Applied Science and Biotechnology Journal for Advanced Research e-ISSN: 2583-553X
Peer Reviewed Journal
Volume-3, Issue-4 (July 2024), Page: 1-9 https://doi.org/10.5281/zenodo.12803769

Adversarial for Sequential Recommendation Walking in the Multi-Latent Space

Ziyi Zhu[1], Zeyu Wang[2*], Zhizhong Wu[3] and Yiqian Zhang[4]
[1]New York University, United States
[2]University of California, Los Angeles, United States
[3]University of California, Berkeley, United States
[4]State Key Laboratory of Biotherapy, West China Hospital, China

[*]**Corresponding Author:** Zeyu Wang

Received: 07-06-2024 Revised: 28-06-2024 Accepted: 22-07-2024

ABSTRACT

Recently, sequential recommendation plays a critical role in our daily life, since it serves as personalized information filters to dis- cover popular users' preferred products over time. Due to the success of the adversarial learning, a mass of research efforts start to strengthen sequential recommendation by the adversarial learning, which is able to learn complex underlying data distribution.

However, existing adversarial sequential recommendation methods suffer from mode collapse and unexplained prediction. To boost the diversity, performance, and interpretability of sequential recommendation system, we propose a novel temporal-aware adversarial framework, namely TSRGAN.

In principle, the input of traditional adversarial-based recommendation system is a noise variable sampled from normal distribution. We argue that it is hard to generate an item cover complex users' preferences(e.g. price, brand and item style) using a single latent space. Therefore, our model employs multiple latent space to generate plausible item which matches user' preferences from multiple views(e.g. Movie style, Movie release date).

Besides, previous adversarial-based recommenders focus on generating active item, but they omits that user's favour is not in- variable. With GANs terminology, the recommenders only will be rewarded when seeking the peak mode, but it neglects minor mode, in other word mode collapse. In order to alleviate this issue, we design a novel diversity reward function and diversify regularization to encourage the model exploring minor mode over time and guarantee generating diversity item with reasonable.

Concretely, we propose multiple learnable latent codes to generate item matching user's preferences from different views, then we leverage the diversity reward signal to shape the distribution of multiple latent space over time. It means that the multiple latent space are sampled form different distribution instead of Gaussian distribution. Such a manipulation of the latent space can be treated as walking from plain distribution latent space to diversity distributions latent space. Further, the reward signal is modified over time, therefore, our methods names "Temporal-aware" adversarial framework.

In short, our model has two sequential stages: encode the user' characteristics and historical behaviours under multiple latent space with the Self Attention-based generator(G), and discriminator(D) try to distinguish the generator's output item from the ground ruth. Besides, discriminator attempt to apply reward signal to shape the latent space distribution time by time. Extensive experiments demonstrate remarkable performance with interpretability improvement against the state-of-the-art baselines.

Keywords: sequential recommendation, adversarial learning, interpretability

I. INTRODUCTION

In the big data era, the large-scale information overwhelms personal knowledge gain. [7, 15, 28, 30] Thus, modern recommendation systems are necessities in our daily life to filter out users' focus.

Mostly, users' behaviors and item' attributes update dynamically and evolving over time within daily recommendation scenario. Thus, modeling the dynamics of sequential user behaviors for pro- viding users' preferred information serves as extremely hot research trend.

Recent years, to better capture sequential dependencies of the user-item interactions, there emerges several work in sequential recommendation task[8, 12, 26]. Notably, most methods treat the user-item interactions as a dynamic sequence and take the sequential dependencies into account to capture the current and recent preference of a user by RNNs[2] or Transformers[27].

Furthermore, researchers have incorporated rich contextual in- formation (such as item attributes) to neural sequential recommenders[10, 11, 16, 33], which has been demonstrated that contextual information servers as a key factor to boost recommenders.

Despite the success of prior methods, they are still hard to match the users' preference distribution and item's attributes distribution over time. For the example of movie recommendation, user usually

