Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/43/20230848

Research on machine learning technology with privacy protection strategy in recommendation field

Yuchen Han

School of Electrical and Information Engineering, Tianjin University, Tianjin, 300072, China

3017234555@tju.edu.cn

Abstract. Technologies such as machine learning can achieve accurate personalized recommendations. However, due to the collection and utilization of a large amount of user information in this process, people are widely worried about data security and privacy issues. This paper first introduces two key issues of privacy protection in the field of machine learning, namely data privacy and model privacy. On this basis, this paper introduces and analyzes homomorphic encryption, differential privacy and federated learning, and compares their advantages and disadvantages. Among them, homomorphic encryption technology has a large computational cost, differential privacy technology has a negative impact on system accuracy, and federated learning technology has a high training and communication cost. Therefore, it will be the future research direction to study more efficient and accurate recommendation models.

Keywords: Recommendation systems, Privacy protection, Machine learning, Differential privacy, Federated learning.

1. Introduction

With the rapid development of the Internet, a huge amount of user data is collected by various application software every day. These data can help companies better analyze user preferences, so as to better serve users. The analysis of user data can also help companies formulate better strategies and get higher profit returns. Therefore, the advertising recommendation model with machine learning technology has become more and more widely used.

However, modern data analysis technology, such as machine learning technology is extremely powerful for the penetration of privacy. In the face of computing power and algorithms, the private information hidden behind massive data is almost transparent. On the one hand, users are worried that their private information is over-collected, and enterprises are generally worried that their models are stolen. On the other hand, many countries have also formulated relevant laws and regulations to protect citizens' privacy [1]. Therefore, how to combine privacy protection technology with the recommendation model to better care for the interests of all parties has become an important research point in the machine learning field.

By consulting relevant literature, this paper studies the key issues of privacy protection in the field of recommendation and sorts out and summarizes the different privacy protection technologies in the recommendation algorithms, promoting researchers to deeply understand the development trends, key issues and the basis of existing knowledge in the research field.

DOI: 10.54254/2755-2721/43/20230848

2. Key privacy issues in recommendation field

The collection and use of user data inevitably comes with privacy concerns. It should be pointed out that the definition of privacy issues will be different for various research projects. According to the types of privacy information to be protected, this study divides the existing research work on privacy protection in the recommendation field into the following two categories:

a. Data privacy, which mainly refers to private attribute information in user data and system interaction information.

b. Model privacy, which mainly refers to the privacy information such as the training algorithm, topological structure and weight parameters in machine learning models.

2.1. Data Privacy

No matter for institutions or individuals, users' personal privacy data includes users' private sensitive attribute information such as identity, gender, age, and user-item interaction information such as users' clicks and ratings on items, which are very precious resources.

However, in the process of personalized recommendation, on the one hand, large-scale collection of such data may lead to direct privacy leakage. Its main manifestations are the unauthorized collection of users' personal information, illegal data sharing and trading, etc. But in recent years, many countries have gradually introduced relevant laws to protect personal privacy data security. For example, the EU officially implemented the General Data Protection Regulation (GDPR) in 2018, which greatly strengthened the protection and supervision of personal information [1].

On the other hand, the lack of generalization ability of the model may lead to indirect privacy leakage. In the model training stage, the more complex the model is, the more powerful the data memory ability will be. Therefore, by interacting with the model, unreliable data analysts can reversely infer the individual sensitive attributes from unknown training data [2]. How to solve the problem of indirect privacy leakage is one of the key research directions in the field of machine learning. For example, in order to make the recommendation model have the ability to protect users' private sensitive attribute information, an adversarial learning framework has been designed by Beigi, G and his team [3]. This adversarial learning framework can combine the recommendation model with the user's private sensitive attribute information classification model, so as to reduce the possibility of the user's private sensitive attribute information leakage and ensure high recommendation accuracy at the same time. Another example is that, according to Lin, G, in order to prevent the server from inferring the items that the user has interacted with through the gradient uploaded by the user, the promoted method FedRec will randomly sample some disturbed items into the interactive item set when calculating the hidden vector gradient of the items, and give the disturbed items an average score or a predicted score to simulate the real interaction, thus achieving the purpose of protecting privacy [4].

