

CSE 640 Project: Exploring AutoEncoders for Collaborative Filtering

Group 27

May 8, 2023

Papers Selected

1. AutoRec: Autoencoders Meet Collaborative Filtering [4]

This paper proposes a new CF model called AutoRec which is based on AutoEncoders. They experiment on two modes: U-AutoRec and I-AutoRec.

2. Training Deep AutoEncoders for Collaborative Filtering [2]

The authors propose a autoencoder based model called DeepAutoRec. They demonstrate that a 6 layered AutoRec generalises better than the shallow ones, use of regularization techniques is beneficial and propose a new training algorithm.

3. Variational Autoencoders for Collaborative Filtering [3]

Authors introduce a generative model with multinomial likelihood and use Bayesian inference for parameter estimation. They call the model VAE.

4. Enhancing VAEs for Collaborative Filtering: Flexible Priors & Gating Mechanisms [1]

Authors propose 4 new models. Vamp: Variational autoencoder with a VampPrior as the prior distribution instead of the original standard normal prior. H + Vamp: Hierarchical VAE with the VampPrior, the difference to the Vamp model is that it has hierarchical stochastic units to model the latent representation H + Vamp (Gated): Additional gating mechanisms are applied to the H + Vamp above. Multi-VAE (Gated): The Multi-VAE model with gating mechanisms

We have chosen these four papers because they all use common machine learning technique of autoencoders. Moreover, they are from core A star conferences. DeepAutorec uses Autorec model as baseline. Similarly, Enhanced VAE paper compares its model's performance with standard VAE of third paper.

Datasets

We ran all the models on Books Crossing and Anime Dataset

1. Booskcrossing

Consists of 272,679 interactions (explicit / implicit) from 2,946 users on 17,384 books. Ratings: 1,295 users and 14,684 books (62,657 ratings applied) History: 2,946 users and 17,384 books (272,679 accesses) Ratings are between 1 - 10. Implicit feedback are represented by 1. Simple demographic info for the users (age, gender, occupation, zip) is also available.

2. Anime

Consists of 520,610 interactions (play / purchase) from 5,000 users on 7,718 animes. History: 5,000 users and 7,390 games (520,610 interactions). Ratings: 4,714 users and 7,157 animes (419,944 interactions). Range of ratings: 1 - 10

Experiments

For each of the models, the genral flow of experiments was as follows- first of all hyperparameter tuning was carried out for a particular model. This was followed by a 5 fold cross validation to obtain best models obtained.

1. Autorec

Tables 1-4 contain the activation hyperparameter tuning results. Tables 5-8 depict the cross validation results on both the datasets for I-AutoRec and U-Autorec.

2. Deep autorec

Tables 9-16 contain the hyperparameter tuning results for layer depth,width,dropout. Tables 17 and 18 depict the cross validation results on both the datasets on the final model.

First	Last	MSE	MSE(clipped)	RMSE(clipped)
Linear	Linear	8.841	4.306	1.192
Linear	Sigmoid	22.0562	21.7444	3.000
Linear	Relu	29.2856	21.744	3.000
Sigmoid	Linear	5.1944	3.4324	1.043
Sigmoid	Sigmoid	22.3122	21.744	3.000
Sigmoid	Relu	5.08596	3.3999	1.0371
Relu	Linear	3.7644	3.007	0.949
Relu	Sigmoid	22.0098	21.7444	3.000
Relu	Relu	6.893	5.276	1.284

Table 1: Activation tuning, U-AutoRec on Books

First	Last	MSE	MSE(clipped)	RMSE(clipped)
Linear	Linear	3.9395	2.869	1.325
Linear	Sigmoid	41.604	41.541	5.641
Linear	Relu	54.128	41.541	5.641
Sigmoid	Linear	2.73198	1.893	1.086
Sigmoid	Sigmoid	41.699	41.541	5.641
Sigmoid	Relu	2.8502	1.930	1.095
Relu	Linear	2.3051	1.973	1.0764
Relu	Sigmoid	41.597	41.541	5.641
Relu	Relu	11.1702	8.801	2.224

