Additional Material for an Evaluation Study of Generative Adversarial Networks for Collaborative Filtering

Fernando B. Pérez Maurera[1,2] (✉)[0000-0001-6578-7404], Maurizio Ferrari Dacrema[1][0000-0001-7103-2788], and Paolo Cremonesi[1][0000-0002-1253-8081]

[1] Politecnico di Milano, Milan, Italy
{fernandobenjamin.perez,maurizio.ferrari,paolo.cremonesi}@polimi.it
[2] ContentWise, Milan, Italy
fernando.perez@contentwise.com

Abstract. This work provides complementary information for the work "An Evaluation Study of Generative Adversarial Networks for Collaborative Filtering". It contains a thorough formulation of Generative Adversarial Networks (GANs) and Conditional Generative Adversarial Networks (CGAN) which are the base for the Collaborative Filtering Adversarial Network (CFGAN) model studied in the main work. Also, this work presents the results of the experiments mentioned in the paper.

1 Generative Adversarial Networks

Generative Adversarial Networks (GANs) have been successfully applied to numerous prediction and classification tasks. In this work, we discuss a family of generative models originated from Collaborative Filtering GAN (CFGAN) used in Recommender Systems. Briefly, a GAN consists of an adversarial setting between two neural networks that are trained together until they reach convergence. The first neural network is called the *generator*, and it is denoted as G. The second network is called the *discriminator* and it is denoted as D [3,4,2].

We use an example to explain the goals of a GAN. Let us suppose that G is a counterfeit organization trying to produce fake bills and that D is the local police department in charge of distinguishing fake from real bills. Let us also suppose that after the classification of bills, both the police department and the counterfeiters can know if the classification was accurate or not. The first goal is that G learns to generate fake bills as realistically as possible to deceive D. The second goal is for D to keep up to date with the counterfeited bills not to enter the economy. On every iteration of this setup, G updates its counterfeiting methods by looking at the number of errors done by the police. D, on the other hand, updates its detection methods, so the identical bills are not misclassified again. This adversarial setting stops when G produces such realistic bills that the discriminator can not classify the source of the bills anymore.

1.1 Theoretical Formulation

Formally, the data is drawn from a distribution $p_{data}(\boldsymbol{x})$, and $\boldsymbol{z}$ is a vector drawn from a prior distribution $p_z(\boldsymbol{z})$. The generator G is defined as a differentiable function $G(\boldsymbol{z}, \theta_g)$ with parameters θ_g such that $\boldsymbol{x} = G(\boldsymbol{z}, \theta_g)$ meaning that it is a function that maps samples $\boldsymbol{z}$ drawn from $p_z(\boldsymbol{z})$ to values $\boldsymbol{x}$ drawn from a distribution $p_g(\boldsymbol{x})$. The learning objective for G is to learn a mapping such that $p_g(\boldsymbol{x}) = p_{data}(\boldsymbol{x})$, i.e., G learns to generate samples drawn from the same distribution as those of the real data.

On the other hand, the discriminator D is a function with parameters θ_d such that $y = D(\boldsymbol{x})$ is a scalar that represents the probability of $\boldsymbol{x}$ to be drawn from $p_{data}(\boldsymbol{x})$ instead of $p_g(\boldsymbol{x})$. The learning objective for D is to learn a mapping that assigns high probabilities to samples from $p_{data}(\boldsymbol{x})$ and low probabilities to samples from $p_g(\boldsymbol{x})$.

Both G and D are set up in an adversarial setting where the former tries to maximize the probability of D to label generated data as real, thus $\max_{\theta_g} D(G(\boldsymbol{z}, \theta_g))$ where $\boldsymbol{z} \sim p_z(\boldsymbol{z})$ and $G(\boldsymbol{z}, \theta_g) \sim p_g(\boldsymbol{x})$. The latter, instead, tries to maximize the probability to distinguish real from generated data, thus $\max_{\theta_d} D(\boldsymbol{x}, \theta_d)$ where $\boldsymbol{x} \sim p_{data}(\boldsymbol{x})$. In Figure 1 we illustrate this adversarial setting. Also in Equation 1 we show the objective function of a GAN [4].

$$\max_{\theta_g} \max_{\theta_d} \; V(G, D) = \mathbb{E}_{\boldsymbol{z} \sim p_z(\boldsymbol{z})} \log D(G(\boldsymbol{z}, \theta_g))) + \mathbb{E}_{\boldsymbol{x} \sim p_{data}(\boldsymbol{x})} \log D(\boldsymbol{x}) \tag{1}$$

In practice, GANs are usually trained to minimize a loss function l_{GAN}, using Stochastic Gradient Descent (SGD) while translating the expected values $\mathbb{E}$ cross-entropy losses [1,4,3].

1.2 Conditional GANs

The main drawback of a GAN is that there is no control over the generated samples, e.g., in [4] a GAN was trained to generate digits from the MNIST dataset. however, it did not control *which* digits were generated by it. Conditional GAN (CGAN) is an extension of the GAN model that solves this issue by including a *condition vector* to the generator and discriminator. This vector represents the features or attributes that the generated and discriminated sample must have [6]. For example, a work on the MNIST dataset showed that by providing the digit class, i.e., "0", "1", up until "9", as a condition to the generator and discriminator a CGAN generates samples of those digits [6]. On [5], a black and white image was used as the condition vector so the generator could generate a colorized version of that image.

