf{\Phi}_u}(\mathbf{r}_u)$ and $\sigma^2_{\mathbf{\Phi}_u}(\mathbf{r}_u)$ in Eq. (4) are k-dimensional outputs of the data-dependent function $g_{\mathbf{\Phi}_u}(\mathbf{r}_u) = [\mu_{\mathbf{\Phi}_u}(\mathbf{r}_u), \sigma_{\mathbf{\Phi}_u}(\mathbf{r}_u)] \in \mathbb{R}^{2k}$, where the parameter $\mathbf{\Phi}_u$ is used to capture the representation of item features on the u-th client $(u = 1, 2, \ldots, n)$.

Methodology

Framework

Fig. 1 shows the overall framework of FedDAE. For each client u, FedDAE inputs the interaction vector $\mathbf{r}_u$ into the global encoder q_φ, the local encoder q_{φ_u}, and the gating network h_{ψ_u}, separately. By multiplying the outputs of the dual encoders by the weights generated by h_{ψ_u} based on the user's interaction data and then summing them, FedDAE achieves adaptive additive personalization. The resulting output is then passed to the global decoder p_θ to reconstruct the user's interaction data $\hat{\mathbf{r}}_u$ $(u = 1, 2, \ldots, n)$. During the training phase, the global encoder and global decoder are collaboratively trained across clients, while the local encoder and gating network are trained locally.

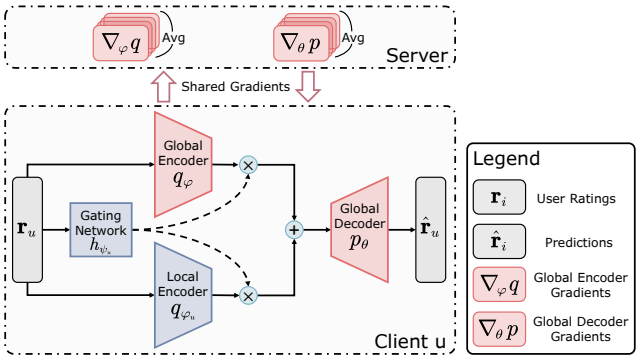

Figure 1: The framework of FedDAE.

Dual Encoders

We propose a dual-encoder mechanism to separately preserve shared knowledge across clients and client-specific personalized knowledge. Specifically, the encoder $q_{\mathbf{\Phi}_u}$ in Eq. 4 is implemented through a dual-encoder structure, consisting of a global encoder q_φ and a local encoder q_{φ_u} for each user u. The parameters φ and φ_u capture the globally shared representation of item features and the personalized representation specific to user u, respectively. This enables FedDAE to achieve personalized item feature representations while maintaining shared information.

Gating Network

To effectively combine the global and local representations, we use a gating network $h_{\psi_u}(\mathbf{r}_u) = [\omega_{u1}, \omega_{u2}] \in \mathbb{R}^2$, parameterized by ψ_u, which dynamically assigns weights ω_{u1} and ω_{u2} to the outputs of g_φ and g_{φ_u} based on the data $\mathbf{r}_u$.

Lemma 1 (Additivity of Gaussian distributions) *Given two independent Gaussian random variables X and Y, distributed as $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. Then $Z = w_1 X + w_2 Y$ follows a new Gaussian distribution $Z \sim N(w_1\mu_1 + w_2\mu_2, w_1^2\sigma_1^2 + w_2^2\sigma_2^2)$.*

Considering Lemma 1, since the latent variable $\mathbf{z}_u$ is drawn from $q_{\mathbf{\Phi}_u}(\mathbf{z}_u|\mathbf{r}_u)$ according to Eq. (4), the combined $\mu_{\mathbf{\Phi}_u}(\mathbf{r}_u)$ and $\sigma^2_{\mathbf{\Phi}_u}(\mathbf{r}_u)$ are then defined as follows:

$$\begin{aligned} \mu_{\mathbf{\Phi}_u}(\mathbf{r}_u) &= \omega_{u1} \cdot \mu_\varphi(\mathbf{r}_u) + \omega_{u2} \cdot \mu_{\varphi_u}(\mathbf{r}_u), \\ \sigma^2_{\mathbf{\Phi}_u}(\mathbf{r}_u) &= \omega^2_{u1} \cdot \sigma^2_\varphi(\mathbf{r}_u) + \omega^2_{u2} \cdot \sigma^2_{\varphi_u}(\mathbf{r}_u), \end{aligned} \tag{5}$$

where $\mathbf{\Phi}_u = \{\varphi, \varphi_u, \psi_u\}$. Through the gating network h_{ψ_u}, FedDAE can adaptively adjust the balance between shared and personalized information based on each client's interaction data, thereby achieving personalized FedCF.

