University of the Peloponnese
Department of Informatics and Telecommunications
Software and Database Systems Laboratory
https://soda.dit.uop.gr

Experimental results for Evaluating User Similarity Metrics in Sparse Collaborative Filtering Datasets

Technical Report TR-24001 (version 2)
Kiriakos Sgardelis, Dionisis Margaris, Dimitris Spiliotopoulos
Costas Vassilakis
k.sgardelis@go.uop.gr, margaris@di.uoa.gr, dspiliot@uop.gr,
costas@uop.gr

February, 2025
Tripoli, Greece

1.   Introduction

In this technical report, we present the experimental findings from evaluating 15 user similarity metrics when applied in sparse collaborative filtering (CF) datasets.

The 15 user similarity metrics evaluated in our work are the following [1-4]:

1. The Jaccard Index (JACC)
2. The Manhattan distance (MANH)
3. The Euclidean distance (EUCL)
4. The Chebyshev distance (CHEB)
5. The Pearson Correlation Coefficient (PCC)
6. The Constrained PCC (CPCC)
7. The Sigmoid Pearson Correlation Coefficient (SPCC)
8. The Cosine similarity (COS)
9. The Adjusted Cosine measure (ACOS)
10. The Spearman rank correlation (SP)
11. The Kendall's Tau correlation (TAU)
12. The mean square difference-based similarity (MSD)
13. The Normalized Sum of Multiplications (NSM)
14. The Adjusted Rand index (ARI)
15. The Adjusted Mutual Information (AMI)

For the NNs selection we use the following two methods:

1. the KNNs: where we use only the K NNs with the higher similarities with the active user [5] (in our experiments we have 2 settings: K=250 and K=500).
2. the similarity threshold: where we use only the NNs whose similarities, with the active user, exceed a specific threshold THR [6] (in our experiments we have 3 settings: THR=0.0, THR=0.25 and THR=0.5), and

For the rating prediction numeric value production, we use two formulas:

1. the weighted sum function [7], and
2. the mean-centered prediction function [8].

Additionally, we report experiments concerning the analysis of metric behaviour in relation to sparsity, as well as for the behaviour of the metrics when ratings are not directly provided by users, but rather computed on the basis of implicit feedback.

2.   Experiment results

In this section, we report on our experiments aiming to find the most effective user similarity metrics in sparse CF datasets, in terms of prediction accuracy (primarily) and prediction coverage. Regarding prediction accuracy, we use 3 metrics, the MAE, the RMSE and the F1-measure (which includes the precision and recall metrics). As far as the rating prediction process is concerned, we exercise the standard 10-fold cross validation (and merge the ten rating prediction results).

The datasets used in the experiment are summarized in Table 1, and the results obtained are listed in the following subsections.

TABLE 1. DATASETS AND THEIR SPARSITY

Name	Sparsity %
R4 - Yahoo! Movies [9]	99.76
Book-crossing [10]	99.99
Amazon Video_Games [11]	99.95
Amazon Cell_Phones_and_Accessories [11]	99.985
Amazon Movies_and_TV [11]	99.98
Amazon Kindle_Store [11]	99.98
Amazon Digital_Music [11]	99.93
Amazon Musical_Instruments [11]	99.93
Epinions [12]	99.99
CiaoDVD [13]	99.91

It has to be mentioned that the rating range of all datasets are [1-5], expect for the Book-crossing which is [0,10].

2.1   Performance evaluation

In the following subsections, we present the detailed rating prediction results, in 2 parts, the first for the similarity threshold NNs selection and the second for the KNN NNs selection.

2.1.1 Experiments using the KNNs neighbor selection with the weighted sum prediction function

Table 2 depicts the prediction MAE results when K=250 and the weighted sum function is used.

