# Contrastive Learning with Frequency-Domain Interest Trends for Sequential Recommendation


Harsh Diwaker
*210108015*
*h.diwaker@iitg.ac.in*

Maharsh Raval
*210108026*
*r.maharsh@iitg.ac.in*

Priyanshu Srivastava
*210108035*
*priyanshu.eee21@iitg.ac.in*



***Abstract*—Sequential recommendation systems face notable challenges in adapting to users' evolving interests and capturing high-frequency interactions that shift over time. Traditional time-domain contrastive learning approaches, while effective to an extent, often struggle with these dynamic changes and are limited by deep learning's inherent low-frequency bias, known as the F-principle. This bias can restrict the model's ability to effectively capture high-frequency patterns crucial for identifying rapid shifts in user preferences.**

**To address these issues, Zhang, Yin, and Dong propose CFIT4SRec [1], a frequency-domain contrastive learning model that leverages frequency-domain transformations for sequential recommendation. By representing historical interaction embeddings as images and applying second-order Fourier transforms, CFIT4SRec performs three targeted frequency augmentations—low-pass, high-pass, and band-stop filtering. These augmentations allow the model to capture a wide range of interest dynamics, accounting for both stable and rapidly changing preferences in user behavior.**

**In this project, we evaluate and simulate CFIT4SRec on four diverse datasets—ML-1M, MIND, Amazon Beauty, and Amazon Baby—to demonstrate its effectiveness in comparison to established models such as DuoRec and CL4SRec. Our analysis highlights CFIT4SRec's superior capability in capturing user preferences across frequency domains, providing a more adaptable and comprehensive model for sequential recommendation tasks. The full implementation is available at *this link*.**


## I. Introduction

Recommendation systems play a crucial role in personalizing content and products and enhancing user experience by predicting and suggesting items aligned with user interests. In recent years, deep learning architectures have significantly advanced the capabilities of recommendation systems, particularly in sequential recommendation, where user interactions over time inform future preferences. Despite these advances, current deep learning-based sequential recommendation models encounter several challenges, particularly in adapting to evolving user interests and capturing high-frequency interest shifts. Traditional models typically operate in the time domain, where they often fall short in effectively identifying rapid changes in user behavior, a limitation exacerbated by deep learning's natural inclination towards low-frequency signals, commonly known as the F-principle. This low-frequency bias restricts models from capturing the high-frequency components that are essential to accurately track short-term shifts in user preferences.

CFIT4SRec, proposed by Zhang, Yin, and Dong [1], introduces a novel approach to address these limitations by employing contrastive learning within the frequency domain for sequential recommendation. Unlike conventional models, CFIT4SRec treats user interaction sequences as images and applies second-order Fourier transforms to derive frequency representations of user preferences. This approach allows the model to perform frequency-based augmentations through three types of filtering—low-pass, high-pass, and band-stop. These filters enable CFIT4SRec to focus on specific frequency components, capturing both stable, long-term interests and transient, short-term shifts in user behavior, thus adapting more flexibly to the dynamic nature of user interests.

## II. Experimental Setup

### A. Datasets

We evaluate the performance of CFIT4SRec on four publicly available datasets from real-world platforms. The selected datasets are Amazon (Beauty and Baby subcategories), MIND (Microsoft News Dataset), and MovieLens-1M (ML1M). These datasets are widely used to evaluate sequential recommendation algorithms and provide diverse interaction records across different domains.

**Amazon Dataset:** The Amazon dataset consists of user interaction records from the Amazon e-commerce platform. It is extensively used in recommendation system research. We utilize two of its sub-datasets:

- **Beauty:** This subset contains user reviews and interactions with beauty products.
- **Baby:** This subset contains user interactions with baby-related products.

**MIND Dataset:** The Microsoft News Dataset (MIND) is a large-scale dataset for news recommendation. It includes anonymized user behavior logs, click history, and metadata of news articles, making it a benchmark for personalized news recommendation systems.

**MovieLens-1M (ML-1M) Dataset:** The ML1M dataset is a widely used benchmark for recommendation systems, consisting of 1 million user ratings for movies. It captures real user interaction information and supports various recommendation tasks.

To ensure data quality and relevance, we ignore duplicate interactions and filter out users and items with fewer than five recorded interactions. The historical interactions are then sorted chronologically. We employ a leave-one-out strategy to partition the datasets, where the last item in the sequence serves as the test set, the second last item as the validation set, and the rest as the training set. The statistics of the processed datasets are summarized in Table 1.