The training procedure of CGANs is similar to that of GANs. The only difference is that both the generator and discriminator receive a new vector $\boldsymbol{c}$, the condition vector, as part of their input.

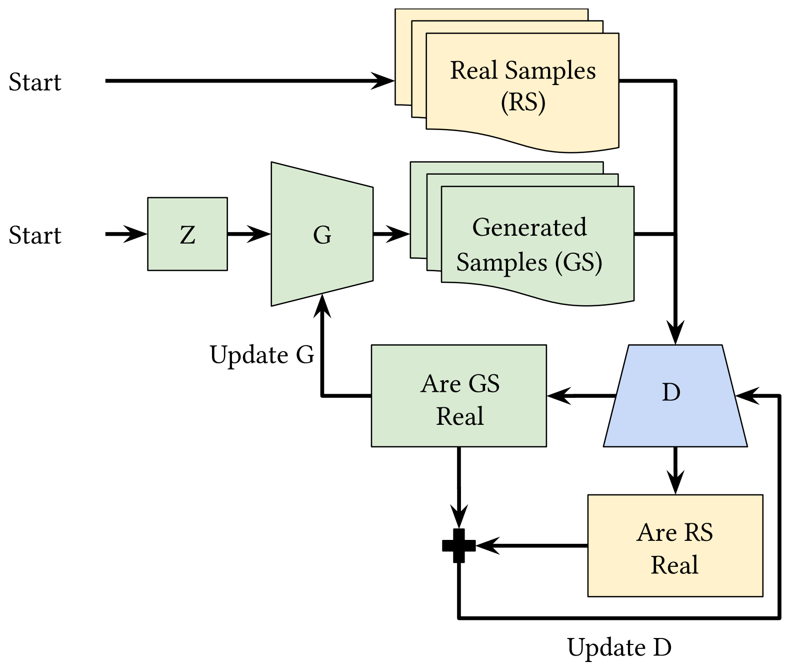

Fig. 1: Training process of a GAN. The generator G receives noise vector $\boldsymbol{z}$ as input and generates a fake sample GS. This sample is then fed to the discriminator D, which outputs the probability of this sample to be the real data (Are GS Real?). The generator is then updated based on this probability. Similarly, the discriminator received two samples: real (RS) and generated (GS), and outputs two probabilities from them being from the real data(Are GS Real? and Are RS Real?). These probabilities are used to update D.

2 Results RQ1: CFGAN Replicability & Numerical Stability

Table 1: Comparison between the accuracy metrics in the reference article [1] and those obtained in the replicability experiment (see Section 5.1 of the paper) at recommendation list length of 5. Statistics calculated over 30 executions, evaluating on the last epoch using recommendation lists of length 20. We consistently obtain *lower* results across two of the three datasets on average. For the Ciao dataset, the original source code trains a different variant (in bold) than the reported in the reference article.

Dataset	Variant	Stats	PREC	REC	MRR	NDCG
Ciao	**iZR**	Mean $\pm$ Std	0.0607 ± 0.0026	0.0693 ± 0.0053	0.1345 ± 0.0059	0.0795 ± 0.0039
	iZP	Reported [1]	0.0720	0.0810	0.1540	0.0920
ML100K	iZP	Mean $\pm$ Std	0.4244 ± 0.0064	0.1434 ± 0.0034	0.6602 ± 0.0081	0.4588 ± 0.0062
	iZP	Reported [1]	0.4440	0.1520	0.6810	0.4760
ML1M	iZP	Mean $\pm$ Std	0.4287 ± 0.0013	0.1060 ± 0.0008	0.6424 ± 0.0025	0.4521 ± 0.0013
	iZP	Reported [1]	0.4320	0.1080	0.6470	0.4550

Table 2: Comparison between the accuracy metrics in the reference article [1] and those obtained in the replicability experiment (see Section 5.1 of the paper) at recommendation list length of 20. Statistics calculated over 30 executions, evaluating on the last epoch using recommendation lists of length 20. We consistently obtain *lower* results across two of the three datasets on average. For the Ciao dataset, the original source code trains a different variant (in bold) than the reported in the reference article.

Dataset	Variant	Stats	PREC	REC	MRR	NDCG
Ciao	**iZR**	Mean $\pm$ Std	0.0402 ± 0.0014	0.1788 ± 0.0071	0.1594 ± 0.0055	0.1135 ± 0.0038
	iZP	Reported [1]	0.0450	0.1940	0.1670	0.1240
ML100K	iZP	Mean $\pm$ Std	0.2851 ± 0.0025	0.3400 ± 0.0050	0.6732 ± 0.0077	0.4207 ± 0.0048
	iZP	Reported [1]	0.2940	0.3600	0.6930	0.4330
ML1M	iZP	Mean $\pm$ Std	0.3079 ± 0.0011	0.2671 ± 0.0020	0.6566 ± 0.0024	0.4035 ± 0.0016
	iZP	Reported [1]	0.3090	0.2720	0.6600	0.4060

