Towards Multi-Platform Recommendation via Cost-Effective Vertical Federated Learning

Wenjie Li*
Tsinghua University
Shenzhen, China
liwj20@mails.tsinghua.edu.cn

Abstract

User behavior data is naturally distributed across multiple platforms, each capturing complementary aspects of user interests. Integrating these signals can enable more comprehensive user modeling and improve recommendation quality. However, direct raw-data sharing across platforms is often infeasible due to data privacy regulations and commercial confidentiality, posing a major barrier to practical multi-platform recommendation. To overcome this barrier, Vertical Federated Learning (VFL) enables platforms to share latent knowledge from user data through privacy-preserving distributed learning, without directly exposing raw data. However, its real-world deployment remains limited by three key challenges: *1) the restricted data scope* of training, *2) the high cost of distributed inference*, and *3) the partial user benefits* for overlapped users only.

This doctoral research aims to address these challenges and move toward practical multi-platform recommendation via cost-effective VFL. Specifically, we develop a series of methods corresponding to these challenges: *1) fMPD*, a fine-grained self-supervised pretraining method that extends data utilization scope to unlabeled data; *2) JPL*, a privileged distillation technique that supports efficient standalone inference while preserving the benefits of federated learning; and *3) ReFer*, a retrieval-enhanced federated recommendation framework that enables full-user-set inference and groupwise user benefits. Experiments on public and industrial datasets validate the effectiveness of these methods and demonstrate their potential in implementing practical multi-platform recommendation. Future work will further connect VFL-based recommendation with generative recommenders and large language models.

Keywords

Recommender Systems, Vertical Federated Learning, Self-supervised Learning, Knowledge Distillation, Retrieval Augmentation

1 Introduction

Recommendation systems are now ubiquitous, capturing user behaviors across various service platforms that reflect diverse user interests. Using this multi-platform data collectively can achieve comprehensive user modeling. However, intrinsic data isolation, privacy laws (GDPR [22]), and commercial confidentiality make direct data sharing impossible, a problem well-known as the "***isolated data island problem***" [26].

To tackle this problem, vertical federated learning [4, 14] has been proposed and explored in various recommendation tasks [23, 25]. However, due to the restricted number of overlapped samples across parties [17] and the distributed nature of the split neural network model, existing works suffer from following challenges:

- ***(Challenge 1)* Limited Training Data Scope**: Aligned users for dissimilar businesses are often limited and constitute only a small portion of the user population. This reduced training set size can increase the risk of overfitting and result in low-quality embeddings and hidden representations, especially in sparse high-dimensional recommendation datasets.
- ***(Challenge 2)* High-Cost Distributed Serving**: The inference process of VFL models incurs extra latency costs (due to cross-party feature transmission and security enhancement operations) and poses new system design challenges (arising from inconsistent network conditions and computational capabilities of different parties). These challenges make it difficult for a federated

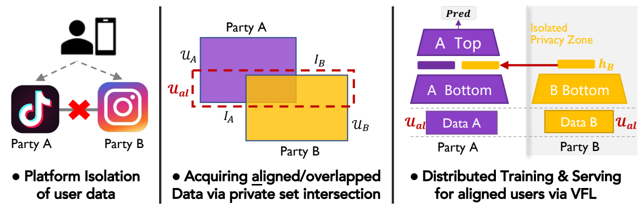

Figure 1: VFL enables multi-platform recommendation without direct data sharing, but remains limited by restricted data perception scope, costly distributed inference, and benefits confined to overlapped users.

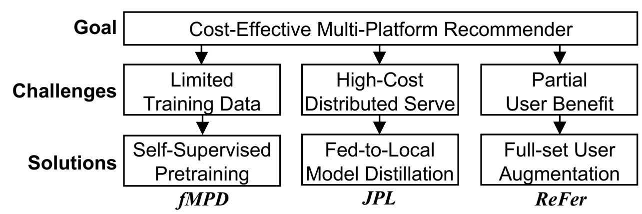

Figure 2: Research map of cost-effective multi-platform recommendation. Three deployment barriers are addressed by fMPD for broader data utilization, JPL for local serving, and ReFer for full-set user benefit.

