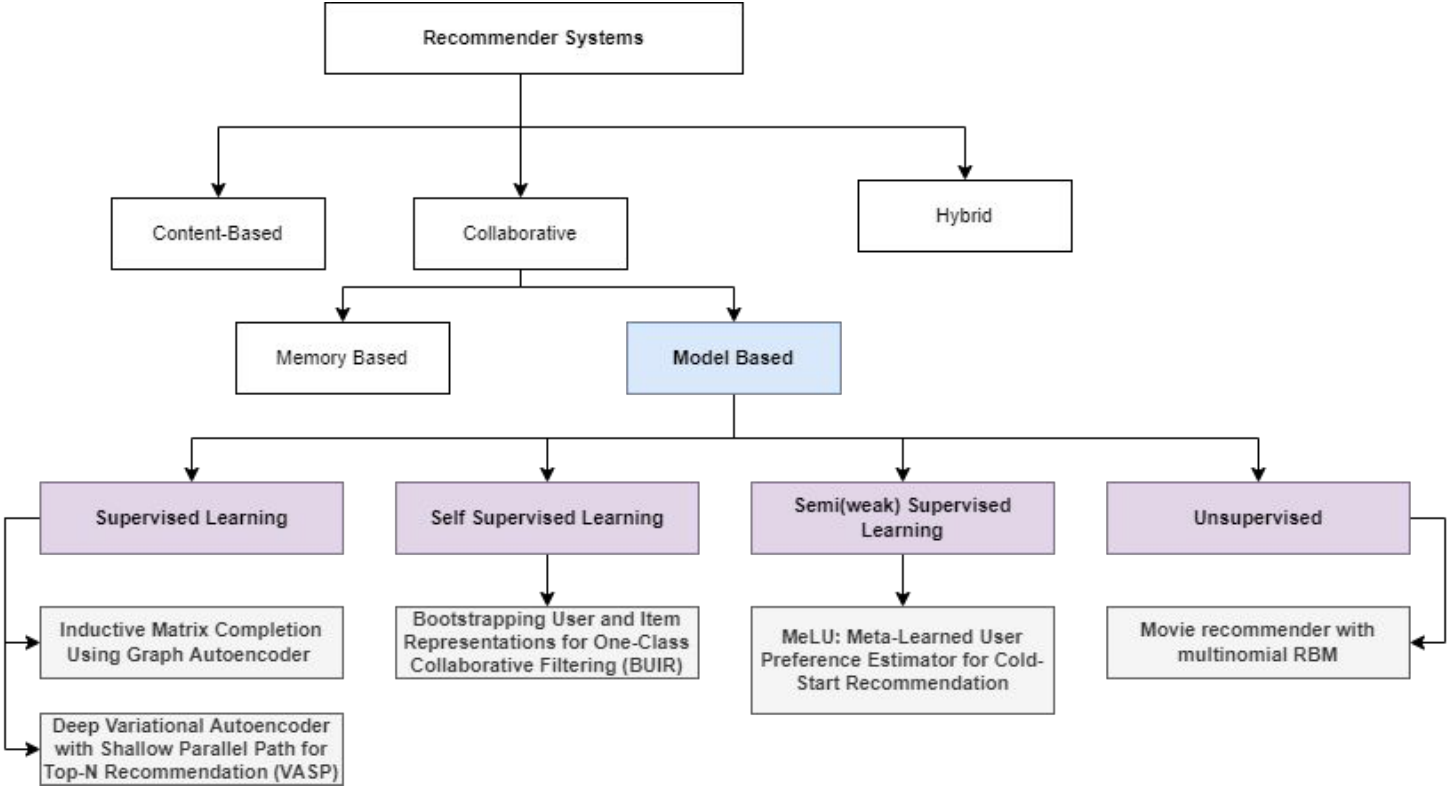
A Review of Deep Learning-Based Movie Recommendation Systems

[Group#22][Project#4]
Chinmay Bhale, Truxten Cook, Kritshekhar Jha,
Kumarage Tharindu Kumarage, Kyle Otstot, Paras Sheth

