One-Class Collaborative Filtering with the Queryable Variational Autoencoder

UNIVERSITY OF TORONTO

Ga Wu, Mohamed Reda Bouadjenek, and Scott Sanner

VECTOR INSTITUTE | INSTITUT VECTEUR

Why do we need Variational Auto-ecnoder?

An Autoencoder based model for CF may be overly sensitive to individual user-item interactions, and hence, it may significantly change the latent representation of a user even with a single interaction update.

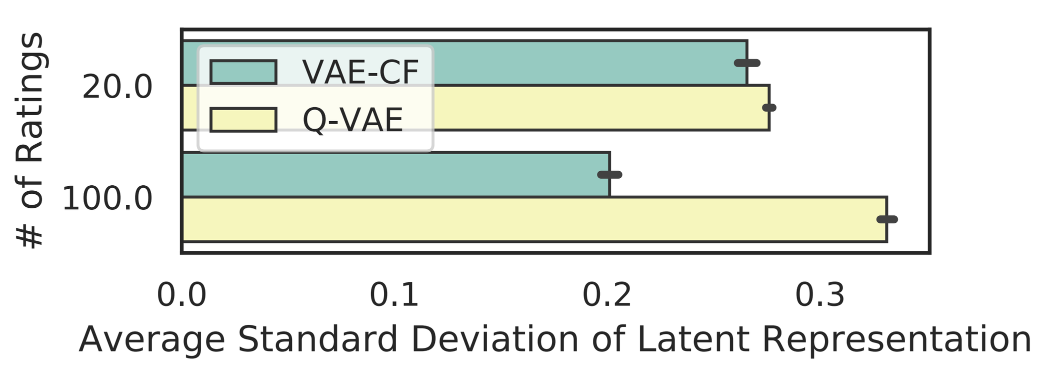
VAEs relax the latent distribution from a (deterministic) Delta function to a Gaussian distribution allowing for explicit representation of user and item uncertainty.

VAEs regularize the latent distribution through Kullback-Leibler (KL) divergence with a tractable standard Gaussian distribution leading to learning stability.

What is Wrong with Existing Works?

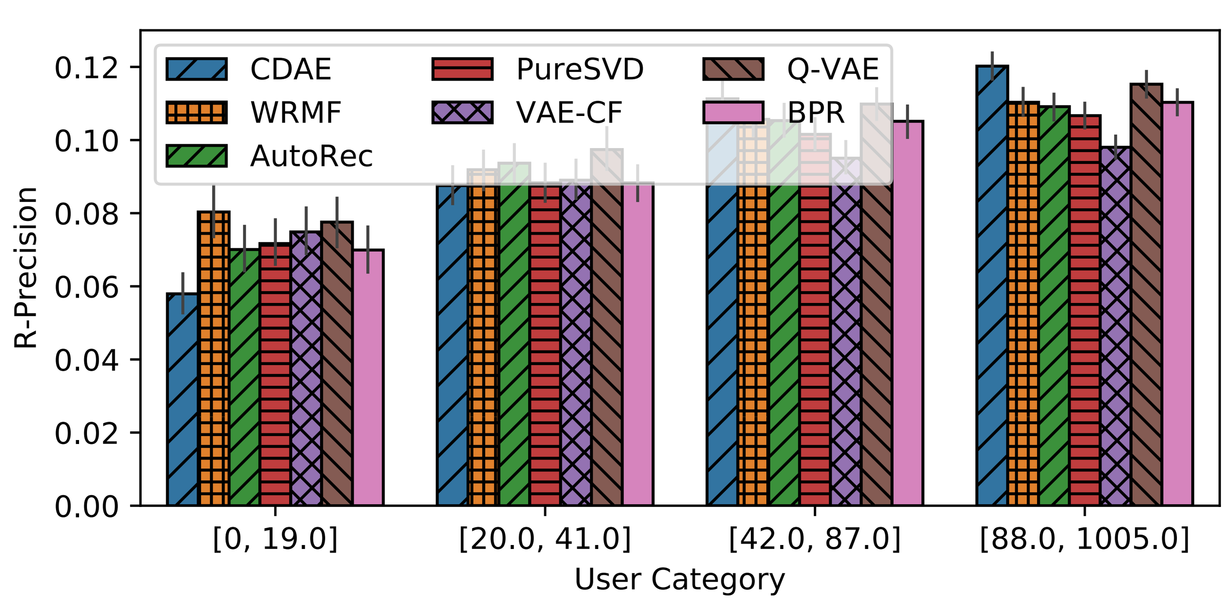
VAE-CF can exhibit suboptimal learning properties; e.g., VAE-CFs will increase their prediction confidence as they receive more preferences per user, even when those preferences may vary widely and create ambiguity in the user representation.