*This doctoral research is supervised by Prof. Shu-Tao Xia.

CIKM '26, TBD

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

inference system to meet the high throughput and real-time latency requirements of advertising systems (million-scale peak QPS and 10 ∼ 100 ms processing time per request [20]). These obstacles may render the federation infeasible or excessively costly for participants.

- ***(Challenge 3)* Partial User Benefits**: The intrinsic field missing in passive parties makes it infeasible for a federated model to train on or make predictions for unaligned users. Thus, vanilla VFL can only bring benefits to aligned users, largely undermining the practicability of VFL. If a participant has more unaligned users, or places greater emphasis on unaligned users in their business, joining the federation is not cost-effective.

Contributions. To address these challenges, we integrate *self-supervised pre-training*[11], *privileged distillation*[10], and *retrieval augmentation*[9] to improve the return on investment of federation. As illustrated in Figure 2, we develop three connected methods: *(i)* fMPD uses fine-grained self-supervised perturbation–detection to learn transferable cross-party representations from unlabeled overlapped data; *(ii)* JPL studies a Semi-VFL setting that uses federated training but supports independent local inference; and *(iii)* ReFer extends vertical federated recommendation to the full user set with retrieval augmentation.

Practical Implications. These methods are suitable for different deployment conditions, while remaining compatible with each other. When abundant unlabeled overlapped data are available, fMPD can further exploit such data through self-supervised pre-training to strengthen federated encoders. When online cross-party serving is too expensive or infrastructure-constrained, JPL can preserve the benefits of federated training while enabling low-cost standalone inference. When sufficient serving resources are available and the goal is to maximize the return of federation, ReFer can robustly extend federated benefits to the full user set through retrieval-augmented federated serving.

Beyond these typical conditions, the three methods can be flexibly combined in practical deployments. For example, pretraining can serve as initialization for subsequent retrieval augmentation or distillation; retrieval augmentation can expand the available data scope for both pretraining and distillation; and distillation can be applied to either pretrained models or retrieval-enhanced models to support efficient inference. Together, these methods form a configurable solution space that adapts VFL to different data availability, infrastructure costs, and deployment objectives.

2 Background and Related Work

2.1 Preliminary: Vertical Split Learning

We focus on a typical two-party vertical split learning setting [21]. Specifically, each party holds a ***bottom model*** (e.g., f_A, f_B) for extracting hidden representations, and the active party additionally holds a ***top model*** g_A to aggregate two sides of representations to make predictions $\hat{y}$ and computes loss $\mathcal{L}(\hat{y}, y)$:

$$\mathbf{h}_{ij}^{A} = f_A(\mathbf{x}_{ij}^{A}; \theta_A^1), \mathbf{h}_{ij}^{B} = f_B(\mathbf{x}_{ij}^{B}; \theta_B) \quad (1)$$

$$\hat{y}_{ij} = f_{\theta_A, \theta_B}^{Fed}(\mathbf{x}_{ij}^{A}, \mathbf{x}_{ij}^{B}) = g_A(\mathbf{h}_{ij}^{A}, \mathbf{h}_{ij}^{B}; \theta_A^2) \quad (2)$$

where $\theta_A = \{\theta_A^1, \theta_A^2\}$. Note that $\mathbf{h}_A$ and $\mathbf{h}_B$ are distributively computed in each party's server. Once $\mathbf{h}_B$ is ready, it will be sent to

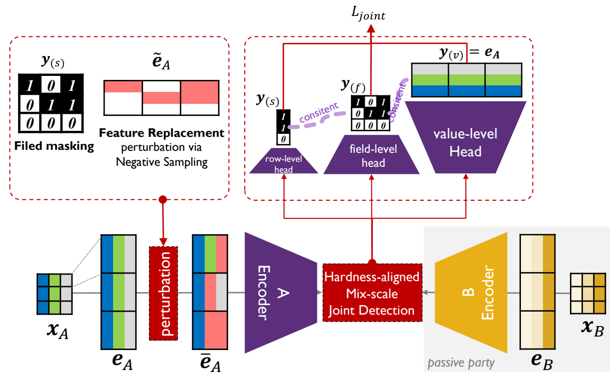

Figure 3: (Solution for C1) fMPD turns unlabeled overlaps into fine-grained perturbation–detection signals for cross-party correlation learning.