Table 2: Activation tuning, U-AutoRec on Anime

3. VAE

Tables 19 through 23 contain the hyperparameter tuning results for bookscrossing for layer depth, initial learning rate, weight decay, anneal steps tuning, sampling layer size. Tables 24 depict the cross validation results on Bookscrossing of the final model. Tables 25 through 30 contain the hyperparameter tuning results for Anime dataset for layer depth, initial learning rate, weight decay, anneal steps tuning, dropout, and anneal cap. Tables 31 depict the cross validation results on Anime dataset of the final model.

- Enhanced VAE Tables 32 and 33 contain the hyperparameter tuning experients on the books dataset. Table 34 shows the final results for

First	Last	MSE	MSE(clipped)	RMSE(clipped)
Linear	Linear	1.831	1.210	0.756
Linear	Sigmoid	16.043	16.004	4.000
Linear	Relu	20.039	12.714	3.189
Sigmoid	Linear	0.0865	0.003	0.045
Sigmoid	Sigmoid	16.089	16.004	4.000
Sigmoid	Relu	0.0875	0.002	0.041
Relu	Linear	0.5574	0.363	0.450
Relu	Sigmoid	16.055	16.004	4.000
Relu	Relu	0.5655	0.393	0.511

Table 3: Activation tuning, I-AutoRec on Books

First	Last	MSE	MSE(clipped)	RMSE(clipped)
Linear	Linear	2.2416	1.760	0.817
Linear	Sigmoid	2.26888	1.765	0.814
Linear	Relu	3.07679	2.240	0.891
Sigmoid	Linear	2.269	1.753	0.8153
Sigmoid	Sigmoid	3.10365	2.253	0.900
Sigmoid	Relu	3.034585	2.224	0.886
Relu	Linear	3.04586	2.230	0.894
Relu	Sigmoid	3.07772	2.2449	0.889
Relu	Relu	3.0385	2.230	0.895

Table 4: Activation tuning, I-AutoRec on Anime

	MSE	MSE(clipped)	RMSE(clipped)
Fold 1	2.169	1.56	0.533
Fold 2	2.031	1.555	0.532
Fold 3	2.135	1.575	0.538
Fold 4	2.051	1.542	0.534
Fold 5	2.104	1.614	0.554
Average	2.098	1.57	0.538

Table 5: Cross validation for U Autorec for books dataset

	MSE	MSE(clipped)	RMSE(clipped)
Fold 1	1.601	0.671	0.324
Fold 2	1.587	0.667	0.326
Fold 3	1.511	0.658	0.322
Fold 4	1.662	0.663	0.322
Fold 5	1.421	0.652	0.32
Average	1.556	0.662	0.322

Table 6: Cross validation for U Autorec for books dataset

	MSE	MSE(clipped)	RMSE(clipped)
Fold 1	1.443	1.115	0.579
Fold 2	1.604	1.218	0.578
Fold 3	1.624	1.227	0.585
Fold 4	1.682	1.34	0.649
Fold 5	1.539	1.183	0.585
Average	1.578	1.216	0.595

Table 7: U Autorec 5 fold cross validation for Anime Dataset

	MSE	MSE(clipped)	RMSE(clipped)
Fold 1	2.284	1.437	0.809
Fold 2	2.16	1.363	0.787
Fold 3	2.29	1.445	0.818
Fold 4	2.325	1.393	0.803
Fold 5	2.223	1.415	0.805
Average	2.256	1.41	0.804

Table 8: I Autorec 5 fold cross validation for anime

Layers	MSE	MSE (clipped)	RMSE (clipped)
2	3464.480	36.551	1.269
4	2022.020	27.891	1.294
6	592.838	15.374	1.245
8	185.701	8.676	1.228
10	29.233	3.421	1.219
12	29.745	3.430	1.217