TABLE 2. MAE RESULTS WHEN K=250 AND THE WEIGHTED SUM FUNCTION IS USED

MAE (K=250 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.247	0.672	0.557	0.407	0.608	0.468	0.825	0.874	3.002	0.792
MANH	0.251	0.685	0.622	0.439	0.658	0.504	0.822	0.892	3.087	0.756
EUCL	0.251	0.683	0.622	0.438	0.657	0.504	0.82	0.891	3.087	0.757
CHEB	0.251	0.682	0.622	0.436	0.657	0.504	0.82	0.889	3.103	0.758
PCC	0.273	0.705	0.63	0.451	0.64	0.5	0.879	0.928	3.101	0.794
CPCC	0.248	0.663	0.623	0.429	0.633	0.494	0.826	0.89	2.869	0.77
SPCC	0.271	0.698	0.61	0.44	0.631	0.498	0.887	0.912	2.947	0.77
COS	0.251	0.68	0.636	0.43	0.655	0.503	0.826	0.897	3.575	0.789
ACOS	0.261	0.686	0.622	0.45	0.606	0.473	0.86	0.933	2.859	0.753
SP	0.219	0.647	0.603	0.409	0.535	0.434	0.85	0.942	2.827	0.789
TAU	0.25	0.692	0.676	0.447	0.647	0.5	0.853	0.919	3.32	0.824
MSD	0.25	0.668	0.618	0.428	0.647	0.498	0.823	0.876	3.032	0.748
NSM	0.251	0.672	0.621	0.429	0.654	0.501	0.823	0.88	3.549	0.753
ARI	0.228	0.657	0.597	0.412	0.504	0.388	0.909	0.938	3.173	0.818
AMI	0.246	0.694	0.673	0.441	0.635	0.486	0.839	0.919	3.098	0.826

Table 3 depicts the prediction RMSE results when K=250 and the weighted sum function is used.

TABLE 3. RMSE RESULTS WHEN K=250 AND THE WEIGHTED SUM FUNCTION IS USED

RMSE (K=250 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.614	1.135	0.985	0.727	1.178	0.95	1.198	1.193	4.222	1.1
MANH	0.618	1.163	1.065	0.769	1.232	0.997	1.198	1.23	4.455	1.121
EUCL	0.617	1.158	1.064	0.767	1.229	0.995	1.195	1.227	4.439	1.121
CHEB	0.618	1.155	1.064	0.764	1.226	0.993	1.194	1.222	4.439	1.118
PCC	0.667	1.202	1.082	0.796	1.224	1.019	1.281	1.282	4.486	1.154
CPCC	0.613	1.137	1.065	0.754	1.2	0.981	1.206	1.232	4.288	1.124
SPCC	0.664	1.191	1.051	0.778	1.216	1.011	1.294	1.255	4.346	1.099
COS	0.616	1.148	1.078	0.754	1.219	0.987	1.199	1.226	4.693	1.144
ACOS	0.68	1.205	1.072	0.817	1.203	1.007	1.279	1.305	4.28	1.118
SP	0.588	1.148	1.054	0.759	1.113	0.923	1.262	1.312	4.303	1.145
TAU	0.619	1.175	1.124	0.784	1.22	0.993	1.235	1.262	4.638	1.188
MSD	0.613	1.132	1.057	0.751	1.211	0.981	1.192	1.202	4.341	1.104
NSM	0.616	1.138	1.062	0.752	1.218	0.983	1.194	1.209	4.686	1.114
ARI	0.614	1.162	1.046	0.757	1.086	0.869	1.306	1.293	4.532	1.158
AMI	0.615	1.171	1.115	0.774	1.209	0.979	1.214	1.253	4.426	1.183

Table 4 depicts the prediction F1-measure results when K=250 and the weighted sum function is used.

TABLE 4. F1 RESULTS WHEN K=250 AND THE WEIGHTED SUM FUNCTION IS USED

F1-measure (K=250 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.965	0.875	0.896	0.939	0.892	0.927	0.842	0.802	0.158	0.863
MANH	0.964	0.873	0.880	0.933	0.884	0.922	0.849	0.807	0.196	0.882
EUCL	0.965	0.874	0.880	0.933	0.884	0.922	0.848	0.808	0.196	0.882
CHEB	0.965	0.873	0.880	0.934	0.884	0.922	0.845	0.806	0.191	0.882
PCC	0.960	0.870	0.878	0.931	0.888	0.922	0.829	0.797	0.223	0.870
CPCC	0.965	0.880	0.881	0.936	0.890	0.924	0.848	0.812	0.230	0.881
SPCC	0.961	0.872	0.886	0.934	0.890	0.923	0.829	0.799	0.222	0.874
COS	0.964	0.874	0.878	0.935	0.884	0.922	0.841	0.804	0.244	0.877
ACOS	0.963	0.878	0.881	0.932	0.896	0.927	0.847	0.798	0.257	0.881
SP	0.970	0.885	0.887	0.940	0.910	0.936	0.848	0.787	0.161	0.871
TAU	0.965	0.872	0.871	0.931	0.885	0.922	0.836	0.798	0.139	0.870
MSD	0.965	0.878	0.881	0.936	0.886	0.924	0.844	0.810	0.210	0.883
NSM	0.965	0.877	0.881	0.936	0.885	0.923	0.843	0.811	0.258	0.882
ARI	0.966	0.881	0.887	0.938	0.913	0.941	0.806	0.784	0.203	0.858
AMI	0.965	0.869	0.867	0.932	0.887	0.923	0.838	0.792	0.173	0.866

