Variational Autoencoders for Recommender Systems

Tianshu Shen (shenti10)
Sijie Tian (tiansiji)

Motivation and Problem Definition

★ Recommender systems (RS):
  - Observe user interactions with items
  - Seek to predict unseen items users will like
  - Integral to the web - help users interact with content effectively

★ Collaborative Filtering (CF):
  - A family of RS algorithms
  - Predicts user preferences using similarity patterns across users and items

Prior Work on Collaborative Filtering

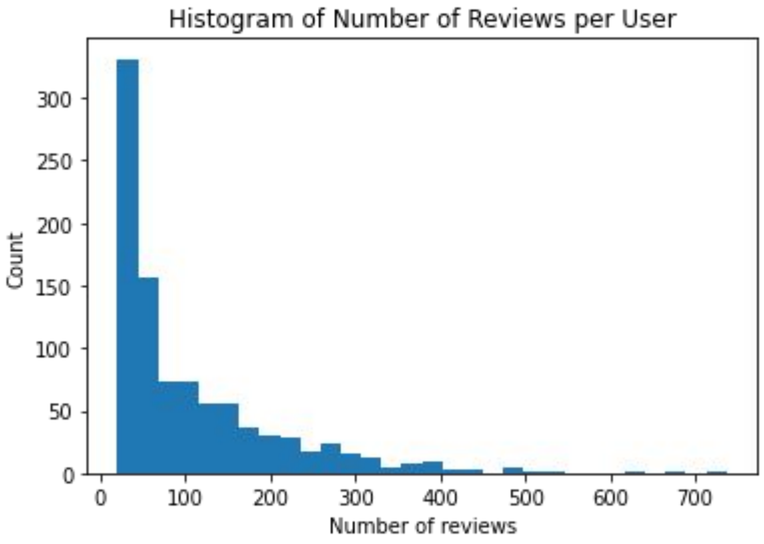
- ★ Neighbourhood methods: k-Nearest-Neighbors (kNN)
  - Filtering techniques: item-based, user-based (Sarwar et al., 2001; Resnick et al., 1994).
- ★ Latent-factor methods:
  - Matrix factorization (Luo et al., 2012)
  - Singular Value Decomposition (SVD) (Vozalis and Margaritis, 2007; Zhang et al., 2005)
  - **The extension of variational autoencoders to CF (VAE-CF)** (Liang et al., 2018)
    - Novel approach; no other prior work done for VAE applied to RS
- ★ Today: Movie recommendation using VAE-CF

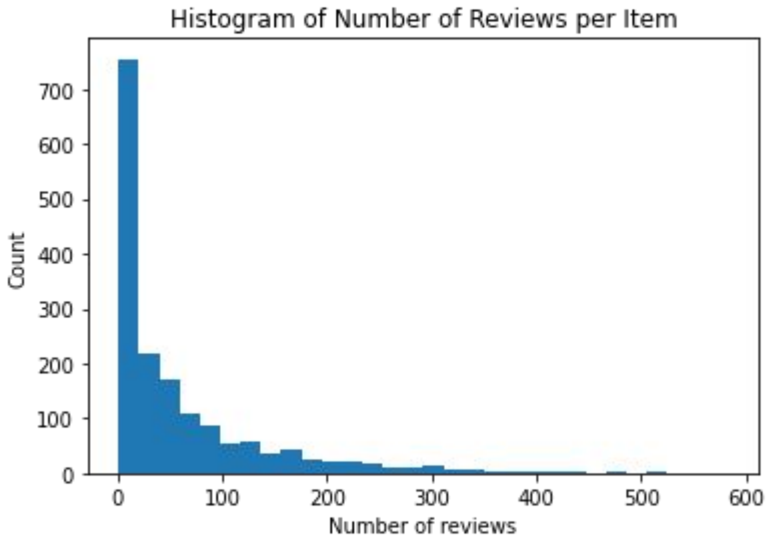
Dataset and Data Processing

- ★ MovieLens 100K dataset - Classic RS dataset
- ★ Describes people's preferences for movies in 1-5 star ratings at some time
  - ○ <userID, itemID, rating, timestamp> tuples
- ★ Since VAE-CF focuses on implicit feedback, change rating values to 1s
  - ○ Raw user behavior > numerical rating values
- ★ Split into 50% train, 20% validation, 30% test sets based on time of review
  - ○ Train on historic data
  - ○ Validate and test on future data

Exploratory Data Analysis

- ★ **100,000** reviews from **943** users and **1682** movies
- ★ Each user has rated **20+** movies, each item has **1+** ratings
- ★ There are **141** movies with only 1 rating
- ★ Item-based data is much more sparse than user-based data