party A across the network. In the backward pass, the gradient $\mathbf{g}_B = \nabla_{\mathbf{h}_B}\mathcal{L}$ of $\mathbf{h}_B$ will be sent to party B for subsequent backward pass and parameter updates. *Only hidden representation* $\mathbf{h}_B$ *and corresponding gradient* $\mathbf{g}_B$ *are shared during training, thus original data is never directly exposed.* We consider another line of security-focused works [1, 8, 27] complementary to us and we focus mainly on the aspect of effectiveness and efficiency.

2.2 Related Work

Despite extensive studies on federated recommendation in horizontal FL settings [25], vertical federated recommendation remains underexplored [14, 23]. Around the key challenge of VFL, existing VFR studies mainly follow three directions: *(1) Pre-training* methods learn reusable encoders from additional unlabeled or non-overlapped data, such as VFed-SSD [11], SS-VFL [3], and HybridSSL [6]; they improve representation quality, but do not directly provide passive-side evidence for non-overlapped users at inference. *(2) Data compensation* methods supplement missing passive information, for example FedCVT [7] in the hidden space via multi-view learning and FedAds [23] through diffusion-based denoising; however, such compensation is often training-oriented and only applicable to aligned samples. *(3) Distillation* methods transfer federated knowledge into a local model [10, 11, 18], enabling efficient inference for all users but sacrificing B-side feature usage in serving.

3 Methods and Results

3.1 Solution 1: Extending Data Utilization via Fine-Grained Self-Supervised Pre-training

Design Motivation. The first direction addresses Challenge 1, where VFL-based recommendation is trained on only a small set of freshly labeled overlapped users. Self-supervised pre-training is a natural remedy, because many historical records have outdated or noisy labels [24] but still preserve *stable profile, item, and cross-platform co-occurrence semantics*. The key is to extract these signals in a way that matches recommendation modeling. CTR models depend on field interactions, while common contrastive SSL in tabular VFL often remains too coarse at the party level or relies on semantically ambiguous feature augmentation. Thus, fMPD is motivated by a simple principle: unlabeled overlaps should provide *fine-grained*

Table 1: fMPD improves downstream CTR prediction under limited labeled data. Δ denotes the difference from the strongest baseline for each metric.

	Avazu		Criteo	
Method	AUC↑	LogLoss↓	AUC↑	LogLoss↓
SplitNN	0.6786	0.3845	0.7438	0.5050
+ SelfTrain	0.7003	0.3662	0.7595	0.4885
+ LocalSSL	0.7046	0.3623	0.7682	0.4763
+ CrossSSL	0.7010	0.3645	0.7700	0.4721
+ HybridSSL	0.7032	0.3647	0.7716	**0.4704**
fMPD	**0.7182**	**0.3559**	**0.7744**	0.4712
Δ	+0.0136	-0.0064	+0.0028	+0.0008

and unambiguous supervision for cross-party correlation learning. The message is not merely to use more data, but to recover the field-level correlations that downstream federated predictors need.

Method Overview. As shown in Figure 3, fMPD follows a *fine-grained multi-scale perturb-then-detect framework.* It perturbs selected active-party feature values and asks the federated model to detect the induced inconsistency conditioned on the remaining active fields and the complete passive-party fields. The generated perturbation label gives explicit supervision, avoiding heuristic positive-pair construction. fMPD organizes detection at sample, field, and feature levels, so the model learns both coarse consistency and fine field dependencies. Active-side perturbation provides controllable corruption, passive-side conditioning supplies cross-view evidence, and multi-scale detection ties the two sources together. Through this design, *underused unlabeled overlaps become structured pre-training signals*, improving transferable representations for downstream VFL-based recommendation.