3 Results RQ2: Reproducibility Evaluation Against Properly-Tuned Baselines

Table 3: Accuracy and beyond-accuracy metrics for tuned baselines and CFGAN on the Ciao dataset at recommendation list length of 20. CFGAN results are different than Table 2 due to the hyper-parameter tuning.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
Random	0.0045	0.0164	0.0028	0.0124	0.0092	0.0070	0.0137	0.1603	0.9852	1.0000	0.8229	10.3215
TopPop	0.0232	0.1091	0.0282	0.0829	0.0645	0.0383	0.1021	0.1199	0.1312	0.0356	0.0170	4.6100
UserKNN CF cosine	0.0416	0.1874	0.0574	0.1650	0.1201	0.0680	0.2087	0.1325	0.8451	0.5078	0.1269	7.7617
UserKNN CF dice	0.0383	0.1747	0.0515	0.1531	0.1102	0.0628	0.1912	0.1317	0.8157	0.4246	0.1030	7.4729
UserKNN CF jaccard	0.0388	0.1786	0.0516	0.1527	0.1112	0.0637	0.1914	0.1308	0.7905	0.3764	0.0884	7.2620
UserKNN CF asymmetric	0.0376	0.1689	0.0489	0.1477	0.1062	0.0615	0.1848	0.1307	0.7818	0.3445	0.0836	7.1874
UserKNN CF tversky	0.0413	0.1850	0.0578	0.1621	0.1193	0.0676	0.2079	0.1323	0.8388	0.5241	0.1274	7.7472
ItemKNN CF cosine	0.0410	0.1807	0.0555	0.1589	0.1165	0.0669	0.2022	0.1313	0.8116	0.4566	0.1063	7.4955
ItemKNN CF dice	0.0387	0.1709	0.0512	0.1495	0.1090	0.0631	0.1898	0.1357	0.8820	0.6585	0.1880	8.2791
ItemKNN CF jaccard	0.0371	0.1611	0.0518	0.1499	0.1069	0.0603	0.1894	0.1356	0.8800	0.6548	0.1858	8.2591
ItemKNN CF asymmetric	0.0409	0.1792	0.0548	0.1587	0.1157	0.0665	0.2012	0.1312	0.8120	0.4410	0.1050	7.4850
ItemKNN CF tversky	0.0375	0.1596	0.0510	0.1511	0.1065	0.0607	0.1899	0.1354	0.8794	0.6578	0.1856	8.2554
RP3beta	0.0442	0.1971	0.0601	0.1730	0.1262	0.0722	0.2194	0.1355	0.8688	0.7060	0.1868	8.2145
PureSVD	0.0366	0.1596	0.0443	0.1428	0.0999	0.0595	0.1784	0.1332	0.8561	0.3875	0.1165	7.7012
SLIM ElasticNet	0.0430	0.1852	0.0582	0.1640	0.1208	0.0698	0.2114	0.1343	0.8758	0.5798	0.1628	8.1157
MF BPR	0.0246	0.1013	0.0323	0.1052	0.0683	0.0396	0.1271	0.1406	0.9320	0.8641	0.3290	9.0762
EASE R	0.0425	0.1825	0.0560	0.1591	0.1176	0.0689	0.2059	0.1343	0.8799	0.5479	0.1596	8.1095
CFGAN iZR	0.0414	0.1783	0.0547	0.1586	0.1149	0.0672	0.2033	0.1409	0.9510	0.6726	0.2955	9.0534
CFGAN iPM	0.0353	0.1548	0.0451	0.1317	0.0976	0.0575	0.1659	0.1391	0.9012	0.5679	0.1844	8.3364
CFGAN iZP	0.0361	0.1635	0.0503	0.1418	0.1041	0.0591	0.1772	0.1383	0.9305	0.5880	0.2209	8.6379
CFGAN uZR	0.0401	0.1719	0.0523	0.1625	0.1124	0.0650	0.2056	0.1312	0.8208	0.2079	0.0801	7.1658
CFGAN uPM	0.0273	0.1116	0.0333	0.1049	0.0732	0.0439	0.1298	0.1306	0.6338	0.0913	0.0377	6.0393
CFGAN uZP	0.0230	0.1054	0.0228	0.0664	0.0579	0.0378	0.0843	0.1204	0.1238	0.0341	0.0169	4.5934