As an evidence, we show the average standard deviation of the diagonal Gaussian latent embeddings for VAE-CF and Q-VAE across 500 users. At the top, we first measure this embedding uncertainty after sampling 20 real interactions from each user's data and at the bottom we add in 80 random (fake) interactions. While Q-VAE increases its uncertainty, VAE-CF oddly becomes more certain in user preferences after observing this incoherent random data.



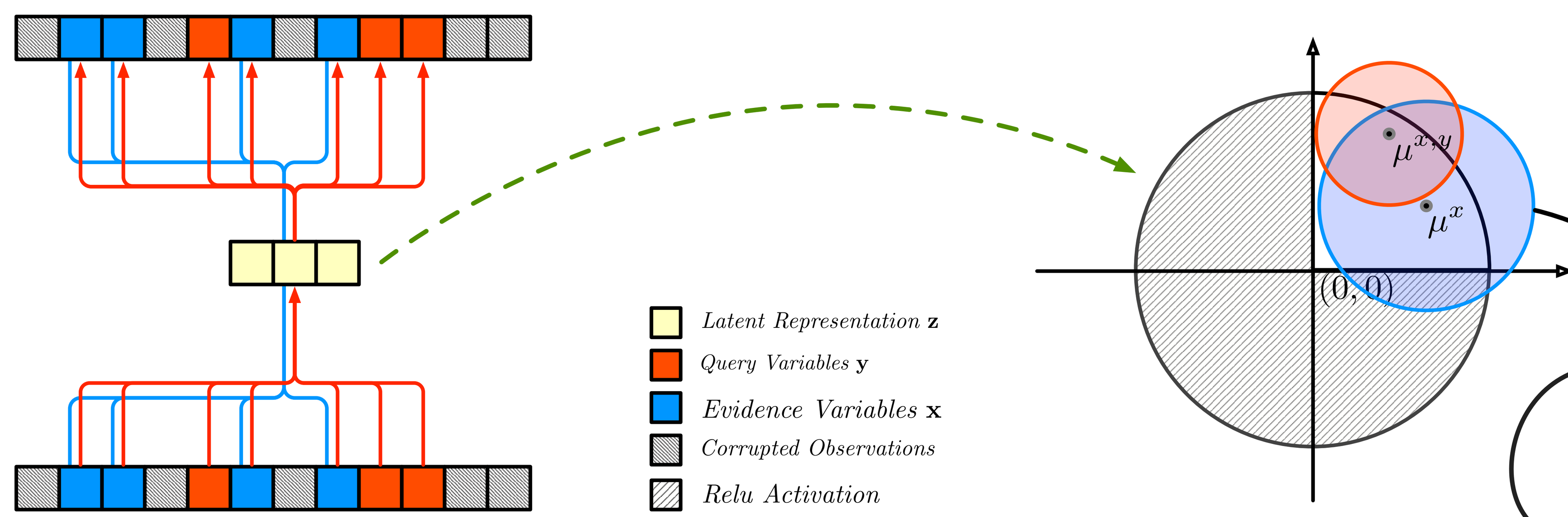
Stable Performance on Multiple User Categories

In comparison, Q-VAE shows relatively stable and good performance over all four user categories and significant prediction performance improvement over VAE-CF, especially with a large number of user interactions.



