Beyond-Accuracy (Sparsed-) coVariance Neural Network Recommender Systems

Type D: Bachelor's Thesis Abstract

Author: Ivan Bozhanin
Supervisors: Andrea Cavallo, Chengen Liu, and Elvin Isufi

Delft University of Technology, Delft, The Netherlands
I.Bozhanin@tudelft.nl

1 Introduction

Accuracy-driven recommender systems risk confining users to "filter bubbles", limiting content discovery and causing recommendation fatigue [8, 6]. Consequently, beyond-accuracy metrics (BAMs) like novelty and diversity are critical for a holistic evaluation of user experience [13]. While Principal Component Analysis (PCA) can model user-item correlations, its $\mathcal{O}(N^3)$ computational complexity is a bottleneck for large-scale systems [7]. coVariance Neural Networks (VNNs) offer a scalable alternative by interpreting the covariance matrix as a graph shift operator (GSO) for graph convolutions [11]. Their sparsified variants (S-VNNs) further employ techniques like hard or soft thresholding to reduce noise and computational complexity [2]. However, their impact on the novelty and diversity of recommendations remains an unexplored research gap.

2 Methodology

Our method uses a VNN, based on the SelectionGNN architecture [4], to predict user ratings on the MovieLens-100K dataset (943 users and 1682 movies). We model users as nodes in a graph, with their rating vectors serving as node features. We benchmark six GSOs: dense, hard-threshold, and soft-threshold variants of both the user-user covariance ($\tilde{\mathbf{C}}$) and precision ($\tilde{\mathbf{P}}$) matrices. These are compared against a PCA, naive, and random baselines.

To directly optimize for beyond-accuracy objectives, we evaluate a **BAM-weighted** compound loss function and compare it against a standard **RMSE-only** training regime. The compound loss is defined as:

$$\mathcal{L} = \lambda_{\text{RMSE}}\mathcal{L}_{\text{RMSE}} + \lambda_{\text{Nov}}(1 - \text{Nov}) + \lambda_{\text{Div}}(1 - \text{Div}) \tag{1}$$

where $\mathcal{L}_{\text{RMSE}}$ is the standard reconstruction error. For evaluation, we measure accuracy with **RMSE** and beyond-accuracy performance with two key metrics on top-N recommendation lists. **Novelty** is quantified based on item popularity; an item is considered novel if it has been rated by few users, thus rewarding recommendations from the "long tail" of the catalog [5]. **Intra-List Diversity (ILD)** is calculated as the average pairwise dissimilarity between all items in

a list [13]; our dissimilarity metric is a **hybrid** of a *collaborative* signal (how differently two items are rated across all users) and a *content* signal (dissimilarity based on movie genres), combined with equal weight.

3 Results and Conclusion

Our key findings show that S-VNNs offer a favorable accuracy-novelty trade-off. The best S-VNN configuration (using a hard-thresholded precision GSO) **increases novelty by 2.8 percentage points** over PCA while incurring only a marginal **0.03 RMSE penalty**. This demonstrates that S-VNNs can tangibly enhance content discovery at a negligible cost to predictive accuracy as visualized in Figure 1. However, diversity improvements were inconsistent across configurations.

Furthermore, the choice of GSO construction is critical: hard-thresholded precision graphs yield the best accuracy (0.952 RMSE), while dense covariance graphs produce the highest diversity (0.868 ILD). A crucial practical insight is that **directly optimizing for BAMs in the loss function proves inefficient**. The BAM-weighted training regime inflated runtime by a factor of $\times 33$ but yielded no statistically significant performance gain over RMSE-only models, suggesting that S-VNNs' inherent graph-based learning **sufficiently captures beyond-accuracy benefits without explicit multi-objective optimization**.

In conclusion, this work provides the first benchmark of S-VNNs for beyond-accuracy recommendation on a small-scale dataset. We show that S-VNNs with a sparsified precision GSO, trained on standard RMSE loss, offer modest improvements in novelty over PCA-based methods, though scalability to larger systems remains to be demonstrated. All code and results are released for full reproducibility. A complete description of the methodology, extended results, and statistical analysis is also available in the full thesis [1].