Table 9: Layer depth tuning (books)

Layers	MSE	MSE (clipped)	RMSE (clipped)
2	54.165	6.251	1.676
4	31.539	4.810	1.683
6	10.456	2.657	1.543
8	5.789	1.914	1.512
10	3.364	1.351	1.296
12	2.400	1.088	1.084

Table 10: Layer depth tuning (anime)

Width	MSE	MSE(clipped)	RMSE(clipped)
128	215006.515	294.851	1.233
256	149637.406	245.263	1.233
512	90557.5	189.858	1.232
1024	38243.683	121.65	1.2449

Table 11: Layer width tuning (books)

Width	MSE	MSE(clipped)	RMSE(clipped)
128	63.632	6.397	2.064
256	46.316	5.413	1.937
512	51.378	5.82	1.723
1024	41.228	5.092	2.516

Table 12: Layer width tuning (anime)

Dropout	MSE	MSE(clipped)	RMSE(clipped)
0	730.733	17.178	1.233
0.5	23.225	2.861	1.213
0.65	108.780	6.539	1.236
0.8	39.396	3.9545	1.223

Table 13: Dropout tuning (books)

Dropout	MSE	MSE(clipped)	RMSE(clipped)
0	5.526	1.811	1.436
0.5	3.025	1.252	1.222
0.65	2.744	1.187	1.182
0.8	2.840	1.200	1.178

Table 14: Dropout tuning (anime)

First	Last	MSE	MSE(clipped)	RMSE(clipped)
elu	linear	18856.894	86.333	1.243
elu	sigmoid	22.046	3.017	3.0001
relu	linear	572024.687	482.482	1.233
relu	sigmoid	22.051	3.017	3
selu	linear	2205.954	29.133	1.308
selu	sigmoid	22.039	3.017	3
tanh	linear	8.524	1.335	1.281
tanh	sigmoid	22.888	3.051	3
sigmoid	linear	8.558	1.336	1.282
Sigmoid	sigmoid	22.904	3.051	3
elu	selu	18295.125	86.082	1.233
relu	selu	555634.25	475.376	1.234
selu	selu	2735.957	32.943	1.24
tanh	selu	8.199	1.308	1.254
sigmoid	selu	8.239	1.31	1.256

Table 15: Deep autorec activation tuning (books)

First	Last	MSE	MSE(clipped)	RMSE(clipped)
elu	linear	69.383	7.193	1.693
elu	sigmoid	41.59	5.642	5.641
elu	selu	123.141	9.659	1.695
relu	linear	783.484	24.411	1.691
relu	sigmoid	41.59	5.642	5.641
relu	selu	632.26	21.92	1.69
selu	linear	28.821	4.558	1.673
selu	sigmoid	41.584	5.642	5.641
selu	selu	39.45	5.392	1.685
tanh	linear	3.478	1.15	1.1461
tanh	sigmoid	42.005	5.654	5.641
tanh	selu	3.37	1.144	1.14
sigmoid	linear	3.485	1.145	1.141
sigmoid	sigmoid	42.02	5.655	5.641
sigmoid	selu	3.366	1.137	1.133

Table 16: Deep autorec activation tuning (anime)

	MSE	MSE(clipped)	RMSE(clipped)	Baseline
Fold 1	4.481	0.447	0.412	0.533
Fold 2	4.182	0.446	0.41	0.532
Fold 3	4.345	0.462	0.422	0.538
Fold 4	4.2	0.449	0.414	0.534
Fold 5	4.48	0.463	0.425	0.554
Average	4.337	0.453	0.416	0.538