Predictive Task, Baselines and Metrics

Predictive Task

Movie clicks predictions for users based on training data

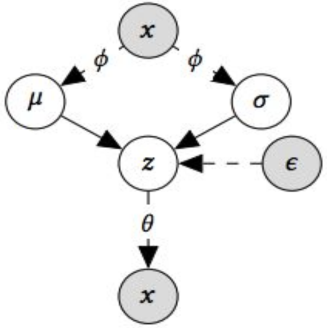
Baselines

★ User & item based Collaborative Filtering using kNN (da Costa et. al.,2018)
★ SVD

Metrics

★ Recall @ K
★ NDCG @ K (Normalized Discounted Cumulative Gain)

VAE-CF

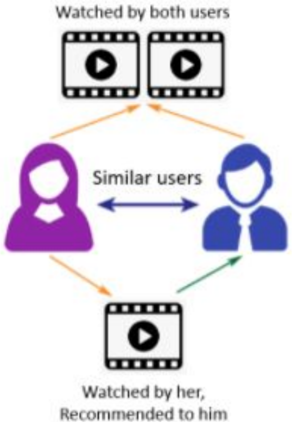


- ★ A probabilistic latent-variable model
- ★ VAE performs variational inference and generation:
  - ○ Inference (encoder): learns latent presentation
    - ■ Follows Gaussian prior $N(\mu_u, \mathrm{diag}\{\sigma^2_u\})$
    - ■ Uses **variational inference** to approximate the intractable posterior distribution for latent embedding given observations ($p(z|x)$)
  - ○ Generation (decoder): produce user click history
    - ■ $X_u \sim \mathrm{Mult}(N_u, \boldsymbol{\pi}(z_u))$

Model Comparison

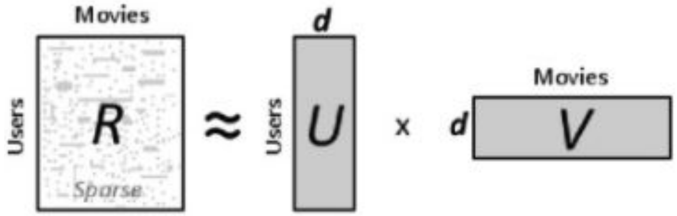
User-based KNN

✓ Introduce serendipity
✓ Easy implementation
× Cold-Start issue



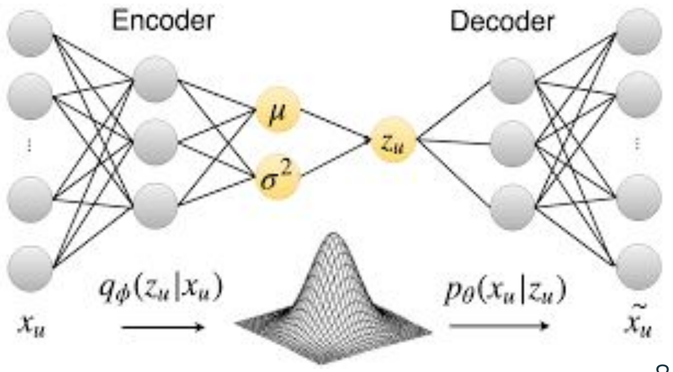
SVD

✓ Handle sparsity
× Uninterpretable latent representations
× Computationally expensive



VAE-CF

✓ Well suited for modelling implicit feedback data
✓ Handle Sparsity issues
× Computationally expensive



VAE-CF: Model Modifications

- VAE-CF was performed using a user-based approach in the original paper
- Item information could be used to augment user-based results
- Combination of predictions from both item-based and user-based VAE-CF model

Additional Hyperparameters:

- ❑ **Weight_u**
- ❑ **Weight_i**

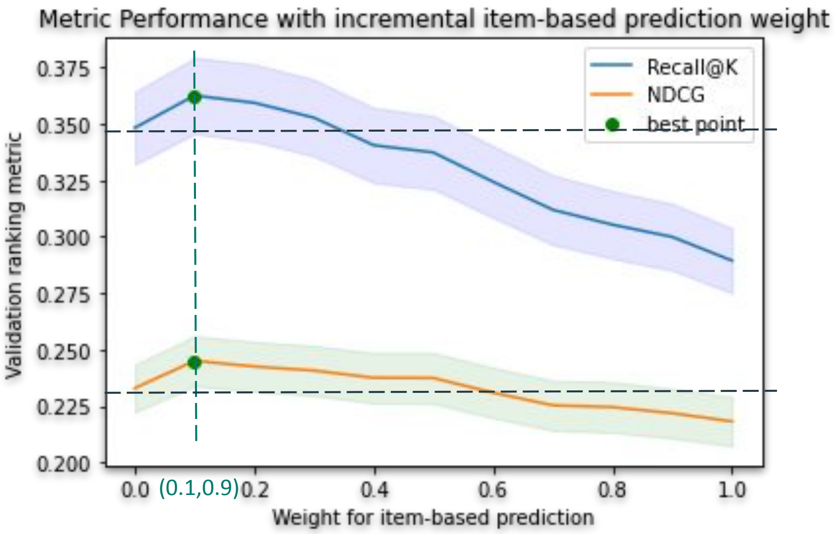
$$\textbf{Weight\_u} \bullet \begin{array}{cc} & I \\ U & \begin{pmatrix} \quad & \quad \\ \quad & \quad \end{pmatrix} \end{array} + \textbf{Weight\_i} \bullet \begin{array}{cc} & I \\ U & \begin{pmatrix} \quad & \quad \\ \quad & \quad \end{pmatrix} \end{array}$$