Reconstruction and Prediction

We define the VAE loss for the FedCF problem as follows. By combining Eq. (3) and Eq. (4) to form a VAE, the evidence lower bound (ELBO) on the u-th client is given by:

$$\mathcal{L}_\beta(\mathbf{r}_u; \mathbf{\Phi}_u, \theta) = \mathbb{E}_{q_{\mathbf{\Phi}_u}(\mathbf{z}_u|\mathbf{r}_u)}[\log p_\theta(\mathbf{r}_u|\mathbf{z}_u)] - \beta \cdot \mathrm{KL}(q_{\mathbf{\Phi}_u}(\mathbf{z}_u|\mathbf{r}_u) \| p_\theta(\mathbf{z}_u)), \tag{6}$$

where $\mathrm{KL}(\cdot)$ is the Kullback-Leibler divergence, and $\beta \in [0, 1]$ is a hyperparameter controlling the strength of regularization, following the β-VAE (Higgins et al. 2017) . The first term in Eq. (6) is the reconstruction loss, which equals to $-\mathcal{L}_{recon}(\hat{\mathbf{r}}_u, \mathbf{r}_u)$, while the second term is for prior matching. By jointly maximizing Eq. (8) of all users, we obtain the reconstructed interactions $\hat{\mathbf{r}}_u$ for each user u $(u = 1, 2, \ldots, n)$. However, relying solely on sharing θ is insufficient for effective information sharing across clients, which may result in suboptimal recommendation performance in a federated setting. To address this issue, we propose a novel personalized FedCF method called FedDAE.

After sampling $\mathbf{z}_u$ from $q_{\mathbf{\Phi}_u}(\mathbf{z}_u|\mathbf{r}_u)$, it is decoded using the decoder p_θ defined in Eq. (3) to obtain the reconstructed interactions $\hat{\mathbf{r}}_u$. Combining Eq. (6), the objective of FedDAE is to maximize the ELBO to approximateas $\log p(\mathbf{r}_u)$ across all clients $(u = 1, 2, \ldots, n)$, which is defined as follows:

$$\max_{\varphi, \{\varphi_u\}, \{\psi_u\}, \theta} \sum_{u=1}^{n} \alpha_u \mathcal{L}_\beta(\mathbf{r}_u; \varphi, \varphi_u, \psi_u, \theta). \tag{7}$$

Here, α_u is the weight of the loss for the u-th client, used to balance the contribution of this client, satisfying $\sum_{u=1}^{n} \alpha_u = 1$. Once the objective function in Eq. (7) is optimized, we obtain the final prediction $\hat{\mathbf{r}}_u = \mu_{\mathbf{\Phi}_u}(\mathbf{r}_u)$, which will be used to personalize recommendations for the u-th user, suggesting items they might be interested in.

Algorithm

To optimize the objective function outlined in Eq. (7), we utilize an alternative optimization algorithm for training the FedDAE method. The overall workflow of this algorithm is summarized in several steps, as demonstrated in the Alg. 1.

The process begin by initializing φ and θ on the server, and all φ_u and ψ_u on their respective clients. For each communication round t, the server randomly selects a subset of clients to participate in the training, denoted as S_t. It is important to note that, as discussed in (Chai et al. 2020; Li, Long, and Zhou 2024), for privacy protection, no client should participate in training for two consecutive rounds. According to the line 4 in the $GlobalProcedure$, the latent variable $\mathbf{z}_u$ is sampled from the normal Gaussian distribution using the reparameterization trick. After receiving φ and θ, the selected clients call the $ClientUpdate$ function to update their local models with the learning rate η.

In the $ClientUpdate$ function, The local model updates are performed over E epochs (line 13). Once the update is completed, the accumulated gradients $\nabla_\varphi^{(u)}$ and $\nabla_\theta^{(u)}$, and the predicted scores $\hat{\mathbf{r}}_u$ are uploaded to the server for global aggregation. The server then apply the updated φ and θ to the next training round. After the training phase is completed, FedDAE use the predicted scores $\hat{\mathbf{R}}$ as the recommendation guide. On each client, FedDAE constructs a global encoder to model the global representation of item features and a local encoder to capture the personalized representation of features influenced by user preferences. The outputs of these encoders are weighted and combined through a gating network, achieving adaptive additive personalization in recommendations. Therefore, despite variations in users' interaction data due to their preferences, FedDAE ensures that data within a user adheres to the i.i.d. assumption. Additionally, the user-related local encoder is stored only locally on the client, protecting user privacy.

Algorithm 1: FedDAE

Input: $\mathbf{R}$, β, k, η, T, E
Initialize: φ, $\{\varphi_u\}$, $\{\psi_u\}$, θ
GlobalProcedure:
1: **for** $t = 1, 2, \ldots, T$ **do**
2: $S_t \leftarrow$ randomly select n_s from n clients
3: **for all** client index $u \in S_t$ **do**
4: Sample $\epsilon \sim \mathcal{N}(0, \mathbf{I}_k)$ and compute $\mathbf{z}_u$ via reparameterization trick;
5: $\hat{\mathbf{r}}_u, \nabla_\varphi^{(u)}, \nabla_\theta^{(u)} \leftarrow$ ClientUpdate(φ, θ);
6: **end for**
7: $\varphi = \varphi - \frac{\eta}{n_s} \sum_{u \in S_t} \nabla_\varphi^{(u)}$;
8: $\theta = \theta - \frac{\eta}{n_s} \sum_{u \in S_t} \nabla_\theta^{(u)}$;
9: **end for**
10: **return:** $\hat{\mathbf{R}} = [\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n]^T$
ClientUpdate:
1: $\nabla_\varphi = 0, \nabla_\theta = 0$;
2: **for** $e = 1, 2, \ldots, E$ **do**
3: Compute gradients $\nabla_\varphi \mathcal{L}_\beta$, $\nabla_{\varphi_u} \mathcal{L}_\beta$, $\nabla_{\psi_u} \mathcal{L}_\beta$, $\nabla_\theta \mathcal{L}_\beta$ according to Eq. (7);
4: Use gradient descent to update φ, φ_u, ψ_u, θ with η;
5: $\nabla_\varphi = \nabla_\varphi + \nabla_\varphi \mathcal{L}_\beta$;
6: $\nabla_\theta = \nabla_\theta + \nabla_\theta \mathcal{L}_\beta$;
7: Compute $\hat{\mathbf{r}}_u$;
8: **end for**
9: **return:** $\hat{\mathbf{r}}_u$**,** ∇_φ**,** ∇_θ