Support Arbitrary Conditional Inference with Queryable Variational Auto-encoder

Instead of fixing the split of evidence variables x and query variables y as in CVAE and BCDE, Q-VAE randomly splits variables during training through a dropout method



Two Kullback–Leibler divergences form "chain" Latent Regularization that corrects latent uncertainty estimation

Alternative Form of Variational Lower-bound

Instead of directly modeling the lower-bound of the log joint probability as other VAE-based recommender systems do, we propose to model the joint probability of any arbitrary partition of x and y as follows:

$$\log p(\mathbf{x}, \mathbf{y}) = \log p(\mathbf{x}) + \log p(\mathbf{y}|\mathbf{x})$$

Q-VAE optimizes two Variational Auto-encoder objective functions on a single VAE network structure:

$$\underbrace{E_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z})d\mathbf{z}] - KL[q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})]}_{\text{Variational Auto-encoder}} + \underbrace{E_{q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})}[\log p_\theta(\mathbf{y}|\mathbf{z})d\mathbf{z}] - KL[q_\phi(z|\mathbf{x},\mathbf{y})||q_\phi(\mathbf{z}|\mathbf{x})]}_{\text{Conditional Variational Auto-encoder}}$$

High Precision and High MAP Recommender

Table 1: Results of Movielens-1M dataset with 95% confidence interval. Hyper-parameters are chosen from the validation set. α: loss-weighting. λ: L2-regularization. ρ: corruption rate.

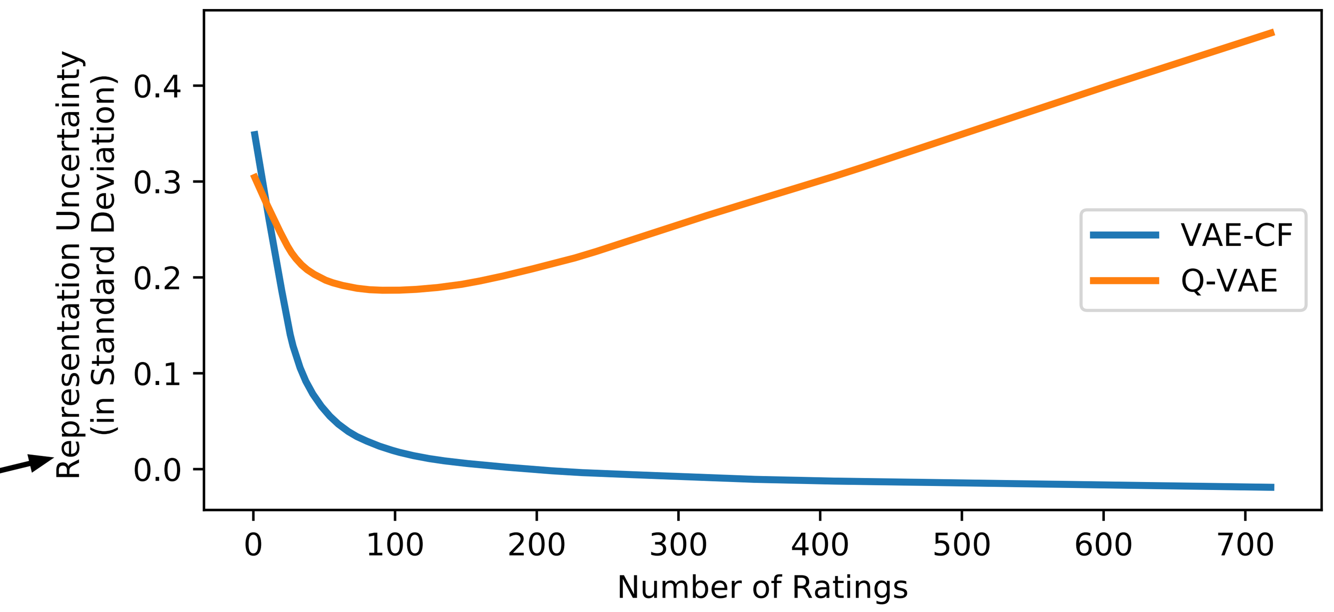
model	rank	α	λ	epochs	ρ	R-Precision	MAP@5	MAP@50	Precision@5	Precision@50	Recall@5	Recall@50
PureSVD	50	0	1	10	0	0.092±0.0024	0.1212±0.0052	0.0987±0.0024	0.116±0.0043	0.0852±0.0018	0.0383±0.002	0.2383±0.0052
BPR	200	0	1e-5	30	0	0.0933±0.0025	0.1192±0.0052	0.1002±0.0025	0.1141±0.0043	0.0875±0.0019	0.0375±0.002	0.2426±0.0052
WRMF	200	10	100	10	0	0.097±0.0026	0.1235±0.0053	0.1039±0.0025	0.1198±0.0045	0.091±0.002	**0.0411±0.0022**	**0.2668±0.0058**
CDAE	200	0	1e-5	300	0.5	0.0941±0.0025	0.1297±0.0056	0.1032±0.0028	0.1226±0.0047	0.0891±0.0021	0.035±0.0018	0.2177±0.0047
VAE-CF	200	0	1e-5	200	0.4	0.0892±0.0025	0.1066±0.0048	0.093±0.0022	0.1054±0.0039	0.0827±0.0017	0.0376±0.002	0.2449±0.0054
AutoRec	200	0	1e-5	300	0	0.0945±0.0025	0.1254±0.0054	0.1017±0.0026	0.1194±0.0045	0.0877±0.0019	0.0377±0.002	0.2398±0.0052
Q-VAE	200	0	0.1	200	0	**0.1±0.0026**	**0.1306±0.0055**	**0.1066±0.0026**	**0.125±0.0046**	**0.0917±0.0020**	0.0404±0.0021	0.2504±0.0054

