CHENGDU UNIVERSITY OF TECHNOLOGY
OXFORD BROOKES UNIVERSITY
Wang Yutong (Ryan)

Book Recommendation using Variational Autoencoders

Chengdu University of Technology and Oxford Brookes College, Computer Science
Supervised by Dr Joojo Walker

Abstract

In this study, four book rating datasets are collected and preprocessed, and three variant VAE models, MultVAE, MacridVAE and RecVAE, are implemented and trained to explore their applicability and efficiency for book recommendation. The project comprehensively evaluated the performance of these models using various evaluation metrics such as normalized discount cumulative gain (NDCG), recall and precision. The results show that these models exhibit significant accuracy and efficiency in handling real user rating data. And the performance on unseen users verified the generalization ability of the models. In particular, the MacridVAE model performs well on multiple datasets, demonstrating its advantages in handling sparse data and improving the quality of long list recommendations.

VAE Models

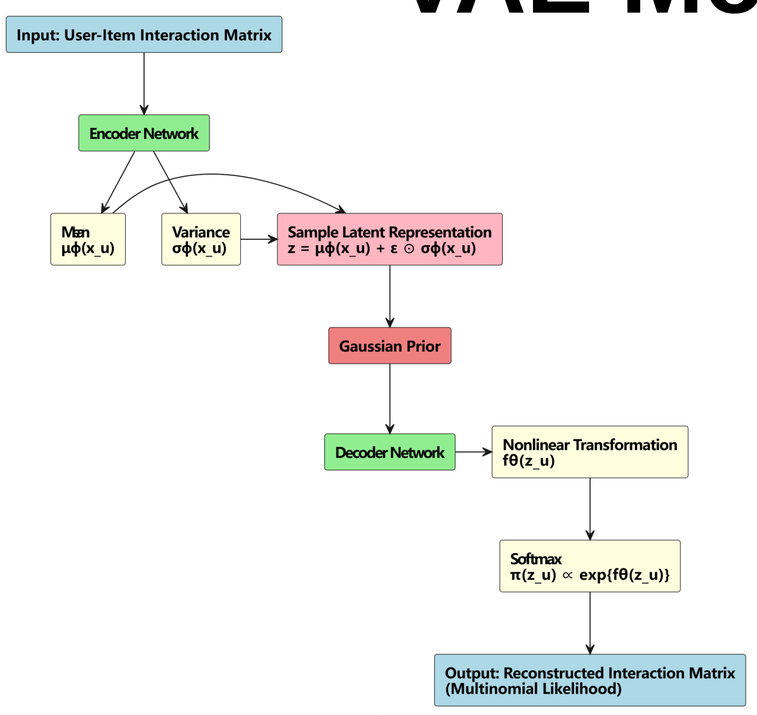

Fig 1 MultVAE

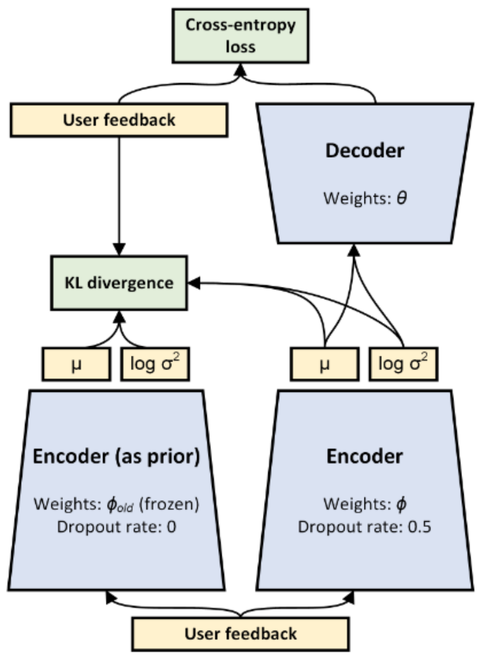

Fig 2 RecVAE

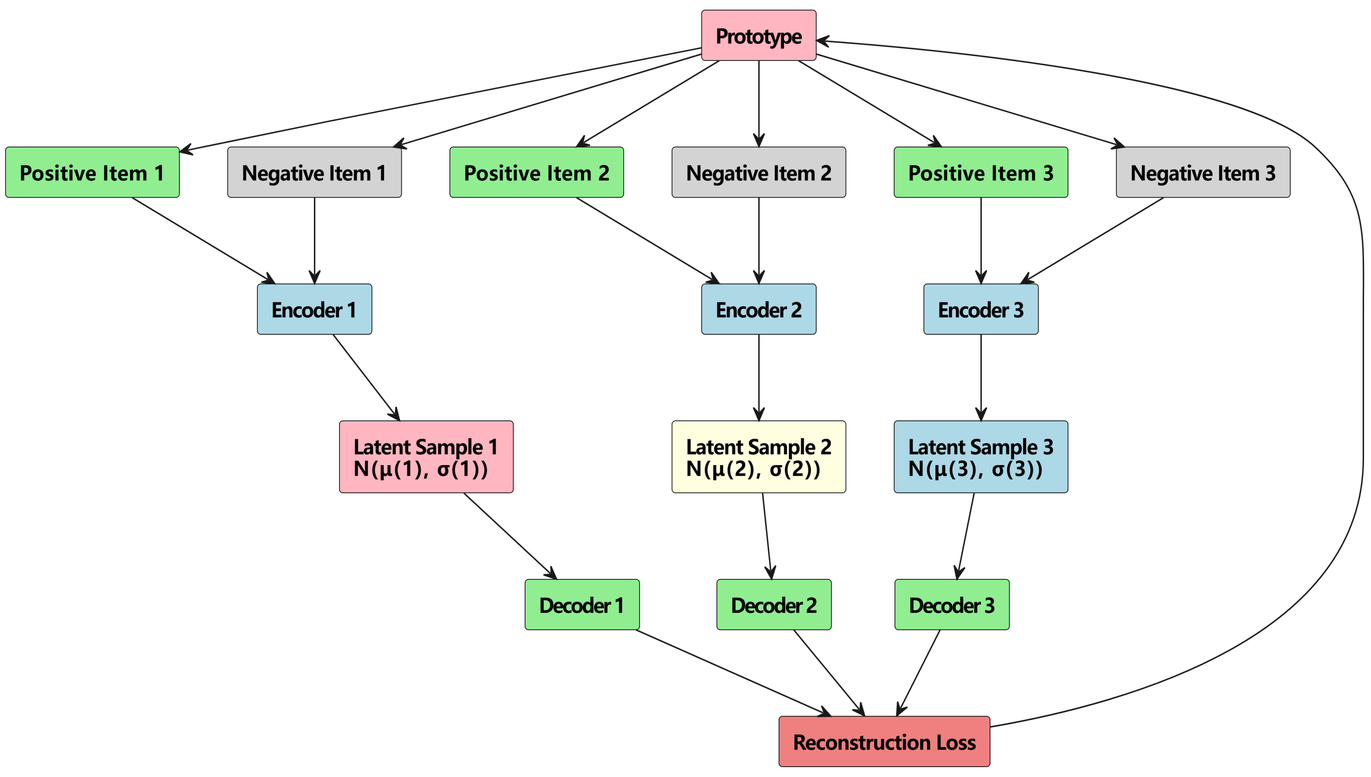

Fig 3 MacridVAE

Results

These figures show the model performance under different evaluation metrics and recommendation list lengths such as Normalized Discounted Cumulative Gain (NDCG), Recall and Precision.