Discussions

Matrix Factorization

In recommendation systems, MF-based CF methods typically decompose the interaction data matrix into an item embedding that preserves shared knowledge and a user embedding that retains personalized knowledge. In our proposed method, we model the data distribution of the interaction matrix and use the weights of the dual encoders to capture both the globally shared knowledge across clients and the personalized knowledge for each user. This allows our method, FedDAE, to retain more complex information, rather than being limited to the form of user-item vectors.

Complexity Analysis

Given n users and m items, the input dimension of the encoder is m. Assuming each encoder on the client side con-

sists of L fully connected layers with a hidden layer dimension of $d >> k$, the time complexity of the encoder is $O(Lmd)$, and the space complexity is $O(L(m+1)d)$. Similarly, assuming the decoder has L' fully connected layers with the same hidden layer dimension d, the time and space complexities of the decoder are $O(L'md)$ and $O(L'(m+1)d)$ respectively. The reparameterization involves generating the mean and variance of the latent subspace $\mathbf{z}_u$ and sampling using a standard normal distribution, with a time complexity of $O(k)$. Additionally, the gating network on each client requires $O(2m)$ time and space complexity. Thus, the overall time complexity is $O(n(2L+L')md)$. Due to each client in FedDAE having an independent gating network, an independent encoder, a globally shared encoder, and a globally shared decoder, the overall space complexity is $O(((n+1)L+L')(m+1)d)$.

Privacy-preservation Enhancement

The proposed framework adopts a decentralized architecture from the FL scheme, which significantly reduces the risk of privacy leaks by maintaining data locality, which generally performs well in a trusted environment. Similar to most FedCF methods (Chai et al. 2020; Zhang et al. 2023a; Li, Long, and Zhou 2024), our approach only transmits the gradients of global model parameters, $\nabla_{\varphi}^{(u)}\mathcal{L}_{\beta}$ and $\nabla_{\theta}^{(u)}p$, without sharing any raw data with third parties. Additionally, if in an untrusted environment, the proposed method can be easily combined with other advanced privacy-preserving techniques, such as differential privacy (Abadi et al. 2016) or randomly cropping gradients, to further strengthen user privacy guarantees. In our experiment, we perform an ablation study to evaluate the effectiveness of privacy protection.

Experiments

Datasets

We perform an extensive experimental analysis to assess the performance of FedDAE on four widely utilized datasets: MovieLens-100K (ML-100K), MovieLens-1M (ML-1M), Amazon-Instant-Video (Video), and QB-article. Table 1 provides the statistical details of the datasets. Each dataset exhibits a high degree of sparsity, with the percentage of observed ratings compared to the total possible ratings (#Users $\times$ #Items) exceeding 90%. The first three datasets contain explicit ratings ranging from 1 to 5. Since FedDAE focuses on generating recommendation predictions for data with implicit feedback, any rating above 0 in these datasets is considered a positive interaction by the user and is assigned 1. The QB-article dataset is an implicit feedback dataset that records user click behavior. In each dataset, we only include users who rated at least 10 items. All datasets used are publicly available, and details are provided in the appendix.

Baselines

The efficacy of FedDAE is comparatively evaluated against several cutting-edge approaches in both centralized and federated environments for validation:

Mult-VAE (Liang et al. 2018): A variational autoencoder model that uses a multinomial distribution as the likelihood function to generate latent representations of users, suitable for collaborative filtering tasks with implicit feedback data in recommendation systems.

RecVAE (Shenbin et al. 2020): significantly improves the performance of Mult-VAE by introducing a composite prior distribution and an alternating training method,

LightGCN (He et al. 2020): enhances recommendation performance to a new level while retaining the advantages of GCN models by simplifying the design of graph convolutional networks (GCNs) and removing nonlinear transformations and weight matrices,

FedVAE (Polato 2021): extends Mult-VAE to the FL framework, achieving collaborative filtering by only transmitting model gradients between clients and the server without exposing the raw user data, ensuring user privacy and security.

PFedRec (Zhang et al. 2023a): achieves the personalization of user information and item features by introducing a dual personalization in the FL framework, enabling the recommendation system to adapt to the needs of different users.

FedRAP (Li, Long, and Zhou 2024): enhances the performance of recommendation systems by applying dual personalization of user and item information in federated learning while retaining the shared parts of item information, particularly excelling in handling heterogeneous data.

The implementations of these baselines are publicly available in their respective papers. Additionally, we implement a central version of FedDAE, named **CentDAE**, to explore the performance of a VAE with dual encoders.

Datasets	#Ratings	#Users	#Items	Sparsity
ML-100k	100,000	943	1,682	93.70%
ML-1M	1,000,209	6,040	3,706	95.53%
Video	23,181	1,372	7,957	99.79%
QB-article	266,356	24,516	7,455	99.81%

Table 1: The statistical information of the used datasets.

