Research Article GU J Sci, Part A, 11(1): 68-79 (2024) 10.54287/gujsa.1393692

Gazi University
Journal of Science
PART A: ENGINEERING AND INNOVATION
http://dergipark.org.tr/gujsa

# A Comparative Study for Privacy-Aware Recommendation Systems

Yavuz CANBAY[1*] Anıl UTKU[2]

[1] *Department of Computer Engineering, Faculty of Engineering and Architecture, Sutcu Imam University, Kahramanmaraş, Türkiye*
[2] *Department of Computer Engineering, Faculty of Engineering, Munzur University, Tunceli, Türkiye*

**Keywords**

Recommender System
Privacy Protection
Differential Privacy

**Abstract**

Recommendation systems are sophisticated processes for information filtering designed to offer users tailored recommendations based on their preferences and interests. Users need help to choose between options as the amount of information on the web grows. As a result, it is critical to deliver personalized recommendations to consumers to promote user loyalty and satisfaction. Because recommender systems use sensitive user information such as ratings, comments, likes, and dislikes, this information can be leaked if no privacy measures are taken. As a result, we presented a comparison of privacy-aware recommendation systems in this paper. Two experiments are carried out. In the first experiment, we examined collaborative filtering algorithms on perturbed ratings and then compared hybrid, collaborative, and content-based algorithms on perturbed ratings in the second experiment. According to the results, the Singular Value Decomposition++ (SVDpp) algorithm presented the lowest Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values for epsilon 100, with 0.8889 and 0.6822, respectively. Furthermore, for epsilon 100, the hybrid filtering technique had the lowest RMSE and MAE rates of 0.90664 and 0.69813, respectively.

**Cite**

Canbay, Y., & Utku, A. (2024). A Comparative Study for Privacy-Aware Recommendation Systems. *GU J Sci, Part A, 11*(1), 68-79. doi:10.54287/gujsa.1393692

**Author ID (ORCID Number)**

0000-0003-2316-7893 Yavuz CANBAY
0000-0002-7240-8713 Anıl UTKU

**Article Process**

**Submission Date** 21.11.2023
**Revision Date** 08.01.2024
**Accepted Date** 26.01.2024
**Published Date** 09.02.2024

## 1. INTRODUCTION

The amount of online data generated nowadays is rising as a result of technical advancements and increased internet usage. Feedback from users, including ratings, comments, likes, and dislikes, generates many data. (Yu et al., 2019). By analyzing data gathered from various users and sources, recommender systems seek to provide users with recommendations for relevant and engaging content (Kunaver & Požrl, 2017). By selecting items that may be helpful for users from vast volumes of data, the algorithms used in these systems seek to offer the most beneficial items to the user. Recommender systems identify patterns in the dataset based on the user's choices and interests after learning about the user's interests (Cai et al., 2018).

The utilization of recommendation systems by e-commerce platforms and social media applications has become increasingly prevalent, resulting in the provision of diverse recommendations to their customers. These recommendations facilitate consumers to obtain valuable recommendations from the information stack, hence providing benefits to companies in terms of sales strategies. One further advantage of recommendation systems is that they enhance customer satisfaction and promote client loyalty (Mehrotra et al., 2018). Recommendation systems also facilitate the presentation of products that may not be prominently shown but are potentially of interest to users.

Besides the advantages of recommender systems, user privacy is an important issue to be considered. Recommender systems utilize digital data, which commonly include personal information such as search

*Corresponding Author, e-mail: yavuzcanbay@ksu.edu.tr

queries, purchase records, preferences, ratings, comments, etc. Nevertheless, the divulgence of this type of information may result in negative consequences for the individuals who own the data. In 2016, the European Parliament and Council established the General Data Protection Regulation (GDPR) to regulate data privacy. GDPR establishes criteria pertaining to the acquisition and processing of private user information, resulting in a fundamental transformation of data processing practices across several sectors. Therefore, ensuring compliance with the GDPR is a significant concern for those responsible for managing data. In the realm of recommender systems, it is crucial to employ privacy protection mechanisms to protect users' privacy (Hu et al., 2021).

Therefore, this study applies a privacy-aware methodology and proposes a recommender system that prioritizes privacy. Furthermore, it does a comparative analysis of the performance of recommender systems using data that has been perturbed. The present paper has been organized in the following manner. The introduction of recommender systems is given in Section II. Section III provides a comprehensive literature review. Differential privacy is presented in Section IV. Experimental studies are conducted in Section V, and the paper is concluded in Section VI.

## 2. RECOMMENDER SYSTEMS

Recommender system can be characterized as an application that endeavors to forecast consumer preferences. The field of recommender systems has witnessed significant growth in research due to its advantages to content creators and users (Milano et al., 2020). These systems facilitate the expeditious acquisition of desired products by providing personalized recommendations to consumers (Guo et al., 2017). For professionals engaged in content production, it is of the most significant importance as it enhances user engagement, raises user traffic and facilitates the development of sales tactics.

