

PD-GAN: Adversarial Learning for Personalized Diversity-Promoting Recommendation

Qiong Wu^{1,2}, Yong Liu^{1,2,*}, Chunyan Miao^{1,2,3,*}, Binqiang Zhao⁴, Yin Zhao⁴, Lu Guan⁴

¹Alibaba-NTU Singapore Joint Research Institute

²The Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly (LILY)

³School of Computer Science and Engineering, Nanyang Technological University

⁴Alibaba Group

{wu.qiong, stephenliu, ascymiao}@ntu.edu.sg, {binqiang.zhao, zhaoyin.zy, guanlu.gl}@alibaba-inc.com

Abstract

This paper proposes Personalized Diversity-promoting GAN (PD-GAN), a novel recommendation model to generate diverse, yet relevant recommendations. Specifically, for each user, a generator recommends a set of diverse and relevant items by sequentially sampling from a personalized Determinantal Point Process (DPP) kernel matrix. This kernel matrix is constructed by two learnable components: the general co-occurrence of diverse items and the user’s personal preference to items. To learn the first component, we propose a novel pairwise learning paradigm using training pairs, and each training pair consists of a set of diverse items and a set of similar items randomly sampled from the observed data of all users. The second component is learnt through adversarial training against a discriminator which strives to distinguish between recommended items and the ground-truth sets randomly sampled from the observed data of the target user. Experimental results show that PD-GAN is superior to generate recommendations that are both diverse and relevant.

1 Introduction

Recommender systems have today become irreplaceable in online world to help users filter through a vast amount of information. A typical recommender system usually selects top items (e.g., products, movies) ranked by their match to users’ personal preferences and interests [Liu *et al.*, 2013; Hu *et al.*, 2014; Liu *et al.*, 2014; 2016; 2017]. However, this strategy may often yield suboptimal recommendations. For example, collaborative filtering favors popular items and thus may recommend popular items that are already known to the user [Ashkan *et al.*, 2015]. Content-based filtering may produce items matching the user’s interests but covering a very narrow scope of topics [Qin and Zhu, 2013].

This led researchers to rethink about the ideal recommendation strategy. Some researchers point out that increasing the diversity of recommendation results to cover a broader

range of users’ interests can improve user satisfaction [Zhang and Hurley, 2008]. Moreover, diversifying recommendation results can play an important role in broadening users’ horizons and helping online service providers to explore users’ potential interests [Cheng *et al.*, 2017].

These thoughts give birth to various diversity-promoting recommendation models [Kunaver and Požrl, 2017]. A common strategy adopted by them is to maintain a trade-off between relevance and diversity of recommendation lists [Zhou *et al.*, 2010]. However, such a strategy often achieves high diversity with a huge sacrifice of relevance. We argue that relevance can be largely preserved or even boosted if personalized diversity is considered. Indeed, it is reasonable to diversify recommendation results for users with broad interests, but blindly pursuing diversity for users with focused interests will definitely hurt the relevance of recommendation results.

In light of this, we propose a novel personalized diversity-promoting recommendation model that considers two aspects of the user’s personal preference: **the personal preference to individual items, and the personal preference to the diversity of a set of items**. Specifically, we randomly sample multiple sets of diverse items from each user’s observed data as the ground-truth. Naturally, the diversity of a ground-truth set sampled from focused users will be smaller than that from eclectic users. The beauty of this approach is that each item in the ground-truth set captures the user’s personal preference to individual items, while the items as a set capture the user’s personal preference to the diversity of items.

What we can do next is to fit a trainable model with the ground-truth and then use the trained model to generate recommendations that are as diverse and relevant as the ground-truth. In this work, we build a novel recommendation model using the GAN framework, called Personalized Diversity-promoting GAN (PD-GAN), which consists of a generative network (generator) and a discriminative network (discriminator) contesting with each other. Specifically, we build the generator of PD-GAN based on a Determinantal Point Process (DPP) model [Chen *et al.*, 2018]. For each user, the generator maintains a personalized DPP kernel matrix, which is constructed from two learnable components: (1) the general co-occurrence of diverse items, which captures the diversity of items, and (2) the user’s personal preference to items, which captures the relevance of items.

*Corresponding author

Firstly, the general co-occurrence of diverse items can be learnt by fitting a DPP kernel matrix with randomly sampled diverse sets from all user’s observed data. However, the kernel matrix learnt in this way favors popular items since such items have greater chance to be sampled in those diverse sets. To mitigate this problem, we propose a novel pairwise learning paradigm using training pairs. Specifically, for each randomly sampled diverse set, we also randomly sample a set of similar items (subsequently referred to as a similar set) as a negative example. The learning objective is to maximize the probability of sampling the diverse sets and minimize the probability of sampling the similar sets. As popular items also get higher chance to be sampled in the similar sets, the effect of item popularity can be reduced.

Secondly, the user’s personal preference to items is learnt through adversarial training against the discriminator which strives to distinguish between the items recommended by the generator and the ground-truth sampled from the target user’s observed data. In this way, the user’s personal preference to an individual item as well as the diversity of a set of items can be learnt from the ground-truth. Through adversarial learning, the generative model will be able to generate diverse and relevant recommendations resembling the ground-truth so much that the discriminator will not distinguish easily.