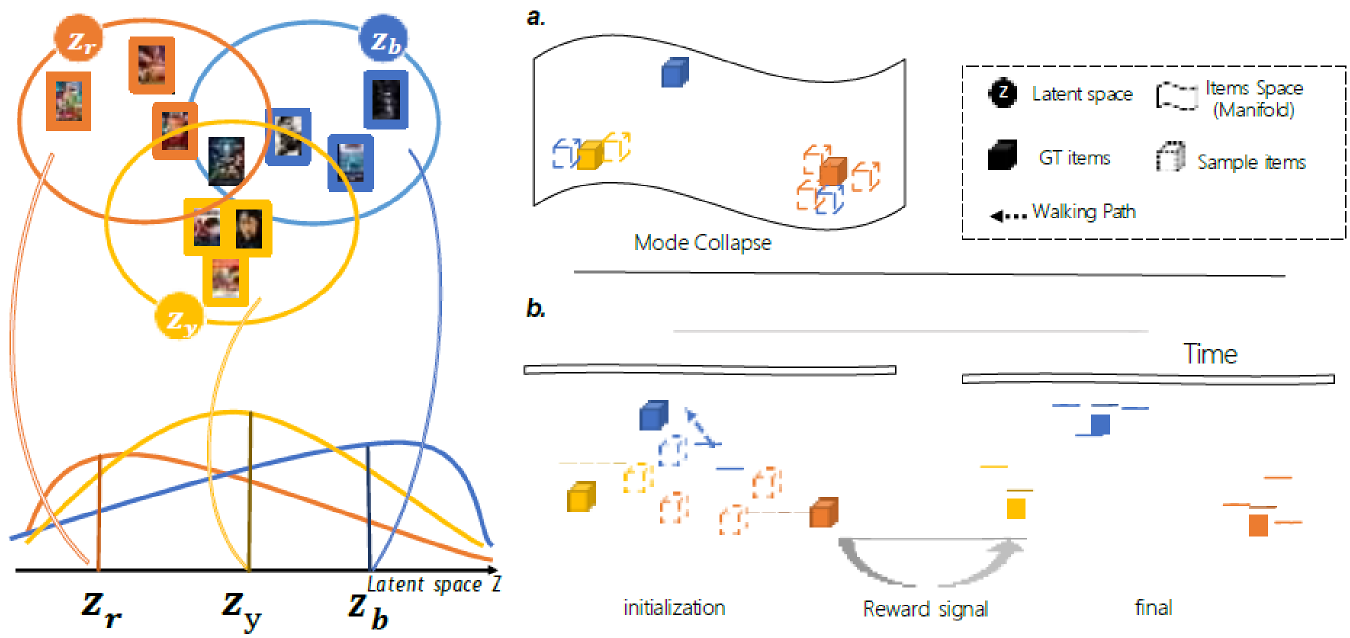

Figure 1: Motivation of our TSRGAN. Given the example of MovieLens-1M dataset, *blue* and *yellow* colors. Additionally, different color frames include movie posters in corresponding genres. Similarly, the different color cubes indicate generated movie in corresponding genres *On the leß part:* we consider that generator attempts to sample different genres items from different distribution latent space. Therefore, we design multiple latent space to model the complex item space *On the right part:* As the (a) illustrated, previous adversarial-based recommender could recommend many item for user but collapses to a few modes (peak modes), which means the recommender is unable to cover completely user's preference. Briefly, those methods are incapable to provide diverse items. Thus, as the (b) illustrated, we propose a reward signal to modify the mean and variance of each latent codes distribution to encourage samples waking to minor modes

interpretability). Basically, within existing adversarial strategy, GANs' incapacity indicates that the vanilla generative process(sampled noise from normal distribution) leads to providing similar item(without diversity). The previous research demonstrates the rationale behind mode collapse: when the distributions of target and generated data is non-overlapping, the loss function convergence to constant, in other words, vanishing gradient. In such a scenario the generator will exhibit poor diversity amongst generated samples, which limits the capability of the GAN. Regarding the without interpretability issue, existing adversarial-based recommendations encode the user-items interactions via black-box DNNs. It means that the black-box recommenders cannot figure out what feature or attribute(e.g. movie style, users' characters) can most match users' preferences . In short, both of the fatal issues hurt recommendation performance disastrously, which they ignore the diversity of recommend item and decisive feature of user-items interactions.

In light of these issues, to deepen the use of adversarial structure for sequential recommendation system, we propose a novel temporal-aware adversarial framework. Notably, we argue that it is hard to generate a item cover every detail of users' preferences(e.g. price, brand and item style) using a single latent space, otherwise, we would have an unbeatable recommendation system. In other words, the expressiveness of the latent space is limited due to its finite dimensionality.

Therefore, in order to recommend a vivid item, we propose to integrate multiple latent space to generate diversity sample, while decouple the latent space to investigate what feature makes effort to recommend the sample with interpretability.

Overall, our framework sets *generator G* and *discriminator D*, following standard GANs' architecture. Same as the most of research on the discrete sequence data, the optimize process of the proposed TSRGAN framework is adopted by policy gradient. The generator utilizes users' characters feature and user-item interactions history to predict the next items for recommendation via transformer component, while the discriminator distinguishes the generated item of from the ground truth users' preference and attempts to guide the latent space of generator via policy gradient reward over time. In detail, we adopt a novel multi-latent space instead of a single normal distribution latent space. The reward signal of discriminator shapes the distribution of the multi-latent space via update the mean and variance. In RL setting, The reward function is through that our framework could draw latent space from a plain sample distribution to diversity sample distributions. To validate the effectiveness of the proposed TSRGAN, we con- duct extensive experiments on two benchmark datasets from different domains. Experimental results show that the proposed TSRGAN is able to achieve better performance compared to several competitive methods. We further show the multi-latent space architecture is indeed useful to stabilize the learning process of adversarial pol- icy. Finally, qualitative analysis demonstrates that our proposed framework can explicitly characterize the effect of various feature distribution over time for sequential recommendation, making the recommendation results highly interpretable. In summary, our contributions are outlined as follows