Recommender systems mostly employ collaborative, content-based and hybrid filtering techniques. Collaborative filtering approach operates under the assumption that individuals who exhibit similar behavioral patterns would possess comparable preferences in subsequent instances (Cunha et al., 2018). The process primarily involves collecting data pertaining to users' behaviors and preferences, followed by analyzing this data to ascertain the degree of similarity amongst users. This type of filtering mainly utilizes user similarities as a basis for generating recommendations without relying on product content. This filtering approach can be implemented using two distinct approaches, namely user-based and item-based methodologies. User-based collaborative filtering is a system that creates recommendations by identifying similarities between users (Kawasaki & Hasuike, 2017). The process of item-based collaborative filtering involves generating recommendations by identifying and analyzing similarities between items. Given the resemblance between the things with which the user engages, the objective is to provide the other item as a recommendation to other users in the event that one of the two adjacent items is engaged (Garanayak et al., 2019). The process of collaborative filtering involves the calculation of similarities between user profiles to generate recommendations. Utilizing cosine similarity and pearson correlation coefficient is prevalent in determining similarities among users. Content-based filtering relies on analyzing the descriptions of items and the content of items that consumers engage with them. Keywords are employed in content-based filtering to describe the items (Javed et al., 2021). Finally, hybrid recommendation systems combine the utilization of collaborative filtering and content-based filtering techniques. Hybrid approaches are utilized to optimize performance while mitigating the drawbacks associated with their constituent methods (Geetha et al., 2018).

SVD is a matrix factorization method commonly used in collaborative filtering. It aims to reduce the dimensionality of a dataset by transforming a space of size $N$ into a lower-dimensional space of size $K$. Within the context of recommender systems, a matrix is commonly employed to describe the relationship between users and items. In this matrix, users and items are represented by rows and columns, respectively. Each element of this matrix corresponds to the ratings that users provide to the respective items (Zhao et al., 2019). SVD++ incorporates an additional factor vector into the SVD method for each item. These item-specific factor vectors are utilized to characterize the features of items, irrespective of whether they have been assessed by users (Xian et al., 2017). The kNNwithZScore algorithm is a collaborative filtering approach incorporating z-score normalization for each user (Sütçü et al., 2021). In the co-clustering algorithm, clusters and co-clusters

are assigned to users and items, respectively. The assignment of clusters is accomplished by the utilization of an optimization technique, specifically the k-means algorithm.

## 3. LITERATURE REVIEW

This section provides a summary of previous researches considering privacy-aware recommender systems.

In their study, Xiong et al. (2020) introduced a framework for a private recommender system that operates on a client-server architecture. This study operates under the assumption that the server side is untrusted and each client uses a differential privacy technique to generate randomized ratings. A novel methodology was devised to represent user ratings using symbolic representations. The server-side algorithm utilizes these symbols to implement a collaborative filtering technique in order to make predictions about user ratings. The researchers utilized the Netflix and MovieLens datasets in conjunction with the exponential mechanism to ensure differential privacy. The algorithm proposed by the researchers offers a suitable methodology for both user-based and item-based approaches. Ultimately, they achieved low error rates, as measured by MAE.

The authors Neera et al. (2021) emphasized the utilization of a local perturbation model with differential privacy, wherein user ratings are randomized at the user's side before being transmitted to the service provider. However, a decrease in accuracy is observed when perturbation is employed. To address this issue, the researchers put out a Matrix Factorization approach based on Local Differential Privacy and incorporated a Gaussian Mixture Model. The researchers employed MovieLens, Jester, and Libimseti datasets in their experimental analysis, thereafter conducting a comparative evaluation of their algorithm against two previously established algorithms documented in the literature. Finally, they achieved superior outcomes compared to their counterparts.

Zhang et al. (2019) proposed a novel probabilistic matrix factorization approach with privacy considerations for recommender systems. To fulfill the consumers' privacy needs, a modified sampling process incorporating differential privacy was used. The researchers conducted a comparative analysis of matrix factorization schemes, examining their performance under both personalized differential privacy and traditional differential privacy frameworks. The researchers employed three distinct datasets, namely MovieLens 100k, MovieLens 1M, and the Netflix datasets. The authors conducted a comparative analysis between the proposed method and current approaches. In the end, the experimental results demonstrated that the proposed method presented superior performance compared to the current approaches.

The paper presented by Selvaraj & Gangadharan (2021) proposed a novel hybrid recommender system that incorporates deep learning techniques while ensuring privacy preservation. Differential privacy was utilized to ensure users' privacy, and a deep neural network was utilized for recommendation. The MovieLens 100k, Film trust, and Book-crossing datasets were utilized in the study, and experiments were undertaken for various epsilon values. Based on the experimental results, their approach demonstrated superior performance in terms of RMSE and MAE metrics compared to alternative approaches.

Liu et al. (2019) utilized autoencoders and differential privacy techniques to develop a recommender system that ensures user privacy. The technique of gradient perturbation is employed to obscure the underlying method, while the performance of the declared approach is evaluated using the MovieLens 100k and MovieLens 1M datasets. The measurement of error rates in this study was conducted using the RMSE metric. The experimental findings consistently demonstrated that the proposed method yielded superior outcomes compared to alternative approaches.

In their study Badsha et al. (2017), introduced a novel recommender system that incorporates collaborative filtering techniques while also ensuring privacy preservation. The utilization of homomorphic encryption was implemented to encrypt the ratings of users, hence enabling the provision of recommendations within an encrypted domain. The GroupLens dataset is utilized for conducting experiments, wherein the performance comparison is made according to computational and communication costs.

Yin et al. (2020) have proposed a novel privacy-preserving collaborative filtering approach for recommender systems. The MovieLens dataset is utilized for conducting experiments, and the RMSE metric is employed for

evaluating performance. To ensure data privacy, the original dataset is subjected to the DiffGen algorithm, which is based on differential privacy principles. The researchers enhanced the traditional collaborative filtering algorithm and conducted a comparative analysis with the traditional approach. As a result, experimental results have demonstrated that the enhanced algorithm yields better outcomes than traditional methods.