Major contributions are summarized below:

- We propose a novel GAN framework to capture users’ personal preferences to both the individual items and the diversity of a set of items, which can significantly improve recommendation diversity while largely preserve or even boost relevance.
- We propose a novel pairwise learning method for capturing the co-occurrence of diverse items with a DPP kernel matrix, which can largely reduce the negative impact of item popularity.
- Experimental results show that PD-GAN is superior to generate items that are both diverse and relevant comparing with several competitive state-of-the-art baselines.

2 Related Works

Diversity of recommendation can be viewed at either an aggregate level or an individual level. The aggregate diversity refers to the recommendation diversity across all users, which generally reflect the ability of a recommender system to recommend long-tail items [Cheng *et al.*, 2017]. Thus, aggregate diversity is usually evaluated by long-tail metrics [Adomavicius and Kwon, 2012]. Individual diversity is targeted at each individual user, which reflect the ability of a recommender system to generate diversified recommendations for each target user. The evaluation methods for individual diversity can be generally classified into two categories: pairwise measures and set-level measures. Pairwise measures usually define a dissimilarity function between items and use the average dissimilarity to characterize the diversity of a recommendation list. The dissimilarity can be defined based on attributes [Ashkan *et al.*, 2015], feature space [Qin and Zhu, 2013], clustering [Lee *et al.*, 2017], or explanation [Yu *et al.*, 2009]. Set-level measures evaluate the diversity of a list as a

whole. Explicit set-level measures model topics and use topic coverage as the diversity measure [Santos *et al.*, 2010]. Implicit set-level measures define diversity in the feature space of the entire list of items, which are often used by determinantal point process models [Chen *et al.*, 2018].

A common approach for diversifying recommendation is to maintain a trade-off between relevance and diversity. Some approaches directly improve diversity by representing relevance and diversity with independent metrics and maximizing the marginal relevance which is a convex combination of the two metrics [Carbonell and Goldstein, 1998; Santos *et al.*, 2010; Yu *et al.*, 2014]. Another group of approaches learn a diversified model by maximizing a submodular objective function [Qin and Zhu, 2013; Ashkan *et al.*, 2015]. However, these approaches usually adopt pairwise measures to characterize the overall diversity of the recommendation list, which may not capture the complex relationships among items [Chen *et al.*, 2018]. An elegant solution is brought by the Determinantal Point Process (DPP) model, which is a probabilistic model for set diversity. The DPP model has recently become popular to model set-level diversity in recommender systems. It has been used to recommend complementary products in shopping basket by learning the kernel matrix of DPP to characterize the relations among items [Gartrell *et al.*, 2017]. It has also been used to generate diverse recommendation lists through Maximum A Posteriori (MAP) inference based on carefully constructed kernel matrices [Chen *et al.*, 2018]. All above mentioned approaches, however, usually do not consider the user’s personal preference to the diversity of items. Our approach takes this information into consideration and strives to generate recommendations that are both diverse and relevant.

3 PD-GAN Model

Without losing generality, suppose there are M users, N items and C item categories in the recommender system. Each item belongs to at least one item category. For each user, PD-GAN will generate top- K recommendations. The objective of PD-GAN is to make top- K recommendations as diverse as possible, and at the same time, as relevant as possible. **Here, diversity is measured by category coverage** [Santos *et al.*, 2010; Wu *et al.*, 2016; 2017].

To achieve this goal, for each user u , we can randomly sample an observed diverse set $T = \{i_1, i_2, \dots\}$ from his/her user-item interaction history, where items in T belong to different categories. The cardinality of T will naturally be larger for eclectic users than that for focused users. For example, for a focused user whose watch history covers 3 categories, the size of T will be no larger than 3; and for a user with broad interest whose watch history covers 20 categories, the size of T will be no larger than 20. As such, for each user u , we can sample multiple diverse sets, denoted by $\mathcal{T}_u = \{T_1, T_2, \dots\}$, which serves as the ground-truth. Each individual item in the ground-truth sets reflects the user’s personal preference to a certain item, while each set of items capture the user’s personal preference to diversity.

To generate diverse and relevant recommendations that resemble the ground-truth, PD-GAN adopt an adversarial learn-

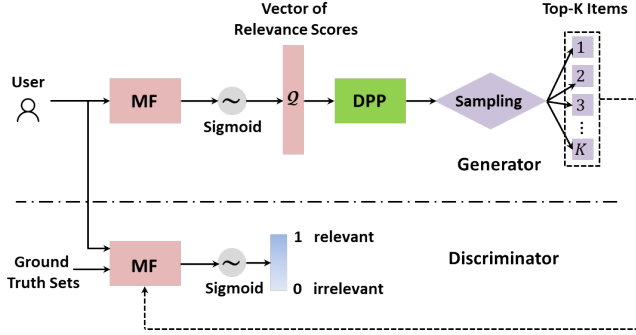


Figure 1: The structure of PD-GAN.

ing framework, as illustrated in Figure 1. For each user, the generator of PD-GAN first evaluates the relevance of all items through Matrix Factorization [Koren *et al.*, 2009], and then incorporate the relevance scores with a pre-learned DPP model. The pre-learned DPP model captures the general co-occurrence of diverse items, or in other words, item diversity. After that, top- K items can be sequentially sampled from the modified DPP model. The discriminator of PD-GAN attempts to distinguish between top- K items generated by the generator and the ground-truth. Through adversarial learning, the generator will be able to generate a set of diverse and relevant items so similar to the ground-truth that the discriminator will not be able to distinguish easily.