Table 5 depicts the prediction coverage results when K=250 and the weighted sum function is used.

TABLE 5. PREDICTION COVERAGE RESULTS WHEN K=250 AND THE WEIGHTED SUM FUNCTION IS USED

coverage (K=250 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	40.17	58.01	15.48	70.03	23.37	44.34	10.99	46.2	46.89	85.93
MANH	39.96	53.83	11.26	63.82	26.2	45.27	10.76	42.37	36.06	79.42
EUCL	39.97	54.08	11.27	63.89	26.29	45.34	10.84	42.57	36.12	79.45
CHEB	39.99	54.47	11.33	64.47	26.45	45.69	10.96	43.37	36.41	79.92
PCC	32.14	47.24	10.87	59	21.06	36.78	7.72	38.86	29.54	76.79
CPCC	39.61	52.83	11.01	64.22	24.94	45.03	10.02	41.41	31.81	79.54
SPCC	32.44	48.83	14.22	63.1	21.62	37.56	7.76	42.64	33.19	82.68
COS	40.15	55.5	11.54	66.07	26.83	46.58	11.02	43.24	31.64	81.36
ACOS	26.48	38.14	10.38	54.51	16.66	30.45	5.52	34.39	28.51	78.4
SP	24.2	36.84	11.63	52.79	15.4	30.5	4.01	31.5	32.89	79.87
TAU	38.12	49.55	10.76	61.4	23.75	42.18	9.36	39.94	35.96	81.25
MSD	40.05	56.14	11.58	65.9	26.3	46.4	10.95	45.32	36.7	80.81
NSM	40.17	56.93	11.57	65.73	27	46.95	11.08	44.66	32.04	80.46
ARI	15.26	30	11.72	45.98	11.3	23.83	2.54	30.44	30.51	77.87
AMI	35.72	48.57	10.15	60.41	21.92	38.76	8.75	39.91	37.17	80.24

Table 6 depicts the prediction MAE results when K=500 and the weighted sum function is used.

TABLE 6. MAE RESULTS WHEN K=500 AND THE WEIGHTED SUM FUNCTION IS USED

MAE (K=500 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.248	0.673	0.572	0.413	0.608	0.47	0.826	0.867	2.999	0.79
MANH	0.253	0.686	0.625	0.44	0.664	0.512	0.826	0.884	3.048	0.755
EUCL	0.253	0.684	0.625	0.439	0.663	0.511	0.824	0.882	3.052	0.755
CHEB	0.252	0.683	0.624	0.437	0.663	0.507	0.824	0.878	3.07	0.755
PCC	0.273	0.71	0.639	0.455	0.641	0.512	0.881	0.924	3.066	0.782
CPCC	0.249	0.662	0.622	0.429	0.636	0.501	0.827	0.879	2.771	0.76
SPCC	0.272	0.706	0.633	0.448	0.638	0.507	0.892	0.913	2.997	0.773
COS	0.253	0.681	0.63	0.432	0.66	0.51	0.826	0.883	3.535	0.785
ACOS	0.264	0.693	0.628	0.458	0.603	0.49	0.86	0.943	2.753	0.744
SP	0.22	0.651	0.609	0.414	0.536	0.446	0.85	0.947	2.632	0.778
TAU	0.251	0.688	0.667	0.443	0.651	0.504	0.857	0.91	3.244	0.817
MSD	0.251	0.67	0.618	0.429	0.65	0.506	0.823	0.865	2.967	0.739
NSM	0.252	0.675	0.62	0.43	0.658	0.508	0.824	0.869	3.496	0.745
ARI	0.228	0.664	0.608	0.42	0.503	0.392	0.91	0.946	3.112	0.81
AMI	0.246	0.69	0.668	0.437	0.636	0.487	0.839	0.905	3.118	0.816

Table 7 depicts the prediction RMSE results when K=500 and the weighted sum function is used.