Table 2: Results of Netflix dataset with 95% confidence interval. Hyper-parameters are chosen from the validation set.

model	rank	α	λ	epochs	ρ	R-Precision	MAP@5	MAP@50	Precision@5	Precision@50	Recall@5	Recall@50
PureSVD	50	0	1	10	0	0.0994±0.0003	0.159±0.0007	0.118±0.0003	0.146±0.0005	0.0953±0.0003	0.0445±0.0003	0.2188±0.0006
BPR	50	0	1e-5	30	0	0.0757±0.0002	0.1197±0.0006	0.096±0.0003	0.115±0.0005	0.0816±0.0002	0.0291±0.0002	0.1859±0.0006
WRMF	200	10	1e4	10	0	0.0985±0.0003	0.1531±0.0007	0.117±0.0003	0.1447±0.0006	0.096±0.0003	0.045±0.0003	**0.2325±0.0007**
CDAE	50	0	1e-5	300	0.2	0.0797±0.0003	0.1251±0.0006	0.0979±0.0003	0.1198±0.0005	0.0832±0.0002	0.0323±0.0002	0.1788±0.0006
VAE-CF	100	0	1e-4	300	0.5	**0.1017±0.0003**	0.1559±0.0007	0.1176±0.0003	0.1465±0.0005	0.0957±0.0003	**0.0467±0.0003**	0.2309±0.0006
AutoRec	50	0	1e-5	300	0	0.0876±0.0003	0.14±0.0006	0.1074±0.0003	0.1324±0.0005	0.0894±0.0003	0.0361±0.0002	0.1958±0.0006
Q-VAE	100	0	1e-5	200	0	0.0976±0.0003	**0.1593±0.0007**	**0.1194±0.0003**	**0.1488±0.0006**	**0.0972±0.0003**	0.0429±0.0003	0.2303±0.0006