- We propose a novel sequential adversarial framework, namely TSRGAN, which largely leverages temporal information to generate diversity item and model more intuitively the distribution of multi-latent space to match complex users' preferences(items' attribute) distribution with interpretable
- To the best of our knowledge, we are the first bring latent space manipulation of adversarial framework into recommendation system task. It opens a novel path towards better understanding sequential recommendation and mitigating existing issues
- Extensive experiments conducted on two benchmark datasets demonstrate the benefits of our proposed TSRGAN beating state-of-the-art methods in top-N ranking metric

II. RELATED WORK

The trend of sequential recommendation research has raised in the last few years. We have surveyed this task and categorized it into three branches namely traditional sequence approach, matching representation methods and neural architecture. We have illustrated the relationship among different prior work in Fig. Furthermore, we also review the literature of adversarial training from three related perspectives: stability, diversity and discrete.

2.1 Sequential Recommendation

Traditional Approach. Traditional approaches can be classified into sequential mining and markov-chain models. Both of these approaches have the ability of capturing sequential dependencies among the user-item interactions. In detail, sequential mining [31] proposes to mine frequently active items on user-item sequence. So that it can provide user most popularity item which attempt to match user's preferences. On the other hand,

Markov Chain-based recommendation [3] adopt Markov chain models to model the transitions over user-item interactions in a sequence, for the prediction of the next interaction. According to the specific technique used, Markov chain-based RSs are divided into ba- sic Markov Chain-based approaches and latent Markov embedding- based approaches. The former one directly calculates the transition probability based on the explicit observations, while the latter first embeds the Markov chains into an Euclidean space and then calculates the transition probabilities between interactions based on their Euclidean distance

2.2 Latent Representation Approach

Latent representation models first learn a latent representation of each user or item, and then predict the subsequent user item interactions by utilizing the learned representations. As a result, more implicit and complex dependencies are captured in a latent space, which greatly benefits the recommendations. Next, we introduce two representative models falling into this taxonomy. Factorization machines. Factorization machine-based Sequential Recommendation System usually utilize the matrix factorization or tensor factorization to factorize the observed user-item interactions into latent factors of users and items for recommendations. Different from collaborative filtering (CF), the matrix or tensor to be factorized is composed of interactions rather than the ratings in CF. Such a model is easily affected by the sparsity of the observed data and thus cannot achieve ideal recommendations. Embedding. Embedding-based Sequential Recommendation System learn a latent representations for each user and item for the subsequent recommendations by encoding all the user-item interactions in a sequence into a latent space. Specifically, some works take the learned latent representations as the input of a network to further calculate an interaction score between users and items, or successive users' actions , while other works directly utilize them to calculate metric like the Euclidean distance as the interaction score . This model has shown great potential in recent years due to its simplicity, efficiency and efficacy

2.3 Deep Neural Network

Deep neural networks have natural strength to model and cap- ture the comprehensive relations over different entities (e.g., users, items, interactions) in a sequence, and thus they nearly dominate Sequential Recommendation System in the past few years. The lat- est progress achieved in Sequential Recommendation System also belongs to this taxonomy. Generally, this taxonomy can be divided into two sub classes: Sequential Recommendation System built on basic deep neural networks and Sequential Recommendation Sys- tem built on deep neural networks with some advanced models incorporated. Basic Deep Neural Networks The most commonly used deep neural networks for Sequential Recommendation System are recurrent neural networks (RNN) due to their natural strength in sequence modelling, but they also have defects. Recently, convolutional neural networks (CNN) and graph neural networks (GNN) have also been applied in Sequential Recommendation System to make up the defects of RNN. Next, we introduce the Sequential Recommendation System built on top of these three deep neural networks respectively. RNN-based Sequential Recommendation System. Given a sequence of historical user-item interactions, an RNN-based SRS tries to predict the next possible interaction by modelling the sequential dependencies over the given interactions. Except for the basic RNN, longshort-term-memory (LSTM)- and gated recurrent unit (GRU)-based RNN have also been developed to capture the long-term dependencies in a sequence. Recent years have witnessed the prosperity of RNN-based Sequential Recommendation System and they dominate the research on the deep learning-based Sequential Recommendation System or even the whole Sequential Recommendation System. Besides the basic RNN structure, some variants are proposed to capture more complex dependencies in a sequence, like hierarchical RNN . However, RNN is not flawless for Sequential Recommendation System, with the shortcomings in two aspects: (1) it is easy to generate fake dependencies due to the overly strong assumption that any adjacent interactions in a