Problem Definition

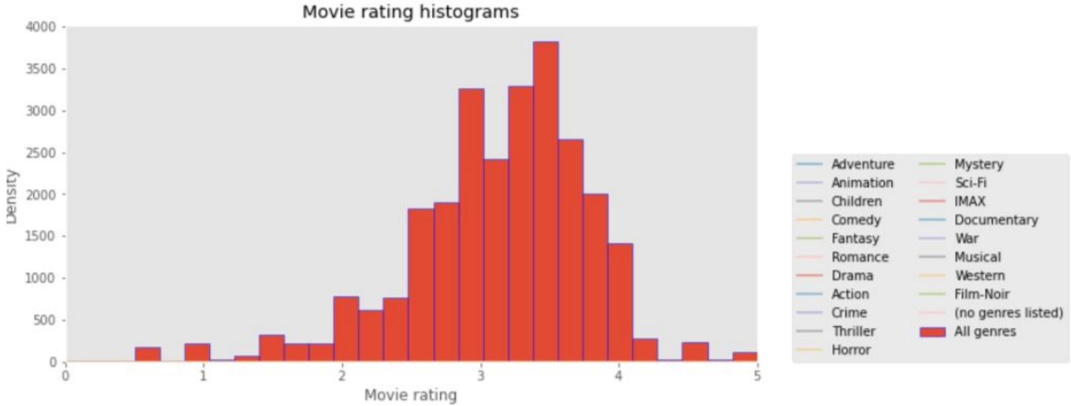
- Recommender systems have proven to be a successful method of reducing information overload in light of the increasing volume of online information.
- In recent years, research on recommender systems and information retrieval has shown that deep learning has a wide-ranging impact. However, they still have issues when working with extremely sparse data or other problems like cold start.
- The purpose of this review paper is to provide a comprehensive and systematic analysis of current research on recommender systems based on deep learning.
- We will provide a detailed taxonomy along with summaries for state-of-the-art algorithms.
- We also plan to provide a perspective on the future trends and research challenges of deep learning in recommender systems.

SOTA Taxonomy



Dataset

MovieLens



1. **1M**:
   - 1 million rankings, in the range of 0.5 to 5.0, over ~4000 movies.
   - Ratings from ~6040 users.
   - Contains user information: Gender::Age::Occupation::Zip-code

2. **20M**:
   - 20 million rankings, in the range of 0.5 to 5.0, over 62,000 movies.
   - Ratings from 138493 users.
   - Movies are associated with genres and user-generated tags.
   - A common benchmark for testing a wide variety of different movie recommendation algorithms.

System Architecture & Evaluations

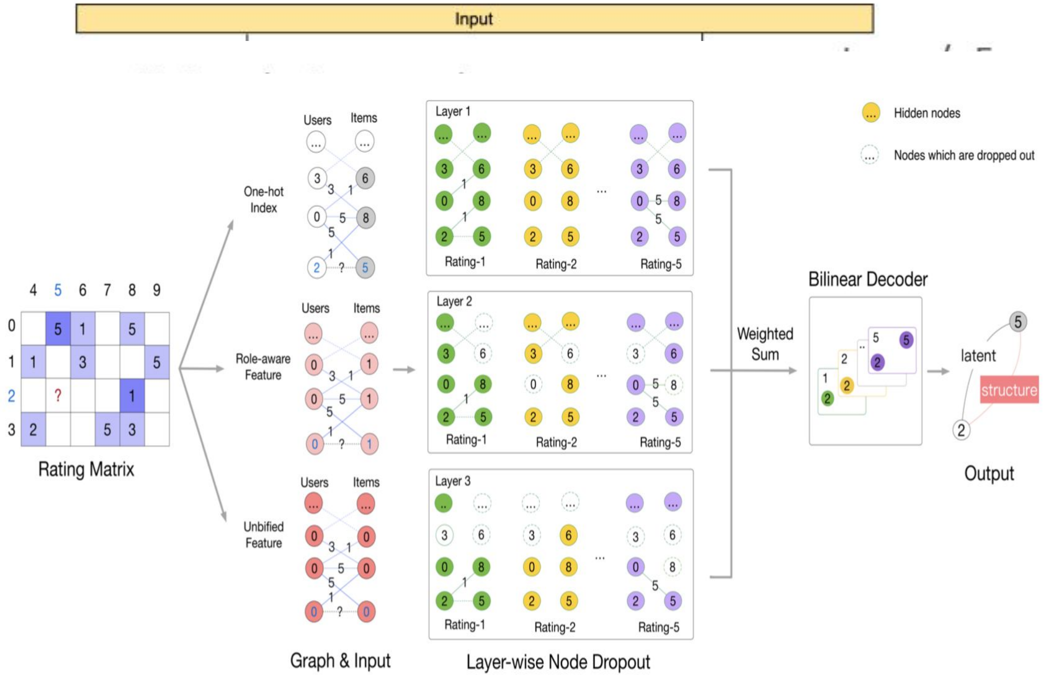
The implemented models in the scope of this project can also be classified based upon the dataset used for training and inference. We will be conducting the comparative analysis of the evaluations results based on this classification.