Experimental Setting

We refer to previous work (He et al. 2017; Zhang et al. 2023a; Li, Long, and Zhou 2024), randomly selecting 4 negative samples for each positive sample in the training set, and using the leave-one-out strategy to validate the methods. In this work, we set $h_{\psi_u}(\mathbf{r}_u) = \mathrm{softmax}(\mathbf{r}_u \cdot \psi_u)$ and thus $w_{u1} + w_{u2} = 1$. We perform hyper-parameter tuning for FedDAE and select the learning rate η from $\{10^i | i = -8, ..., -1\}$. The Adam optimizer (Kingma and Ba 2014) is applied to update FedDAE's parameters. Referring to previous work (Liang et al. 2018; Polato 2021), we gradually anneal $\beta = 1$. To ensure fairness, we fix the latent embedding dimension to 256 and the training batch size to 2048 for all methods. The number of layers for methods with hidden layers is fixed at 3. We set $\alpha_u = 1/n$ for the average schemes of all userd federated methods. Following Mult-VAE, FedDAE also employs a dropout (Srivastava et al. 2014) layer before inputting $\mathbf{r}_u$ into the encoders to reduce the risk of over-fitting and uses the reparametrization trick (Kingma and Welling 2013) to eliminate the stochasticity caused by sampling $\mathbf{z}_u$, which allows the model to be optimized using gradient descent. In the main experiments,

	Method	ML-100K		ML-1M		Video		QB-Article	
		HR@20	NDCG@20	HR@20	NDCG@20	HR@20	NDCG@20	HR@20	NDCG@20
Central	MultiVAE	0.1093 ± 0.0143	0.0436 ± 0.0065	0.0682 ± 0.0016	0.0276 ± 0.0005	0.0689 ± 0.0088	0.0245 ± 0.0028	0.0129 ± 0.0025	0.0061 ± 0.0008
	RecVAE	0.1526 ± 0.0046	0.0590 ± 0.0017	0.1067 ± 0.0029	0.0409 ± 0.0010	0.0693 ± 0.0041	0.0257 ± 0.0028	0.0195 ± 0.0075	0.0064 ± 0.0020
	LightGCN	0.1761 ± 0.0166	0.0953 ± 0.0128	0.1207 ± 0.0015	0.0677 ± 0.0014	0.0761 ± 0.0056	0.0323 ± 0.0013	0.0204 ± 0.0067	0.0070 ± 0.0039
	CentDAE	0.1119 ± 0.0024	0.0453 ± 0.0010	0.0702 ± 0.0008	0.0275 ± 0.0004	0.0702 ± 0.0084	0.0254 ± 0.0034	0.0185 ± 0.0033	0.0064 ± 0.0007
Federated	FedVAE	0.1200 ± 0.0040	0.0501 ± 0.0032	0.0639 ± 0.0012	0.0250 ± 0.0006	0.0634 ± 0.0128	0.0227 ± 0.0036	0.0116 ± 0.0031	0.0045 ± 0.0013
	PFedRec	0.0541 ± 0.0054	0.0115 ± 0.0026	0.0600 ± 0.0002	0.0231 ± 0.0000	0.0138 ± 0.0064	0.0026 ± 0.0012	0.0065 ± 0.0002	0.0033 ± 0.0000
	FedRAP	0.0626 ± 0.0032	0.0122 ± 0.0017	0.0625 ± 0.0015	0.0240 ± 0.0003	0.0142 ± 0.0097	0.0028 ± 0.0019	0.0116 ± 0.0026	0.0035 ± 0.0010
	FedDAE	**0.1232 ± 0.0023**	**0.0503 ± 0.0014**	**0.0650 ± 0.0039**	**0.0261 ± 0.0011**	**0.0654 ± 0.0080**	**0.0230 ± 0.0029**	**0.0124 ± 0.0030**	**0.0052 ± 0.0014**

Table 2: Experimental results on HR@20 and NDCG@20 shown in percentages on four real-world datasets. **Central** and **Federated** represent centralized and federated methods, respectively. The best results are highlighted in boldface.

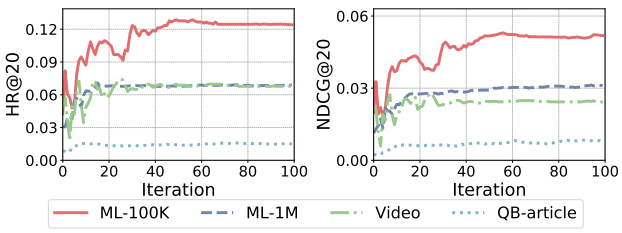

Figure 2: The visualization of FedDAE's convergence and efficiency across the four used datasets.

	FedDAE	FedDAE$_{w=0.25}$	FedDAE$_{w=0.5}$	FedDAE$_{w=0.75}$
HR@20	0.1287	0.1213 (5.75%↓)	0.1245 (3.26%↓)	0.1191 (7.46%↓)
NDCG@20	0.0524	0.0506 (3.44%↓)	0.0513 (2.10%↓)	0.0469 (10.50%↓)

Table 3: Performance Comparison of FedDAE and its fixed weight variants FedDAE$_w$ ($w = 0.25, 0.5, 0.75$) on the ML-100K dataset. The values in parentheses indicate the percentage decrease compared to the FedDAE method.

all clients participated in the training for FedVAE, PFedRec, FedRAP, and FedDAE, and no additional privacy protection measures are applied. In addition, the number of local epoch is set to 5 for these federated methods, and we do not use any pre-training strategies for any methods.