Correct Handling of Uncertainty

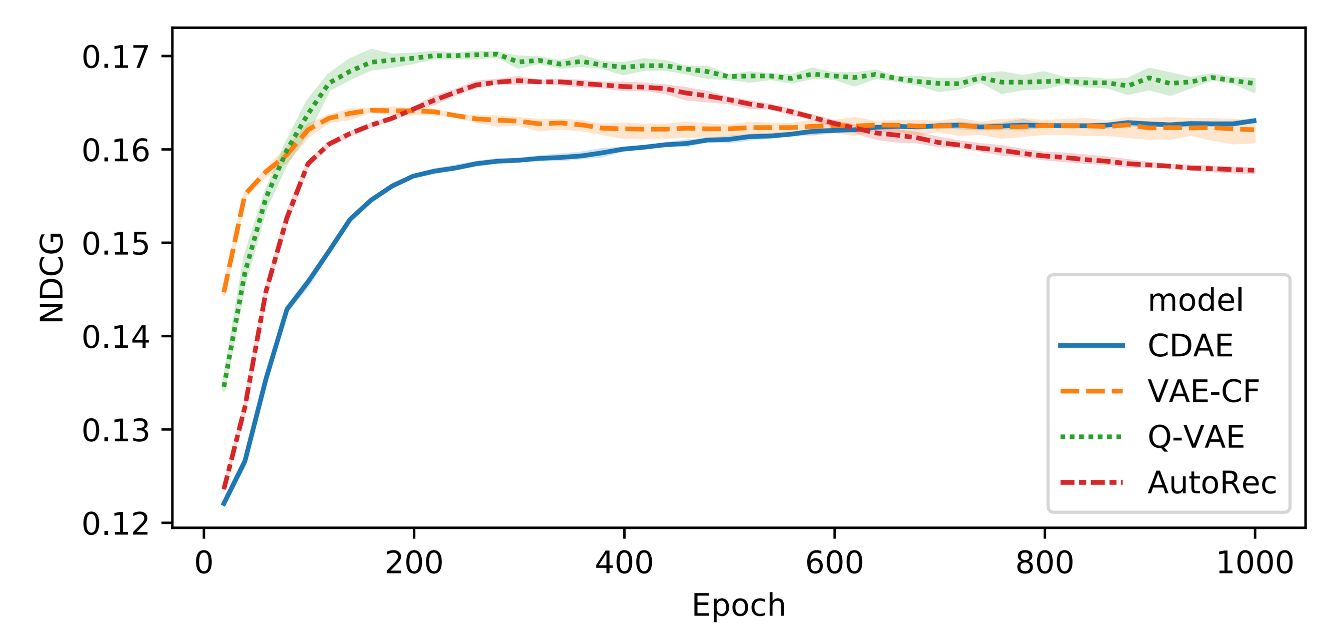
Both VAE-CF and Q-VAE explicitly model the user latent representation distributions. Hence, in this experiment, we analyze the latent representation uncertainty of users vs. their number of ratings.



Fast and Stable Convergence

VAE-based algorithms converge faster than the original Autoencoder approaches (which tend to overfit).

Q-VAE benefits from relatively fast and smooth convergence without overfitting due to the mutually structured regularization of its two objectives.



Conclusion

We proposed the Queryable Variational Auto-encoder (Q-VAE) as a way to explicitly condition recommendations in one-class collaborative filtering on observed user preferences to better model latent uncertainty of user preferences.

Our experiments show that the Q-VAE not only converges faster, but also outperforms several state-of-the-art Auto-encoder based recommendation models.

Also, we showed that Q-VAE avoids over-confidence with a large number of user preferences leading to strong recommendation performance across the user preference density spectrum.

42nd International ACM SIGIR Conference on Research and Development in Information Retrieval