The improved collaborative filtering recommendation system that takes into account differential privacy was introduced by Wang & Wang (2020). Differential privacy was employed as a means to ensure privacy in the process of matrix factorization. The suggested method is tested using the MovieLens 100k and Netflix datasets, and the performance evaluation is conducted using the RMSE metric. The researchers conducted a comparative analysis of the proposed approach and other established methods, including user-based k-nearest neighbor (kNN), item-based kNN, SVD++, and alternating least squares. Based on the study results, their approach showed better results than other methods.

The current literature presents clear evidence that matrix factorization, autoencoders, deep neural networks, kNN, and SVD++ are among the techniques commonly employed for recommendation systems. The datasets commonly employed in experimental studies are MovieLens, Netflix, GroupLens, and Libimseti. The error metrics utilized for performance evaluations include RMSE and MAE. In contrast to the existing works, this study is the first use of differential privacy in both hybrid and collaborative recommendation systems and presenting a comprehensive comparison.

## 4. DIFFERENTIAL PRIVACY

The concept of privacy can be described as "the right to be let alone" according to Warren and Brendis (1890), and as "the selective control of access to the self" as proposed by Altman (1976). In time, privacy has been classified into various categories (Banisar & Davies, 1999);

- Information privacy; refers to the set of regulations and guidelines that govern collecting, processing, and managing individuals' personal data,
- Bodily privacy; refers to the regulations of protecting individuals' physical well-being from harmful or unwelcome actions,
- Communication privacy; describes the act of ensuring the privacy of many forms of communication, including but not limited to email and telephone conversations,
- Territorial privacy; means the regulations that establish the boundaries of permissible intrusion into the household and other physical spaces.

The protection of information privacy or data privacy is of utmost importance in an individual's life, as personal data has the ability to directly or indirectly identify a person. Furthermore, it is worth noting that data, particularly raw data, serves as the fundamental source for digital world. Therefore, researchers are currently placing significant emphasis on the topic of information privacy or data privacy.

Within the existing body of literature, other privacy protection methods have been established, including k-anonymity (Sweeney, 2002), l-diversity (Machanavajjhala et al., 2006), t-closeness (Li et al., 2007) and δ-presence (Nergiz et al., 2007). In addition, Dwork (2006) suggested the concept of differential privacy as a potential solution to the issue of background attacks, which all aforementioned protection models suffer. This technique facilitates the execution of privacy-preserving procedures on raw datasets. The assurance is that the outcome of any query performed on a dataset will remain consistent, regardless of whether a record is included in the dataset or not. This implies that the addition or removal of any record will have no impact on the result of any query. Therefore, this scenario allows for the mitigation of background information attacks.

**Definition:** a mechanism *M* is said to satisfy ε-differential privacy if, for any neighboring databases *D* and *D'*, and for any subset *S* of all potential outputs;

$$Pr[M(D) \in S] \leq \exp(\varepsilon) * Pr[M(D') \in S] \quad (1)$$

The sensitivity of a query on a dataset is an essential property. Sensitivity refers to the extent to which the outcome of a query is influenced by the presence or absence of a particular record.

**Definition:** the sensitivity of a query can be described as in Equation 2;

$$\Delta f = \max_{DD'} \|f(D) - f(D')\| \tag{2}$$

**Definition:** the given mechanism *M* satisfies e-differential privacy for a function *f*;

$$M(D) = f(D) + Laplace(\frac{\Delta f}{e}) \tag{3}$$

Within the existing body of literature, there are five distinct techniques that encompass the provision of a guarantee of differential privacy (Mirshghallah et al., 2020);

- Input perturbation; involves the addition of noise to the raw dataset.
- Gradient perturbation; refers the addition of noise to gradients as a means of perturbation.
- Objective perturbation; involves the injection of noise into the results of the objective function of an optimization problem.
- Label perturbation; the integration of noise into the voting mechanism that decides the assignment of labels
- Output perturbation; the outputs of a function running on a database are made noisy.

In this research, the method of input perturbation was selected as the preferred technique. The main reasons are outlined below:

- Among these perturbation techniques, only the input perturbation recognizes the presence of untrusted third parties,
- Additionally, it facilitates the sharing of sensitive data while maintaining privacy.

Differential privacy can be classified into two categories, namely global and local, as stated in reference (Arachchige et al., 2019). In the context of global differential privacy, a curator deemed trustworthy is responsible for data collection. Subsequently, this curator generates noisy responses to queries posed by untrusted third parties. In the context of local differential privacy, data perturbation occurs at the data owner's side, after which the perturbed data is shared with untrusted third parties or servers. Figure 1 illustrates the operational mechanisms of global and local differential privacy.

The proliferation of digital platforms and smart televisions has facilitated the consumption of movies and films by viewers, who can then express their preferences through voting mechanisms. Therefore, this scenario enables platform businesses to gather data through the utilization of local perturbation, similar to the approach employed by Apple in reference (Differential Privacy). The utilization of local differential privacy was motivated by the observation made in (Fukuchi et al. 2017) that input perturbation adheres to the principles of local differential privacy. Consequently, local differential privacy was also employed in this study.

**Definition**; Let *D* represent an input database, *x* denote a d-dimensional vector representing any data instance in *D*, *i* and *j* symbolize row number and attribute number, respectively. Input perturbation can be defined as the following transformation:

$$x_i^j\_private = x_i^j + Laplace(\frac{\Delta f}{\epsilon}) \tag{4}$$

where;

$$\Delta f = \|\max(x^j) - \min(x^j)\| \quad (5)$$

$\Delta f$ is calculated for each attribute and then a differentially private version of $D$ can be obtained by using Equation 4.