TABLE 7. RMSE RESULTS WHEN K=500 AND THE WEIGHTED SUM FUNCTION IS USED

RMSE (K=500 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.615	1.131	0.994	0.731	1.178	0.95	1.199	1.177	4.165	1.095
MANH	0.621	1.154	1.062	0.763	1.234	0.996	1.199	1.207	4.375	1.105
EUCL	0.62	1.147	1.061	0.76	1.23	0.992	1.196	1.202	4.359	1.104
CHEB	0.619	1.143	1.06	0.757	1.228	0.985	1.195	1.195	4.356	1.099
PCC	0.667	1.204	1.086	0.799	1.225	1.024	1.285	1.274	4.414	1.124
CPCC	0.613	1.123	1.058	0.746	1.2	0.977	1.208	1.209	4.159	1.097
SPCC	0.665	1.199	1.068	0.788	1.222	1.018	1.303	1.256	4.348	1.094
COS	0.618	1.135	1.065	0.748	1.22	0.982	1.196	1.197	4.594	1.127
ACOS	0.688	1.221	1.077	0.834	1.201	1.028	1.282	1.319	4.152	1.093
SP	0.588	1.157	1.054	0.769	1.114	0.931	1.262	1.32	4.105	1.121
TAU	0.62	1.16	1.108	0.773	1.22	0.988	1.24	1.245	4.523	1.17
MSD	0.614	1.121	1.049	0.743	1.21	0.978	1.191	1.176	4.226	1.074
NSM	0.617	1.127	1.053	0.745	1.217	0.98	1.193	1.182	4.579	1.088
ARI	0.614	1.179	1.05	0.775	1.087	0.878	1.307	1.31	4.458	1.14
AMI	0.615	1.155	1.106	0.762	1.206	0.97	1.213	1.224	4.373	1.16

Table 8 depicts the prediction F1-measure results when K=500 and the weighted sum function is used.

TABLE 8. RMSE RESULTS WHEN K=500 AND THE WEIGHTED SUM FUNCTION IS USED

F1-measure (K=500 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.965	0.876	0.895	0.938	0.892	0.927	0.842	0.807	0.145	0.867
MANH	0.964	0.874	0.882	0.934	0.883	0.921	0.846	0.810	0.184	0.884
EUCL	0.965	0.875	0.883	0.934	0.883	0.922	0.845	0.810	0.182	0.884
CHEB	0.965	0.875	0.883	0.934	0.883	0.922	0.843	0.809	0.175	0.884
PCC	0.960	0.870	0.879	0.931	0.888	0.921	0.831	0.798	0.227	0.872
CPCC	0.965	0.882	0.885	0.937	0.890	0.925	0.848	0.816	0.223	0.886
SPCC	0.961	0.871	0.884	0.933	0.889	0.922	0.830	0.800	0.224	0.873
COS	0.964	0.876	0.882	0.935	0.884	0.922	0.840	0.809	0.222	0.880
ACOS	0.962	0.879	0.884	0.932	0.897	0.925	0.848	0.798	0.258	0.882
SP	0.970	0.885	0.891	0.940	0.910	0.934	0.848	0.786	0.156	0.873
TAU	0.965	0.875	0.875	0.933	0.885	0.922	0.835	0.801	0.130	0.873
MSD	0.965	0.879	0.885	0.937	0.886	0.924	0.844	0.813	0.193	0.886
NSM	0.965	0.878	0.885	0.936	0.885	0.923	0.843	0.814	0.236	0.886
ARI	0.966	0.879	0.889	0.937	0.913	0.940	0.806	0.784	0.206	0.860
AMI	0.965	0.872	0.869	0.934	0.887	0.925	0.838	0.797	0.167	0.869

Table 9 depicts the prediction coverage results when K=500 and the weighted sum function is used.