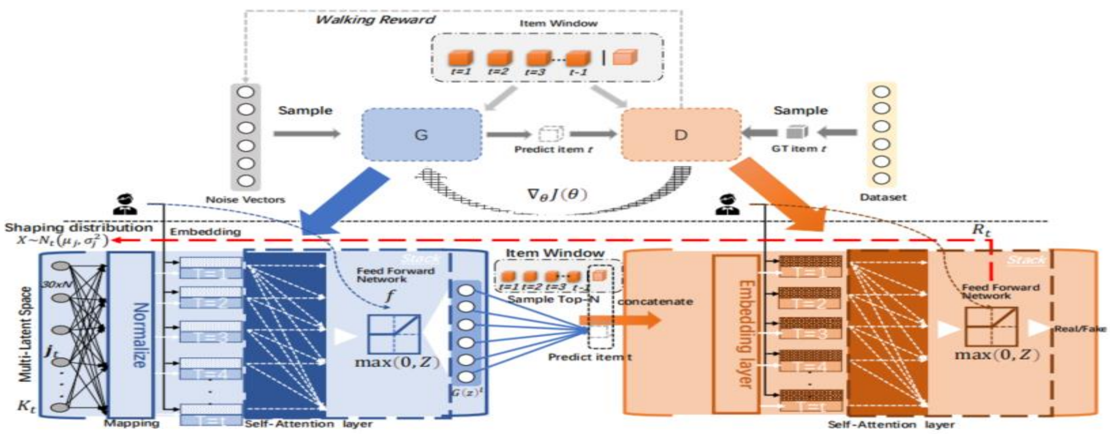

Figure 2: Illustrated the relationship among different previous research in GANs and sequential recommendation

sequence must be dependent, which may not be the cases in the real world because there are usually irrelevant or noisy interactions inside a sequence; and (2) it is likely to capture the point-wise dependencies only while ignoring the collective dependencies (e.g., several interactions collaboratively affect the next one). CNN-based Sequential Recommendation System. Different from RNN, given a sequence of user-item interactions, a CNN first puts all the embedding's of these interactions into a matrix, and then treats such a matrix as an "image" in the time and latent spaces. Finally, a CNN learns sequential patterns as local features of the image using convolutional filters for the subsequent recommendations. Since a CNN does not have strong order assumptions over the interactions in a sequence, and they learn patterns models are in their early stages.

2.4 Generative Adversarial Network

Initial GANs. As previous work demonstrates CNN that have the ability to extract feature from image data. However, for the Generative Adversarial Network, it is extremely tricky to facilities **G** and **D** to capture meaningful feature. Also, In this model, a real image is converted into a multi scale pyramid image, and a convolutional GAN is trained to produce multi- scale and multi-level feature maps where the final feature map can be derived by combining all of them. The Laplacian pyramid is a linear invertible image demonstration containing band-pass images and a low frequency residual

Conditional GANs. In previous GANs, the authors proposed conditional GANs as a solution for image-to-image translation problems. The proposed model not only learns the mapping from input image to output image, but also adopted a loss function to train this map- ping. This approach provides the opportunity to apply the same generic method to the problems that traditionally would need com- plex loss formulations. The architecture is shown in before. As com- pared to the other GAN architectures, the conditional GANs have significant performance on the multi-modal data in comparison with baselines. On the other hand, InfoGAN was another development that uses the mutual information between a small subset of the latent variables to gain semantic information. The architecture is presented in before. Such model can be applied to determining different objects in an unsupervised way and also all the produced samples by InfoGAN are semantically well meaningful.

Zhou et al. introduced a normalization technique with conditional GAN that limits the searching space of the weights in a low-dimensional manifold. In [110, 111], the authors proposed a conditional adversarial network for energy management systems. Their method is demonstrated to converge faster in term of number of epoch, but the authors did not highlight the model complexity. Odena et al. [48] proposed a novel GAN classifier (ACGAN)in which the architecture is similar to Infogan. In this model, the condition variable will not be added to the discriminator, and an external classifier is applied to predicting the probability over the class labels.

The loss function is optimized to improve the class prediction. In [49], the authors proposed a data aug- mentation with balancing GAN (BAGAN) the architecture shows in Figure 3(d). Class conditioning is applied in the hidden space to run the generation procedure towards the objected class. The in the BAGAN is adjusted with the encoder module that enables it to learn in the hidden space. The structure of BAGAN is similar to InfoGAN and ACGAN. However, BAGAN only generate a single output but, InfoGAN and ACGAN have two outputs. In [112], the author presented a deep conditional GAN model that takes its strength from the semantic layout and scene attributes integrated as conditioning variables. This approach able to produces realistic images under different situations, with clear object edges. Figure 5 compares the generated images by InfoGAN and ACGAN on CIFAR-10.

III. EXPERIMENT

In this section, we first introduce some preliminaries of experiments dataset and metrics, and then analyse experiments results with several representative and state-of-the-art baselines. Secondly, we explore the interpretability of multiple latent space. To this end, we show some implementation details and our code of model will be published upon paper accepted.