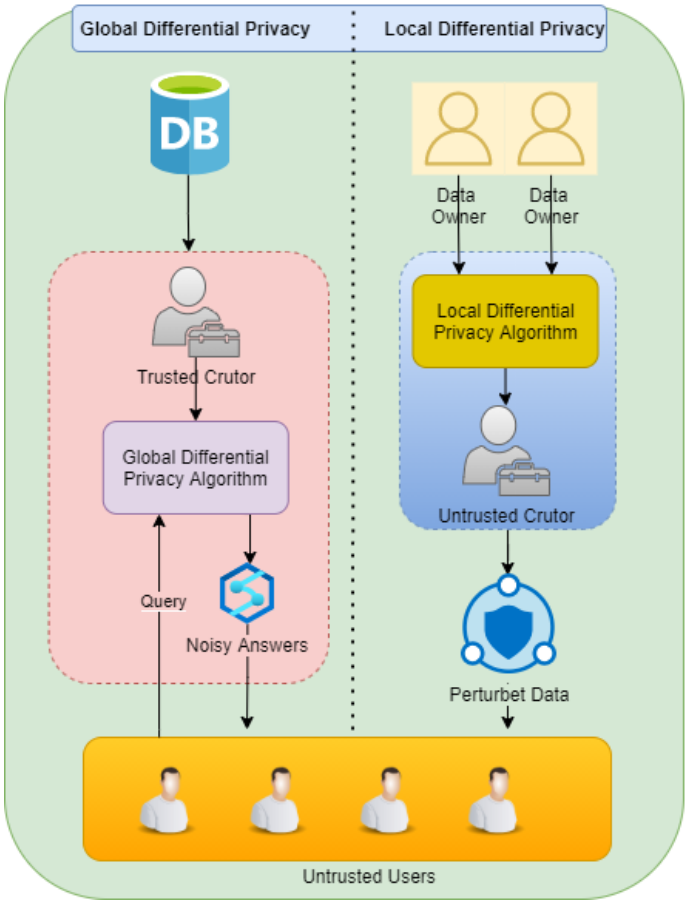


***Figure 1***. *Global and local differential privacy*

## 5. EXPERIMENTS

Recommender systems utilize behavioral and preference data, regarded as personal data. Given the potential lack of trustworthiness associated with such systems, it is imperative to provide privacy measures. Firstly, this work presents an analysis of the outcomes of a recommendation system that utilizes collaborative filtering, specifically focusing on the effects of perturbed input data. Secondly, this study aims to examine the outcomes of hybrid, collaborative, and content-based filtering algorithms when applied to input data that has been perturbed by noise.

### 5.1 Performance evaluation metrics

Let, $t_x$ be the actual value, $t'_x$ be the predicted value, and $n$ is the number of samples.

RMSE is calculated by using Equation 6;

$$RMSE = \sqrt{\frac{1}{n}\sum_{x=1}^{n}(t_x - t'_x)^2} \quad (6)$$

MAE is calculated using Equation 7 by averaging the absolute values of the differences between the predicted values and the actual values.

$$MAE = \frac{1}{n}\sum_{x=1}^{n}|t_x - t'_x| \quad (7)$$

### 5.2 Experiment-1

In conducting the first experiment, the Movies Dataset is utilized (The Movies Dataset). The dataset consists of 100,000 evaluation points for 9,066 movies made by 671 users.

Figure 2 illustrates the whole process of the system as applied in Experiment-1. In the traditional recommendation system, denoted in path 1, the original dataset is provided to the collaborative filtering algorithm. Subsequently, the resulting outputs are assessed using specific metrics, leading to the final outcomes. In the context of the differentially private recommender system, as depicted in path 2, the process involves the application of input perturbation to the original data, resulting in the acquisition of perturbed data. The perturbed data is subjected to collaborative filtering, and the resulting outputs are assessed using several metrics to achieve the final results.

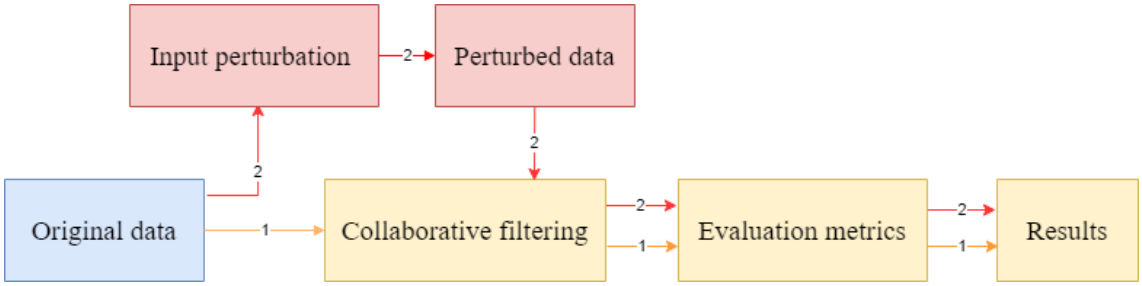


***Figure 2**. The general workflow in Experiment-1*

The first experiment yielded the performance outcomes of a recommender system that respects privacy. The most often used collaborative filtering algorithms include SVD, kNNBasic, SVDpp, kNNwithZscore, and CoClustering. This experiment utilized the aforementioned techniques to assess the performances on perturbed data. As previously stated, input perturbation was employed on the raw data, and the outcomes were assessed across several epsilon values, specifically 5, 10, 50 and 100. The performance criteria are selected as RMSE and MAE (Canbay & Taş, 2022). The results that were acquired were presented in Table 1. Furthermore, Figure 3 presents a visual representation that allows for a comparative analysis of the outcomes.