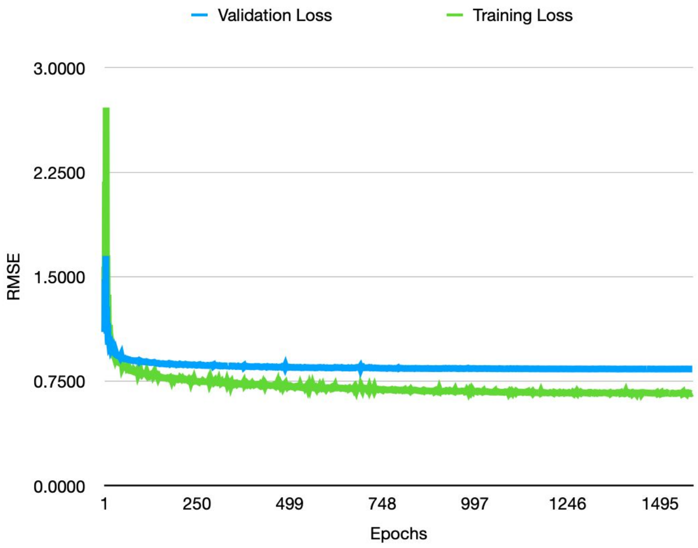
- **Dataset: Movielens 20M - Use only the rating information**
  - Deep Variational Autoencoder with Shallow Parallel Path for Top-N Recommendation (VASP)

- **Dataset: MovieLens 1M - Rating information and user auxiliary information**
  - MeLU: Meta-Learned User Preference Estimator for Cold-Start Recommendation
  - Inductive Matrix Completion Using Graph Autoencoder
  - MultiNominal Restricted Boltzman Machine
  - Bootstrapping User and Item Representations for One-Class Collaborative Filtering (BUIR)

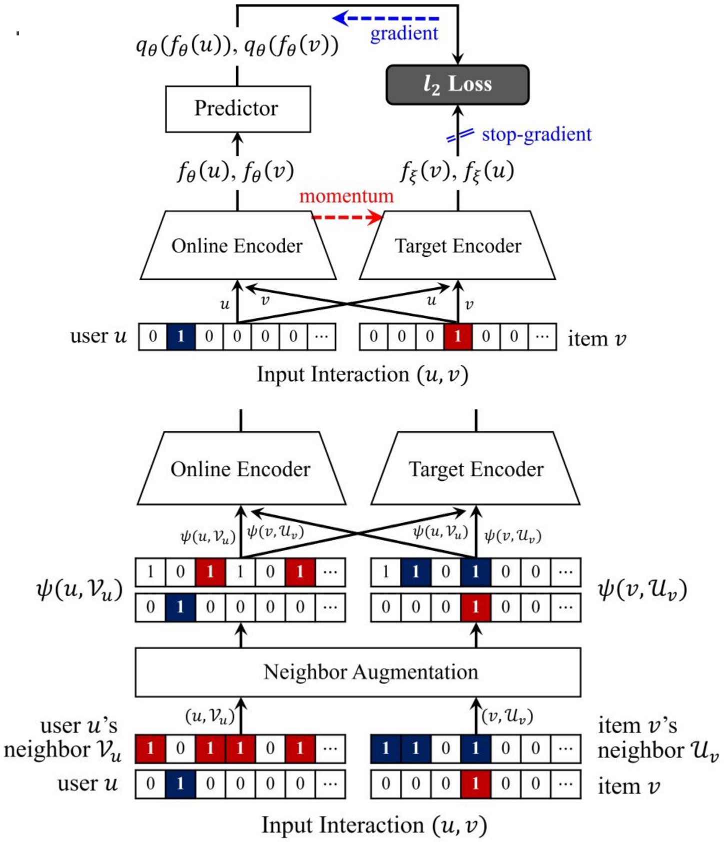
[Evaluation/Visualisation] Supervised Learning: "Deep Variational Autoencoder with Shallow Parallel Path for Top-N Recommendation (VASP)"