2.2. Model Privacy

For a company or organization, the data needed for model training and the process of training all consume a lot of time and money [5]. Therefore, in addition to protecting the original data, when faced with a large number of malicious users or attackers, how to protect the machine learning model's own parameters, model structure, data sets used by the training model, etc. from being stolen is another field worthy of study.

Taking MLaas (machine-learning as-a service) as an example, services such as labeling of machine learning data, model training and prediction have been uploaded to the cloud. Therefore, the attackers can reversely infer the specific parameters or structure of the model from the predicted value through limited access to the API of the service, or combine the sample and the predicted value to train a replacing model [6]. Once the model is leaked, it may bring huge losses to the company or organization. For example, Wang et al. put forward a method for hyperparameter attacks. Experiments show that this method uses linear regression, logistic regression, support vector machine and neural network to successfully obtain the hyperparameter of the model [7].

DOI: 10.54254/2755-2721/43/20230848

In order to solve the above problems, some literature uses differential privacy technology to protect the model. That is, adding random noise to sensitive information during model training or in the final model parameters makes it impossible for attackers to detect the influence of the change in original training data on model output, thus increasing the difficulty of model extraction. For instance, according to Abadi M, by adding noise to the gradient before using it to update the parameters of the deep neural network, the proposed algorithm can still train the deep neural network of non-convex targets with a modest privacy budget and training efficiency [8].

3. Key privacy protection technologies in recommendation field

In order to meet different privacy protection requirements, a variety of privacy protection technologies have been proposed and applied to recommendation systems. It is mainly divided into three categories. The first category is cryptography technology represented by homomorphic encryption; the second category is perturbation method represented by differential privacy technology; and the third category is distributed machine learning method represented by federated learning. The basic idea of privacy protection strategy in the machine learning field is to apply the above privacy protection methods to the model training process. Table 1 summarizes the advantages and disadvantages of these privacy protection technologies.

Table 1. Comparison of privacy protection technologies in recommendation field.

Category	Advantages	Disadvantages
Homomorphic Encryption	Strong data protection ability.	Complex to construct and implement; large calculation and storage cost.
Differential Privacy	Low computational cost; Widely applicable; mature	Randomized noise leads to the decrease of recommendation accuracy.
Federated Learning	Avoids the isolated data issue, suitable for multi-party computing, avoids data leakage on the server side.	High communication stability and bandwidth requirements; difficult to train model.

3.1. Cryptography technology

Cryptography technology is a data encryption algorithm based on mathematical theory, among which homomorphic encryption is the most classic and widely used. Its mathematical definition is as follows:

$$\mathrm{E}(\mathrm{m1}) \star \mathrm{E}(\mathrm{m2}) = \mathrm{E}(\mathrm{m1} \star \mathrm{m2}), \forall \mathrm{m1}, \mathrm{m2} \in \mathrm{M} \quad (1)$$

Where E is the encryption algorithm and M is the set of all possible information. If the encryption algorithm E satisfies Formula (1), then E conforms to the homomorphic encryption property of ★ operation. At present, homomorphic encryption mainly supports operation in addition and multiplication. The main meaning of Formula (1) is that after the data is homomorphic encrypted, the ciphertext calculation result can be obtained by a specific calculation. Then the ciphertext calculation result can be homomorphic decrypted to obtain plaintext, and this calculation process is equivalent to directly calculating plaintext data.

In the field of privacy-preserving recommendation systems, homomorphic encryption can be used to encrypt user characteristics or gradient results, so that the recommendation system platform can perform recommendation algorithms without seeing the specific information of users. However, it is worth noting that the data encryption process often involves a large number of operations, which will generate huge computational costs in complex situations [9]. For example, according to Jinsu and his team, they proposed a privacy protection matrix factorization algorithm based on fully homomorphic encryption that uses encrypted users' rating data and returns encrypted outputs. And a new data structure is introduced in this article to calculate the encryption vector necessary for matrix factorization, in an attempt to overcome the performance degradation caused by homomorphic encryption [10].