***Table 1**. The results of Experiment-1*

| Epsilon | SVD | kNNBasic | SVDpp | kNNwithZscore | CoClustering | Metrics |
|---|---|---|---|---|---|---|
| 100 | 0.8983 | 0.9702 | 0.8889 | 0.9199 | 1.0705 | RMSE |
| | 0.6923 | 0.7466 | 0.6822 | 0.7012 | 0.8611 | MAE |
| 50 | 0.9066 | 0.9771 | 0.8959 | 0.9271 | 1.0754 | RMSE |
| | 0.6986 | 0.7530 | 0.6876 | 0.7069 | 0.8642 | MAE |
| 10 | 1.1184 | 1.1775 | 1.1212 | 1.1372 | 1.2567 | RMSE |
| | 0.8696 | 0.9142 | 0.8702 | 0.8788 | 1.0010 | MAE |
| 5 | 1.6292 | 1.6469 | 1.6702 | 1.6156 | 1.6983 | RMSE |
| | 1.2458 | 1.2585 | 1.2807 | 1.2294 | 1.3144 | MAE |
| no privacy | 0.8969 | 0.9675 | 0.8863 | 0.9174 | 0.9618 | RMSE |
| | 0.6907 | 0.7439 | 0.6794 | 0.6984 | 0.7452 | MAE |

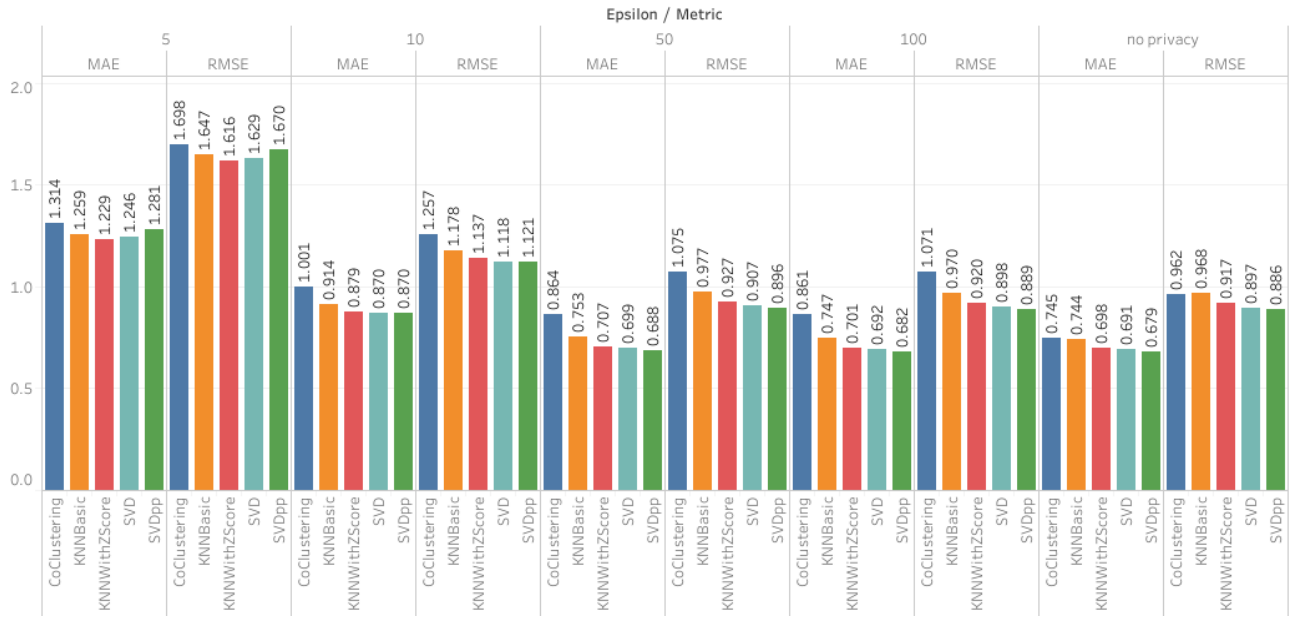


***Figure 3.*** *Graphical comparison of the results of Experiment-1*

As depicted in Figure 3, RMSE and MAE values present a decreasing trend, whereas the epsilon values increase. Furthermore, RMSE and MAE show optimal performance for epsilon 100. The results indicate that our technique demonstrates successful applicability in privacy-preserving recommender systems.

### 5.3 Experiment-2

In this experiment, Goodbooks dataset is employed (Goodbooks Dataset). The dataset encompasses the ratings provided by 53,424 individuals for a collection of 10,000 widely-read novels.

The workflow for Experiment-2 is illustrated in Figure 4. Traditional recommendation system, denoted by path 1, processes the original raw data. Following this, the generated outputs are assessed using specific metrics, and ultimately, the results are acquired. Nevertheless, the perturbed data is provided as input for algorithms that perform collaborative filtering, content-based filtering, and hybrid filtering. The assessments are conducted utilizing a variety of evaluation metrics, and a comparison is made of the results.