3.1 Preliminaries

Datasets. In order to measure the performance of our proposed, our model is conducted in two benchmark real-life datasets: MovieLens 1M movie[5], and Yahoo music. For fair comparison, we construct experiment setting similar with the state-of-the-art[19]. Specifically, Yahoo! and MovieLens-1M datasets are too miscellaneous, so we filter items and users according to interaction number which are greater than ten and five, respectively. It means that selecting item with popularity and active users. Note that we divide each user-item interaction sequence into three subset following[12]: the last item in the interaction sequence serves as test set, while penultimate item is treated as validation set, and the rest items are fed into network as training set. The statistics of our two pre-process datasets as list in Tab 1:

Table 1: Summarize of two pre-process datasets

Detest	#Users	#Items	#Interactions
Movielens-1M	6,040	3,416	0.98M
Yahoo! Music[1]	15,280	140,782	3.68M

[1] https://webscope.sandbox.yahoo.com/catalog.php?datatype=r

Evaluate Metrics. We adopt two widespread Top-N metrics to mea- sure our TSRGAN performance, Hit Rate@10 and NDCG@10, which is fair comparison with previous work[12, 19]. Clearly, Hit@10 is a metric of counting how many interactions user involve in a ten-sized list of ranked items, while NDCG@10 is normalized dis- counted cummulative gain which can measure relevant position of item. Additionally, we follow the training tricks[12, 19] which can reduce computation, that randomly sample 100 inactive items as negative sample to initialize D_ϕ . Eventually, we present the average performance over both metrics next section.

3.2 Results and Analysis

Let us show our develop adversarial framework TSRGAN beating against several competitive baselines. To better understand the characterises of our picking baselines, we illustrate it as a table (Tab.2):

Baselines.

- PopRec: PopRec serves as a trivial baseline that recommend items based on their the number of interaction(popularity).
- BPR[21]: Bayesian Personalized Ranking is a superior work that learning personalized rankings from implicit

e-ISSN: 2583-553X
Peer Reviewed Journal
https://doi.org/10.5281/zenodo.12803769

feedback.

- MF[20]: Matrix factorization is a classical methods which utilizes coefficients of different factor in a matrix, and captures the correlation of different features
- IRGAN[29]: IRGAN is the first using GANs framework in information retrieval. This work unifies generative and dis- criminative methods to boost matrix factorization technology and optimize the gradient in discrete data.
- FPMC[22]: It jointly takes advantage of first-order Markov Chain and Matrix Factorization, which extracts both dynamic user-item interactions and static user preference.
- GRU4Rec[9]: GRU4Rec is a founder recommender that open- ing RNNs to simulate user-item interaction in session-based scenario.
- Self-Attentive[12]: It is the first Transformer-based sequential recommendation, which can capture long-term semantics feature and pay attention to relatively few actions.
- PLASTIC[35]: This methods is particularly designed for sequential movie recommendation. PLASTIC considers both user and movie information in long and short-term perspectives to model the movie preferences via adversarial training. Note that PLASTIC is implemented in MovieLens-100K, and we do our best reproduce it in our experiment setting.
- MFGAN[19]: MFGAN is an adversarial framework and it also leverages the Transformer architecture similar to [12]. This framework build more than one discriminators to decouple implied factors in context.

Comparison. The comparisons are illustrated in Table 3. Thus, we can observe that:

1) Regarding the traditional recommendation system baselines(non- sequential), which omits items' time attribute, it can be clearly found that they performs weaker than sequential methods. Notably, BRR[21] has the best result among the traditional methods, because of its strong assumption between two similar items. In contrast, MF[20] cannot get used to matching sequential data, due to regression loss. Though, PopRec serves as a simple baseline just ranking according to item' popularity, it has robust performance in sequential scenario. We suppose that the k-core pre-processing policy affects the data distribution, while inadvertently boost the popularity ranking methods.

2) Regarding Transformer-based sequential recommendation baselines, both Self-Attentive[12] and MFGAN[19] achieve remark- able performance rather than other sequential baselines. From the short review table above, key finding emerges: such an Transformer architecture[27] is inimitably suitable for capturing feature from sequence data, which also make sense to sequential recommendation. On the other hands, non-Transformer architecture methods still performs better than traditional recommenders. Markov Chain wise method[22] could capture the both dynamic user-item relations and static user preference. RNN-based architecture [9, 35] used to be general comparison baselines in sequential data, which it can encoder context information to boost recommendation sys- tem. Further, PLASTIC[35] has better modeling performance than GRU[9] baseline, because of PLASTIC leveraging temporal-aware user' interests.

3) Regarding incorporating user characterises baselines, it can be observed that user characterises generally can enhance the model matching ability. FPMC[22] encoder the user features in a matrix as the input of Markov Chain, which contributes to learning long- term user preference. Additionally, PLASTIC encodes both user and movie feature via RNNs to modelling long- and short-term dynamic feature. However, it is very tricky to incorporate the user feature to the user-items interactions. Thus, we set the user profile as global feature to match the complex item space via a simple attention mechanism. That policy also achieves substantial performance than user feature matrix policy.