Evaluation

We evaluate the prediction performance in the experiments using two widely used metrics: *Hit Rate* (HR@K) and *Normalized Discounted Cumulative Gain* (NDCG@K). These criteria were formally defined in the work (He et al. 2015). HR measures hit rate to show whether relevant items are recommended, while NDCG accounts for ranking quality, emphasizing user experience by ranking items of greater interest higher. In this study, we set K = 20 and repeat all experiments five times to report the mean and standard deviations.

Comparison Analysis

Table 2 presents the experimental results of all baseline methods and FedDAE on four widely used recommendation datasets, demonstrating that FedDAE outperforms all federated methods and achieves performance close to centralized methods. This is attributed to FedDAE's ability to adaptively and individually model each user's data distribution, retaining shared information about item features while better integrating user-specific item representations. This indicates that FedDAE has learned fine-grained personalized features, effectively leveraging user preferences.

Additionally, to study FedDAE's convergence, we evaluated its performance over 100 iterations on all datasets and visualized the HR@20 and NDCG@20 results, as shown in Fig. 2. Insights from Table 2 and Fig. 2 reveal that the sparsity and scale of the datasets significantly affect FedDAE's performance. On datasets with lower sparsity and moderate scale, such as ML-100K and ML-1M, FedDAE shows rapid convergence and better results. In contrast, on datasets with higher sparsity, like Video and QB-article, the model's convergence speed and final performance are noticeably slower. This indicates that additional strategies may be needed to improve the model's learning ability when dealing with highly sparse datasets. Furthermore, while the Adam optimizer helps quickly find optimal model parameters and enhances recommendation performance, it also causes more noticeable fluctuations in FedDAE's convergence curves. These fluctuations might be due to the heterogeneity and noise in the datasets. A theoretical convergence analysis of the convergence is provided in our appendix.

Ablation Study

To investigate the impact of each component on the model's performance, we propose a variant of the FedDAE method, named **FedDAE**$_w$. It uses a fixed weight ww for each client to weigh the output of the global encoder. Since the weight of the local encoder is complementary to w (i.e., $1 - w$), the local encoder's weight is also fixed. For more ablation studies, please refer to our appendix.

Abalation Study on Adaptive Weights. To explore the impact of the gating network output weight on each client in FedDAE, we compared the performance of FedDAE$_w$ with different fixed weights ($w = 0.25, 0.5, 0.75$) and FedDAE on the ML-100K dataset. Table 3 shows that when the weight $w = 0.5$, FedDAE$_w$ performs most similarly to FedDAE. However, as the weight increases to 0.75, performance significantly decreases, indicating that personalized information plays a crucial role in FedDAE. The lack of personalized information results in a greater decline in recommendation performance compared to the lack of shared information. Overall, this ablation study demonstrates the effectiveness of the gating network in FedDAE.

Ablation Study on Item Representation. To visually demonstrate the differentiated capability of FedDAE en-

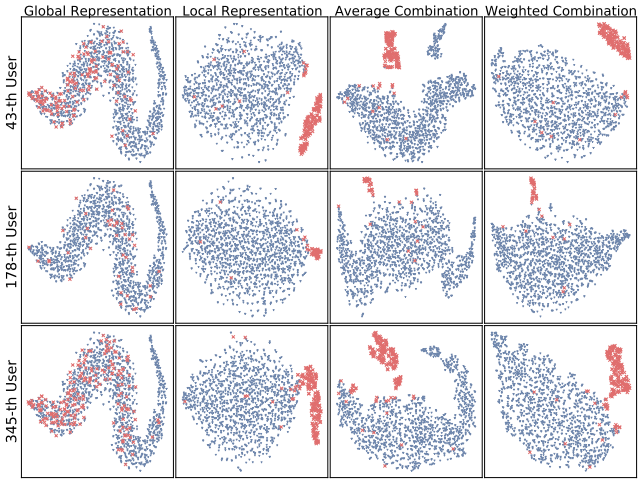

Figure 3: The t-SNE visualization of item features learned by FedDAE on the ML-100K dataset illustrates the representations across different users. In the visualization, red indicates items that users have interacted with, while blue indicates items they have not interacted with. The global representation is generated by FedDAE's global encoder, and the local representation is produced by the user-specific encoder. **Average Combination** refers to the simple average of the global and local representations, while **Weighted Combination** reflects the weighted combination based on the gating network outputs tailored to the client data. FedDAE's adaptive personalization enhances its ability to distinguish between items users have interacted with and those they haven't, leading to improved recommendation performance.

coders in modeling item features, we selected three local encoders $\{q_{\varphi_u}\}_{u=43,178,345}$ from different users and the global encoder q_φ, all learned from the ML-100K dataset. The weights of all layers of each encoder were multiplied and then visualized. Since this study primarily focuses on implicit feedback recommendation, where each item is either interacted with by the user or not, We employe t-SNE (Van der Maaten and Hinton 2008) to map these item embeddings into a two-dimensional space, with the normalized results shown in Fig. 3. In these visualizations, blue represents items not interacted with by the user, while red represents items that have been interacted with. From the first column of Fig. 3, we can see that in the global encoder's modeled item representation, interacted and non-interacted items are mixed together, indicating that q_φ only models the shared features of items. The second column illustrates that each user's local encoder can divide the item features into two distinct clusters, proving that the local encoder q_{φ_u} can learn user-specific personalized features. The third column shows the average combination of global and local representations, but its boundaries are not as distinct compared to the weighted combination of FedDAE shown in the fourth column. The weighted combination, based on the output of the gating network h_ψ from user interaction data $\mathbf{r}_u$, more clearly differentiates the two clusters. This demonstrates that FedDAE's adaptive additive personalization can more effectively represent item features, supporting its superior performance in personalized recommendation.