Table 17: Deep Autorec 5 fold cross validation on Books dataset

	MSE	MSE(clipped)	RMSE(clipped)	Baseline
Fold 1	1.833	0.417	0.415	0.579
Fold 2	2.355	0.424	0.423	0.578
Fold 3	3.079	0.417	0.415	0.585
Fold 4	3.316	0.416	0.414	0.649
Fold 5	2.948	0.423	0.421	0.585
Average	2.706	0.4194	0.417	0.595

Table 18: Deep Autorec 5 fold cross validation for Anime Dataset

Layers	NDCG@100	Recall@20	Recall@50
3	0.125	0.100	0.153
5	0.112	0.078	0.134
7	0.097	0.072	0.114
9	0.088	0.057	0.102

Table 19: VAE layer depth tuning (books)

Initial LR	NDCG@100	Recall@20	Recall@50
1e-4	0.116	0.083	0.14
1e-3	0.125	0.100	0.153
1e-2	0.097	0.072	0.108
1e-1	0.05	0.034	0.064

Table 20: VAE learning rate tuning (books)

Weight Decay	NDCG@100	Recall@20	Recall@50
0	0.125	0.100	0.153
1e-4	0.125	0.95	0.144
1e-3	0.125	0.094	0.143
1e-2	0.083	0.062	0.085
1e-1	0.056	0.036	0.063

Table 21: VAE weight decay tuning (books)

Steps	NDCG@100	Recall@20	Recall@50
400,000	0.125	0.095	0.143
200,000	0.125	0.100	0.153
20,000	0.127	0.104	0.154
2,000	0.127	0.094	0.152

Table 22: VAE anneal tuning (books)

Size	NDCG@100	Recall@20	Recall@50
150	0.131	0.092	0.145
300	0.143	0.110	0.160
600	0.142	0.107	0.162
1200	0.138	0.105	0.156
2400	0.135	0.110	0.155

Table 23: VAE sampling layer size tuning (books)

Fold	NDCG@100	NDCG@20	Recall@50	Recall@20
1	0.141 (0.010)	0.111 (0.010)	0.159 (0.012)	0.116 (0.010)
2	0.142 (0.009)	0.112 (0.009)	0.160 (0.012)	0.121 (0.010)
3	0.140 (0.010)	0.107 (0.009)	0.158 (0.012)	0.113 (0.010)
4	0.141 (0.010)	0.110 (0.009)	0.161 (0.012)	0.116 (0.010)
5	0.144 (0.010)	0.113 (0.009)	0.161 (0.012)	0.117 (0.010)

Table 24: 5 fold cross validation of VAE on books dataset

Layers	NDCG@100	Recall@20	Recall@50
2	0.454	0.429	0.518
3	0.397	0.322	0.478
4	0.405	0.369	0.473

Table 25: VAE number of layers tuning (anime)

Initial LR	NDCG@100	Recall@20	Recall@50
1e-4	0.292	0.259	0.35
1e-3	0.405	0.369	0.473
1e-2	0.295	0.247	0.378

Table 26: VAE learning rate tuning (anime)

Weight Decay	NDCG@100	Recall@20	Recall@50
1e-4	0.393	0.359	0.349
1e-3	0.392	0.358	0.348
1e-2	0.375	0.348	0.331
1e-1	0.366	0.327	0.311

Table 27: VAE weight decay tuning (anime)

Steps	NDCG@100	Recall@20	Recall@50
400,000	0.366	0.327	0.312
200,000	0.366	0.327	0.311
20,000	0.360	0.325	0.306
2,000	0.322	0.295	0.271

Table 28: VAE Anneal tuning (anime)

Dropout	NDCG@100	Recall@20	Recall@50
0	0.424	0.399	0.370
0.25	0.423	0.394	0.37
0.5	0.422	0.395	0.371
0.75	0.422	0.397	0.367

Table 29: VAE dropout tuning (anime)

Anneal Cap	NDCG@100	Recall@20	Recall@50
2e-1	0.422	0.395	0.371
4e-1	0.421	0.394	0.371
6e-1	0.421	0.394	0.371
8e-1	0.421	0.395	0.371