TABLE I
STATISTICS OF THE DATASETS.

| Dataset | #users | #items | #actions | avg. act./user | avg. act./item | Sparsity (%) |
|---|---|---|---|---|---|---|
| Beauty | 22364 | 12102 | 198502 | 8.9 | 16.4 | 99.93 |
| Baby | 19446 | 7051 | 160792 | 8.3 | 22.8 | 99.88 |
| MIND | 46919 | 12260 | 5817882 | 124.0 | 474.6 | 98.99 |
| ML-1M | 6041 | 3417 | 999611 | 165.5 | 292.6 | 95.16 |

### B. Evaluation Metrics

To evaluate the performance of all methods, we adopt the popular ranking metrics, including hit rate (HR) and normalized discounted cumulative gain (NDCG). HR@K focuses on the presence of positive samples, while NDCG@K further considers position ranking information. We evaluate the ranking results over the whole item set for a fair comparison. In this work, we report HR and NDCG with K = {5,10}.

### C. Competing Models

To demonstrate the effective performance of the CFIT4SRec model, we have evaluated it against several state-of-the-art competing models, including CL4SRec and DuoRec. These models are widely used in the domain of sequential recommendation systems and serve as benchmarks for assessing the efficacy of novel approaches.

- **CL4SRec** [2] combines contrastive learning and sequential recommendation models to learn high-quality sequence encoder.
- **DuoRec** [3] proposes a supervised data augmentation strategy, where sequences with the same target items have similar semantics to alleviate representation degradation problems.

### D. Implementation Details

For each baseline model, we adopt the official implementations provided by the respective authors to ensure a fair comparison. We set the embedding size to 64 and the batch size to 256 for training data and 512 for validation data. Following the instructions from the original papers, we configure other hyperparameters accordingly.

Our model is implemented using the popular **RecBole** recommendation system framework, PyTorch 1.8.2. We utilize the Adam optimizer with a learning rate of 0.001. An early stopping strategy is applied, where training stops if HR@10 does not improve within 10 training epochs. The parameter $\gamma$, controlling $L_2$ regularization, is searched within the range $\{0, 0.001, 0.0001\}$. The number of layers and heads in the Transformer is set to 2. The Dropout parameter is tuned is set to 0.3, and the parameter $\lambda$ in Eq. (25) is fixed at 0.1. Similarly, the parameter $\tau$ is set to 1. The thresholds for LPA and HPA are selected from $\{6, 12, 18, 24\}$.

Our model is trained on a Colab environment using a G4 TPU. This ensures high computational efficiency and compatibility with the training pipeline.

## III. Analysis of CFIT4SRec Across Models and Hyperparameters

### A. Overall Performance Comparison

Table 2 presents the overall performance of all methods on four public datasets, where the best performance is shown in bold. Compared with conventional methods, sequential recommendations with contrastive learning improve the performance by a significant margin in most cases on all datasets. This demonstrates that contrastive learning can effectively learn a high-quality sequence encoder to provide more accurate recommendations.

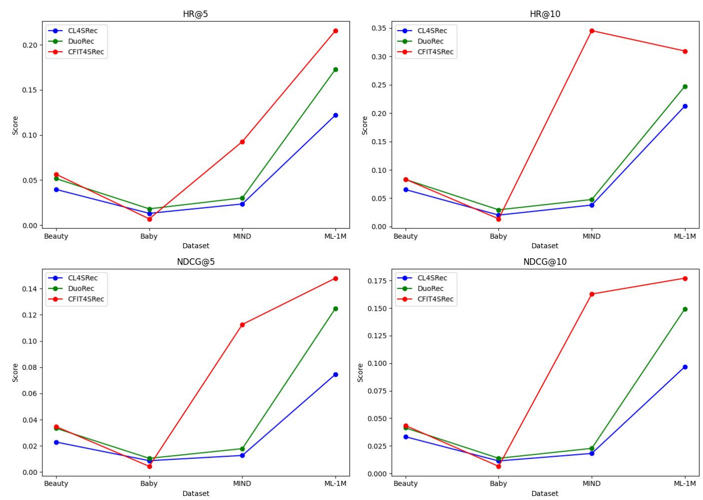


Fig. 1. Overall Performance Comparison

As shown in Figure 1, CFIT4SRec shows significant improvements over the state-of-the-art models on all datasets. Specifically, our model achieves a 12.75% to 24.75% improvement in HR and NDCG compared to the strongest baseline on ML-1M. These results clearly demonstrate the superiority of our solution. We attribute such improvement to the proposed data augmentation operations—low-pass augmentation, high-pass augmentation, and band-stop augmentation—which enable the user representation encoder to accommodate different frequency components (i.e., varying degrees of interest trends). This enhances the model's inference power and overall performance.