4) Regarding adversarial-based recommendation system base- lines, it can be demonstrated that the capacity of negative samples. IRGAN opens the GAN-based recommendation system research, which utilizes the policy gradient to optimize the generator and the negative sample can boost the personalize recommender. Thus, IRGAN achieves the best performance among the non-sequential baselines. Further, PLASTIC considers extra information(e.g. user characterises, movie poster) to improve the judgment of discriminator. Our state-of-the-art baseline MFGAN leverages multiple discriminators to determine which factor servers as the key feature when generating item. The reason why MFGAN has better performance than PLASTIC is credited to encoder difference. Although, both PLASTIC and MFGAN are adversarial framework, MFGAN is based on Transformer architecture rather than RNNs encoder. Besides, overall adversarial-based suggests that negative sample provides strong signal for generating good quality items for users. Notably, the key idea of previous research using adversarial frame- work is feeding extra information to help discriminator judgement ability.

Eventually, our work is demonstrated that achieve remarkable performance rather than other competitive baselines. Clearly, our methods models the interactions taking advantages of above perspectives(e.g. Transformer, user characterise and adversarial frame- work). However, our motivation is orthogonal to previous studies that we focus building comprehensive recommendation system with diversity and interpretability. More specific, we establish multiple latent space that models different views of complex item space over time. Furthermore, we leverage a novel diversity reward to encourage the network exploring the item space, and cover the minor modes so that recommend diversity but reasonable item.

Table 2: Comparisons of Different Methods

Movielens-1M	HR@10	NDCG@10	MRR	Diversity
PopRec	0.4053	0.2211	0.1829	
BPR[21]	0.5381	0.2956	0.2362	
MF[20]	0.3399	0.1867	0.1593	
IRGAN[29]	0.4054	0.2187	0.1794	
FPMC[22]	0.5782	0.3625	0.3052	
GRU4Rec[9	0.6415	0.3875	0.3548	
]	0.7847	0.5644	0.4916	
Self-Attentive[12]	0.7245	0.5489	0.4599	
PLASTIC[35]	0.8026	0.6192	0.5251	
MFGAN[19]				
Our	**0.8244**	**0.6320**	**0.5421**	
w/o reward	0.7920	0.6208	0.5018	
Our w/o latent space	0.7413	0.5684	0.4754	
Yahoo! Music	**HR@10**	**NDCG@10**	**MRR**	**Diversity**
PopRec	0.6132	0.3864	0.3305	
BPR[21]	0.6494	0.4401	0.3872	
MF[20]	0.4040	0.2130	0.1757	
IRGAN[29]	0.4966	0.2771	0.2262	
FPMC[22]	0.5306	0.4148	0.3893	
GRU4Rec[9	0.7155	0.4639	0.3923	
]	0.8549	0.8005	0.7815	
Self-Attentive[12]	0.8038	0.7489	0.4789	
PLASTIC[35]	0.8663	0.8196	0.8020	
MFGAN[19]				
Our	**0.8712**	**0.8321**	**0.8392**	
w/o reward	0.8537	0.7712	0.7921	
w/o latent space	0.7329	0.7343	0.7239	

IV. CONCLUSION

In this paper, we have proposed a Temporal-aware Generative Adversarial Network (TSRGAN) for sequential recommendation. In our framework, the generator taking user behavior sequences and user' characterises as input is used to generate possible next items, and multiple latent space are used to cover users' favourite distribution. We have constructed extensive experiments on two real-world datasets. Experimental results have shown that our approach outperforms several competitive baselines. Especially, we have found that using multiple latent space is useful to enhance the interpretability of recommendation algorithms. Currently, we consider a simple setting where multiple latent codes are separately designed. As future work, we will investigate how to design a more principled way to share relational feature across different layers. We will also consider incorporating explicit to control the recommend types and items.

REFERENCES

1. Wang, Z., Yan, H., Wang, Y., Xu, Z., Wang, Z., & Wu, Z. (2024). *Research on autonomous robots navigation based on reinforcement learning*. arXiv preprint arXiv:2407.02539.
2. Wantlin, K., Wu, C., Huang, S. C., Banerjee, O., Dadabhoy, F., Mehta, V. V., ... & Rajpurkar, P. (2023). *Benchmd: A benchmark for modality-agnostic learning on medical images and sensors*. arXiv preprint arXiv:2304.08486.
3. Li, Zhenglin, et al. (2023). Stock market analysis and prediction using LSTM: A case study on technology stocks. *Innovations in Applied Engineering and Technology*, 1-6.
4. Liu, Y., Yang, H., & Wu, C. (2023). Unveiling patterns: A study on semi-supervised classification of strip surface defects. *IEEE Access, 11*, 119933-119946.
5. Restrepo, D., Wu, C., Vásquez-Venegas, C., Nakayama, L. F., Celi, L. A., & López, D. M. (2024). DF-DM: A foundational process model for multimodal data fusion in the artificial intelligence era. *Research Square*.