Table 30: VAE anneal cap (anime)

Fold	NDCG@100	NDCG@20	Recall@50	Recall@20
1	0.451	0.471	0.521	0.394
2	0.449	0.469	0.429	0.422
3	0.441	0.471	0.474	0.394
4	0.440	0.472	0.465	0.425
5	0.448	0.467	0.421	0.395

Table 31: 5 fold cross validation of VAE of anime dataset

Size	NDCG@100	Recall@50	Recall@20
100	0.08225	0.08954	0.0685
200	0.10807	0.12451	0.08452
300	0.11643	0.13008	0.08692
500	0.15861	0.17519	0.13047
1000	0.15792	0.17452	0.12987

Table 32: Enhanced VAE layer size tuning(books)

No. of Latent Features	NDCG@100	Recall@50	Recall@20
(100,100)	0.06934	0.07329	0.04483
(200,200)	0.06637	0.06919	0.04740
(300,300)	0.06342	0.06527	0.05031
(400,400)	0.06078	0.06251	0.05098

Table 33: EnhancedVAE number of latent features tuning(books)

books dataset. Tables 35 and 36 contain the hyperparameter tuning experiments on the anime dataset. Table 37 shows the final results for anime dataset.

Conclusion

Deep Autorec was able to beat the U-Autorec baseline and Enhanced VAE also showed improvements over the standard VAE model for both the datasets. While H+Vamp model (one of the enhanced VAE model) showed superior performance than standard VAE and other 3 enhanced models on both the datasets.

Models	NDCG@100	NDCG@20	Recall@50	Recall@20
Baseline (Multi-VAE)	0.14160	0.11060	0.15980	0.11660
Vamp	0.14824	0.11550	0.16368	0.12258
HVamp	0.14973	0.11612	0.16401	0.12462
Multi-VAE (Gated)	0.14789	0.11673	0.16947	0.12493
HVamp (Gated)	0.15861	0.12136	0.17519	0.13047

Table 34: Results of paper 4 on books dataset

Size	NDCG@100	Recall@50	Recall@20
100	0.456	0.478	0.408
300	0.466	0.489	0.42
500	0.459	0.481	0.411
750	0.449	0.472	0.404
1000	0.44	0.466	0.396

Table 35: EnhancedVAE layer size tuning(anime)

Layers	NDCG@100	Recall@50	Recall@20
1	0.466	0.489	0.42
2	0.461	0.485	0.417
3	0.454	0.474	0.407
4	0.442	0.463	0.399

Table 36: Enhanced VAE layer depth tuning(anime)

Models	NDCG@100	NDCG@20	NDCG@10	Recall@50	Recall@20
Baseline(Multi-VAE)	0.446	0.47	0.455	0.462	0.406
Multi-VAE (gated)	0.451	0.473	0.459	0.471	0.41
HVamp	0.455	0.479	0.465	0.48	0.414
Vamp (gated)	0.46	0.484	0.469	0.485	0.418
H+Vamp (gated)	0.466	0.491	0.473	0.489	0.42

Table 37: Results of paper 4 on anime dataset

References

- [1] KIM, D., AND SUH, B. Enhancing vaes for collaborative filtering: flexible priors & gating mechanisms. In *Proceedings of the 13th ACM Conference on Recommender Systems* (2019), pp. 403–407.
- [2] KUCHAIEV, O., AND GINSBURG, B. Training deep autoencoders for collaborative filtering. *arXiv preprint arXiv:1708.01715* (2017).
- [3] LIANG, D., KRISHNAN, R. G., HOFFMAN, M. D., AND JEBARA, T. Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 world wide web conference* (2018), pp. 689–698.
- [4] SEDHAIN, S., MENON, A. K., SANNER, S., AND XIE, L. Autorec: Autoencoders meet collaborative filtering. In *Proceedings of the 24th international conference on World Wide Web* (2015), pp. 111–112.