TABLE 9. PREDICTION COVERAGE RESULTS WHEN K=500 AND THE WEIGHTED SUM FUNCTION IS USED

coverage (K=500 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	40.44	59.42	17.15	72.42	23.37	44.74	11.08	49.93	50.29	89.53
MANH	40.55	58.88	14.17	69.51	28.01	48.04	11.19	48.31	42.42	85.21
EUCL	40.56	59.13	14.18	69.59	28.06	48.18	11.19	48.58	42.49	85.25
CHEB	40.57	59.42	14.36	70.11	28.15	48.58	11.19	49.54	42.96	86.13
PCC	32.5	49.92	13.69	63.47	22.11	38.43	7.63	43.54	34.04	82.86
CPCC	40.06	56.93	14.09	69.65	26.46	47.53	10.07	47.17	37.68	85.76
SPCC	32.6	50.37	16.18	65.27	22.27	38.7	7.67	45.77	36.09	85.98
COS	40.62	60.02	14.6	70.98	28.47	49	11.19	49.58	37.9	86.98
ACOS	26.36	37.8	12.91	56.56	17.13	30.47	5.38	36.19	32.65	83.85
SP	24.25	36.49	14.55	53.8	15.37	30.44	4.01	32.82	39.57	84.58
TAU	38.7	54.66	13.23	67.02	25.14	44.72	9.36	45.18	42.06	86.77
MSD	40.47	59.8	14.99	71.13	27.86	48.7	11.03	51.57	44.09	87.47
NSM	40.6	60.68	14.96	71.07	28.65	49.27	11.19	51.13	38.24	87.09
ARI	15.25	29.15	14.17	45.91	11.21	23.53	2.54	30.97	32.59	83.05
AMI	36.15	53.37	12.85	65.94	23.38	41.5	8.77	45.7	41.97	85.55

2.1.2 Experiments using the KNNs neighbor selection with the mean-centered prediction function

Table 10 depicts the prediction MAE results when K=250 and the mean-centered function is used.

TABLE 10. MAE RESULTS WHEN K=250 AND THE MEAN-CENTERED FUNCTION IS USED

MAE (K=250 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.298	0.675	0.557	0.441	0.643	0.493	0.689	0.846	2.667	0.713
MANH	0.304	0.701	0.638	0.473	0.694	0.538	0.69	0.879	2.902	0.756
EUCL	0.304	0.699	0.638	0.472	0.693	0.537	0.688	0.876	2.898	0.756
CHEB	0.304	0.697	0.638	0.469	0.693	0.537	0.685	0.874	2.898	0.754
PCC	0.329	0.723	0.638	0.484	0.683	0.538	0.769	0.903	2.935	0.773
CPCC	0.302	0.689	0.642	0.463	0.674	0.529	0.691	0.877	2.86	0.757
SPCC	0.328	0.714	0.612	0.472	0.673	0.534	0.765	0.884	2.812	0.732
COS	0.303	0.692	0.648	0.461	0.689	0.536	0.689	0.873	3.068	0.754
ACOS	0.348	0.738	0.644	0.503	0.671	0.522	0.774	0.93	2.84	0.758
SP	0.273	0.673	0.612	0.451	0.582	0.472	0.754	0.92	2.687	0.75
TAU	0.305	0.709	0.686	0.475	0.688	0.538	0.724	0.895	2.94	0.769
MSD	0.303	0.685	0.636	0.461	0.684	0.531	0.69	0.862	2.892	0.748
NSM	0.304	0.688	0.638	0.463	0.689	0.534	0.687	0.866	3.068	0.751
ARI	0.284	0.675	0.592	0.449	0.545	0.418	0.798	0.912	2.855	0.765
AMI	0.299	0.698	0.675	0.467	0.67	0.516	0.717	0.892	2.822	0.767

Table 11 depicts the prediction RMSE results when K=250 and the mean-centered function is used.