VAE-CF

- Model was trained by the user history click ({0,1}) data
- Hyperparameter tuning was performed using grid search
- Metrics used: Recall @ K, NDCG @ K

Hyperparameters:

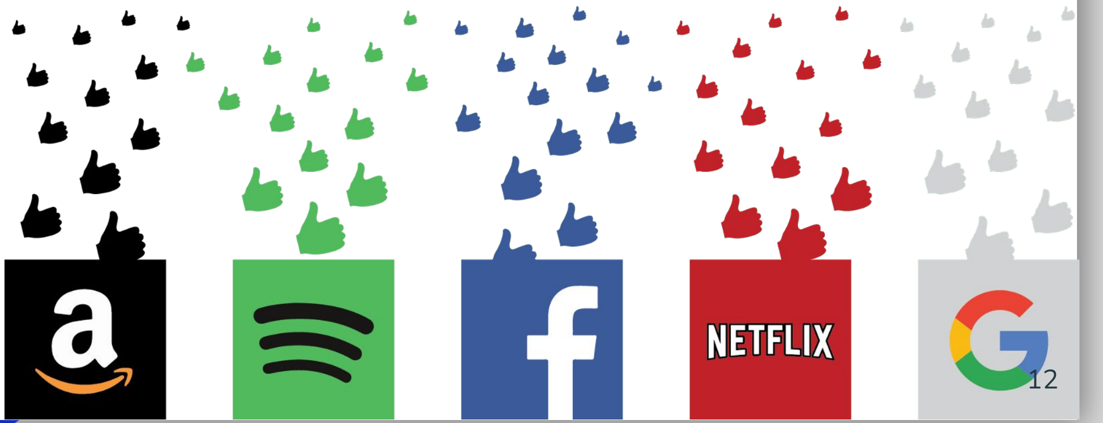
- ❑ Epoch
- ❑ Learning rate
- ❑ Lambda (regularization)
- ❑ Embedding latent dimension
- ❑ Corruption (1 - drop-out rate)
- ❑ Optimizer: RMSProp
- ❑ **Weight_u**
- ❑ **Weight_i**



Result

Method	Recall@50	NDCG@100
VAE-CF	0.31857	0.47528
VAE-CF Modified	**0.36250**	0.47637
User-based kNN	0.33932	**0.50341**
Item-based kNN	0.34181	0.50310
SVD	IPR	IPR

Thank you!



References

da Costa, Arthur, Eduardo Fressato, Fernando Neto, Marcelo Manzato, Ricardo Campello. 2018. Case recommender: A flexible and extensible python framework for recommender systems. Proceedings of the 12th ACM Conference on Recommender Systems. RecSys '18, ACM, 494–495.

Liang, Dawen, Rahul G Krishnan, Matthew D Hoffman, Tony Jebara. 2018. Variational autoencoders for collaborative filtering. Proceedings of the 2018 World Wide Web Conference . 689–698.

Luo, Xin, Yunni Xia, Qingsheng Zhu. 2012. Incremental collaborative filtering recommender based on regularized matrix factorization. Knowledge-Based Systems 27 271–280.

Resnick, Paul, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, John Riedl. 1994. Grouplens: an open architecture for collaborative filtering of netnews. Proceedings of the 1994 ACM conference on Computer supported cooperative work. 175–186.

Sarwar, Badrul, George Karypis, Joseph Konstan, John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. Proceedings of the 10th international conference on World Wide Web. 285–295.

Vozalis, Manolis G, Konstantinos G Margaritis. 2007. Using svd and demographic data for the enhancement of generalized collaborative filtering. Information Sciences 177 3017–3037.

Zhang, Sheng, Weihong Wang, James Ford, Fillia Makedon, Justin Pearlman. 2005. Using singular value decomposition approximation for collaborative filtering. Seventh IEEE International Conference on E-Commerce Technology (CEC'05). IEEE, 257–264.