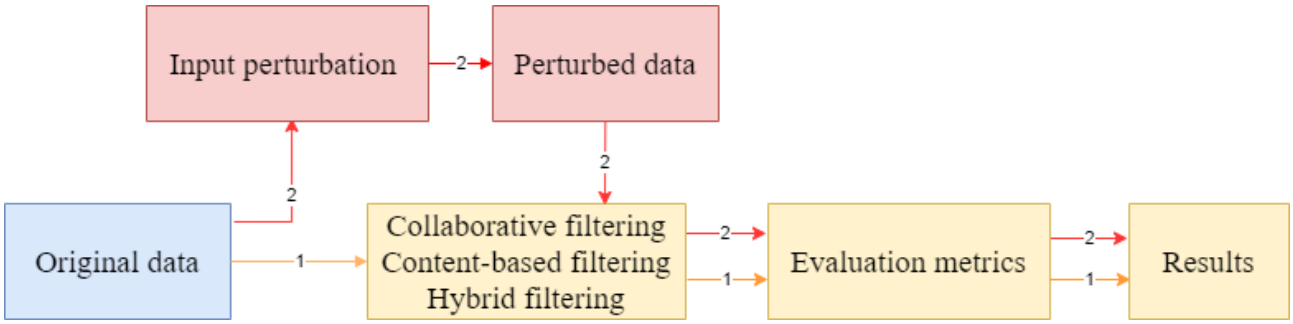


***Figure 4****. The general workflow in Experiment-2*

In this experiment, we conducted a comparison of hybrid, collaborative filtering, and content-based filtering methodologies. Experiments are undertaken for various epsilon values. Based on the findings depicted in Table 2 and Figure 5, it can be observed that the optimal performance in terms of RMSE and MAE is achieved for epsilon 100. Therefore, it is evident that hybrid recommendation systems yield the lowest error rates.

***Table 2**. The results of Experiment-2*

| Epsilon | Hybrid filtering | Collaborative filtering | Content-based filtering | Metrics |
|---|---|---|---|---|
| 100 | 0.90664 | 1.08282 | 0.95007 | RMSE |
| | 0.69813 | 0.86515 | 0.72732 | MAE |
| 50 | 0.91604 | 1.0905 | 0.96019 | RMSE |
| | 0.70704 | 0.87385 | 0.73434 | MAE |
| 10 | 1.10979 | 1.26876 | 1.15695 | RMSE |
| | 0.85954 | 1.00469 | 0.89377 | MAE |
| 5 | 1.57873 | 1.72447 | 1.6389 | RMSE |
| | 1.19621 | 1.32805 | 1.25 | MAE |
| no privacy | 0.90327 | 0.97872 | 0.93991 | RMSE |
| | 0.69786 | 0.75621 | 0.71738 | MAE |