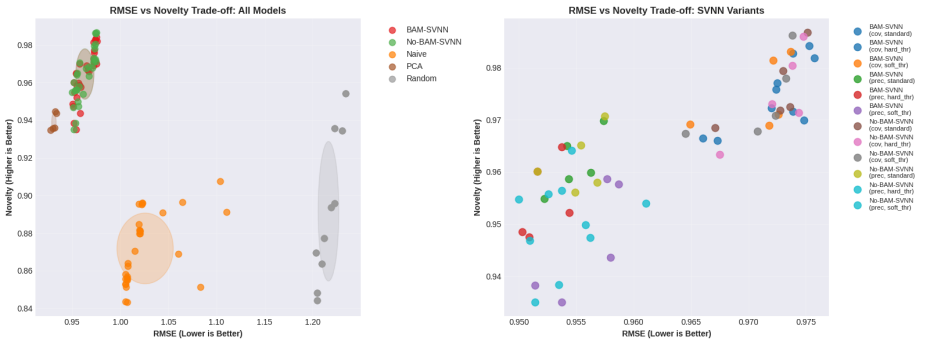

Fig. 1. Accuracy (RMSE) vs. Novelty trade-off across model families. (S-)VNN variants (ellipses on the left) outperform PCA in novelty while maintaining comparable accuracy. Ellipses denote 95% confidence regions.

References

1. Bozhanin, I.: Beyond-Accuracy (Sparsed-) coVariance Neural Network Recommender Systems. Bachelor Thesis, Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science (2025). https://resolver.tudelft.nl/uuid:06cb24c7-f380-4ea7-92ce-07762c066b74
2. Cavallo, A., Gao, Z., Isufi, E.: Sparse Covariance Neural Networks. arXiv preprint arXiv:2410.01669 (2024)
3. Gewers, F.L., Ferreira, G.R., de Arruda, H.F., et al.: Principal Component Analysis: A Natural Approach to Data Exploration. ACM Comput. Surv. **54**(4), 1–34 (2022)
4. Gama, F., Marques, A.G., Leus, G., Ribeiro, A.: Convolutional neural network architectures for signals supported on graphs. IEEE Trans. Signal Process. **67**(4), 1034–1049 (2019)
5. Jadon, A., Patil, A.: A Comprehensive Survey of Evaluation Techniques for Recommendation Systems. arXiv preprint arXiv:2312.16015 (2024)
6. Nian, L., Xin, B., Cheng, L., et al.: Modeling User Fatigue for Sequential Recommendation. arXiv preprint arXiv:2405.11764 (2024)
7. Pan, V.Y., Chen, Z., Zheng, A., et al.: The complexity of the algebraic eigenproblem. In: Mathematical Sciences Research Institute, Berkeley, pp. 1998–71 (1998)
8. Pariser, E.: The Filter Bubble: What The Internet Is Hiding From You. Penguin Books Limited, London (2011)
9. Castells, P., Hurley, N.J., Vargas, S.: Novelty and Diversity in Recommender Systems. In: Ricci, F., Rokach, L., Shapira, B. (eds.) Recommender Systems Handbook, pp. 881–918. Springer, Boston, MA (2015)
10. Sharma, K., Lee, Y.-C., Nambi, S., et al.: A Survey of Graph Neural Networks for Social Recommender Systems. arXiv preprint arXiv:2212.04481 (2024)
11. Sihag, S., Mateos, G., McMillan, C., Ribeiro, A.: coVariance Neural Networks. In: Advances in Neural Information Processing Systems, vol. 35, pp. 17003–17016 (2023)
12. Montag, C., Yang, H., Elhai, J.D.: On the Psychology of TikTok Use: A First Glimpse From Empirical Findings. Front. Public Health **9**, 641673 (2021)
13. Vargas, S., Castells, P.: Rank and relevance in novelty and diversity metrics for recommender systems. In: Proceedings of the fifth ACM conference on Recommender systems, pp. 109–116. ACM, Chicago (2011)