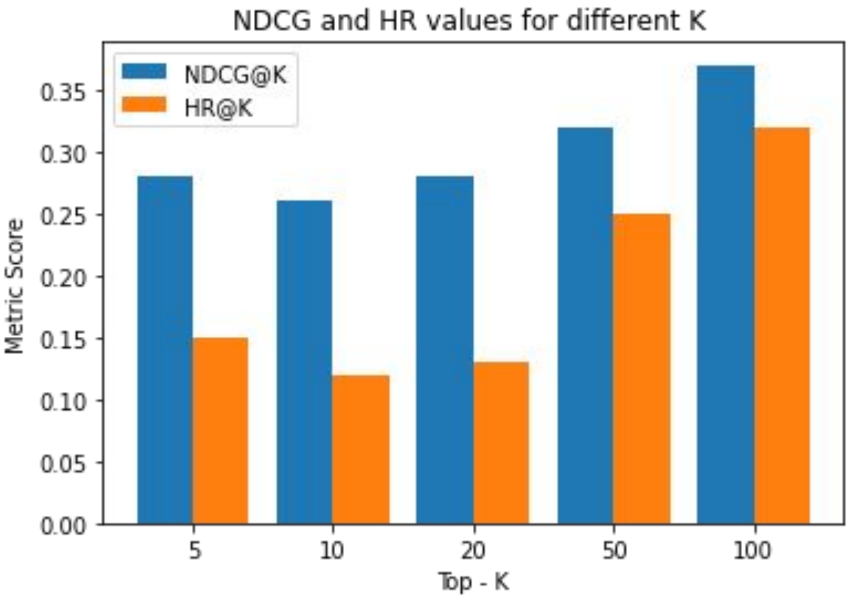
[Overview] Supervised Learning: "Deep Variational Autoencoder with Shallow Parallel Path for Top-N Recommendation (VASP)"



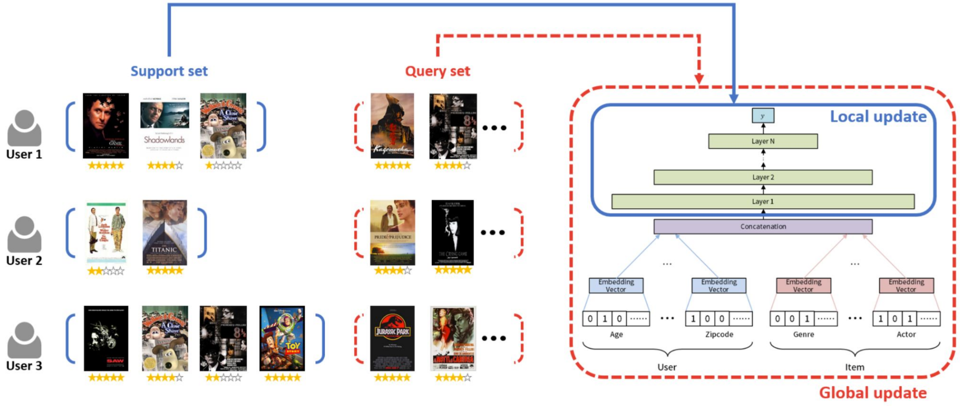


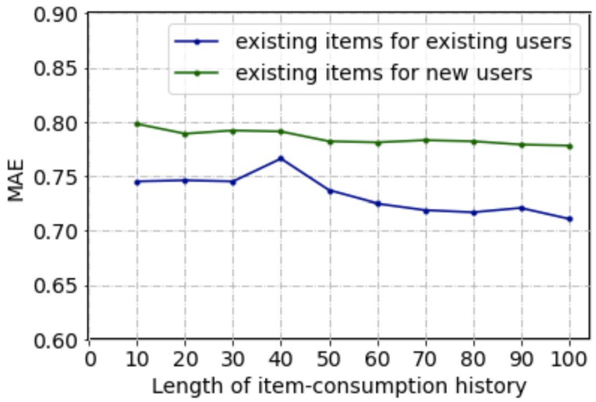
Self-Supervised Learning: **"Bootstrapping User and Item Representations for One-Class Collaborative Filtering (BUIR)"**





Semi-supervised Learning: **"MeLU: Meta-Learned User Preference Estimator for Cold-Start Recommendation"**