κ^2	Metric	FedDAE	FedDAE$_{w=0.25}$	FedDAE$_{w=0.5}$	FedDAE$_{w=0.75}$
0.2	HR@20	0.1255	0.1213 (3.35%↓)	0.1245 (0.80%↓)	0.1149 (8.45%↓)
	NDCG@20	0.0505	0.0483 (4.36%↓)	0.0486 (3.76%↓)	0.0476 (5.74%↓)
0.5	HR@20	0.1245	0.1181 (5.14%↓)	0.1213 (2.57%↓)	0.1170 (6.02%↓)
	NDCG@20	0.0494	0.0479 (3.04%↓)	0.0481 (2.63%↓)	0.0470 (4.86%↓)
0.8	HR@20	0.1191	0.1170 (1.76%↓)	0.1181 (0.84%↓)	0.1160 (2.60%↓)
	NDCG@20	0.0489	0.0473 (3.27%↓)	0.0486 (0.61%↓)	0.0464 (5.11%↓)
1	HR@20	0.1191	0.1085 (8.90%↓)	0.1096 (7.98%↓)	0.1085 (8.90%↓)
	NDCG@20	0.0479	0.0453 (5.43%↓)	0.0466 (2.71%↓)	0.0450 (6.05%↓)

Table 4: Performance comparison of FedDAE and FedDAE$_w$ variants under Different Noise Variance Levels on the ML-100K Dataset, showing the performance metrics HR@20 and NDCG@20 for FedDAE and its variants with fixed weights ($w = 0.25, 0.5, 0.75$) under different noise variance levels (0.2, 0.5, 0.8, 1). The values in parentheses indicate the percentage decrease compared to FedDAE.

Abalation Study on Privacy Protection. To evaluate the impact of additional privacy protection measures on the recommendation performance of FedDAE and FedDAE$_w$, we introduce noise into the gradients uploaded by users and test these methods on the ML-100K dataset across various noise levels ($\kappa^2 = 0, 0.2, 0.5, 0.8, 1$). Table 4 shows that while performance declines as noise increases, FedDAE consistently outperforms other variants at all noise levels, likely due to its adaptive balancing of global and local encoder outputs. FedDAE$_{w=0.5}$ ranks second, demonstrating strong robustness across different noise levels. In contrast, FedDAE$_{w=0.25}$ performs worse than both FedDAE and FedDAE$_{w=0.5}$ but better than FedDAE$_{w=0.75}$. As the weight parameter w increases, FedDAE becomes more reliant on the global encoder, which may explain why FedDAE$_{w=0.75}$ performs the worst under all noise conditions. These results highlight the importance of balancing the weights of the dual encoders, especially when implementing privacy protections.

Conclusion

This paper first revisits FedCF from the perspective of VAEs and proposes a novel personalized FedCF method called FedDAE. FedDAE constructs a VAE model with dual encoders and a global decoder for each client, capturing user-specific representations of item features while preserving globally shared information. The outputs of the two encoders are dynamically weighted through a gating network based on user interaction data, achieving adaptive additive personalization. Experimental results on four widely used recommendation datasets demonstrate that FedDAE outperforms existing FedCF methods and various ablation baselines, showcasing its ability to provide efficient personalized recommendations while protecting user privacy. Additionally, our research reveals that the sparsity and scale of the datasets significantly impact the performance of FedDAE, suggesting that highly sparse datasets require additional strategies to enhance the model's learning capability.

References

Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H. B.; Mironov, I.; Talwar, K.; and Zhang, L. 2016. Deep learning with differential privacy. In *SIGSAC*, 308–318.

Aggarwal, C. C.; and Aggarwal, C. C. 2016. Model-based collaborative filtering. *Recommender systems: the textbook*, 71–138.

Ammad-Ud-Din, M.; Ivannikova, E.; Khan, S. A.; Oyomno, W.; Fu, Q.; Tan, K. E.; and Flanagan, A. 2019. Federated collaborative filtering for privacy-preserving personalized recommendation system. *arXiv preprint arXiv:1901.09888*.

Aneja, J.; Schwing, A.; Kautz, J.; and Vahdat, A. 2021. A contrastive learning approach for training variational autoencoder priors. *NeuRIPS*, 34: 480–493.

Arivazhagan, M. G.; Aggarwal, V.; Singh, A. K.; and Choudhary, S. 2019. Federated learning with personalization layers. *arXiv preprint arXiv:1912.00818*.

Chai, D.; Wang, L.; Chen, K.; and Yang, Q. 2020. Secure federated matrix factorization. *IEEE IS*, 36(5): 11–20.

Chen, S.; Long, G.; Jiang, J.; and Zhang, C. 2024. Personalized Adapter for Large Meteorology Model on Devices: Towards Weather Foundation Models. *arXiv preprint arXiv:2405.20348*.