Table 4: Accuracy and beyond-accuracy metrics for tuned baselines and CFGAN on the ML100K dataset at recommendation list length of 20. CFGAN results are different than Table 2 due to the hyper-parameter tuning.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
Random	0.0153	0.0141	0.0044	0.0543	0.0185	0.0147	0.0600	0.1476	0.9878	1.0000	0.8303	10.6237
TopPop	0.1560	0.1675	0.0971	0.4130	0.2132	0.1616	0.7151	0.0972	0.4526	0.0502	0.0207	5.3840
UserKNN CF cosine	0.2641	0.3239	0.2292	0.6581	0.3962	0.2910	1.3349	0.1031	0.8365	0.2432	0.0730	7.3278
UserKNN CF dice	0.2629	0.3218	0.2282	0.6514	0.3943	0.2894	1.3248	0.1029	0.8367	0.2335	0.0727	7.3199
UserKNN CF jaccard	0.2627	0.3227	0.2281	0.6545	0.3946	0.2896	1.3268	0.1029	0.8369	0.2353	0.0729	7.3238
UserKNN CF asymmetric	0.2582	0.3204	0.2220	0.6550	0.3891	0.2859	1.3059	0.1026	0.8237	0.2226	0.0663	7.1813
UserKNN CF tversky	0.2683	0.3257	0.2350	0.6609	0.4015	0.2942	1.3519	0.1035	0.8515	0.2662	0.0819	7.4997
ItemKNN CF cosine	0.2591	0.3189	0.2263	0.6469	0.3901	0.2859	1.3043	0.1031	0.8455	0.1984	0.0724	7.3031
ItemKNN CF dice	0.2454	0.3029	0.2124	0.6317	0.3719	0.2711	1.2497	0.1021	0.8165	0.1712	0.0613	7.0631
ItemKNN CF jaccard	0.2401	0.2943	0.2077	0.6334	0.3652	0.2644	1.2348	0.1014	0.7882	0.1525	0.0536	6.8701
ItemKNN CF asymmetric	0.2652	0.3266	0.2337	0.6496	0.3978	0.2927	1.3185	0.1062	0.8935	0.3479	0.1113	7.9516
ItemKNN CF tversky	0.2779	0.3372	0.2455	0.6615	0.4121	0.3047	1.3759	0.1036	0.8598	0.2329	0.0833	7.5266
RP3beta	0.2603	0.3204	0.2286	0.6566	0.3928	0.2872	1.3184	0.1022	0.8174	0.1887	0.0629	7.1103
PureSVD	0.2863	0.3451	0.2543	0.6667	0.4225	0.3130	1.4022	0.1053	0.8885	0.3037	0.1079	7.9044
SLIM ElasticNet	0.2915	0.3563	0.2683	0.6952	0.4375	0.3207	1.4585	0.1037	0.8597	0.2595	0.0860	7.5762
MF BPR	0.2263	0.2829	0.1746	0.5620	0.3296	0.2515	1.0800	0.1017	0.7774	0.1682	0.0524	6.8514
EASE R	0.2929	0.3530	0.2688	0.6909	0.4368	0.3202	1.4630	0.1042	0.8713	0.2783	0.0938	7.7031
CFGAN iZR	0.2415	0.3025	0.1923	0.5866	0.3546	0.2685	1.1513	0.1073	0.9017	0.3400	0.1235	8.0988
CFGAN iPM	0.2171	0.2350	0.1641	0.5247	0.3023	0.2257	1.0393	0.1037	0.8417	0.2390	0.0731	7.3341
CFGAN iZP	0.2757	0.3283	0.2394	0.6580	0.4041	0.2997	1.3595	0.1040	0.8572	0.2589	0.0848	7.5584
CFGAN uZR	0.2754	0.3387	0.2409	0.6500	0.4078	0.3038	1.3454	0.1060	0.8844	0.2371	0.1018	7.8011
CFGAN uPM	0.2326	0.3057	0.1864	0.5809	0.3498	0.2642	1.1201	0.1058	0.8699	0.2607	0.0951	7.7164
CFGAN uZP	0.2555	0.3163	0.2194	0.6399	0.3834	0.2827	1.2712	0.1055	0.8855	0.2529	0.1055	7.8565

Table 5: Accuracy and beyond-accuracy metrics for tuned baselines and CFGAN on the ML1M dataset at recommendation list length of 20. CFGAN results are different than Table 2 due to the hyper-parameter tuning.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
Random	0.0099	0.0056	0.0024	0.0326	0.0108	0.0072	0.0364	0.0732	0.9946	1.0000	0.8977	11.8223
TopPop	0.1552	0.1146	0.0917	0.3852	0.1938	0.1319	0.7065	0.0473	0.4529	0.0299	0.0095	5.4298
UserKNN CF cosine	0.2883	0.2525	0.2265	0.6505	0.3855	0.2693	1.3785	0.0520	0.9110	0.4329	0.0828	8.6076
UserKNN CF dice	0.2877	0.2551	0.2252	0.6556	0.3863	0.2704	1.3802	0.0512	0.8890	0.3194	0.0644	8.2557
UserKNN CF jaccard	0.2893	0.2565	0.2273	0.6571	0.3884	0.2719	1.3866	0.0514	0.8971	0.3528	0.0702	8.3766
UserKNN CF asymmetric	0.2891	0.2570	0.2273	0.6595	0.3888	0.2721	1.3894	0.0513	0.8921	0.3286	0.0655	8.2848
UserKNN CF tversky	0.2890	0.2561	0.2272	0.6569	0.3881	0.2715	1.3850	0.0513	0.8948	0.3408	0.0683	8.3392
ItemKNN CF cosine	0.2785	0.2392	0.2138	0.6279	0.3688	0.2574	1.3205	0.0515	0.9018	0.3197	0.0687	8.3666
ItemKNN CF dice	0.2566	0.2110	0.1913	0.5940	0.3376	0.2316	1.2253	0.0503	0.8519	0.2531	0.0455	7.7630
ItemKNN CF jaccard	0.2556	0.2108	0.1908	0.5925	0.3369	0.2310	1.2226	0.0503	0.8453	0.2596	0.0445	7.7211
ItemKNN CF asymmetric	0.2600	0.2196	0.1985	0.6254	0.3490	0.2381	1.2744	0.0497	0.8148	0.2097	0.0362	7.4341
ItemKNN CF tversky	0.2657	0.2180	0.2005	0.6147	0.3496	0.2395	1.2710	0.0513	0.9005	0.1814	0.0560	8.1173
RP3beta	0.2758	0.2385	0.2146	0.6425	0.3700	0.2558	1.3346	0.0506	0.8565	0.3427	0.0528	7.9254
PureSVD	0.2913	0.2421	0.2234	0.6333	0.3783	0.2644	1.3555	0.0516	0.9142	0.2439	0.0712	8.4463
SLIM ElasticNet	0.3119	0.2695	0.2508	0.6724	0.4123	0.2892	1.4658	0.0514	0.8984	0.3153	0.0696	8.3690
MF BPR	0.2485	0.2103	0.1762	0.5753	0.3242	0.2278	1.1594	0.0512	0.8855	0.3126	0.0631	8.2195
EASE R	0.3171	0.2763	0.2560	0.6795	0.4192	0.2953	1.4853	0.0518	0.9146	0.3338	0.0803	8.5897
CFGAN iZR	0.2862	0.2547	0.2146	0.6312	0.3770	0.2696	1.3288	0.0542	0.9583	0.4123	0.1459	9.4737
CFGAN iPM	0.2505	0.1950	0.1734	0.5454	0.3138	0.2193	1.1354	0.0523	0.9218	0.3669	0.0901	8.7458
CFGAN iZP	0.2407	0.1742	0.1661	0.5230	0.2972	0.2021	1.0929	0.0530	0.9256	0.4894	0.0901	8.7580
CFGAN uZR	0.2955	0.2473	0.2241	0.6222	0.3799	0.2692	1.3541	0.0523	0.9205	0.2167	0.0837	8.6304
CFGAN uPM	0.2367	0.1928	0.1629	0.5513	0.3054	0.2125	1.1064	0.0516	0.8962	0.1782	0.0550	8.0858
CFGAN uZP	0.2764	0.2342	0.2074	0.6208	0.3620	0.2536	1.3010	0.0513	0.9062	0.1833	0.0617	8.2408