6. Li, Shaojie, Yuhong Mo, & Zhenglin Li. (2022). Automated pneumonia detection in chest x-ray images using deep learning model. *Innovations in Applied Engineering and Technology*, 1-6.
7. Shovestul, B., Saxena, A., Reda, S., Dudek, E., Wu, C., Lamberti, J. S., & Dodell-Feder, D. (2022). Social affective forecasting and social anhedonia in schizophrenia-spectrum disorders: a daily diary study. *Schizophrenia, 8*(1), 97.
8. Restrepo, D., Wu, C., Cajas, S. A., Nakayama, L. F., Celi, L. A. G., & Lopez, D. M. (2024). *Multimodal deep learning for low-resource settings: A vector embedding alignment approach for healthcare applications*. medRxiv.
9. Wu, C., Yang, X., Gilkes, E. G., Cui, H., Choi, J., Sun, N., ... & Nakayama, L. (2023, October). De-identification and obfuscation of gender attributes from retinal scans. in *Workshop on Clinical Image-Based Procedures*, pp. 91-101. Cham: Springer Nature, Switzerland.
10. Mo, Y., Tan, C., Wang, C., Qin, H., & Dong, Y. (2024). Make scale invariant feature transform "Fly" with CUDA. *International Journal of Engineering and Management Research, 14*(3), 38-45.
11. Nakayama, L. F., Choi, J., Cui, H., Gilkes, E. G., Wu, C., Yang, X., ... & Celi, L. A. (2023). Pixel snow and differential privacy in retinal fundus photos de-identification. *Investigative Ophthalmology & Visual Science, 64*(8), 2399-2399.
12. Cajas, S. A., Restrepo, D., Moukheiber, D., Kuo, K. T., Wu, C., Chicangana, D. S. G., ... & Celi, L. A. *A multi-modal satellite imagery dataset for public health analysis in Colombia*.
13. Moukheiber, D., Restrepo, D., Cajas, S. A., Montoya, M. P. A., Celi, L. A., Kuo, K. T., ... & Kuo, P. C. (2024). A multimodal framework for extraction and fusion of satellite images and public health data. *Scientific Data, 11*(1), 634.
14. Restrepo, D., Wu, C., Vásquez-Venegas, C., Matos, J., Gallifant, J., & Nakayama, L. F. (2024). *Analyzing diversity in healthcare LLM research: A scientometric perspective*. medRxiv.
15. He, Shuyao, et al. (2024). Lidar and monocular sensor fusion depth estimation. *Applied Science and Engineering Journal for Advanced Research, 3*(3), 20-26.
16. Lin, Z., & Xu, F. (2023, July). Simulation of robot automatic control model based on artificial intelligence algorithm. in *2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS)*, pp. 535-539. IEEE.
17. Lin, Z., Wang, Z., Zhu, Y., Li, Z., & Qin, H. (2024). Text sentiment detection and classification based on integrated learning algorithm. *Applied Science and Engineering Journal for Advanced Research, 3*(3), 27-33.
18. Mo, Yuhong, et al. (2024). Large Language Model (LLM) AI text generation detection based on transformer deep learning algorithm. *International Journal of Engineering and Management Research, 14*(2), 154-159.
19. Lin, Z., Wang, Z., Zhu, Y., Li, Z., & Qin, H. (2024). Text sentiment detection and classification based on integrated learning algorithm. *Applied Science and Engineering Journal for Advanced Research, 3*(3), 27-33.
20. Dai, S., Dai, J., Zhong, Y., Zuo, T., & Mo, Y. (2024). The cloud-based design of unmanned constant temperature food delivery trolley in the context of artificial intelligence. *Journal of Computer Technology and Applied Mathematics, 1*(1), 6-12.
21. Lin, Z., Wang, C., Li, Z., Wang, Z., Liu, X., & Zhu, Y. (2024). Neural radiance fields convert 2d to 3d texture. *Applied Science and Biotechnology Journal for Advanced Research, 3*(3), 40-44.
22. Zhan, D., Yi, S., Xu, D., Yu, X., Jiang, D., Yu, S., ... & Zhang, W. (2019). *Adaptive transfer learning of multi-view time series classification*. arXiv preprint arXiv:1910.07632.
23. Zhan, D., Yi, S., & Jiang, D. (2018, November). Small-scale demographic sequences projection based on time series clustering and lstm. in *IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 803-809. IEEE
24. Yi, S., Zhan, D., Zhang, W., Jiang, D., An, K., & Wang, H. (2019). *Fisgan: Gan with flow-based importance sampling*. arXiv preprint arXiv:1910.02519.
25. Mo, Yuhong, et al. (2024). Password complexity prediction based on roberta algorithm. *Applied Science and Engineering Journal for Advanced Research 3*(3), 1-5.
26. Gao, H., Li, Y., Long, K., Yang, M., & Shen, Y. (2024). *A survey for foundation models in autonomous driving*. arXiv preprint arXiv:2402.01105.
27. Zhao, Y., & Gao, H. (2024). *Utilizing large language models for information extraction from real estate transactions*. arXiv preprint arXiv:2404.18043.
28. Liu, Tianrui, et al. (2024). Spam detection and classification based on distilbert deep learning algorithm. *Applied Science and Engineering Journal for Advanced Research, 3*(3), 6-10.
29. Zhan, D., Yi, S., Xu, D., Yu, X., Jiang, D., Yu, S., ... & Zhang, W. (2019). *Adaptive transfer learning of multi-view time series classification*. arXiv preprint arXiv:1910.07632.
30. Zhan, D., Yi, S., & Jiang, D. (2018, November). Small-scale demographic sequences projection based on time series clustering and lstm-rnn. in *IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 803-809.