Next, we will first introduce how to learn the general co-occurrence of diverse items through a DPP model, and then detail the adversarial learning framework for capturing users' personal preferences from the ground-truth.

3.1 Learning Co-occurrence of Diverse Items

Conditional DPP

DPP is an elegant probabilistic model for modeling set diversity [Kulesza *et al.*, 2012]. DPP on a discrete set $\mathcal{I} = \{i_1, i_2, \dots, i_N\}$ is a probability measure \mathcal{P} on $2^{\mathcal{I}}$, the set of all subsets of \mathcal{I} . In the context of item recommendation, \mathcal{I} is all items. When \mathcal{P} gives nonzero probability to the empty set, there exists a kernel matrix $\mathbf{L} \in \mathbb{R}^{N \times N}$ such that for every subset $T \subseteq \mathcal{I}$, the probability of T is $\mathcal{P}(T) \propto \det(\mathbf{L}_T)$, where \mathbf{L} is a real, positive semi-definite kernel matrix indexed by the elements of \mathcal{I} , and \mathbf{L}_T is \mathbf{L} restricted to only those rows and columns which are indexed by T , i.e., $\mathbf{L}_T \equiv [\mathbf{L}_{ij}]_{i,j \in T}$. Intuitively, the diagonal entry L_{ii} of the kernel matrix \mathbf{L} captures the quality of item i , while the off-diagonal entry L_{ij} measures the similarity between items i and j . Following the observation that $\sum_{T' \subseteq \mathcal{I}} \det(\mathbf{L}_{T'}) = \det(\mathbf{L} + \mathbf{I})$, where \mathbf{I} is the identity matrix, the conditional probability of observing T is :

$$\mathcal{P}(T) = \frac{\det(\mathbf{L}_T)}{\det(\mathbf{L} + \mathbf{I})}. \quad (1)$$

Learning of L-kernel

L-kernel can be learnt from observed diverse sets. However, when the number of items N is large, learning a nonparametric full-rank \mathbf{L} is very computationally expensive. In this work, we follow [Gartrell *et al.*, 2017]'s approach, and use

low-rank factorization of the $N \times N$ \mathbf{L} matrix, which is much more computationally efficient, as follows:

$$\mathbf{L} = \mathbf{A}\mathbf{A}^T, \quad (2)$$

where \mathbf{A} is a $N \times D$ matrix and $D \ll N$.

To learn the L-kernel, we can randomly sample a collection of observed diverse sets from all user-item interaction history, denoted by $\mathcal{T} = \bigcup_{u=1}^M \mathcal{T}_u = \{T_1, T_2, \dots, T_X\}$, and fit \mathbf{L} by maximizing $\log \mathcal{P}(\mathcal{T}|\mathbf{A})$ through gradient descent. In this way, L-kernel is optimized towards giving higher conditional probability to those sets of frequently co-occurred diverse items. However, a problem with this training method is that the L-kernel is biased towards popular items, as popular items have higher chance to be sampled in the observed diverse sets. As such, popular items that are similar to each other may be sampled from the trained L-kernel due to high perceived quality (reflected by the diagonal entries of \mathbf{L}).

To mitigate this problem, we propose a novel pairwise training method for learning the L-kernel. For each observed diverse set T , we also sample a similar set \tilde{T} from the observation, which consists of items from the same category. As such, we can get X training pairs $\langle T, \tilde{T} \rangle = \{\langle T_1, \tilde{T}_1 \rangle, \langle T_2, \tilde{T}_2 \rangle, \dots, \langle T_X, \tilde{T}_X \rangle\}$. Our learning objective becomes maximizing the probability of sampling \mathcal{T} and minimizing the probability of sampling $\tilde{\mathcal{T}}$, formally:

$$\begin{aligned} \max_{\mathbf{A}} \mathcal{J}_{dpp} &= \log \mathcal{P}(\mathcal{T}|\mathbf{A}) - \log \mathcal{P}(\tilde{\mathcal{T}}|\mathbf{A}) \\ &= \sum_{x=1}^X \log \mathcal{P}(T_x|\mathbf{A}) - \log \mathcal{P}(\tilde{T}_x|\mathbf{A}) \\ &= \sum_{x=1}^X \log \det(\mathbf{L}_{[T_x]}) - \log \det(\mathbf{L}_{[\tilde{T}_x]}), \end{aligned} \quad (3)$$

where $[T_x]$ and $[\tilde{T}_x]$ index the observations in \mathcal{T} and $\tilde{\mathcal{T}}$, respectively. In this way, the effect of popularity will be mitigated by minimizing $\log \mathcal{P}(\tilde{\mathcal{T}}|\mathbf{A})$, as popular items also get higher chance to be sampled in the similar sets.