4 Results RQ3: Impact of Theoretical and Methodological Concerns

Table 6: Accuracy and beyond-accuracy values for different item-based CFGAN models for the Ciao dataset at recommendation list length of 5. The suffix RN-X means that the model uses random noise of size X. The suffix Class indicates that the model uses the user/item class as the condition vector. The suffix NO-ES indicates that the model does not use early-stopping. Hyper-parameter sets of variants are the same as those in Table 3.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
CFGAN iZR NO-ES	0.0594	0.0679	0.0489	0.1370	0.0791	0.0633	0.1549	0.0351	0.9858	0.4870	0.2452	8.7517
CFGAN iZR CC	0.0078	0.0078	0.0038	0.0146	0.0081	0.0078	0.0150	0.0354	0.1065	0.0119	0.0047	2.6102
CFGAN iZR RN-331	0.0591	0.0686	0.0492	0.1245	0.0775	0.0635	0.1432	0.0343	0.9795	0.4061	0.1797	8.3154
CFGAN iZR RN-662	0.0584	0.0659	0.0427	0.1131	0.0710	0.0620	0.1331	0.0342	0.9772	0.3935	0.1706	8.2311
CFGAN iZR RN-1324	0.0566	0.0638	0.0458	0.1234	0.0735	0.0599	0.1406	0.0344	0.9784	0.4551	0.1933	8.3853
CFGAN iPM NO-ES	0.0516	0.0615	0.0395	0.1068	0.0650	0.0561	0.1221	0.0331	0.9554	0.3370	0.1233	7.6628
CFGAN iPM CC	0.0103	0.0116	0.0082	0.0261	0.0139	0.0109	0.0269	0.0362	0.9773	0.5754	0.2512	8.6491
CFGAN iPM RN-331	0.0537	0.0600	0.0435	0.1206	0.0703	0.0567	0.1361	0.0336	0.9509	0.2762	0.0965	7.3690
CFGAN iPM RN-662	0.0537	0.0614	0.0407	0.1095	0.0670	0.0573	0.1277	0.0334	0.9460	0.2806	0.0954	7.3164
CFGAN iPM RN-1324	0.0481	0.0519	0.0380	0.1114	0.0627	0.0500	0.1243	0.0336	0.9500	0.3356	0.1103	7.4920
CFGAN iZP NO-ES	0.0434	0.0454	0.0337	0.1002	0.0558	0.0444	0.1123	0.0345	0.9771	0.3875	0.1723	8.2420
CFGAN iZP CC	0.0047	0.0045	0.0025	0.0108	0.0052	0.0046	0.0108	0.0371	0.9534	0.3051	0.1088	7.5188
CFGAN iZP RN-331	0.0506	0.0632	0.0414	0.1115	0.0675	0.0562	0.1245	0.0338	0.9730	0.3430	0.1435	7.9888
CFGAN iZP RN-662	0.0491	0.0524	0.0389	0.1173	0.0643	0.0507	0.1320	0.0336	0.9670	0.3118	0.1213	7.7437
CFGAN iZP RN-1324	0.0491	0.0594	0.0440	0.1212	0.0693	0.0537	0.1342	0.0338	0.9709	0.3341	0.1367	7.9143