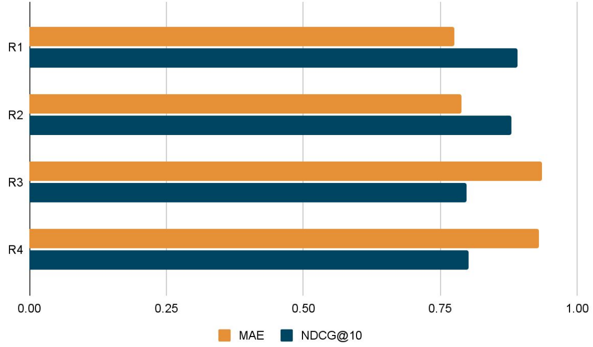




Recommendation Scenarios:

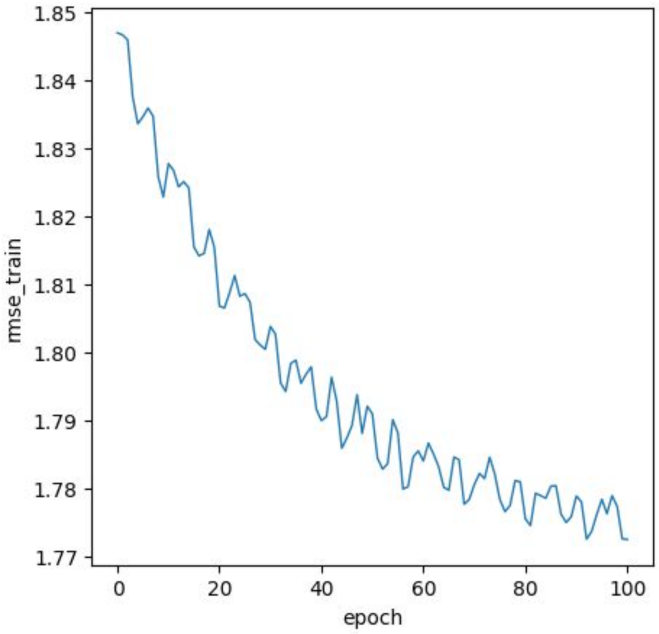
- R1- Recommendation of existing items for existing users
- R2- Recommendation of existing items for new users
- R3- Recommendation of new items for existing users
- R4- Recommendation of new items for new users

Cold Start



Unsupervised Learning: **MultiNomial Restricted Boltzmann Machine**

- Generative neural network
- Attempt to learn the distribution of *tastes* (latent factors) and use this approximated distribution to generative new rankings.
- Make guesses about the distribution and try to approximate it.
  - Gibbs Sampling to sample from the distribution and calculate forward and backward gradients
  - Minimize *Free Energy* in the system.



Comparative Study - Performance

Dataset	Learning Paradigm	Algorithm	NCDG@100	Recall@20	Recall@50
MovieLens **20M**	Supervised Learning	VASP	**0.444**	**0.411**	**0.548**

Dataset	Learning Paradigm	Algorithm	RMSE	MAE	NCDG@10	HR@10
MovieLens **1M**	Supervised Learning	IMC-GAE	0.8299	0.738	.779	**0.689**
MovieLens **1M**	Semi-Supervised Learning	MeLU	0.847	0.7653	**0.843**	-
MovieLens **1M**	Self-Supervised Learning	BUIR	-	-	0.11	0.25
MovieLens **1M**	Unsupervised Learning	Multinominal Restricted Boltzman Machine	**0.828**	**0.538**	0.685	-

Comparative Study - Resources

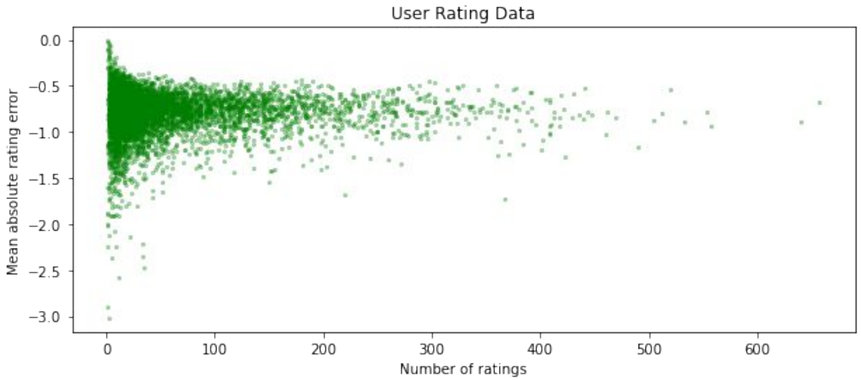
Dataset	Learning Paradigm	Algorithm	Running Time (Min)	GPU Config
MovieLens **20M**	Supervised Learning	VASP	680	Tesla V100, 16GB
Dataset	**Learning Paradigm**	**Algorithm**	**Running Time (Min)**	GPU config
MovieLens **1M**	Supervised Learning	IMC-GAE	25	Nvidia K80 12GB
MovieLens **1M**	Semi-Supervised Learning	MeLU	240	Tesla V100, 16GB
MovieLens **1M**	Self-Supervised Learning	BUIR	25	Nvidia GeForce RTX 3090 24GB
MovieLens **1M**	Unsupervised Learning	MultiNominal Restricted Boltzmann Machine	1/6	Nvidia GeForce RTX 3080 16GB