DOI: 10.54254/2755-2721/43/20230848

3.2. Differential Privacy technology

Differential privacy is a privacy protection technology proposed by Dwork in 2006 [11]. It is based on strict mathematical theory and aims to ensure that attackers cannot infer individual sensitive information from output differences. The typical mathematical definition of differential privacy is as follows:

$$\Pr[A(D) = O] \leq e^{\varepsilon} \Pr[A(D') = O] \tag{2}$$

The Formula (2) implies that given two adjacent data sets D and D' (that is, the difference between the two data sets is only one record), when a random function A acts on the two adjacent data sets, the obtained output is indistinguishable and very similar. The parameter ε is the privacy budget. The smaller it is, the higher the level of privacy protection is, and the greater the amount of noise that needs to be added.

In the recommendation field, user information can be perturbed by differential privacy technology to prevent the identification of a single user. It is worth noting that compared with homomorphic encryption, differential privacy can be achieved only by noise addition mechanism, so there is no additional computational cost. However, the added noise will make the data inaccurate, which will affect the accuracy of model prediction to some extent. Therefore, some literature will also weigh the privacy and usability of the algorithm by designing a reasonable mechanism. For example, Xiao Y's team combine the differential privacy algorithm with the reinforcement learning framework. Through reinforcement learning, a reasonable privacy budget is obtained, which makes the recommendation algorithm strike a balance between data availability and privacy [12].

3.3. Federated Learning technology

Traditional machine learning algorithms need to gather data to train the model. In this situation, it is almost impossible for different organizations to share data, and the gathered data may also lead to the problem of privacy leakage. Therefore, in 2016, Google proposed a distributed machine learning framework, namely federated learning. The training data is distributed across multiple devices or organizations, and each participant processes it locally to form a local model. The local models are aggregated and updated through a federated learning algorithm, ultimately resulting in a global model. This not only protects privacy, but also avoids the trust problem of data sharing. Moreover, in the federated learning mode, participants can also protect the privacy of data and models through encryption technology and perturbation technology. For example, on the basis of federated learning, the FebMF algorithm uses a homomorphic encryption algorithm to encrypt the gradient, so that the server can update the hidden vector of items but cannot restore the original information of users [13].

In the field of privacy-preserving recommendation systems, horizontal federated learning is the most widely used. Its basic process is shown in Figure 1:

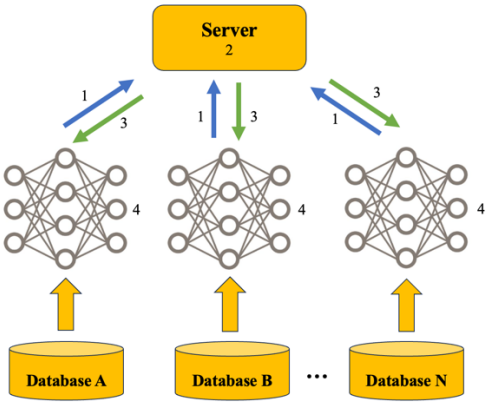

Figure 1. The learning process of Horizontal Federal Learning.

DOI: 10.54254/2755-2721/43/20230848

It can be explained in the following four steps:

Step 1: Each participant downloads the latest model from the server, and uses local data to train the model.

Step 2: Each participant encrypts (or uses technologies such as differential privacy) the gradient and uploads it to the server. Then the server updates the model parameters after aggregating the gradient of each user.

Step 3: The server returns the updated model to each participant.

Step 4: Each participant updates its own model.

To be more specific, the server can be seen as the recommendation system, and users only provide encrypted data (such as gradient) to the platform side, while browsing records and other information are stored locally, thus realizing privacy protection. However, it is noteworthy that although distributed storage and calculation protect users' privacy, they also bring greater training difficulty and more communication costs. Therefore, how to optimize the efficiency of a federated learning recommendation system, speed up convergence, and reduce communication times is also one of the current research hotspots.