Chen, S.; Long, G.; Shen, T.; Jiang, J.; and Zhang, C. 2023. Federated Prompt Learning for Weather Foundation Models on Devices. *arXiv preprint arXiv:2305.14244*.

Collins, L.; Hassani, H.; Mokhtari, A.; and Shakkottai, S. 2021. Exploiting shared representations for personalized federated learning. In *ICML*, 2089–2099. PMLR.

Deng, Y.; Kamani, M. M.; and Mahdavi, M. 2020. Adaptive personalized federated learning. *arXiv preprint arXiv:2003.13461*.

Flanagan, A.; Oyomno, W.; Grigorievskiy, A.; Tan, K. E.; Khan, S. A.; and Ammad-Ud-Din, M. 2020. Federated multi-view matrix factorization for personalized recommendations. In *ECML-KDD*, 324–347. Springer.

Gao, R.; Hou, X.; Qin, J.; Chen, J.; Liu, L.; Zhu, F.; Zhang, Z.; and Shao, L. 2020. Zero-VAE-GAN: Generating unseen features for generalized and transductive zero-shot learning. *TIP*, 29: 3665–3680.

He, X.; Chen, T.; Kan, M.-Y.; and Chen, X. 2015. Trirank: Review-aware explainable recommendation by modeling aspects. In *CIKM*, 1661–1670.

He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *SIGIR*, 639–648.

He, X.; Liao, L.; Zhang, H.; Nie, L.; Hu, X.; and Chua, T.-S. 2017. Neural collaborative filtering. In *WWW*, 173–182.

Hegedűs, I.; Danner, G.; and Jelasity, M. 2019. Decentralized recommendation based on matrix factorization: A comparison of gossip and federated learning. In *ECML-KDD*, 317–332. Springer.

Higgins, I.; Matthey, L.; Pal, A.; Burgess, C. P.; Glorot, X.; Botvinick, M. M.; Mohamed, S.; and Lerchner, A. 2017. beta-vae: Learning basic visual concepts with a constrained variational framework. *ICLR*, 3.

Jordan, M. I.; Ghahramani, Z.; Jaakkola, T. S.; and Saul, L. K. 1999. An introduction to variational methods for graphical models. *Machine learning*, 37: 183–233.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kingma, D. P.; and Welling, M. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Klushyn, A.; Chen, N.; Kurle, R.; Cseke, B.; and van der Smagt, P. 2019. Learning hierarchical priors in vaes. *NeuRIPS*, 32.

Ko, H.; Lee, S.; Park, Y.; and Choi, A. 2022. A survey of recommendation systems: recommendation models, techniques, and application fields. *Electronics*, 11(1): 141.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer*, 42(8): 30–37.

Koren, Y.; Rendle, S.; and Bell, R. 2021. Advances in collaborative filtering. *Recommender systems handbook*, 91–142.

Lee, W.; Song, K.; and Moon, I.-C. 2017. Augmented variational autoencoders for collaborative filtering with auxiliary information. In *CIKM*, 1139–1148.

Li, T.; Hu, S.; Beirami, A.; and Smith, V. 2021. Ditto: Fair and robust federated learning through personalization. In *ICML*, 6357–6368. PMLR.

Li, Z.; and Long, G. 2024. Navigating the Future of Federated Recommendation Systems with Foundation Models. *arXiv preprint arXiv:2406.00004*.

Li, Z.; Long, G.; Jiang, J.; and Zhang, C. 2024. Personalized Item Representations in Federated Multimodal Recommendation. *arXiv preprint arXiv:2410.08478*.

Li, Z.; Long, G.; and Zhou, T. 2024. Federated Recommendation with Additive Personalization. In *ICLR*.

Liang, D.; Krishnan, R. G.; Hoffman, M. D.; and Jebara, T. 2018. Variational autoencoders for collaborative filtering. In *WWW*, 689–698.

Liang, F.; Pan, W.; and Ming, Z. 2021. Fedrec++: Lossless federated recommendation with explicit feedback. In *AAAI*, volume 35, 4224–4231.

Liang, S.; Pan, Z.; liu, w.; Yin, J.; and de Rijke, M. 2024. A Survey on Variational Autoencoders in Recommender Systems. *ACM Computing Surveys*.

Lin, G.; Liang, F.; Pan, W.; and Ming, Z. 2020a. Fedrec: Federated recommendation with explicit feedback. *IEEE IS*, 36(5): 21–30.

Lin, Y.; Ren, P.; Chen, Z.; Ren, Z.; Yu, D.; Ma, J.; Rijke, M. d.; and Cheng, X. 2020b. Meta matrix factorization for federated rating predictions. In *SIGIR*, 981–990.

Luo, S.; Xiao, Y.; and Song, L. 2022. Personalized Federated Recommendation via Joint Representation Learning, User Clustering, and Model Adaptation. In *CIKM*, 4289–4293.

Luo, S.; Xiao, Y.; Zhang, X.; Liu, Y.; Ding, W.; and Song, L. 2023. Perfedrec++: Enhancing personalized federated recommendation with self-supervised pre-training. *arXiv preprint arXiv:2305.06622*.

McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; and y Arcas, B. A. 2017. Communication-efficient learning of deep networks from decentralized data. In *AISTATS*, 1273–1282. PMLR.