Sampling

Once the L-kernel is determined, DPP can sequentially sample top- K diverse items. Here, we adopt the fast greedy MAP inference [Chen *et al.*, 2018] for sequentially sampling from the L-kernel, which has been shown to be very computationally efficient. Subsequently in this paper, the sequential sampling process of DPP is denoted by $\mathcal{S}_{K-DPP}(\mathbf{L})$.

3.2 Adversarial Learning

Generator

As introduced in the previous section, the L-kernel of the DPP model captures the general co-occurrence of diverse items in all user-item interactions, and the log-probability of sampling any subset of items $T \subseteq \mathcal{I}$ is:

$$\log \mathcal{P}(T) \propto \log \det(\mathbf{L}_T). \quad (4)$$

However, this kernel does not reflect users' personal preferences to items. To incorporate users' personal preferences, we modify the log-probability of T to:

$$\log \mathcal{P}(T) \propto \alpha \cdot \sum_{i \in T} Q(u, i) + (1 - \alpha) \cdot \log \det(\mathbf{L}_T), \quad (5)$$

where $Q(u, i)$ is a quality-evaluating function that evaluates the perceived relevance of i from u , and $\alpha \in [0, 1]$ is a trade-off parameter between the user's personal preference to individual items and item diversity.

Equation (5) corresponds to a DPP with the following kernel [Chen *et al.*, 2018]:

$$\mathbf{L}_u = \text{Diag}(\exp(\beta Q(u))) \cdot \mathbf{L} \cdot \text{Diag}(\exp(\beta Q(u))), \quad (6)$$

where $\beta = \frac{\alpha}{2(1-\alpha)}$. Hence, we can simply re-construct the \mathbf{L} -kernel and do sequential sampling once the quality-evaluating function is defined.

There are many quality-evaluating methods available in literature, such as Matrix Factorization (MF) [Koren *et al.*, 2009], Factorization Machine (FM) [Rendle, 2010], as well as Deep Neural Network-based methods [He and Chua, 2017]. Without losing generality, in this work, we adopt logistic MF [Johnson, 2014] as an example, formally:

$$Q(u, i) = \sigma(\mathbf{v}_G^u \mathbf{v}_G^i{}^T), \quad (7)$$

where $\sigma(a) = \frac{1}{1+\exp(-a)}$ is the sigmoid function, and \mathbf{v}_G^u and \mathbf{v}_G^i are $1 \times D$ embedding vectors for u and i , respectively.

Now, a diverse set of K items, denoted by T_G , can be generated by the generator $\mathcal{G}_\theta(u)$ for each user u with \mathbf{L}_u , formally:

$$\mathcal{G}_\theta(u) \sim \mathcal{S}_{K\text{-DPP}}(\mathbf{L}_u), \quad (8)$$

where the trainable parameter θ includes \mathbf{v}_G^u and \mathbf{v}_G^i .

Discriminator

The discriminator $\mathcal{D}_\eta(u, T_G)$ evaluates the overall relevance of the generated set T_G for user u : it outputs a value in $[0, 1]$, with 1 meaning the highest relevance and 0 meaning the lowest relevance. To be consistent with the generator, we also adopt logistic MF as the quality-evaluating method for the discriminator. Specifically, we formulate $\mathcal{D}_\eta(u, T_G)$ as:

$$\mathcal{D}_\eta(u, T_G) = \sigma\left(\frac{1}{|T_G|} \sum_{i \in T_G} \mathbf{v}_D^u \mathbf{v}_D^i\right), \quad (9)$$

where the trainable parameter η includes \mathbf{v}_D^u and \mathbf{v}_D^i .

Optimization

The goal of the generator is to generate top- K items that could deceive the discriminator. Hence, given \mathcal{D}_η , we learn \mathcal{G}_θ by minimizing the following objective:

$$\min_{\theta} \sum_{u=1}^M \log(1 - \mathcal{D}_\eta(u, \mathcal{G}_\theta)). \quad (10)$$

Due to the discrete sampling process $\mathcal{S}_{K\text{-DPP}}(\mathbf{L}_u)$ in \mathcal{G}_θ , it cannot be directly optimized through gradient descent. Such problem can be solved with policy gradient method originated from reinforcement learning. In this work, we adopt the REINFORCE [Wang *et al.*, 2017] method to optimize \mathcal{G}_θ . Given T_G , the objective can be reformulated as follows:

$$\max_{\theta} \mathcal{J}_G(T_G) = \sum_{i=1}^K \mathcal{R}(u, T_G) \log Q(u, i), \quad (11)$$