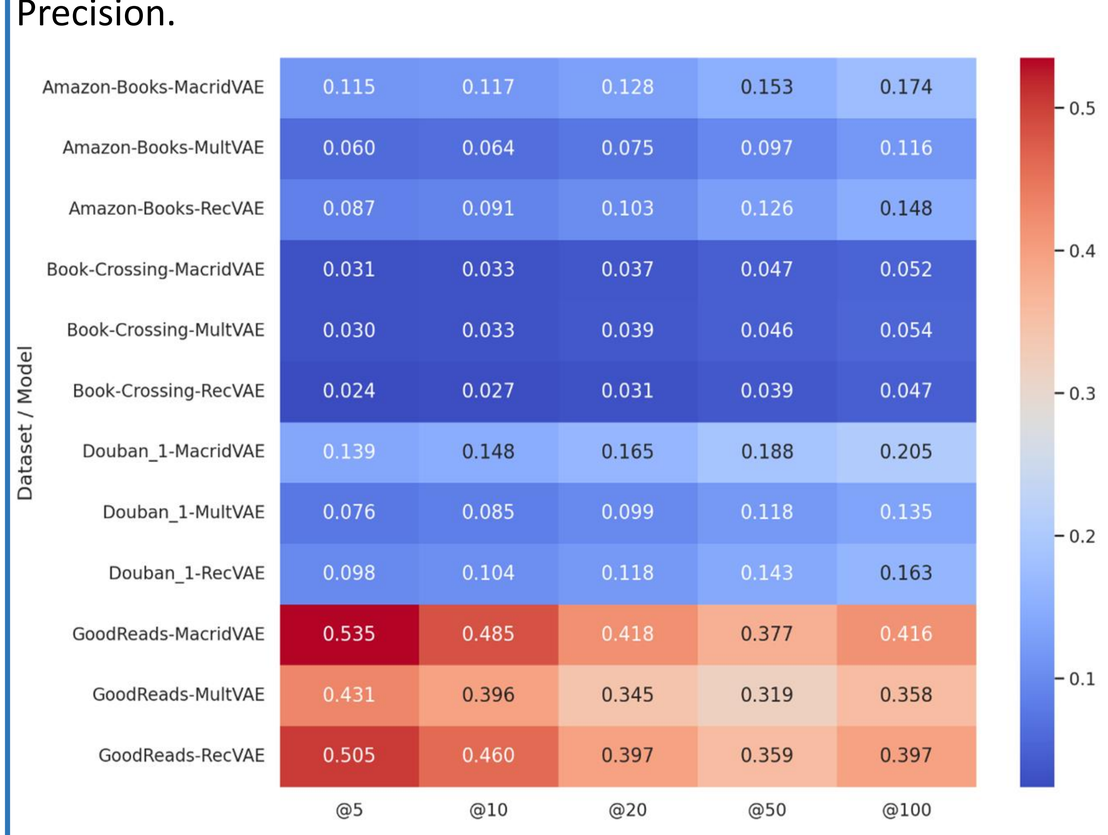

Fig 4 Heatmap of NDCG Performance

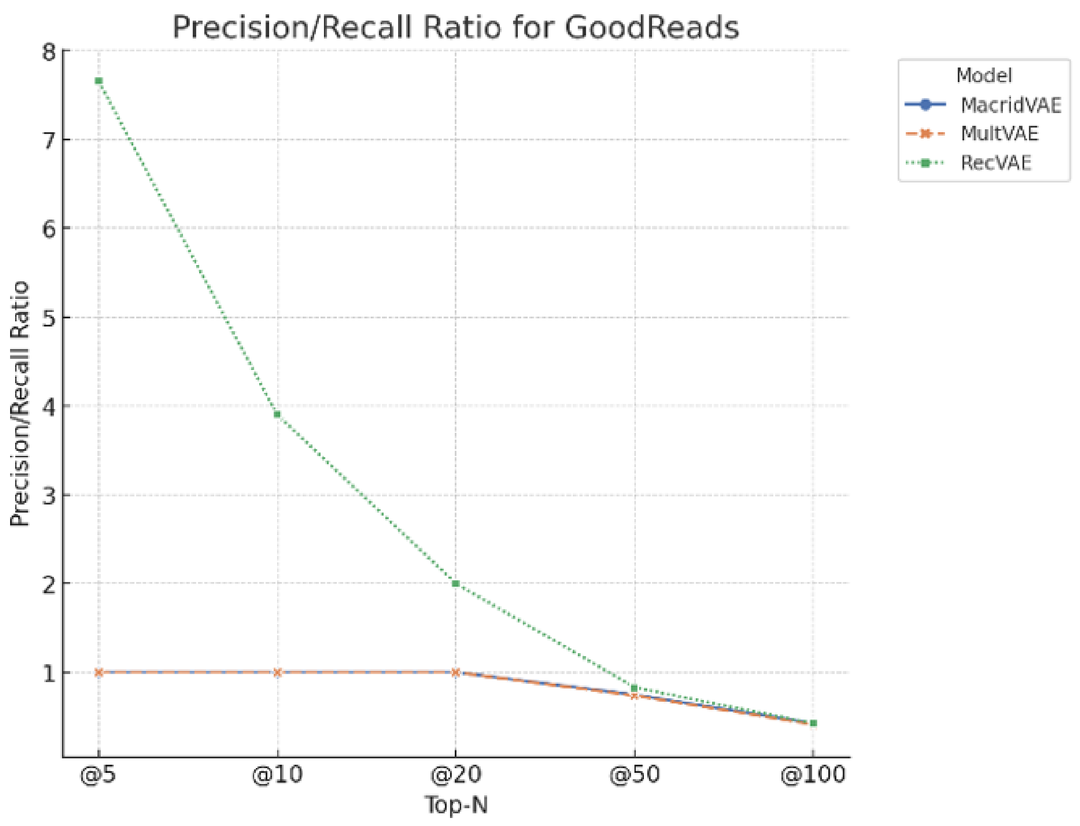

Fig 5 Precision/Recall Ratio

Significance of this work:

I. The studies here indicate that VAE - a generative model can be used for recommendations on book-related e-commerce platforms.
II. VAE models can alleviate the data sparsity and cold start problems effectively.

Conclusion

In this study, we have successfully explored the application of three VAE models (MultVAE, MacridVAE, RecVAE) in personalized book recommendation systems using the variational autoencoder (VAE) technique. By carefully preprocessing and analyzing four book rating datasets, we addressed data sparsity, improved recommendations for new users and books, and tackled the cold-start problem. The experimental results show that the MacridVAE model, in particular, achieves the best NDCG@100 of 0.416 on all datasets, which is significantly better than the other models, proving its advantages in long-list recommendation and handling sparse data.

Future Work

- Hyperparameters Tuning
- Algorithm Exploration
- Data Extension
- Cross-domain Recommendations

Reference