Results. We report results of fMPD and relevant baselines on vertically split Avazu and Criteo CTR benchmarks [5, 11]. The baselines include supervised SplitNN [4, 21], semi-supervised SelfTrain [2, 15, 19], and SSL-VFL methods LocalSSL [3], CrossSSL [6, 12], and HybridSSL [6]. Table 1 reports downstream AUC and LogLoss, where AUC measures ranking quality and LogLoss measures probabilistic calibration. We observe that *(1)* fMPD improves the strongest baseline by +0.0136 AUC on Avazu and +0.0028 on Criteo; *(2)* the gains hold on both datasets, indicating that the learned representations transfer to downstream CTR prediction under limited labeled supervision. These results demonstrate that fMPD satisfies the design goal of converting unlabeled overlapped records into fine-grained cross-party supervision, thereby addressing Challenge 1 on limited labeled-data utilization.

3.2 Solution 2: Reducing Inference Cost with Federated Privileged Distillation

Design Motivation. The second direction addresses Challenge 2, where the distributed serving cost of standard VFL blocks deployment in latency-sensitive recommendation systems. Offline federation can be scheduled and amortized, but online federated inference requires cross-party computation, communication, and often security operations for every request. This cost is hard to reconcile

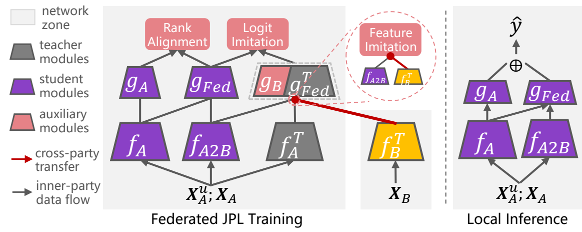

Figure 4: (Solution for C2) JPL distills full-field federated knowledge into a local-serving student for local inference.

with high-throughput advertising services. However, removing the passive party at inference would collapse the model back to a local recommender and lose the value of federation. This motivates Semi-VFL: *learn from full-field federation during training, but serve independently with active-party fields at inference.* The main challenge is to preserve passive-side knowledge under *field missing* while adapting across aligned and unaligned samples. This setting is more practical than requiring every request to enter federation, but harder than ordinary distillation because knowledge must generalize beyond aligned users.

Method Overview. As shown in Figure 4, JPL uses a privileged distillation pipeline. It first trains a federated teacher on aligned users with full fields, then trains a local-serving student on the full active-party sample space. The student contains a local branch for active-side patterns and a federated branch for distilled cross-party knowledge, sharing the active encoder to connect aligned and unaligned samples. Two regularizers provide the main transfer path: *federated equivalence imitation* preserves passive-side effects under local inputs, and *cross-head rank alignment* coordinates the local and federated heads. The module details are auxiliary to this goal: transfer federated knowledge into a single local predictor. After training, the deployed student performs standalone active-party inference, reducing serving cost while retaining federated benefits.

Results. We report results of JPL and relevant baselines on Avazu and Criteo under Semi-VFL local-serving inference [10], where full-field federated knowledge is used during training but only active-party fields are available at inference. The baselines include Fed [4, 21], the local-only control Local [10], FPD [11, 18], FedUD* [16], and FedCVT* [7]. Table 2 reports local-serving AUC on overall, aligned, and unaligned groups; blank entries denote methods that cannot serve the corresponding group under this protocol. We observe that *(1)* JPL obtains the best AUC on all reported groups for both datasets; *(2)* its advantage holds across overall, aligned, and unaligned users, indicating that the local student benefits from federated training over the full active-party sample space. These results demonstrate that JPL satisfies the local-serving design requirement by retaining federated benefits under standalone active-party inference, thereby addressing Challenge 2 on online VFL serving cost.

3.3 Solution 3: Achieving Full-Set User Benefits with Retrieval Augmentation

Design Motivation. The third direction addresses Challenge 3, where standard VFL benefits only aligned users. Even if federated serving is available, unaligned active-party users lack passive-party

Table 2: JPL preserves federated benefits under local-serving inference. Δ: gain over the strongest local-serving baseline.