### B. Effect of Data Augmentation Techniques on Model Performance

In our project, we enhanced the model by applying three frequency-based augmentation techniques—low-pass, high-pass, and band-stop augmentation—directly to the input data. Below is an overview of how each technique was implemented and how it improved our model:

1) **Low-pass Augmentation**: We applied low-pass filtering to the input sequences to remove high-frequency noise and focus on smooth, long-term trends. This enabled the model to learn stable, enduring patterns in user behavior, ignoring short-term fluctuations that could distract from the user's core preferences.

TABLE II
OVERALL PERFORMANCE OF DIFFERENT METHODS. ALL IMPROVEMENTS ARE STATISTICALLY SIGNIFICANT.

| Dataset | Metric | CL4SRec | DuoRec | CFIT4SRec | Improv. (%) |
|---|---|---|---|---|---|
| Beauty | HR@5 | 0.0396 | 0.0516 | **0.0564** | 9.30 |
| | HR@10 | 0.0652 | 0.0831 | **0.0834** | 5.06 |
| | NDCG@5 | 0.0229 | 0.0334 | **0.0347** | 3.58 |
| | NDCG@10 | 0.0332 | 0.0413 | **0.0434** | 5.08 |
| Baby | HR@5 | 0.0059 | 0.0068 | **0.0070** | 2.94 |
| | HR@10 | 0.0125 | 0.0130 | **0.0136** | 4.62 |
| | NDCG@5 | 0.0038 | 0.0043 | **0.0044** | 2.33 |
| | NDCG@10 | 0.0058 | 0.0062 | **0.0064** | 3.23 |
| MIND | HR@5 | 0.1738 | 0.1740 | **0.1902** | 9.27 |
| | HR@10 | 0.3158 | 0.3163 | **0.3455** | 9.23 |
| | NDCG@5 | 0.1070 | 0.1089 | **0.1126** | 3.35 |
| | NDCG@10 | 0.1538 | 0.1545 | **0.1627** | 5.29 |
| ML-1M | HR@5 | 0.1221 | 0.1729 | **0.2157** | 24.75 |
| | HR@10 | 0.2127 | 0.2476 | **0.3096** | 12.75 |
| | NDCG@5 | 0.0746 | 0.1249 | **0.1479** | 18.41 |
| | NDCG@10 | 0.0968 | 0.1493 | **0.1771** | 18.62 |

2) **High-pass Augmentation**: High-pass filtering highlighted sudden changes in user interests by emphasizing high-frequency components. This allowed the model to become more responsive to recent shifts in user behavior, enabling it to adapt quickly when users exhibited new, emerging interests.
3) **Band-stop Augmentation**: Band-stop filtering was introduced to randomly remove specific frequency components randomly, preventing the model from overfitting to certain patterns. This forced the model to generalize better, making it more robust and adaptable to real-world data variability.

**Multi-task Learning Framework:** We incorporated these augmented datasets within a multi-task learning framework, where the model was trained simultaneously on both the original and frequency-augmented data. For each user sequence, we created low-pass, high-pass, and band-stop versions, using each as an additional "task" for the model. This approach enabled the model to learn generalizable representations across different frequency variations.

Through this multi-task setup, our model became more versatile, improving its ability to generalize and perform well on traditional recommendation tasks. The result was a recommendation system that not only understood stable, long-term preferences but could also adapt to sudden changes in user interests, achieving greater overall accuracy and robustness.

### *C. Robustness Analysis*

We conducted extensive experiments to evaluate the robustness of the proposed model, CFIT4SRec, against data sparsity and noisy interactions.

**Data Sparsity Analysis:** To simulate data sparsity scenarios, we kept the test dataset unchanged and varied the percentage of the training dataset (25%, 50%, 75%, and 100%). The performance of the proposed model was compared with the strongest baseline, DuoRec. Figure 2 & Figure 3 illustrate the NDCG@10 scores on the ML-1M and Beauty datasets.