[1] P. Kerudi, T. Srujan, and R. C K, 'BOOK RECOMMENDATION SYSTEM: A SYSTEMATIC REVIEW AND RESEARCH ISSUES', *Huazhong Keji Daxue Xuebao Ziran Kexue BanJournal Huazhong Univ. Sci. Technol. Nat. Sci. Ed.*, vol. 50, p. HST-0621, Jun. 2021.

[2] H. Alharthi, D. Inkpen, and S. Szpakowicz, 'A survey of book recommender systems', *J. Intell. Inf. Syst.*, vol. 51, no. 1, pp. 139–160, Aug. 2018, doi: 10.1007/s10844-017-0489-9.

[3] D. Liang, R. G. Krishnan, M. D. Hoffman, and T. Jebara, 'Variational Autoencoders for Collaborative Filtering', arXiv.org. Accessed: Oct. 24, 2023. [Online]. Available: https://arxiv.org/abs/1802.05814v1

[4] J. Ma, C. Zhou, P. Cui, H. Yang, and W. Zhu, 'Learning disentangled representations for recommendation', in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Red Hook, NY, USA: Curran Associates Inc., 2019, pp. 5711–5722.

[5] I. Shenbin, A. Alekseev, E. Tutubalina, V. Malykh, and S. I. Nikolenko, 'RecVAE: a New Variational Autoencoder for Top-N Recommendations with Implicit Feedback', in *Proceedings of the 13th International Conference on Web Search and Data Mining*, Jan. 2020, pp. 528–536. doi: 10.1145/3336191.3371831.