Table 7: Accuracy and beyond-accuracy values for different user-based CFGAN models for the Ciao dataset at recommendation list length of 5. The suffix RN-X means that the model uses random noise of size X. The suffix Class indicates that the model uses the user/item class as the condition vector. The suffix NO-ES indicates that the model does not use early-stopping. Hyper-parameter sets of variants are the same as those in Table 3.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
CFGAN uZR NO-ES	0.0462	0.0502	0.0323	0.0935	0.0555	0.0481	0.1074	0.0351	0.9749	0.3177	0.1449	8.0094
CFGAN uZR CC	0.0294	0.0331	0.0234	0.0746	0.0392	0.0311	0.0809	0.0293	0.1563	0.0097	0.0049	2.7006
CFGAN uZR RN-673	0.0625	0.0662	0.0474	0.1254	0.0768	0.0643	0.1484	0.0319	0.9002	0.1039	0.0374	6.0583
CFGAN uZR RN-1347	0.0541	0.0572	0.0408	0.1180	0.0675	0.0556	0.1354	0.0315	0.8742	0.0787	0.0284	5.6410
CFGAN uZR RN-2694	0.0531	0.0562	0.0427	0.1186	0.0685	0.0546	0.1370	0.0319	0.8972	0.0935	0.0351	5.9642
CFGAN uPM NO-ES	0.0394	0.0444	0.0273	0.0902	0.0492	0.0417	0.0959	0.0318	0.7946	0.0416	0.0182	4.9494
CFGAN uPM CC	0.0091	0.0057	0.0052	0.0161	0.0096	0.0070	0.0188	0.0338	0.0711	0.0104	0.0046	2.5305
CFGAN uPM RN-673	0.0312	0.0330	0.0251	0.0723	0.0403	0.0321	0.0822	0.0322	0.7183	0.0319	0.0139	4.5567
CFGAN uPM RN-1347	0.0322	0.0347	0.0217	0.0717	0.0388	0.0334	0.0767	0.0316	0.6906	0.0208	0.0115	4.1955
CFGAN uPM RN-2694	0.0375	0.0427	0.0265	0.0770	0.0457	0.0399	0.0862	0.0319	0.6621	0.0223	0.0106	4.0856
CFGAN uZP NO-ES	0.0331	0.0392	0.0226	0.0693	0.0402	0.0359	0.0756	0.0288	0.1684	0.0104	0.0050	2.7277
CFGAN uZP CC	0.0250	0.0275	0.0200	0.0637	0.0334	0.0262	0.0685	0.0292	0.1586	0.0097	0.0049	2.6959
CFGAN uZP RN-673	0.0300	0.0345	0.0226	0.0674	0.0381	0.0321	0.0744	0.0289	0.1662	0.0111	0.0050	2.7300
CFGAN uZP RN-1347	0.0325	0.0381	0.0258	0.0747	0.0424	0.0351	0.0824	0.0288	0.1705	0.0097	0.0050	2.7425
CFGAN uZP RN-2694	0.0312	0.0361	0.0224	0.0690	0.0389	0.0335	0.0755	0.0291	0.1596	0.0089	0.0049	2.7153

Table 8: Accuracy and beyond-accuracy values for different item-based CFGAN models for the Ciao dataset at recommendation list length of 10. The suffix RN-X means that the model uses random noise of size X. The suffix Class indicates that the model uses the user/item class as the condition vector. The suffix NO-ES indicates that the model does not use early-stopping. Hyper-parameter sets of variants are the same as those in Table 4.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
CFGAN iZR NO-ES	0.0458	0.1023	0.0471	0.1482	0.0871	0.0633	0.1759	0.0709	0.9773	0.6570	0.3165	9.1324
CFGAN iZR CC	0.0070	0.0139	0.0039	0.0186	0.0103	0.0093	0.0190	0.0736	0.0674	0.0230	0.0086	3.5335
CFGAN iZR RN-331	0.0506	0.1147	0.0500	0.1406	0.0922	0.0703	0.1715	0.0695	0.9681	0.5501	0.2338	8.7084
CFGAN iZR RN-662	0.0494	0.1115	0.0426	0.1288	0.0847	0.0685	0.1594	0.0694	0.9651	0.5679	0.2257	8.6499
CFGAN iZR RN-1324	0.0469	0.1041	0.0458	0.1373	0.0853	0.0646	0.1651	0.0697	0.9669	0.6288	0.2488	8.7747
CFGAN iPM NO-ES	0.0442	0.1060	0.0411	0.1235	0.0802	0.0624	0.1467	0.0672	0.9366	0.4818	0.1615	8.0937
CFGAN iPM CC	0.0105	0.0207	0.0088	0.0320	0.0179	0.0139	0.0338	0.0736	0.9706	0.7958	0.3485	9.1739
CFGAN iPM RN-331	0.0447	0.0934	0.0426	0.1322	0.0797	0.0604	0.1589	0.0682	0.9287	0.4053	0.1337	7.8476
CFGAN iPM RN-662	0.0447	0.1019	0.0408	0.1235	0.0792	0.0621	0.1503	0.0679	0.9257	0.4172	0.1333	7.8268
CFGAN iPM RN-1324	0.0406	0.0860	0.0384	0.1236	0.0727	0.0552	0.1464	0.0682	0.9312	0.4855	0.1554	8.0156
CFGAN iZP NO-ES	0.0398	0.0807	0.0348	0.1154	0.0681	0.0533	0.1363	0.0694	0.9632	0.5019	0.2095	8.5461
CFGAN iZP CC	0.0045	0.0101	0.0028	0.0134	0.0072	0.0063	0.0134	0.0741	0.9366	0.4447	0.1480	8.0022
CFGAN iZP RN-331	0.0408	0.0973	0.0416	0.1254	0.0782	0.0575	0.1452	0.0682	0.9555	0.4751	0.1807	8.3285
CFGAN iZP RN-662	0.0428	0.0903	0.0388	0.1313	0.0758	0.0581	0.1563	0.0681	0.9508	0.4529	0.1643	8.1924
CFGAN iZP RN-1324	0.0433	0.1047	0.0453	0.1371	0.0840	0.0612	0.1590	0.0684	0.9547	0.4581	0.1747	8.2850