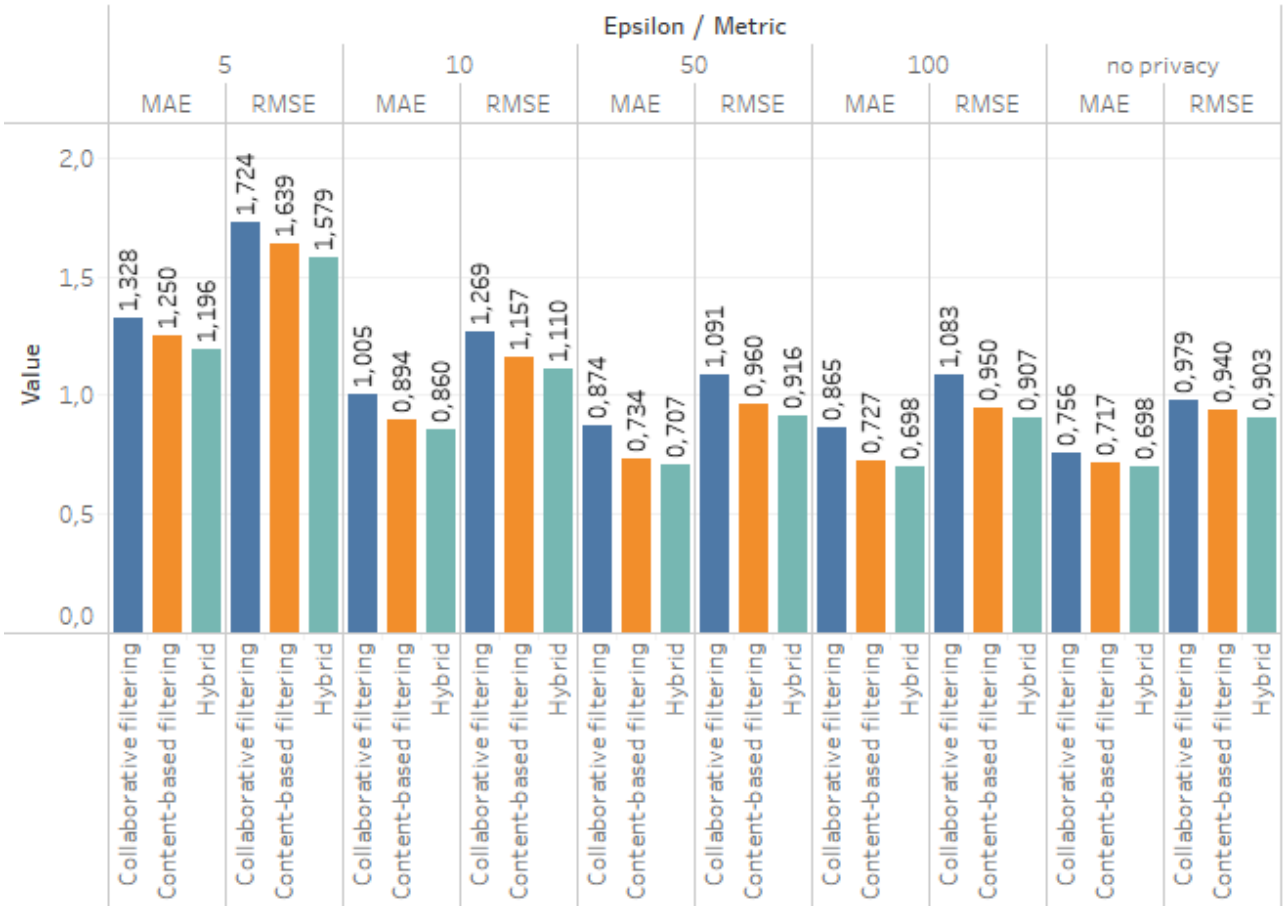


***Figure 5.** Graphical comparison of the results of Experiment-2*

## 6. CONCLUSION

The widespread adoption of smartphones has significantly contributed to the increased utilization of online shopping platforms. As the shift towards online retail continues to gain momentum, consumers are increasingly experiencing a sense of inundation due to the abundance of choices available to them. Consequently, they encounter difficulties in locating a product or service that aligns with their specific requirements and desired outcomes. Most online shopping platforms employ recommendation systems to assist customers in discovering things that may be of interest to them.

Recommendation systems heavily depend on users' behavioral and preference data, such as ratings, likes and dislikes, to generate precise recommendations. Nevertheless, customers often encounter privacy concerns as a result of unethical data gathering and analytical practices conducted by service providers. This study presents a comparative study for a recommender system that prioritizes privacy considerations. Input perturbation is utilized to perturb the raw data.

The experiments were conducted on two distinct cases. The first experiment is the comparison of collaborative filtering approaches on perturbed data. It has been observed that an increase in epsilon leads to a decrease in error. Among the algorithms, SVD++ gave the minimum error rates for all metrics. The second experiment investigates the efficacy of hybrid, collaborative, and content-based filtering techniques. The results showed that the hybrid filtering strategy demonstrated the most favorable outcomes in terms of error rates, regardless of the epsilon values. Both experiments have shown that the implementation of differential privacy effectively protects the privacy of users in recommender systems while providing higher utility.

## AUTHOR CONTRIBUTIONS

Methodology and writing-reviewing, Y.C and A.U; editing, Y.C and A.U; conceptualization and software, Y.C and A.U. All authors have read and legally accepted the final version of the article published in the journal.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

Altman, I. (1976). A conceptual analysis. *Environment and behavior*, *8*(1), 7-29. https://doi.org/10.1177/001391657600800102

Arachchige, P. C. M., Bertok, P., Khalil, I., Liu, D., Camtepe, S., & Atiquzzaman, M. (2019). Local Differential Privacy for Deep Learning. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2019.2952146

Badsha, S., Yi, X., Khalil, I., & Bertino, E. (2017). *Privacy preserving user-based recommender system.* IEEE 37th International Conference on Distributed Computing Systems (ICDCS). https://doi.org/10.1109/ICDCS.2017.248

Banisar, D., & Davies, S. (1999). Privacy and human rights: An international survey of privacy laws and practice. *Global Internet Liberty Campaign*. (Accessed:17/04/2023) https://gilc.org/privacy/survey/intro.html

Cai, G., Lee, K., & Lee, I. (2018). Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos. *Expert Systems with Applications, 94*, 32-40. https://doi.org/10.1016/j.eswa.2017.10.049

Canbay, P., & Hüseyin, T. (2022). Yapıların Isıtma ve Soğutma Yükünün Yapay Zeka ile Tahmini. *International Journal of Pure and Applied Sciences, 8*(2), 478-489. https://doi.org/10.29132/ijpas.1166227

Cunha, T., Soares, C., & de Carvalho, A. C. (2018). Metalearning and Recommender Systems: A literature review and empirical study on the algorithm selection problem for Collaborative Filtering. *Information Sciences, 423*, 128-144. https://doi.org/10.1016/j.ins.2017.09.050

Differential Privacy. (Accessed:23/05/2023) https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf

Dwork, C. (2006). *Differential Privacy*. International Colloquium on Automata, Languages and Programming, Berlin, Heidelberg. https://doi.org/10.1007/11786986

Fukuchi, K., Tran, Q. K., & Sakuma, J. (2017). *Differentially private empirical risk minimization with input perturbation.* International Conference on Discovery Science. https://doi.org/10.1007/978-3-319-67786-6_6

Garanayak, M., Mohanty, S. N., Jagadev, A. K., & Sahoo, S. (2019). Recommender system using item based collaborative filtering (CF) and K-means. *International Journal of Knowledge-based and Intelligent Engineering Systems*, *23*(2), 93-101. https://doi.org/10.3233/KES-190402

Geetha, G., Safa, M., Fancy, C., & Saranya, D. (2018). *A hybrid approach using collaborative filtering and content based filtering for recommender system.* Journal of Physics: Conference Series. https://doi.org/10.1088/1742-6596/1000/1/012101

Goodbooks Dataset. (Accessed:02/06/2023) https://www.kaggle.com/datasets/zygmunt/goodbooks-10k

Guo, Y., Wang, M., & Li, X. (2017). Application of an improved Apriori algorithm in a mobile e-commerce recommendation system. *Industrial Management & Data Systems*, *117*(2), 287-303. https://doi.org/10.1108/IMDS-03-2016-0094