Bonus Section

- **Task:** “Ranking expert users and top movies by genre”

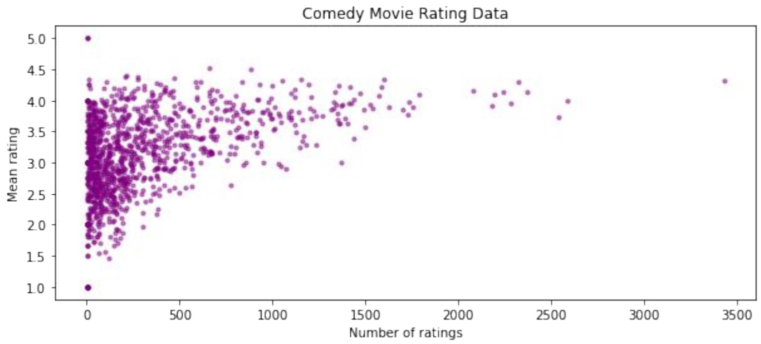
How to quantify user ***expertise***?

- Consistency with public opinion
- Number of watched movies

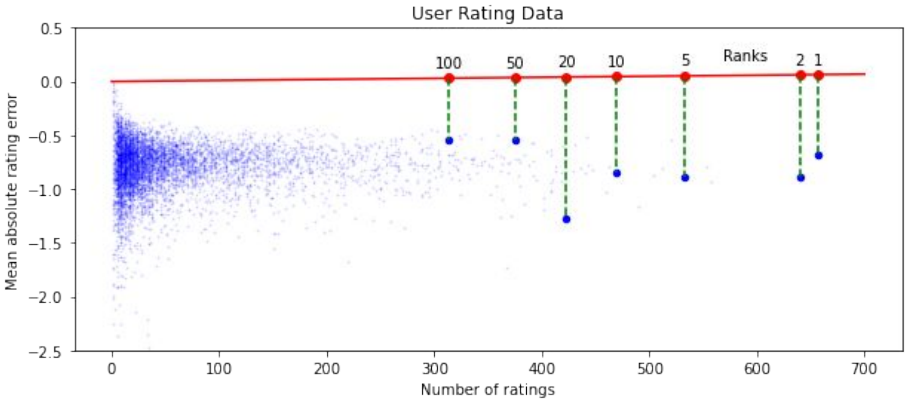


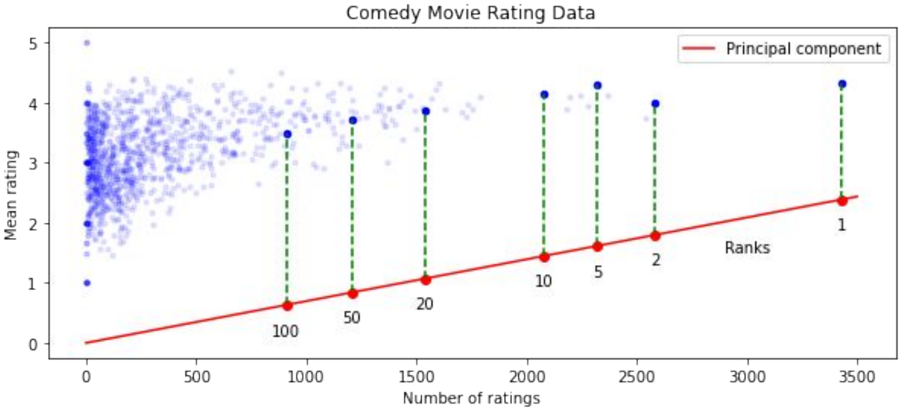
How to quantify ***top*** movies?