Algorithm 1 Adversarial Training for PD-GAN

Input: generator \mathcal{G}_θ ; discriminator \mathcal{D}_η ; pre-trained DPP-kernel \mathbf{L} ; training data \mathcal{T}

```

1: Random initialize  $\mathcal{G}_\theta, \mathcal{D}_\eta$ 
2: Pre-train  $\mathcal{G}_\theta, \mathcal{D}_\eta$  using  $\mathcal{T}$ 
3: while PD-GAN not converged do
4:   for g-steps do
5:     for each  $u$  do
6:       Calculate personalized DPP kernel  $\mathbf{L}_u$  by (6)
7:       Sequentially sample top- $K$  items by (8)
8:       Update generator parameters via policy gradient
         by (11)
9:     end for
10:   end for
11:   for d-steps do
12:     for each  $u$  do
13:       Generator sequentially sample top- $K$  items by (8)
14:       Update discriminator parameters via stochastic
         gradient descent by (13)
15:     end for
16:   end for
17: end while

```

where $Q(u, i)$ gives the posterior probability, and the reward function $\mathcal{R}(u, T_G)$ reflects how much the generator can deceive the discriminator, which is defined as follows:

$$\mathcal{R}(u, T_G) = -\log(1 - \mathcal{D}_\eta(u, T_G)). \quad (12)$$

As such, the more the generator can deceive the discriminator, the higher relevant score ($\mathcal{D}_\eta(u, T_G)$) will be given by the discriminator to T_G , and thus higher reward value will be credited to the generator.

The discriminator \mathcal{D} serves as an adversary of \mathcal{G} , aiming to distinguish between the generated set T_G from the ground-truth ones. Hence, given $\mathcal{G}_\theta(u)$, we learn \mathcal{D}_η by maximizing the following objective:

$$\max_{\eta} \mathcal{J}_D = \frac{1}{|\mathcal{T}_u|} \sum_{T \in \mathcal{T}_u} \log \mathcal{D}_\eta(u, T) + \log(1 - \mathcal{D}_\eta(u, T_G)). \quad (13)$$

We can then compute the gradient of (13) with respect to η and use stochastic gradient descent to update η .

Following the general adversarial training procedure, we learn PD-GAN by iterating two steps: (1) fixing the discriminator \mathcal{D}_η , and optimizing the generator \mathcal{G}_θ using (11); (2) fixing \mathcal{G}_θ and optimizing \mathcal{D}_η using (13). The overall training logic is summarized in Algorithm 1. For each training epoch, the generator is updated as follows: firstly, calculate personalized DPP kernel and sample top- K items. As \mathbf{L} can be pre-trained, according to [Chen *et al.*, 2018], the time complexity of this step is $O(K^2N)$. Secondly, update generator parameters via policy gradient, and the time complexity is $O(TKD)$. The discriminator is updated as follows: generator sequentially sample top- K items ($O(K^2N)$) and update discriminator parameter via stochastic gradient descent ($O(TKD)$). Hence, the time complexity of one training epoch is $O(K^2N + TKD)$. Through adversarial learning, both the generator and the discriminator will grow stronger,

and when the model converges, the generator will be able to generate diverse and relevant recommendations resembling the ground-truth so much that the discriminator will not be able to distinguish easily.

4 Experiments

4.1 Experimental Settings

Datasets

We experiment with two public datasets: **Movielens (100k)**¹ and **Anime**². The Movielens dataset consists of 100k ratings (1 to 5) from users to movies. It contains 18 explicit categories and each movie may belong to more than one category. The Anime dataset consists of 1 million ratings (1 to 10) from users to animes. It contains 44 explicit categories and each anime may belong to more than one category. The statistics of the two datasets are summarized in Table 1.

Table 1: Statistics of the Movielens datasets.

Dataset	#users	#items	#interactions	#categories
Movielens	943	1,683	100,000	18
Anime	73,516	12,294	1,048,575	44

Model Training

In this experiment, we focus on implicit feedback. Following the setting of [Liu *et al.*, 2015; Wang *et al.*, 2017; Liu *et al.*, 2018], we treat 5-star ratings of Movielens dataset and 10-star ratings of Anime as positive feedback and all other ratings as unknown feedback. For Anime dataset, we exclude users and items with less than 300 ratings. For training and testing data splitting, we apply a 4:1 random splitting on the two datasets. Then, we randomly sample ground-truth diverse sets from each user’s positive examples in the training data. We do sequential sampling for each diverse set as follows: add a randomly sampled item if it adds new category information to the current set. The number of sampled ground-truth diverse sets are 80,000 for the movielens dataset and 1,392,010 for the anime dataset. To do pairwise training for the L-kernel of DPP, we also randomly sample the same amount of similar sets (consisting of items belonging to the same category from each user’s positive examples) to form training pairs.

Before adversarial training, we first train the L-kernel of DPP with all sampled diverse-similar training pairs. As shown in [Wang *et al.*, 2017], GAN performance can be boosted by initializing generator with a conventional model, and thus we also initialize the parameters of PD-GAN generator with MF-BPR [Rendle *et al.*, 2009] or IRGAN [Wang *et al.*, 2017]. After we initialize the parameters of PD-GAN generator and the DPP kernel for sampling, we do adversarial training by fixing either the generator or the discriminator to train the alternative part. For both datasets, we use embedding size of 30. The learning rate is set to 0.01. The parameter α is set to 0.9.