Table 9: Accuracy and beyond-accuracy values for different user-based CFGAN models for the Ciao dataset at recommendation list length of 10. The suffix RN-X means that the model uses random noise of size X. The suffix Class indicates that the model uses the user/item class as the condition vector. The suffix NO-ES indicates that the model does not use early-stopping. Hyper-parameter sets of variants are the same as those in Table 4.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
CFGAN uZR NO-ES	0.0403	0.0857	0.0330	0.1071	0.0670	0.0548	0.1300	0.0708	0.9611	0.4209	0.1882	8.3950
CFGAN uZR CC	0.0252	0.0567	0.0248	0.0824	0.0474	0.0348	0.0945	0.0598	0.1385	0.0163	0.0089	3.6257
CFGAN uZR RN-673	0.0520	0.1112	0.0472	0.1410	0.0900	0.0709	0.1758	0.0646	0.8645	0.1500	0.0547	6.6173
CFGAN uZR RN-1347	0.0477	0.1031	0.0427	0.1336	0.0828	0.0652	0.1630	0.0637	0.8328	0.1218	0.0432	6.2696
CFGAN uZR RN-2694	0.0481	0.1041	0.0425	0.1366	0.0833	0.0658	0.1650	0.0644	0.8597	0.1359	0.0514	6.5233
CFGAN uPM NO-ES	0.0339	0.0777	0.0284	0.1002	0.0594	0.0472	0.1136	0.0645	0.7647	0.0646	0.0298	5.6721
CFGAN uPM CC	0.0095	0.0128	0.0051	0.0204	0.0121	0.0109	0.0251	0.0680	0.0672	0.0163	0.0085	3.5044
CFGAN uPM RN-673	0.0278	0.0605	0.0264	0.0835	0.0502	0.0381	0.0986	0.0658	0.7010	0.0564	0.0233	5.3303
CFGAN uPM RN-1347	0.0270	0.0588	0.0231	0.0793	0.0470	0.0370	0.0908	0.0647	0.6081	0.0356	0.0177	4.8165
CFGAN uPM RN-2694	0.0291	0.0618	0.0263	0.0850	0.0513	0.0395	0.0996	0.0648	0.5935	0.0408	0.0174	4.8640
CFGAN uZP NO-ES	0.0278	0.0715	0.0252	0.0793	0.0512	0.0401	0.0905	0.0591	0.1424	0.0186	0.0090	3.6485
CFGAN uZP CC	0.0247	0.0556	0.0218	0.0749	0.0439	0.0342	0.0839	0.0596	0.1353	0.0186	0.0089	3.6314
CFGAN uZP RN-673	0.0259	0.0602	0.0244	0.0765	0.0476	0.0363	0.0886	0.0588	0.1458	0.0200	0.0090	3.6595
CFGAN uZP RN-1347	0.0278	0.0657	0.0276	0.0839	0.0521	0.0391	0.0968	0.0591	0.1420	0.0186	0.0089	3.6507
CFGAN uZP RN-2694	0.0250	0.0565	0.0231	0.0764	0.0456	0.0347	0.0878	0.0590	0.1416	0.0193	0.0089	3.6484

Table 10: Accuracy and beyond-accuracy values for different item-based CFGAN models for the Ciao dataset at recommendation list length of 20. The suffix RN-X means that the model uses random noise of size X. The suffix Class indicates that the model uses the user/item class as the condition vector. The suffix NO-ES indicates that the model does not use early-stopping. The suffix Reference is the model in the reference article. Hyper-parameter sets of variants are the same as those in Table 5.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
CFGAN iZR NO-ES	0.0359	0.1647	0.0498	0.1573	0.1058	0.0589	0.1933	0.1437	0.9628	0.7869	0.3842	9.4161
CFGAN iZR CC	0.0064	0.0246	0.0045	0.0219	0.0143	0.0102	0.0231	0.1473	0.0662	0.0423	0.0164	4.5218
CFGAN iZR RN-331	0.0393	0.1761	0.0536	0.1489	0.1116	0.0643	0.1907	0.1414	0.9519	0.7290	0.3094	9.1185
CFGAN iZR RN-662	0.0395	0.1718	0.0465	0.1371	0.1046	0.0643	0.1796	0.1410	0.9480	0.7491	0.2983	9.0630
CFGAN iZR RN-1324	0.0377	0.1650	0.0495	0.1465	0.1052	0.0613	0.1845	0.1417	0.9500	0.8018	0.3236	9.1666
CFGAN iPM NO-ES	0.0356	0.1662	0.0453	0.1322	0.1005	0.0587	0.1654	0.1370	0.9120	0.6644	0.2192	8.5638
CFGAN iPM CC	0.0083	0.0309	0.0093	0.0348	0.0214	0.0131	0.0380	0.1498	0.9602	0.9480	0.4495	9.5812
CFGAN iPM RN-331	0.0357	0.1533	0.0458	0.1415	0.0990	0.0579	0.1771	0.1389	0.9036	0.5865	0.1892	8.3721
CFGAN iPM RN-662	0.0368	0.1641	0.0448	0.1329	0.1003	0.0601	0.1701	0.1387	0.9023	0.5939	0.1926	8.3843
CFGAN iPM RN-1324	0.0346	0.1489	0.0420	0.1339	0.0938	0.0562	0.1656	0.1394	0.9106	0.6726	0.2189	8.5553
CFGAN iZP NO-ES	0.0336	0.1426	0.0384	0.1246	0.0881	0.0544	0.1550	0.1401	0.9404	0.6533	0.2566	8.8508
CFGAN iZP CC	0.0055	0.0210	0.0034	0.0175	0.0117	0.0088	0.0177	0.1482	0.9097	0.6177	0.1932	8.4167
CFGAN iZP RN-331	0.0345	0.1576	0.0450	0.1350	0.0980	0.0565	0.1637	0.1383	0.9303	0.6110	0.2299	8.6842
CFGAN iZP RN-662	0.0349	0.1492	0.0426	0.1407	0.0953	0.0566	0.1751	0.1382	0.9257	0.6095	0.2174	8.6045
CFGAN iZP RN-1324	0.0338	0.1538	0.0484	0.1448	0.1007	0.0554	0.1753	0.1386	0.9292	0.6236	0.2263	8.6638