TABLE 11. RMSE RESULTS WHEN K=250 AND THE MEAN-CENTERED FUNCTION IS USED

RMSE (K=250 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.601	1.061	0.928	0.709	1.126	0.912	0.977	1.149	3.712	1.014
MANH	0.607	1.093	1.009	0.747	1.176	0.957	0.977	1.193	3.945	1.069
EUCL	0.606	1.089	1.009	0.745	1.173	0.955	0.973	1.19	3.937	1.069
CHEB	0.606	1.085	1.008	0.743	1.171	0.954	0.971	1.185	3.932	1.066
PCC	0.659	1.146	1.021	0.779	1.181	0.993	1.099	1.236	4.052	1.087
CPCC	0.605	1.085	1.013	0.736	1.155	0.947	0.987	1.197	3.954	1.069
SPCC	0.659	1.135	0.992	0.762	1.172	0.985	1.094	1.21	3.941	1.042
COS	0.604	1.077	1.015	0.732	1.164	0.949	0.975	1.182	4.094	1.064
ACOS	0.737	1.198	1.032	0.832	1.19	1.008	1.131	1.279	3.993	1.076
SP	0.584	1.105	1.002	0.754	1.079	0.904	1.073	1.268	3.778	1.072
TAU	0.608	1.106	1.048	0.758	1.168	0.958	1.025	1.214	3.922	1.083
MSD	0.603	1.07	1.005	0.731	1.159	0.944	0.976	1.169	3.915	1.058
NSM	0.604	1.071	1.006	0.733	1.164	0.945	0.973	1.173	4.105	1.062
ARI	0.606	1.105	0.98	0.743	1.047	0.846	1.145	1.244	4.007	1.081
AMI	0.604	1.095	1.043	0.746	1.155	0.94	1.003	1.206	3.875	1.084

Table 12 depicts the prediction F1-measure results when K=250 and the mean-centered function is used.

TABLE 12. F1-MEASURE RESULTS WHEN K=250 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (K=250 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.965	0.882	0.903	0.939	0.893	0.928	0.861	0.813	0.283	0.875
MANH	0.964	0.876	0.883	0.933	0.884	0.921	0.865	0.811	0.309	0.872
EUCL	0.965	0.877	0.883	0.934	0.884	0.921	0.865	0.811	0.311	0.872
CHEB	0.965	0.877	0.883	0.934	0.885	0.921	0.865	0.811	0.312	0.872
PCC	0.960	0.873	0.883	0.931	0.887	0.920	0.841	0.803	0.312	0.866
CPCC	0.965	0.881	0.883	0.936	0.889	0.923	0.862	0.813	0.277	0.873
SPCC	0.960	0.875	0.891	0.933	0.889	0.922	0.844	0.806	0.310	0.874
COS	0.965	0.879	0.882	0.936	0.885	0.922	0.863	0.812	0.336	0.873
ACOS	0.957	0.875	0.883	0.929	0.892	0.924	0.853	0.798	0.303	0.872
SP	0.969	0.887	0.892	0.939	0.909	0.934	0.855	0.793	0.278	0.872
TAU	0.964	0.876	0.875	0.933	0.886	0.921	0.852	0.805	0.280	0.869
MSD	0.965	0.881	0.884	0.936	0.886	0.923	0.861	0.813	0.303	0.873
NSM	0.965	0.880	0.883	0.936	0.885	0.922	0.863	0.813	0.340	0.872
ARI	0.965	0.883	0.894	0.938	0.913	0.941	0.815	0.792	0.277	0.865
AMI	0.965	0.876	0.874	0.934	0.888	0.924	0.853	0.804	0.298	0.869

As far as the prediction coverage when K=250 is concerned, it is the exact same, when the weighted sum prediction function was used (TABLE 5).

Table 13 depicts the prediction MAE results when K=500 and the mean-centered function is used.