The results show that CFIT4SRec consistently outperforms DuoRec across all datasets and all levels of data sparsity. Furthermore, the performance degradation of CFIT4SRec is significantly slower than that of DuoRec under sparse training conditions. For instance, with only 50% of the training data on ML-1M, DuoRec experiences a 25.59% drop in performance, while CFIT4SRec shows a much smaller degradation of 13.33%.

Similarly, on the Beauty dataset with 50% of the training data, DuoRec suffers a 48.91% performance drop, whereas CFIT4SRec drops by only 44.70%. These results highlight the robustness of CFIT4SRec and its ability to handle data sparsity effectively. We attribute this improvement to the proposed model's design, which mitigates the effects of data sparsity by leveraging its advanced data augmentation techniques and frequency-aware encoder.

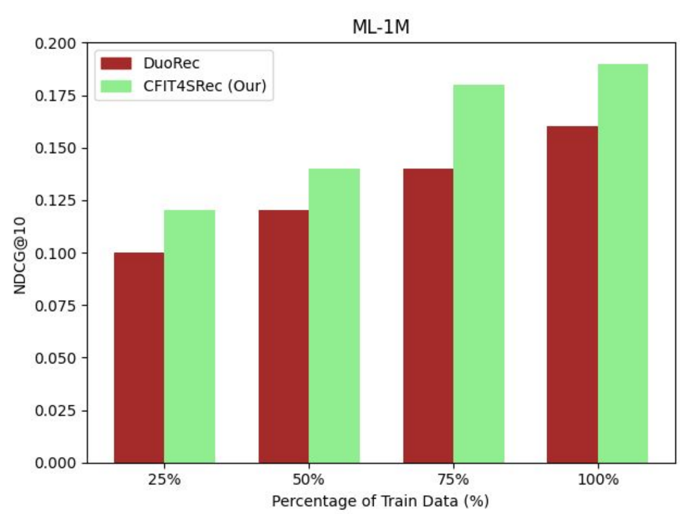


Fig. 2. Performance comparison w.r.t. train ratio in HDCG@10 for ML-1M.

**Noise Interactions Analysis**

To evaluate the effect of noisy interactions on the model's

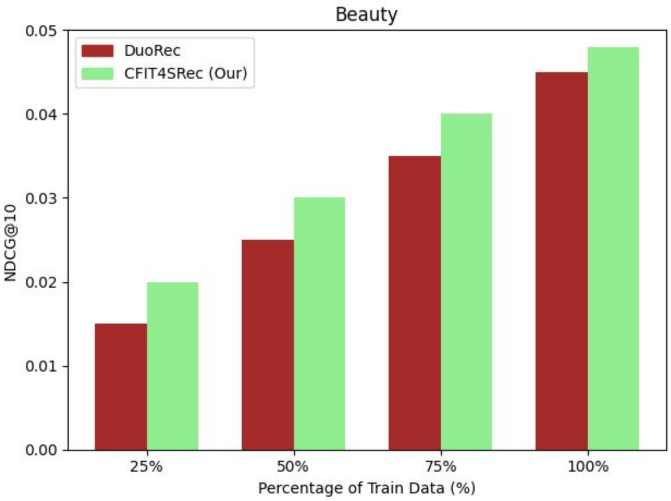


Fig. 3. Performance comparison w.r.t. train ratio in HDCG@10 for Beauty.

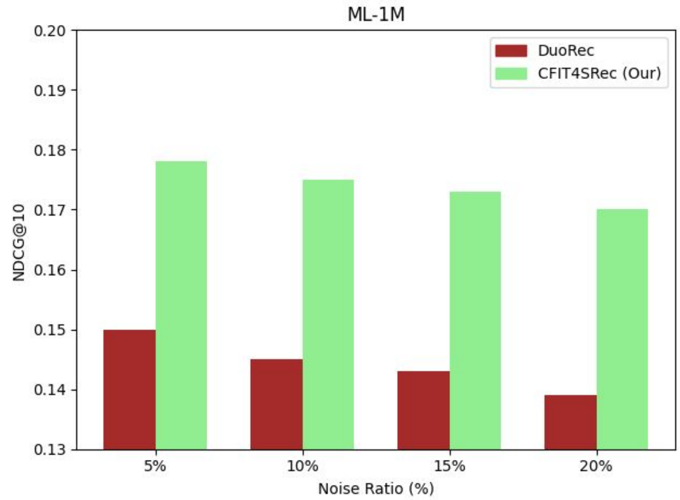


Fig. 4. Performance comparison w.r.t. noise ratio in HDCG@10 for ML-1M.

performance, we conducted an experiment where a certain percentage of items (5%, 10%, 15%, and 20%) in each test sequence were randomly replaced with negative user-item interactions. The training dataset remained unchanged, and the maximum sequence length was preserved to maintain consistency in model input.