- Average user rating
- Number of watches/reviews



- **Solution:** project the 2D metric space onto vector that best preserves variance (PCA) and establishes natural ranking system





Rank	UserID
1	4163
2	1677
3	1938

Rank	Movie Title
1	American Beauty (1999)
2	Back to the Future (1985)
3	Men in Black (1997)

Project Plan: Tasks, Deadlines, Division of Work

#	Task Description	Task Owner	Deadline	Status
1	Background Study & Literature Survey	Everyone	25-Oct	Completed
2	Project Proposal	Everyone	12-Oct	Completed
3	Brainstorm the taxonomy for the survey	Everyone	25-Oct	Completed
4	Understanding of the selected SOTA for Supervised Learning	Chinmay Bhale	25-Oct	Completed
5	Implement the selected SOTA for Supervised Learning	Chinmay Bhale	15-Nov	Completed
6	Understanding of the selected SOTA for Supervised Learning	Kritshekhar Jha	25-Oct	Completed
7	Implement the selected SOTA for Supervised Learning	Kritshekhar Jha	15-Nov	Completed
8	Understanding of the selected SOTA for Self Supervised Learning	Paras Sheth	25-Oct	Completed
9	Implement the selected SOTA for Self Supervised Learning	Paras Sheth	15-Nov	Completed
10	Understanding of the selected SOTA for Semi Supervised Learning	Tharindu Kumarage	25-Oct	Completed
11	Implement the selected SOTA for Semi Supervised Learning	Tharindu Kumarage	15-Nov	Completed
12	Understanding of the selected SOTA for Unsupervised Learning	Truxten Cook	25-Oct	Completed
13	Implement the selected SOTA for Unsupervised Learning	Truxten Cook	15-Nov	Completed
14	Understanding of the selected SOTA for Unsupervised Learning	Kyle Otsot	25-Oct	Completed
15	Implement the selected SOTA for Unsupervised Learning	Kyle Otsot	15-Nov	Completed
16	Comparative analysis & discussion	Everyone	20-Nov	Completed
17	Project Presentation	Everyone	14-Nov	Completed
18	Group Demo	Everyone	30-Nov	Completed
19	Final Report	Everyone	2-Dec	Completed

Project Repository

https://github.com/kotstot6/MovieRecommendation

README.md

[Fall2022] [CSE573] [Group Project] Semantic Web Mining

[Group:22][Project#4] - A Review of Deep Learning-Based Movie Recommendation Systems

Group Members

- Chinmay Bhale
- Truxten Cook
- Kritshekhar Jha
- Kumarage Tharindu Kumarage
- Kyle Otstot
- Paras Sheth

Problem Definition

- Recommender systems have proven to be a successful method of reducing information overload in light of the increasing volume of online information.
- In recent years, research on recommender systems and information retrieval has shown that deep learning has a wide-ranging impact. However, they still have issues when working with extremely sparse data or other problems like cold start
- The purpose of this review paper is to provide a comprehensive and systematic analysis of current research projects on recommender systems based on deep learning
- We will provide a detailed taxonomy along with summaries for state-of-the-art algorithms.
- We also plan to provide a perspective on the future trends and research challenges of deep learning in recommender systems

[kotstot6/MovieRecommendation (github.com)](https://github.com/kotstot6/MovieRecommendation)