Table 11: Accuracy and beyond-accuracy values for different user-based CFGAN models for the Ciao dataset at recommendation list length of 20. The suffix RN-X means that the model uses random noise of size X. The suffix Class indicates that the model uses the user/item class as the condition vector. The suffix NO-ES indicates that the model does not use early-stopping. The suffix Reference is the model in the reference article. Hyper-parameter sets of variants are the same as those in Table 5.

	PREC	REC	MAP	MRR	NDCG	F1	ARHR	Novelty	Div. MIL	Cov. Item	Div. Gini	Div. Shannon
CFGAN uZR NO-ES	0.0324	0.1386	0.0360	0.1147	0.0840	0.0525	0.1469	0.1430	0.9402	0.5598	0.2418	8.7633
CFGAN uZR CC	0.0217	0.1019	0.0278	0.0892	0.0624	0.0358	0.1062	0.1221	0.1153	0.0341	0.0168	4.5713
CFGAN uZR RN-673	0.0385	0.1635	0.0499	0.1485	0.1069	0.0623	0.1932	0.1315	0.8228	0.2138	0.0825	7.2085
CFGAN uZR RN-1347	0.0376	0.1596	0.0457	0.1428	0.1017	0.0608	0.1819	0.1298	0.7864	0.1782	0.0676	6.9228
CFGAN uZR RN-2694	0.0386	0.1640	0.0460	0.1449	0.1033	0.0625	0.1850	0.1311	0.8172	0.2027	0.0785	7.1371
CFGAN uPM NO-ES	0.0289	0.1279	0.0315	0.1094	0.0767	0.0472	0.1303	0.1320	0.7352	0.1121	0.0524	6.5022
CFGAN uPM CC	0.0084	0.0240	0.0051	0.0233	0.0153	0.0125	0.0301	0.1395	0.0565	0.0312	0.0161	4.4677
CFGAN uPM RN-673	0.0227	0.0967	0.0287	0.0912	0.0631	0.0367	0.1107	0.1370	0.6737	0.0891	0.0415	6.1492
CFGAN uPM RN-1347	0.0224	0.1012	0.0259	0.0868	0.0613	0.0367	0.1028	0.1321	0.5781	0.0616	0.0323	5.6931
CFGAN uPM RN-2694	0.0223	0.0970	0.0282	0.0914	0.0626	0.0362	0.1101	0.1324	0.5298	0.0683	0.0294	5.6474
CFGAN uZP NO-ES	0.0227	0.1088	0.0283	0.0848	0.0646	0.0375	0.1021	0.1212	0.1209	0.0371	0.0168	4.5927
CFGAN uZP CC	0.0216	0.0989	0.0254	0.0812	0.0593	0.0354	0.0966	0.1220	0.1159	0.0312	0.0167	4.5733
CFGAN uZP RN-673	0.0223	0.1009	0.0278	0.0826	0.0621	0.0365	0.1010	0.1205	0.1238	0.0349	0.0169	4.5925
CFGAN uZP RN-1347	0.0233	0.1079	0.0310	0.0902	0.0670	0.0383	0.1096	0.1204	0.1240	0.0364	0.0169	4.5988
CFGAN uZP RN-2694	0.0226	0.1052	0.0265	0.0838	0.0621	0.0372	0.1011	0.1206	0.1227	0.0371	0.0169	4.5945

Table 12: Number of training epochs when using the early-stopping approach. We adjust our obtained number of epochs to make them comparable to the ones in the reference article and reference source code.

Variant	# of Epochs	Variant	# of Epochs
CFGAN iZP Code	$1,500$	CFGAN iZP Article	$1,000$
CFGAN iZR	240	CFGAN uZR	355
CFGAN iPM	$1,200$	CFGAN uPM	$1,040$
CFGAN iZP	430	CFGAN uZP	540

References

1. Chae, D.K., Kang, J.S., Kim, S.W., Lee, J.T.: CFGAN: A generic collaborative filtering framework based on generative adversarial networks. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management. p. 137–146. CIKM '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3269206.3271743
2. Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A.: Generative adversarial networks: An overview. IEEE Signal Processing Magazine **35**(1), 53–65 (Jan 2018). https://doi.org/10.1109/MSP.2017.2765202
3. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. Commun. ACM **63**(11), 139–144 (Oct 2020). https://doi.org/10.1145/3422622
4. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. NIPS'14, vol. 2, p. 2672–2680. MIT Press, Cambridge, MA, USA (2014)
5. Isola, P., Zhu, J., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5967–5976 (July 2017). https://doi.org/10.1109/CVPR.2017.632
6. Mirza, M., Osindero, S.: Conditional generative adversarial nets. CoRR **abs/1411.1784** (2014), http://arxiv.org/abs/1411.1784