¹<https://grouplens.org/datasets/movielens/100k/>

²<https://www.kaggle.com/CooperUnion/anime-recommendations-database>

Evaluation Metrics

After adversarial training, a recommendation list is generated for each user by PD-GAN generator with the best performing parameters. To evaluate the relevance of recommended items, we use *Precision* (P) and *Normalized Discounted Cumulative Gain* (NDCG, a position sensitive metric which assigns higher score to hits at higher positions) as used in [Wang *et al.*, 2017]. To evaluate the diversity of recommended items, we use *Category Coverage* (CC), which is calculated by the number of categories covered by top-K items divided by the total number of categories available in the dataset. A higher category coverage means the top-K items are more diverse.

Baselines

We compare PD-GAN with the following baseline methods:

-**MF-BRP** [Rendle *et al.*, 2009]. This method optimizes MF with the BPR objective. It is a highly competitive approach for implicit feedback problems.

-**IRGAN** [Wang *et al.*, 2017]. This method combines generative and discriminative information retrieval via adversarial training. We initialize the parameters of IRGAN generator with MF-BPR.

-**MMR** [Carbonell and Goldstein, 1998]. This method is a canonical baseline for diversified ranking problems.

-**DPP** [Chen *et al.*, 2018]. This is a state-of-the-art method for diversified recommendation. We use one-hot encoding of categories as item features and MF-BRP predicted ratings as item scores.

4.2 Experimental Results

We report the experimental results from the following three aspects: performance comparison between PD-GAN and baselines, impact of α on all diversity-promoting methods, and impact of adversarial training on PD-GAN.

Comparison with Baselines

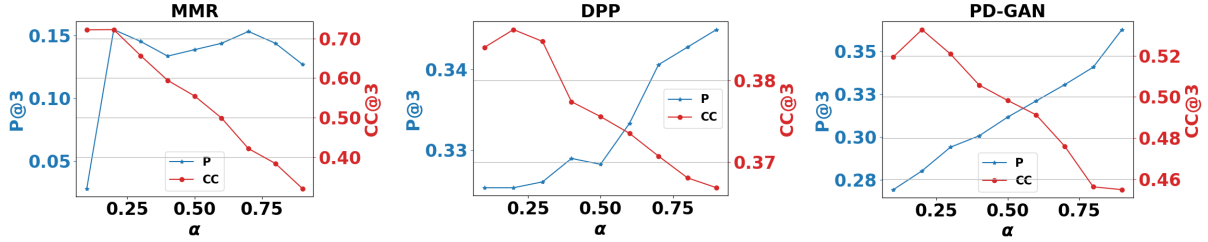
The comparison between PD-GAN and four baselines is summarized in Table 2. The performance is compared in terms of both relevance (P@3, P@5, NDCG@3 and NDCG@5) and diversity (CC@3 and CC@5). The performance improvement of PD-GAN over baselines are highlighted in bold. Overall, these experimental results indicate that PD-GAN can achieve superior performance in terms of balancing accuracy and diversity, comparing with all four baselines.

Comparing with relevance-promoting baselines, i.e., MF-BPR and IRGAN, PD-GAN achieves a significant improvement in terms of diversity on both datasets. For Anime dataset, PD-GAN not only outperform the two relevance-promoting baselines in terms of diversity, but also outperform them in terms of relevance. For Movielens dataset, although IRGAN achieves the best performance of relevance, but with a much inferior performance of diversity comparing to PD-GAN. For top-5 recommendations, with a small loss of 0.27% of NDCG, PD-GAN can achieve a much more significant improvement of 20.49% of CC than IRGAN.

Comparing with diversity-promoting baselines, i.e., MMR and DPP, PD-GAN achieves a significant improvement in terms of both relevance and diversity on Movielens dataset. For Anime dataset, PD-GAN achieved significantly better relevance than MMR and significantly better diversity than DPP.

Table 2: Performance comparison with the state-of-the-art methods.

Movielens												
Methods	P@3		P@5		NDCG@3		NDCG@5		CC@3		CC@5	
MF-BPR	0.3290	+10.15%	0.3111	+5.05%	0.3906	+9.93%	0.4484	+7.40%	0.3664	+24.18%	0.4772	+22.76%
IRGAN	0.3776	-4.03%	0.3434	-4.83%	0.4319	-5.79%	0.4829	-0.27%	0.3601	+26.35%	0.4862	+20.49%
MMR	0.1532	+136.55%	0.1486	+119.91%	0.1898	+126.24%	0.2359	+104.15%	0.4216	+7.92%	0.5068	+11.37%
DPP	0.3450	+5.04%	0.3207	+1.90%	0.4058	+5.82%	0.4631	+3.99%	0.3669	+24.01%	0.5068	+15.59%
PD-GAN	0.3624	-	0.3268	-	0.4294	-	0.4816	-	0.4550	-	0.5858	-
Anime												
Methods	P@3		P@5		NDCG@3		NDCG@5		CC@3		CC@5	
MF-BPR	0.4904	+1.90%	0.4539	+0.55%	0.5273	+0.78%	0.5595	+0.29%	0.2593	+6.17%	0.3499	+5.26%
IRGAN	0.4917	+1.63%	0.4537	+0.60%	0.5279	+0.66%	0.5597	+0.25%	0.2601	+5.84%	0.3499	+5.26%
MMR	0.3252	+53.66%	0.3205	+42.40%	0.3594	+47.86%	0.4271	+31.37%	0.3170	-13.15%	0.3973	-7.30%
DPP	0.4947	+1.01%	0.4591	-0.59%	0.5301	+0.25%	0.5629	-0.32%	0.2646	+4.04%	0.3622	+1.68%
PD-GAN	0.4997	-	0.4564	-	0.5314	-	0.5611	-	0.2753	-	0.3683	-

Figure 2: Impact of parameter α on Movielens dataset.