Mehta, R.; and Rana, K. 2017. A review on matrix factorization techniques in recommender systems. In *CSCITA*, 269–274. IEEE.

Nazabal, A.; Olmos, P. M.; Ghahramani, Z.; and Valera, I. 2020. Handling incomplete heterogeneous data using vaes. *PR*, 107: 107501.

Perifanis, V.; and Efraimidis, P. S. 2022. Federated neural collaborative filtering. *KBS*, 242: 108441.

Polato, M. 2021. Federated variational autoencoder for collaborative filtering. In *IJCNN*, 1–8. IEEE.

Shen, J.; Zhou, T.; and Chen, L. 2020. Collaborative filtering-based recommendation system for big data. *IJCSE*, 21(2): 219–225.

Shenbin, I.; Alekseev, A.; Tutubalina, E.; Malykh, V.; and Nikolenko, S. I. 2020. Recvae: A new variational autoencoder for top-n recommendations with implicit feedback. In *WSDM*, 528–536.

Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; and Salakhutdinov, R. 2014. Dropout: a simple way to prevent neural networks from overfitting. *JMLR*, 15(1): 1929–1958.

T Dinh, C.; Tran, N.; and Nguyen, J. 2020. Personalized federated learning with moreau envelopes. *NeuRIPS*, 33: 21394–21405.

Tan, A. Z.; Yu, H.; Cui, L.; and Yang, Q. 2022. Towards personalized federated learning. *TNNLS*.

Tan, Y.; Liu, Y.; Long, G.; Jiang, J.; Lu, Q.; and Zhang, C. 2023. Federated learning on non-iid graphs via structural knowledge sharing. In *AAAI*, volume 37, 9953–9961.

Tomczak, J.; and Welling, M. 2018. VAE with a VampPrior. In *AISTATS*, 1214–1223. PMLR.

Tran, N.-T.; and Lauw, H. W. 2024. Learning Multi-Faceted Prototypical User Interests. In *The Twelfth International Conference on Learning Representations*.

Van der Maaten, L.; and Hinton, G. 2008. Visualizing data using t-SNE. *JMLR*, 9(11).

Voigt, P.; and Von dem Bussche, A. 2017. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 10(3152676): 10–5555.

Wang, Y.; Zhang, H.; Liu, Z.; Yang, L.; and Yu, P. S. 2022. Contrastvae: Contrastive variational autoencoder for sequential recommendation. In *CIKM*, 2056–2066.

Xie, Z.; Liu, C.; Zhang, Y.; Lu, H.; Wang, D.; and Ding, Y. 2021. Adversarial and contrastive variational autoencoder for sequential recommendation. In *WWW*, 449–459.

Yan, P.; and Long, G. 2023. Personalization disentanglement for federated learning. In *ICME*, 318–323. IEEE.

Yang, L.; Tan, B.; Zheng, V. W.; Chen, K.; and Yang, Q. 2020. Federated recommendation systems. *Federated Learning: Privacy and Incentive*, 225–239.

Yang, Y.; Long, G.; Shen, T.; Jiang, J.; and Blumenstein, M. 2024. Dual-Personalizing Adapter for Federated Foundation Models. *arXiv preprint arXiv:2403.19211*.

Yu, X.; Zhang, X.; Cao, Y.; and Xia, M. 2019. VAEGAN: A collaborative filtering framework based on adversarial variational autoencoders. In *IJCAI*, volume 19, 4206–4212.

Yuan, W.; Qu, L.; Cui, L.; Tong, Y.; Zhou, X.; and Yin, H. 2023. HeteFedRec: Federated recommender systems with model heterogeneity. *arXiv preprint arXiv:2307.12810*.

Zhang, C.; Long, G.; Zhou, T.; Yan, P.; Zhang, Z.; Zhang, C.; and Yang, B. 2023a. Dual personalization on federated recommendation. *arXiv preprint arXiv:2301.08143*.

Zhang, C.; Long, G.; Zhou, T.; Zhang, Z.; Yan, P.; and Yang, B. 2024a. When Federated Recommendation Meets Cold-Start Problem: Separating Item Attributes and User Interactions. In *WWW*, 3632–3642.

Zhang, H.; Li, H.; Chen, J.; Cui, S.; Yan, K.; Wuerkaixi, A.; Zhou, X.; Shen, Z.; and Li, Y. 2024b. Beyond Similarity: Personalized Federated Recommendation with Composite Aggregation. *arXiv preprint arXiv:2406.03933*.

Zhang, H.; Liu, H.; Li, H.; and Li, Y. 2024c. Transfr: Transferable federated recommendation with pre-trained language models. *arXiv preprint arXiv:2402.01124*.

Zhang, H.; Luo, F.; Wu, J.; He, X.; and Li, Y. 2023b. LightFR: Lightweight federated recommendation with privacy-preserving matrix factorization. *TIS*, 41(4): 1–28.

Zhang, H.; Zhou, X.; Shen, Z.; and Li, Y. 2024d. PrivFR: Privacy-Enhanced Federated Recommendation With Shared Hash Embedding. *TNNLS*.

Zhang, J.; and Jiang, Y. 2021. A vertical federation recommendation method based on clustering and latent factor model. In *EIECS*, 362–366. IEEE.

Zhu, Z.; Si, S.; Wang, J.; and Xiao, J. 2022. Cali3f: Calibrated fast fair federated recommendation system. In *IJCNN*, 1–8. IEEE.