TABLE 13. MAE RESULTS WHEN K=500 AND THE MEAN-CENTERED FUNCTION IS USED

MAE (K=500 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.299	0.674	0.568	0.445	0.643	0.494	0.691	0.839	2.652	0.711
MANH	0.305	0.696	0.637	0.471	0.697	0.542	0.695	0.866	2.843	0.746
EUCL	0.305	0.693	0.637	0.469	0.696	0.541	0.693	0.863	2.839	0.746
CHEB	0.305	0.691	0.637	0.466	0.696	0.537	0.691	0.859	2.838	0.742
PCC	0.33	0.726	0.642	0.486	0.682	0.547	0.767	0.897	2.871	0.752
CPCC	0.302	0.683	0.638	0.46	0.676	0.533	0.693	0.864	2.779	0.743
SPCC	0.329	0.721	0.629	0.478	0.679	0.543	0.768	0.885	2.809	0.728
COS	0.304	0.686	0.636	0.46	0.692	0.539	0.691	0.857	2.978	0.744
ACOS	0.353	0.759	0.647	0.515	0.67	0.544	0.774	0.942	2.746	0.746
SP	0.274	0.681	0.616	0.456	0.584	0.483	0.754	0.924	2.548	0.731
TAU	0.305	0.702	0.674	0.472	0.689	0.54	0.729	0.885	2.873	0.757
MSD	0.303	0.68	0.633	0.459	0.684	0.536	0.69	0.846	2.819	0.731
NSM	0.304	0.683	0.634	0.46	0.69	0.538	0.69	0.85	2.974	0.735
ARI	0.284	0.688	0.602	0.458	0.547	0.425	0.799	0.923	2.791	0.753
AMI	0.299	0.689	0.661	0.463	0.668	0.514	0.716	0.876	2.808	0.756

Table 14 depicts the prediction RMSE results when K=500 and the mean-centered function is used.

TABLE 14. RMSE RESULTS WHEN K=500 AND THE MEAN-CENTERED FUNCTION IS USED

RMSE (K=500 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.602	1.058	0.937	0.712	1.126	0.912	0.978	1.136	3.666	1.012
MANH	0.608	1.084	1.006	0.742	1.177	0.956	0.981	1.17	3.878	1.052
EUCL	0.607	1.078	1.005	0.739	1.174	0.953	0.978	1.166	3.869	1.052
CHEB	0.606	1.074	1.003	0.736	1.172	0.947	0.977	1.159	3.857	1.047
PCC	0.66	1.154	1.025	0.784	1.183	1.001	1.096	1.231	4.002	1.061
CPCC	0.605	1.074	1.007	0.729	1.155	0.944	0.988	1.178	3.866	1.046
SPCC	0.66	1.147	1.008	0.773	1.18	0.995	1.099	1.214	3.943	1.034
COS	0.605	1.064	1.001	0.727	1.164	0.944	0.977	1.155	3.985	1.048
ACOS	0.748	1.235	1.04	0.859	1.195	1.038	1.134	1.3	3.915	1.061
SP	0.584	1.118	1.006	0.765	1.082	0.914	1.073	1.28	3.672	1.05
TAU	0.608	1.095	1.037	0.751	1.169	0.955	1.03	1.201	3.846	1.065
MSD	0.602	1.058	0.997	0.725	1.157	0.941	0.976	1.143	3.813	1.032
NSM	0.604	1.059	0.997	0.726	1.162	0.942	0.975	1.146	3.995	1.036
ARI	0.606	1.13	0.988	0.766	1.051	0.86	1.144	1.266	3.965	1.066
AMI	0.603	1.08	1.033	0.737	1.152	0.932	1.002	1.181	3.84	1.067

Table 15 depicts the prediction F1-measure results when K=500 and the mean-centered function is used.

TABLE 15. F1-MEASURE RESULTS WHEN K=500 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (K=500 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.965	0.882	0.902	0.938	0.893	0.928	0.861	0.817	0.275	0.878
MANH	0.964	0.878	0.886	0.934	0.884	0.921	0.863	0.813	0.301	0.874
EUCL	0.965	0.879	0.886	0.935	0.884	0.921	0.863	0.814	0.302	0.874
CHEB	0.965	0.879	0.886	0.935	0.884	0.922	0.862	0.815	0.301	0.875
PCC	0.960	0.873	0.885	0.931	0.887	0.919	0.844	0.804	0.303	0.871
CPCC	0.965	0.884	0.887	0.937	0.890	0.923	0.862	0.817	0.262	0.876
SPCC	0.960	0.874	0.889	0.932	0.888	0.921	0.845	0.806	0.303	0.875
COS	0.965	0.881	0.887	0.936	0.885	0.922	0.861	0.815	0.322	0.876
ACOS	0.957	0.873	0.886	0.928	0.892	0.920	0.854	0.796	0.288	0.875
SP	0.969	0.885	0.895	0.938	0.909	0.933	0.855	0.792	0.272	0.875
TAU	0.965	0.879	0.879	0.934	0.886	0.922	0.852	0.807	0.274	0.872
MSD	0.965	0.883	0.888	0.937	0.887	0.923	0.862	0.817	0.291	0.877
NSM	0.965	0.882	0.888	0.937	0.885	0.923	0.862	0.817	0.327	0.877
ARI	0.965	0.881	0.896	0.936	0.913	0.939	0.816	0.790	0.280	0.867
AMI	0.965	0.878	0.878	0.935	0.889	0.925	0.853	0.807	0.291	0.872

As far as the prediction coverage when K=500 is concerned, it is the exact same, when the weighted sum prediction function was used (TABLE 9).