As MMR maintains an explicit trade-off between relevance and diversity, we can observe that MMR usually achieves good diversity with a huge sacrifice of relevance. Unlike MMR, the L-kernel of DPP captures both relevance (diagonal elements) and diversity (off-diagonal elements). Hence, DPP-based models can improve diversity without a significant sacrifice of relevance.

Impact of α

PD-GAN and the two diversity-promoting baselines, i.e., MMR and DPP, all involve a trade-off parameter α between relevance and diversity. We illustrate the impact of α with the performance of P@3 and CC@3 on Movielens dataset in Figure 2. It can be observed that for all three methods, the general trend of diversity is decreasing when α increases. For DPP and PD-GAN, a reverse trend is observed for relevance. For MMR, relevance fluctuates when α increases and achieves the best performance when $\alpha = 0.7$. Moreover, comparing with MMR, PD-GAN has a much superior performance of relevance across all α values, and comparing with DPP, PD-GAN has a much superior performance of diversity across all α values.

Effect of Adversarial Training

To understand the impact of adversarial training on PD-GAN's performance, we illustrate the performance of relevance and diversity with P@3 and CC@3 over training epochs on Movielens datasets in Figure 3. It can be observed that through adversarial training, diversity increases monotonously as training epochs grow, while relevance first increases and then decreases. This trend shows that with the capability of the generator and the discriminator grow, both relevance and diversity can be improved as the generated sets more resemble the ground-truth sets. Moreover, when relevance reaches its full capacity at around 140 epoches, it starts to degenerate when diversity continues to grow.

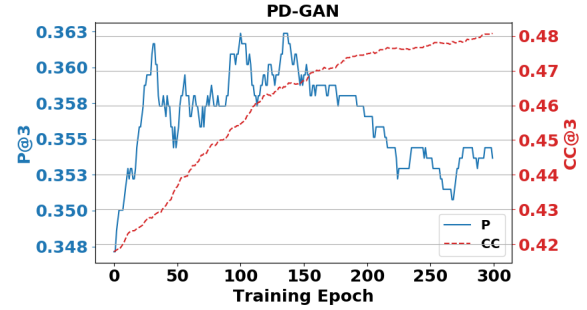


Figure 3: The learning curve of PD-GAN on Movielens datasets.

5 Conclusion

In this work, we propose a novel recommendation model called PD-GAN to learn users' personal preferences to individual items as well as item diversity through an adversarial learning process. Specifically, in the generator of PD-GAN, we adopt a DPP model to learn the general co-occurrence of diverse items and combine it with users' personal preferences to items to generate recommendations that are both diverse and relevant. The discriminator of PD-GAN serves as an adversarial to the generator and strives to distinguish between the recommended items and the ground-truth sampled from the target user's observed data. Through adversarial training, the generator will be able to generate diverse and relevant recommendations so similar to the ground-truth that the discriminator will not be able to distinguish easily. We compare PD-GAN with several competitive state-of-the-art methods on two public datasets. The experimental results show that PD-GAN has superior performance in terms of generating recommendations that are both relevant and diverse.

Acknowledgments

This research is supported by the Alibaba-NTU Singapore Joint Research Institute, Nanyang Technological University, Singapore.