Method	Avazu Overall	Avazu Aligned	Avazu Unaligned	Criteo Overall	Criteo Aligned	Criteo Unaligned
Fed	–	0.7082	–	–	0.7844	–
Local	0.7082	0.6992	0.7172	0.7750	0.7734	0.7768
FPD	0.7103	0.7037	0.7166	0.7764	0.7746	0.7785
FedUD*	0.7055	0.6982	0.7129	0.7711	0.7694	0.7728
FedCVT*	0.7145	0.7089	0.7195	0.7653	0.7634	0.7674
JPL	**0.7207**	**0.7127**	**0.7279**	**0.7775**	**0.7758**	**0.7794**
Δ	+0.0062	+0.0038	+0.0084	+0.0011	+0.0012	+0.0009

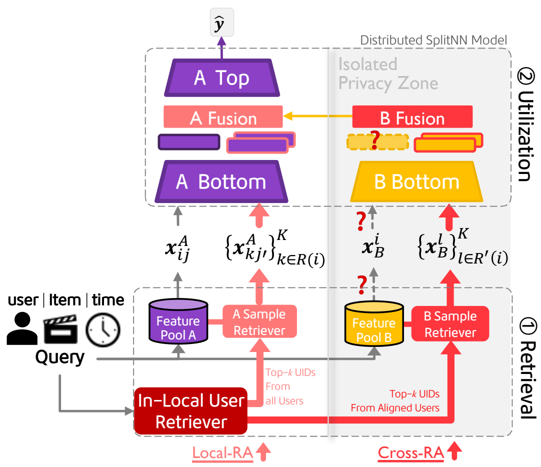

Figure 5: (Solution for C3) ReFer retrieves cross-party and local evidence to extend VFL benefits from aligned users to the full user set and meanwhile keeps privacy-preserving and inference efficiency

fields, so a conventional split model cannot make full-field predictions for them. This weakens the return on federation, because real platforms care about the full active-party population and often have many unaligned users. Discarding these users wastes active-side behavior data, while serving only aligned users makes collaboration useful to a narrow group. Naive fixes such as zero filling or hidden-space imputation are limited because *they do not provide concrete raw-space evidence from the passive side*. This motivates Fully-VFR, whose goal is *full-set user training and inference* rather than aligned-only collaboration. The core issue is therefore not only missing fields, but whether federation creates visible value for every user group.

Method Overview. Figure 5 summarizes ReFer as a retrieval-enhanced retrieval-and-utilization framework for Fully-VFR [9]. The key idea is to replace fixed missing-field defaults with relevant auxiliary evidence. *Cross-RA* retrieves similar aligned users and uses their passive-side records as a bridge to the missing passive domain, while *Local-RA* retrieves active-side neighbors from the full user pool to reduce aligned–unaligned group bias. A hierarchical federated retrieval process preserves data ownership, and query-aware fusion modules utilize the retrieved evidence for prediction. Cross-RA supplies passive-domain evidence, Local-RA improves group mutual understanding, and fusion turns retrieved evidence into prediction signals. Through this design, *passive-side knowledge can support unaligned users through retrieval*, and federation benefits extend from the aligned subset to the full user population.

Table 3: ReFer extends VFL benefits to the full user set. Δ denotes the difference from the strongest baseline.

Method	Douban, Book-Movie Overall	Douban, Book-Movie Aligned	Douban, Book-Movie Unaligned	Meituan, Search-Rec Overall	Meituan, Search-Rec Aligned	Meituan, Search-Rec Unaligned
UN-Local	0.6289	0.6092	0.6681	0.6204	0.5976	0.6412
AL-Local	0.6744	0.6893	0.6241	0.6106	0.6108	0.6330
Local	0.6784	0.6841	0.6547	0.6305	0.6130	0.6441
Fed	0.6532	0.6924	0.6282	0.6001	0.6296	0.6237
Fed-Fill	0.6795	0.6855	0.6547	0.6439	0.6308	0.6518
FTL	0.6179	0.6451	0.6516	0.6018	0.5987	0.6260
FPD	0.6823	0.6890	0.6564	0.6321	0.6153	0.6445
FedCVT	0.6818	0.6862	0.6653	0.6316	0.6141	0.6453
ReFer	**0.6989**	**0.7099**	**0.6705**	**0.6499**	**0.6357**	**0.6605**
Δ	+0.0166	+0.0175	+0.0024	+0.0060	+0.0049	+0.0087