31. Yi, S., Zhan, D., Zhang, W., Jiang, D., An, K., & Wang, H. (2019). *Fisgan: Gan with flow-based importance sampling*. arXiv preprint arXiv:1910.02519.
32. Hu, X., Sun, Z., Nian, Y., Wang, Y., Dang, Y., Li, F., ... & Tao, C. (2024). Self-explainable graph neural network for alzheimer disease and related dementias risk prediction: Algorithm development and validation study. *JMIR Aging, 7*(1), e54748.
33. Song, Jintong, et al. (2024). A comprehensive evaluation and comparison of enhanced learning methods. *Academic Journal of Science and Technology, 10*(3), 167-171.
34. Li, F., Rasmy, L., Xiang, Y., Feng, J., Abdelhameed, A., Hu, X., ... & Tao, C. (2024). Dynamic prognosis prediction for patients on DAPT after drug‐eluting stent implantation: Model development and validation. *Journal of the American Heart Association*, *13*(3), e029900.
35. Liu, Tianrui, et al. (2024). Spam detection and classification based on distilbert deep learning algorithm. *Applied Science and Engineering Journal for Advanced Research*, *3*(3), 6-10.
36. He, J., Li, F., Hu, X., Li, J., Nian, Y., Wang, J., ... & Tao, C. (2022, June). Chemical-protein relation extraction with pre-trained prompt tuning. in *IEEE 10th International Conference on Healthcare Informatics (ICHI)*, pp. 608-609.
37. He, J., Li, F., Li, J., Hu, X., Nian, Y., Xiang, Y., ... & Tao, C. (2024). Prompt tuning in biomedical relation extraction. *Journal of Healthcare Informatics Research*, *8*(2), 206-224.
38. Mo, Yuhong, et al. (2024). Prediction of heart failure patients based on multiple machine learning algorithms. *Applied and Computational Engineering, 75*, 1-7. doi:10.54254/2755-2721/75/20240498.
39. Zhang, Jingyu, et al. (2024). Research on detection of floating objects in river and lake based on AI image recognition. *Journal of Artificial Intelligence Practice, 7*(2), 97-106.
40. Xiang, Ao, et al. (2024). Research on splicing image detection algorithms based on natural image statistical characteristics. *Journal of Image Processing Theory and Applications*, *7*(1), 43-52.
41. Yu Cheng, Qin Yang, Liyang Wang, Ao Xiang, & Jingyu Zhang. (2024). Research on credit risk early warning model of commercial banks based on neural network algorithm. *Financial Engineering and Risk Management, 7,* 11-19. doi:10.23977/ferm.2024.070402.
42. Xiang, Ao, et al. (2024). *A neural matrix decomposition recommender system model based on the multimodal large language model*. arXiv preprint arXiv:2407.08942.
43. Zhu, Armando, et al. (2024). Cross-task multi-branch vision transformer for facial expression and mask wearing classification. *Journal of Computer Technology and Applied Mathematics*, *1*(1), 46-53.
44. Li, Keqin, et al. (2024). Utilizing deep learning to optimize software development processes. *Journal of Computer Technology and Applied Mathematics*, *1*(1), 70-76.
45. Li, Keqin, et al. (2024). The application of augmented reality (AR) in remote work and education. *Journal of Computer Technology and Applied Mathematics*, *1*(1), 33-39.
46. Hong, Bo, et al. (2024). The application of artificial intelligence technology in assembly techniques within the industrial sector. *Journal of Artificial Intelligence General Science (JAIGS), 5*(1), 1-12.
47. Dai, Shuying, et al. (2024). AI-based NLP section discusses the application and effect of bag-of-words models and TF-IDF in NLP tasks. *Journal of Artificial Intelligence General Science (JAIGS), 5*(1), 13-21.
48. Zhao, Peng, et al. (2024). Task allocation planning based on hierarchical task network for national economic mobilization. *Journal of Artificial Intelligence General Science (JAIGS), 5*(1), 22-31.