Figure 4 & Figure 5 presents the results of this experiment, comparing the performance of the proposed CFIT4SRec model with DuoRec on the ML-1M and Beauty datasets. The results demonstrate that CFIT4SRec consistently outperforms DuoRec across all datasets, reaffirming the robustness and superiority of the proposed solution.

As the percentage of noise interactions increased, CFIT4SRec showed average performance degradation rates of 10.27% on ML-1M and 16.86% on Beauty. In contrast, DuoRec exhibited higher degradation rates of 16.32% and 26.54% on the respective datasets. These results highlight the ability of CFIT4SRec to handle noisy data effectively, maintaining better performance even under challenging conditions.

We attribute this improvement to the proposed frequency-based augmentation techniques, which generate more confident positive samples for contrastive learning. This enhances the model's inference ability and enables it to generalize better in the presence of noise.

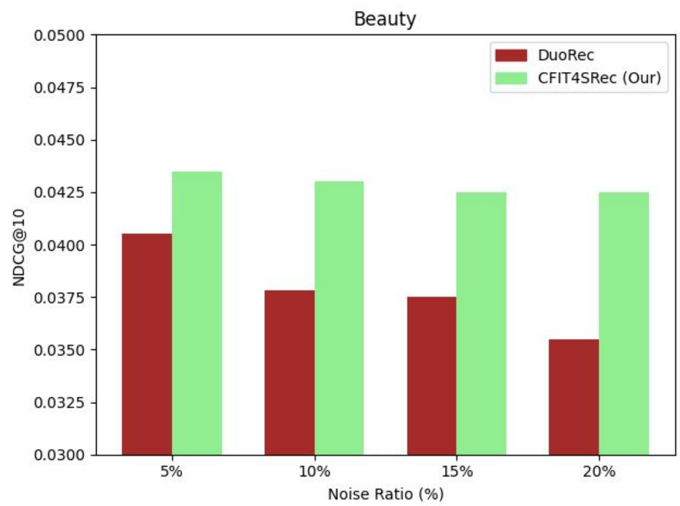


Fig. 5. Performance comparison w.r.t. noise ratio in HDCG@10 for Beauty.

## IV. Conclusion

This work introduces Contrastive Learning with Frequency-Domain Interest Trends for Sequential Recommendation (CFIT4SRec), a novel approach to address the challenges posed by time-series complexity and low-frequency preferences. CFIT4SRec leverages second-order frequency domain representations to construct positive samples for contrastive learning. This enables the sequence encoder to effectively handle varying frequency components, enhancing its inference ability.

The frequency components capture interest trends by reflecting relationships between attributes and their surroundings in the hidden space. This improves the model's capability to distinguish features with varying degrees of interest trends. Extensive experiments on four public benchmark datasets demonstrate that CFIT4SRec achieves substantial performance improvements, consistently outperforming state-of-the-art baselines.

## References


[1] Yichi Zhang, Guisheng Yin, and Yuxin Dong. Contrastive Learning with Frequency-Domain Interest Trends for Sequential Recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems*, pages 141–150, 2023. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3604915.3608790.

[2] Ruihong Qiu, Zi Huang, Hongzhi Yin, and Zijian Wang, "Contrastive Learning for Representation Degeneration Problem in Sequential Recommendation," in *WSDM '22: The Fifteenth ACM International Conference on Web Search and Data Mining*, Virtual Event / Tempe, AZ, USA, February 21-25, 2022, pp. 813–823. doi: https://doi.org/10.1145/3488560.3498433.

[3] Xu Xie, Fei Sun, Zhaoyang Liu, Shiwen Wu, Jinyang Gao, Bolin Ding, and Bin Cui, "Contrastive Learning for Sequential Recommendation," 2020. arXiv: https://doi.org/10.48550/arXiv.2010.14395.