4. Conclusion

This study discusses data privacy and model privacy in the field of machine learning, and summarizes three privacy protection methods: homomorphic encryption, differential privacy technology and federated learning. To sum up, the recommendation algorithm, which considers privacy protection strategy, still has the following shortcomings.

First, most privacy protection methods in the recommended field basically revolve around cryptography, anonymous technologies such as differential privacy, and federated learning. Although these methods are effective, they also have some inherent and intractable defects and a lack of essential innovation.

Secondly, because the goals of the recommendation algorithm and the privacy protection algorithm are mutually exclusive, adding a privacy protection module to the recommendation algorithm will reduce the performance of the recommendation task, so it is difficult to balance the accuracy and privacy of the recommendation model.

Because of the limited scope of this study, future research can comprehensively consider the demand for data privacy and model privacy, and study a more robust machine learning method for privacy protection to reduce the risks and losses caused by attacks on models and data. Researchers can also use a combination of various technologies to study a more efficient and accurate privacy protection scheme.

References

[1] Han W. (2023). Research on legal issues of personal information protection in the context of big data (Master's thesis, Shanxi University of Finance and Economics). https://kns.cnki.net/kcms2/article/abstract?v=-YY6Aedvp4YHQm_xrDOUEAaBHxP5Dl1U4lj1ITLjZZbIRMflZEZHG8UfDKWfaV-r2MnYHZMJv LHaM1Ng2TgTeiW0yrbrDWoEI4hufrGEj77-wWk1gXWV7iM1iCPtKLKr&uniplatform=NZKPT&language=CHS

[2] Fredrikson Matthew, Lantz Eric, Jha Somesh, Lin Simon & Ristenpart Thomas(2014). Privacy in Pharmacogenetics: An End-to-End Case Study of Personalized Warfarin Dosing.. Proceedings of the ... USENIX Security Symposium. UNIX Security Symposium, 201417-32.

[3] Beigi, G., Mosallanezhad, A., Guo, R., Alvari, H., & Liu, H.. (2019). Privacy-aware recommendation with private-attribute protection using adversarial learning.

[4] Lin, G., Liang, F., Pan, W., & Ming, Z.. (2020). Fedrec: federated recommendation with explicit feedback. Intelligent Systems, IEEE, PP(99), 1-1.

[5] He YiZ, Hu XB, He JW, Meng Guozhu & Chen Kai. (2019). A review of privacy and security issues in machine learning systems. Computer Research and Development (10), 2049-2070.

DOI: 10.54254/2755-2721/43/20230848

[6] Florian Tramèr, Zhang, F., Juels, A., Reiter, M. K., & Ristenpart, T.. (2016). Stealing machine learning models via prediction apis.
[7] Wang, B., & Gong, N. Z. . (2018). Stealing hyperparameters in machine learning. IEEE.
[8] Abadi, M., Chu, A., Goodfellow, I., Mcmahan, H. B., & Zhang, L.. (2016). Deep Learning with Differential Privacy. the 2016 ACM SIGSAC Conference. ACM.
[9] Feng H, Yi HW, Li XH & Li R. (2023). A review of privacy protection research in recommender systems. Computer Science and Exploration (08), 1814-1832.
[10] Jinsu, K., Dongyoung, K., Yuna, K., Hyunsoo, Y. , Junbum, S. , & Sungwook, K.. (2018). Efficient privacy-preserving matrix factorization for recommendation via fully homomorphic encryption. Acm Transactions on Privacy & Security, 21(4), 1-30.
[11] Dwork, C., Mcsherry, F., Nissim, K., & Smith, A.. (2006). Calibrating noise to sensitivity in private data analysis. Proceedings of the VLDB Endowment.
[12] Xiao, Y., Xiao, L., Lu, X., Zhang, H., & Poor, H. V.. (2021). Deep-reinforcement-learning-based user profile perturbation for privacy-aware recommendation. IEEE Internet of Things Journal, 8(6), 4560-4568.
[13] Chai, D., Wang, L., Chen, K., & Yang, Q.. (2020). Secure federated matrix factorization. Intelligent Systems, IEEE, PP(99), 1-1.