Results. We report results of ReFer and relevant baselines on Douban Book-Movie [28] and industrial Meituan Search-Rec under full-user-set inference [9]. The baselines include local controls (UN-Local, AL-Local, Local) [9], aligned/full-set VFL variants (Fed, Fed-Fill) [4, 9, 21], and transfer or imputation methods FTL [13], FPD [11, 18], and FedCVT [7]. Table 3 reports full-user-set AUC on overall, aligned, and unaligned groups, evaluating both global recommendation quality and group-wise user benefits. We observe that *(1)* ReFer achieves the best overall AUC, with gains of +0.0166 on Douban Book-Movie and +0.0060 on Meituan Search-Rec; *(2)* it also improves both aligned and unaligned users, indicating group-wise benefits in addition to overall gains. These results demonstrate that ReFer satisfies the full-user-set design goal by using retrieval augmentation to extend VFL benefits beyond the overlapped subset, thereby addressing Challenge 3.

4 Future Work

In the future, we plan to extend cost-effective VFL-based recommendation toward frontier recommender systems while maintaining practical deployment constraints. Motivated by the potential of VFL to support privacy-preserving multi-platform collaboration and multi-domain data utilization, we will explore VFL from three directions: federated semantic representations for generative recommendation, lightweight adapters that inject partner-side knowledge into local backbones, and constrained LLM-assisted completion for missing cross-domain signals. We also plan to evaluate these extensions under practical benchmarks that jointly consider recommendation quality, unaligned-user coverage, communication cost, inference latency, and privacy risks.

GenAI Usage Disclosure

The authors used generative AI tools to assist with manuscript editing, wording refinement, and LaTeX formatting. The tools were not used to generate experimental results or alter reported data. The authors reviewed and verified the final manuscript content and remain fully responsible for the work.

References

[1] Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. 2016. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*. 308–318.

[2] Avrim Blum and Tom Mitchell. 1998. Combining labeled and unlabeled data with co-training. In *COLT*. 92–100.

[3] Timothy Castiglia, Shiqiang Wang, and Stacy Patterson. 2022. Self-supervised vertical federated learning. In *Workshop on Federated Learning: Recent Advances and New Challenges (in Conjunction with NeurIPS 2022)*.

[4] Iker Ceballos, Vivek Sharma, Eduardo Mugica, Abhishek Singh, Alberto Roman, Praneeth Vepakomma, and Ramesh Raskar. 2020. SplitNN-driven Vertical Partitioning. *CoRR* abs/2008.04137 (2020). arXiv:2008.04137 https://arxiv.org/abs/2008.04137

[5] CriteoLabs. 2014. The Criteo dataset for kaggle display advertising challenge. https://labs.criteo.com/2014/02/download-kaggle-display-advertising-challenge-dataset/.

[6] Yuanqin He, Yan Kang, Xinyuan Zhao, Jiahuan Luo, Lixin Fan, Yuxing Han, and Qiang Yang. 2024. A hybrid self-supervised learning framework for vertical federated learning. *IEEE Transactions on Big Data* (2024).

[7] Yan Kang, Yang Liu, and Xinle Liang. 2022. FedCVT: Semi-supervised Vertical Federated Learning with Cross-view Training. *ACM Transactions on Intelligent Systems and Technology (TIST)* 13, 4 (2022), 1–16.

[8] Oscar Li, Jiankai Sun, Xin Yang, Weihao Gao, Hongyi Zhang, Junyuan Xie, Virginia Smith, and Chong Wang. 2021. Label Leakage and Protection in Two-party Split Learning. In *International Conference on Learning Representations*.

[9] Wenjie Li, Zhongren Wang, Jinpeng Wang, Shu-Tao Xia, Jile Zhu, Mingjian Chen, Jiangke Fan, Jia Cheng, and Jun Lei. 2024. ReFer: Retrieval-Enhanced Vertical Federated Recommendation for Full Set User Benefit. In *SIGIR*.

[10] Wenjie Li, Qiaolin Xia, Hao Cheng, Kouyin Xue, and Shu-Tao Xia. 2022. Vertical semi-federated learning for efficient online advertising. *arXiv preprint arXiv:2209.15635* (2022).