Hu, M., Wu, D., Wu, R., Shi, Z., Chen, M., & Zhou, Y. (2021). RAP: A Light-Weight Privacy-Preserving Framework for Recommender Systems. *IEEE Transactions on Services Computing*, *15*(5), 2969-2981. https://doi.org/10.1109/TSC.2021.3065035

Javed, U., Shaukat, K., Hameed, I. A., Iqbal, F., Alam, T. M., & Luo, S. (2021). A review of content-based and context-based recommendation systems. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(3), 274-306. https://doi.org/10.3991/ijet.v16i03.18851

Kawasaki, M., & Hasuike, T. (2017). *A recommendation system by collaborative filtering including information and characteristics on users and items.* 2017 ieee symposium series on computational intelligence. https://doi.org/10.1109/SSCI.2017.8280983

Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems–A survey. *Knowledge-based systems*, *123*, 154-162. https://doi.org/10.1016/j.knosys.2017.02.009

Li, N., Li, T., & Venkatasubramanian, S. (2007). *t-Closeness: Privacy Beyond k-Anonymity and l-Diversity*. IEEE International Conference on Data Engineering, Istanbul, Turkey. https://doi.org/10.1109/ICDE.2007.367856

Liu, X., Li, Q., Ni, Z., & Hou, J. (2019). *Differentially private recommender system with autoencoders.* 2019 International Conference on Internet of Things and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). https://doi.org/10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00094

Machanavajjhala, A., Gehrke, J., Kifer, D., & Venkitasubramaniam, M. (2006). *l-Diversity: Privacy Beyond k-Anonymity*. International Conference on Data Engineering, Atlanta, USA. https://doi.org/10.1109/ICDE.2006.1

Mehrotra, R., McInerney, J., Bouchard, H., Lalmas, M., & Diaz, F. (2018). *Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems.* Acm international conference on information and knowledge management. https://doi.org/10.1145/3269206.3272027

Milano, S., Taddeo, M., & Floridi, L. (2020). Recommender systems and their ethical challenges. *AI & society*, *35*, 957-967. https://doi.org/10.1007/s00146-020-00950-y

Mirshghallah, F., Taram, M., Vepakomma, P., Singh, A., Raskar, R., & Esmaeilzadeh, H. (2020). Privacy in Deep Learning: A Survey. *arXiv preprint arXiv:2004.12254*. https://doi.org/10.48550/arXiv.2004.12254

The Movies Dataset. (Accessed:13/05/2023) https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset

Neera, J., Chen, X., Aslam, N., Wang, K., & Shu, Z. (2021). Private and Utility Enhanced Recommendations with Local Differential Privacy and Gaussian Mixture Model. *IEEE Transactions on Knowledge and Data Engineering*. https://doi.org/10.1109/TKDE.2021.3126577

Nergiz, M. E., Atzori, M., & Clifton, C. (2007). *Hiding the Presence of Individuals from Shared Databases*. ACM SIGMOD International Conference on Management of Data, Beijing, China. https://doi.org/10.1145/1247480.1247554

Selvaraj, S., & Gangadharan, S. S. (2021). Privacy preserving hybrid recommender system based on deep learning. *Turkish Journal of Electrical Engineering and Computer Sciences, 29*(5), 2385-2402. https://doi.org/10.3906/elk-2010-40

Sütçü, M., Ecem, K., & Erdem, O. (2021). Movie Recommendation Systems Based on Collaborative Filtering: A Case Study on Netflix. *Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi, 37*(3), 367-376.

Sweeney, L. (2002). k-Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10*(5), 557-570. https://doi.org/10.1142/S0218488502001648

Wang, J., & Wang, A. (2020). *An improved collaborative filtering recommendation algorithm based on differential privacy*. 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS). https://doi.org/10.1109/ICSESS49938.2020.9237702

Warren, S. D., & Brandeis, L. D. (1890). The Right to Privacy. *Harvard law review*, 193-220. https://doi.org/10.2307/1321160

Xian, Z., Li, Q., Li, G., & Li, L. (2017). New collaborative filtering algorithms based on SVD++ and differential privacy. *Mathematical Problems in Engineering, 2017*. https://doi.org/10.1155/2017/1975719

Xiong, P., Zhang, L., Zhu, T., Li, G., & Zhou, W. (2020). Private collaborative filtering under untrusted recommender server. *Future Generation Computer Systems, 109*, 511-520. https://doi.org/10.1016/j.future.2018.05.077

Yin, C., Shi, L., Sun, R., & Wang, J. (2020). Improved collaborative filtering recommendation algorithm based on differential privacy protection. *The Journal of Supercomputing, 76*, 5161-5174. https://doi.org/10.1007/s11227-019-02751-7

Yu, T., Shen, Y., & Jin, H. (2019). *A visual dialog augmented interactive recommender system*. ACM SIGKDD international conference on knowledge discovery & data mining. https://doi.org/10.1145/3292500.3330991

Zhang, S., Liu, L., Chen, Z., & Zhong, H. (2019). Probabilistic matrix factorization with personalized differential privacy. *Knowledge-based systems, 183*, 104864. https://doi.org/10.1016/j.knosys.2019.07.035

Zhao, J., Geng, X., Zhou, J., Sun, Q., Xiao, Y., Zhang, Z., & Fu, Z. (2019). Attribute mapping and autoencoder neural network based matrix factorization initialization for recommendation systems. *Knowledge-based systems, 166*, 132-139. https://doi.org/10.1016/j.knosys.2018.12.022