References

- [Adomavicius and Kwon, 2012] Gediminas Adomavicius and YoungOk Kwon. Improving aggregate recommendation diversity using ranking-based techniques. *TKDE*, 24(5):896–911, 2012.
- [Ashkan *et al.*, 2015] Azin Ashkan, Branislav Kveton, Shlomo Berkovsky, and Zheng Wen. Optimal greedy diversity for recommendation. In *IJCAI*, pages 1742–1748, 2015.
- [Carbonell and Goldstein, 1998] Jaime Carbonell and Jade Goldstein. The use of mmr, diversity-based reranking for reordering documents and producing summaries. In *SIGIR*, pages 335–336, 1998.
- [Chen *et al.*, 2018] Laming Chen, Guoxin Zhang, and Eric Zhou. Fast greedy map inference for determinantal point process to improve recommendation diversity. In *NIPS*, pages 5623–5634, 2018.
- [Cheng *et al.*, 2017] Peizhe Cheng, Shuaiqiang Wang, Jun Ma, Jiankai Sun, and Hui Xiong. Learning to recommend accurate and diverse items. In *WWW*, pages 183–192, 2017.
- [Gartrell *et al.*, 2017] Mike Gartrell, Ulrich Paquet, and Noam Koenigstein. Low-rank factorization of determinantal point processes for recommendation. In *AAAI*, pages 1912–1918, 2017.
- [He and Chua, 2017] Xiangnan He and Tat-Seng Chua. Neural factorization machines for sparse predictive analytics. In *SIGIR*, pages 355–364, 2017.
- [Hu *et al.*, 2014] Longke Hu, Aixin Sun, and Yong Liu. Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. In *SIGIR*, pages 345–354. ACM, 2014.
- [Johnson, 2014] Christopher C Johnson. Logistic matrix factorization for implicit feedback data. *NIPS*, 27, 2014.
- [Koren *et al.*, 2009] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, (8):30–37, 2009.
- [Kulesza *et al.*, 2012] Alex Kulesza, Ben Taskar, et al. Determinantal point processes for machine learning. *Foundations and Trends® in Machine Learning*, 5(2–3):123–286, 2012.
- [Kunaver and Požrl, 2017] Matevž Kunaver and Tomaž Požrl. Diversity in recommender systems—a survey. *Knowledge-Based Systems*, 123:154–162, 2017.
- [Lee *et al.*, 2017] Sang-Chul Lee, Sang-Wook Kim, Sunju Park, and Dong-Kyu Chae. A single-step approach to recommendation diversification. In *WWW*, pages 809–810, 2017.
- [Liu *et al.*, 2013] Xin Liu, Yong Liu, Karl Aberer, and Chunyan Miao. Personalized point-of-interest recommendation by mining users’ preference transition. In *CIKM*, pages 733–738. ACM, 2013.
- [Liu *et al.*, 2014] Yong Liu, Wei Wei, Aixin Sun, and Chunyan Miao. Exploiting geographical neighborhood characteristics for location recommendation. In *CIKM*, pages 739–748. ACM, 2014.
- [Liu *et al.*, 2015] Yong Liu, Peilin Zhao, Aixin Sun, and Chunyan Miao. A boosting algorithm for item recommendation with implicit feedback. In *IJCAI*, 2015.
- [Liu *et al.*, 2016] Xin Liu, Yong Liu, and Xiaoli Li. Exploring the context of locations for personalized location recommendations. In *IJCAI*, pages 1188–1194, 2016.
- [Liu *et al.*, 2017] Yong Liu, Peilin Zhao, Xin Liu, Min Wu, and Xiao-Li Li. Learning user dependencies for recommendation. In *IJCAI*, 2017.
- [Liu *et al.*, 2018] Yong Liu, Lifan Zhao, Guimei Liu, Xinyan Lu, Peng Gao, Xiao-Li Li, and Zhihui Jin. Dynamic bayesian logistic matrix factorization for recommendation with implicit feedback. In *IJCAI*, pages 3463–3469, 2018.
- [Qin and Zhu, 2013] Lijing Qin and Xiaoyan Zhu. Promoting diversity in recommendation by entropy regularizer. In *IJCAI*, pages 2698–2704, 2013.
- [Rendle *et al.*, 2009] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *UAI*, pages 452–461, 2009.
- [Rendle, 2010] Steffen Rendle. Factorization machines. In *ICDM*, pages 995–1000, 2010.
- [Santos *et al.*, 2010] Rodrygo LT Santos, Craig Macdonald, and Iadh Ounis. Exploiting query reformulations for web search result diversification. In *WWW*, pages 881–890, 2010.
- [Wang *et al.*, 2017] Jun Wang, Lantao Yu, Weinan Zhang, Yu Gong, Yinghui Xu, Benyou Wang, Peng Zhang, and Dell Zhang. Irgan: A minimax game for unifying generative and discriminative information retrieval models. In *SIGIR*, pages 515–524, 2017.
- [Wu *et al.*, 2016] Qiong Wu, Siyuan Liu, Chunyan Miao, Yuan Liu, and Cyril Leung. A social curiosity inspired recommendation model to improve precision, coverage and diversity. In *WI*, pages 240–247, 2016.
- [Wu *et al.*, 2017] Qiong Wu, Siyuan Liu, and Chunyan Miao. Modeling uncertainty driven curiosity for social recommendation. In *WI*, pages 790–798, 2017.
- [Yu *et al.*, 2009] Cong Yu, Laks Lakshmanan, and Sihem Amer-Yahia. It takes variety to make a world: diversification in recommender systems. In *EDBT*, pages 368–378, 2009.
- [Yu *et al.*, 2014] Jun Yu, Sunil Mohan, Duangmanee Pew Putthividhya, and Weng-Keen Wong. Latent dirichlet allocation based diversified retrieval for e-commerce search. In *WSDM*, pages 463–472, 2014.
- [Zhang and Hurley, 2008] Mi Zhang and Neil Hurley. Avoiding monotony: improving the diversity of recommendation lists. In *RecSys*, pages 123–130, 2008.
- [Zhou *et al.*, 2010] Tao Zhou, Zoltán Kuscik, Jian-Guo Liu, Matúš Medo, Joseph Rushton Wakeling, and Yi-Cheng Zhang. Solving the apparent diversity-accuracy dilemma of recommender systems. *PNAS*, 107(10):4511–4515, 2010.