[11] Wenjie Li, Qiaolin Xia, Junfeng Deng, Hao Cheng, Jiangming Liu, Kouying Xue, Yong Cheng, and Shu-Tao Xia. 2022. Semi-Supervised Cross-Silo Advertising with Partial Knowledge Transfer. *arXiv preprint arXiv:2205.15987* (2022).

[12] Wenjie Li, Qiaolin Xia, Junfeng Deng, Hao Cheng, Jiangming Liu, Kouying Xue, Yong Cheng, and Shu-Tao Xia. 2022. VFed-SSD: Towards practical vertical federated advertising. *arXiv preprint arXiv:2205.15987* (2022).

[13] Yang Liu, Yan Kang, Chaoping Xing, Tianjian Chen, and Qiang Yang. 2020. A secure federated transfer learning framework. *IEEE Intelligent Systems* 35, 4 (2020), 70–82.

[14] Yang Liu, Yan Kang, Tianyuan Zou, Yanhong Pu, Yuanqin He, Xiaozhou Ye, Ye Ouyang, Ya-Qin Zhang, and Qiang Yang. 2022. Vertical Federated Learning. *arXiv preprint arXiv:2211.12814* (2022).

[15] Kamal Nigam and Rayid Ghani. 2000. Analyzing the effectiveness and applicability of co-training. In *CIKM*. 86–93.

[16] Wentao Ouyang, Rui Dong, Ri Tao, and Xiangzheng Liu. 2024. FedUD: exploiting unaligned data for cross-platform federated click-through rate prediction. In *SIGIR 2024*.

[17] Daniel Peterson, Pallika Kanani, and Virendra J Marathe. 2019. Private federated learning with domain adaptation. *arXiv preprint arXiv:1912.06733* (2019).

[18] Zhenghang Ren, Liu Yang, and Kai Chen. 2022. Improving Availability of Vertical Federated Learning: Relaxing Inference on Non-overlapping Data. *ACM Transactions on Intelligent Systems and Technology (TIST)* (2022).

[19] Chuck Rosenberg, Martial Hebert, and Henry Schneiderman. 2005. Semi-supervised self-training of object detection models. (2005).

[20] Jianqiang Shen, Burkay Orten, Sahin Cem Geyik, Daniel Liu, Shahriar Shariat, Fang Bian, and Ali Dasdan. 2015. From 0.5 million to 2.5 million: Efficiently scaling up real-time bidding. In *ICDM*. IEEE, 973–978.

[21] Jiankai Sun, Xin Yang, Yuanshun Yao, Aonan Zhang, Weihao Gao, Junyuan Xie, and Chong Wang. 2021. Vertical Federated Learning without Revealing Intersection Membership. *arXiv preprint arXiv:2106.05508* (2021).

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

[23] Penghui Wei, Hongjian Dou, Shaoguo Liu, Rongjun Tang, Li Liu, Liang Wang, and Bo Zheng. 2023. FedAds: A Benchmark for Privacy-Preserving CVR Estimation with Vertical Federated Learning. *arXiv preprint arXiv:2305.08328* (2023).

[24] Xin Wu, Qing Liu, Jiarui Qin, and Yong Yu. 2022. PeerRank: Robust Learning to Rank With Peer Loss Over Noisy Labels. *IEEE Access* 10 (2022), 6830–6841. doi:10.1109/ACCESS.2022.3142096

[25] Liu Yang, Ben Tan, Vincent W. Zheng, Kai Chen, and Qiang Yang. 2020. *Federated Recommendation Systems*. Springer International Publishing, Cham, 225–239. doi:10.1007/978-3-030-63076-8_16

[26] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. 2019. Federated machine learning: Concept and applications. *TIST* 10, 2 (2019), 1–19.

[27] Chengliang Zhang, Suyi Li, Junzhe Xia, Wei Wang, Feng Yan, and Yang Liu. 2020. {BatchCrypt}: Efficient homomorphic encryption for {Cross-Silo} federated learning. In *2020 USENIX annual technical conference (USENIX ATC 20)*. 493–506.

[28] Chuang Zhao, Hongke Zhao, Ming He, Jian Zhang, and Jianping Fan. [n. d.]. Cross-domain recommendation via user interest alignment. In *WWW 2023*.