References

[1] J. Dong, X. Li, and B. Fang, “A recommendation system based on multi-attribute,” in 2016 9th International Conference on Service Science (ICSS) IEEE, 2016, pp. 165–169.
[2] X. Zhang, W. Dou, Q. He, R. Zhou, C. Leckie, R. Kotagiri, and Z. Salcic, “Lshiforest: A generic framework for fast tree isolation based ensemble anomaly analysis,” in 2017 IEEE 33rd international conference on data engineering (ICDE) IEEE, 2017, pp. 983–994.
[3] M. Wang, M. Liu, J. Liu, S. Wang, G. Long, and B. Qian, “Safe medicine recommendation via medical knowledge graph embedding,” arXiv preprint arXiv:1710.05980, 2017.
[4] Y.-L. Chen, Y.-H. Yeh, and M.-R. Ma, “A movie recommendation method based on users’ positive and negative profiles,” Information Processing & Management, vol. 58, no. 3, p. 102531, 2021.
[5] A. Y. Mohamad, S. R. Harun, N. A. A. Shahidan, A. Nanthaamornphong, A. Mustapha, and M. H. A. Wahab, “Collaborative filtering approach for movie recommendations,” in 2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON) IEEE, 2022, pp. 1–6.
[6] B. M. Sarwar, Sparsity, scalability, and distribution in recommender systems. University of Minnesota, 2001.
[7] J. Liu, C. Shi, C. Yang, Z. Lu, and S. Y. Philip, “A survey on heterogeneous information network based recommender systems: Concepts, methods, applications and resources,” AI Open, 2022.
[8] B. Walek and V. Fojtik, “A hybrid recommender system for recommending relevant movies using an expert system,” Expert Systems with Applications , vol. 158, p. 113452, 2020.
[9] K. Bougiatiotis and T. Giannakopoulos, “Enhanced movie content similarity based on textual, auditory and visual information,” Expert Systems with Applications, vol. 96, pp. 86–102, 2018.
[10] B. Bhatt, J. P. Premal, and H. Gaudani, “A review paper on machine learning based recommendation system 1,” 2014.
[11] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, “Recommender system application developments: a survey,” Decision Support Systems , vol. 74, pp. 12–32, 2015.
[12] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, “Recommendation systems: Principles, methods and evaluation,” Egyptian informatics journal, vol. 16, no. 3, pp. 261–273, 2015.
[13] N. Sachdeva, M. P. Dhaliwal, C.-J. Wu, and J. McAuley, “Infinite recommendation networks: A data-centric approach,” arXiv preprint arXiv:2206.02626 , 2022.
[14] M. Wu and N. Goodman, “Multimodal generative models for scalable weakly-supervised learning,” Advances in Neural Information Processing Systems, vol. 31, 2018.
[15] X. Liu, F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang, and J. Tang, “Self-supervised learning: Generative or contrastive,” IEEE Transactions on Knowledge and Data Engineering, 2021.
[16] Y. Wu, C. DuBois, A. X. Zheng, and M. Ester, “Collaborative denoising auto-encoders for top-n recommender systems,” in Proceedings of the ninth ACM international conference on web search and data mining, 2016, 153–162.
[17] S. Li, J. Kawale, and Y. Fu, “Deep collaborative filtering via marginalized denoising auto-encoder,” in Proceedings of the 24th ACM international on conference on information and knowledge management, 2015, pp. 811–820.
[18] M. Gao, L. Chen, X. He, and A. Zhou, “Bine: Bipartite network embedding,” in The 41st international ACM SIGIR conference on research & development in information retrieval, 2018, pp. 715–724.
[19] N. CC and A. Mohan, “A social recommender system using deep architecture and network embedding,” Applied Intelligence, vol. 49, no. 5, pp. 1937–1953, 2019
[20] J. E. Van Engelen and H. H. Hoos, “A survey on semi-supervised learning,” Machine Learning , vol. 109, no. 2, pp.373–440, 2020.
[21] A. H. Khan, J. Siddqui, and S. S. Sohail, “A survey of recommender systems based on semi-supervised learning,” in International Conference on Innovative Computing and Communications Springer, 2022, pp. 319–327.
[22] Z. Du, X. Wang, H. Yang, J. Zhou, and J. Tang, “Sequential scenario-specific meta learner for online recommendation,” in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 2895–2904.
[23] M. Vartak, A. Thiagarajan, C. Miranda, J. Bratman, and H. Larochelle, “A meta-learning perspective on cold-start recommendations for items,” in Advances in neural information processing systems , 2017, pp. 6904–6914.
[24] H. Bharadhwaj, “Meta-learning for user cold-start recommendation,” in 2019 International Joint Conference on Neural Networks (IJCNN) IEEE, 2019, pp. 1–8.
[25] T. Aditya, K. Rajaraman, and M. Subashini, “Comparative analysis of clustering techniques for movie recommendation,” MATEC Web of Conferences , vol. 225, p. 02004, 01 2018.
[26]F. M. Harper and J. A. Konstan, “The movielens datasets: History and context,”ACM Trans. Interact. Intell. Syst.vvol. 5, no. 4, dec 2015. [Online]. Available: https://doi-org.ezproxy1.lib.asu.edu/10.1145/2827872
[27] J. Bennett, S. Lanning et al. , “The netflix prize,” in Proceedings of KDD cup and workshop , vol. 2007. Citeseer, 2007, p. 35