2.1.3 Experiments using the similarity threshold neighbor selection with the weighted sum prediction function

Table 16 depicts the prediction MAE results when THR=0.0 and the weighted sum function is used.

TABLE 16. MAE RESULTS WHEN THR=0.0 AND THE WEIGHTED SUM FUNCTION IS USED

MAE (THR=0.0 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.248	0.673	0.6	0.417	0.608	0.47	0.826	0.861	2.995	0.8
MANH	0.253	0.686	0.661	0.441	0.663	0.517	0.826	0.863	2.993	0.778
EUCL	0.253	0.684	0.66	0.438	0.662	0.517	0.824	0.859	3.014	0.782
CHEB	0.252	0.683	0.661	0.437	0.662	0.517	0.824	0.858	3.042	0.786
PCC	0.261	0.681	0.66	0.438	0.628	0.495	0.853	0.867	3.117	0.788
CPCC	0.248	0.663	0.646	0.429	0.637	0.516	0.814	0.85	2.738	0.77
SPCC	0.26	0.678	0.655	0.434	0.626	0.491	0.852	0.866	3.102	0.788
COS	0.253	0.684	0.673	0.436	0.664	0.527	0.826	0.861	3.431	0.809
ACOS	0.241	0.651	0.628	0.419	0.581	0.462	0.823	0.865	2.677	0.736
SP	0.217	0.628	0.62	0.392	0.53	0.433	0.822	0.871	2.834	0.784
TAU	0.25	0.676	0.659	0.431	0.644	0.512	0.847	0.86	3.022	0.792
MSD	0.251	0.675	0.663	0.434	0.653	0.522	0.822	0.856	2.972	0.79
NSM	0.252	0.681	0.666	0.435	0.662	0.525	0.824	0.858	3.399	0.797
ARI	0.226	0.643	0.62	0.401	0.5	0.385	0.89	0.899	3.143	0.802
AMI	0.246	0.688	0.674	0.435	0.635	0.495	0.839	0.873	3.084	0.799

Table 17 depicts the prediction RMSE results when THR=0.0 and the weighted sum function is used.

TABLE 17. RMSE RESULTS WHEN THR=0.0 AND THE WEIGHTED SUM FUNCTION IS USED

RMSE (THR=0.0 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.615	1.13	1.004	0.732	1.178	0.95	1.199	1.162	4.106	1.096
MANH	0.621	1.144	1.046	0.754	1.229	0.988	1.199	1.165	4.197	1.09
EUCL	0.619	1.135	1.037	0.748	1.224	0.984	1.196	1.157	4.164	1.085
CHEB	0.619	1.133	1.035	0.746	1.222	0.982	1.195	1.154	4.151	1.085
PCC	0.633	1.147	1.053	0.752	1.2	0.974	1.228	1.173	4.279	1.084
CPCC	0.61	1.113	1.023	0.735	1.194	0.976	1.185	1.15	3.956	1.061
SPCC	0.632	1.143	1.047	0.748	1.197	0.969	1.227	1.171	4.264	1.082
COS	0.618	1.129	1.043	0.742	1.218	0.985	1.196	1.153	4.425	1.101
ACOS	0.61	1.126	1.031	0.737	1.155	0.949	1.206	1.182	3.931	1.037
SP	0.577	1.101	1.021	0.708	1.099	0.898	1.207	1.19	4.105	1.085
TAU	0.615	1.132	1.041	0.741	1.206	0.976	1.218	1.162	4.168	1.097
MSD	0.614	1.12	1.034	0.74	1.209	0.981	1.19	1.148	4.05	1.079
NSM	0.617	1.126	1.036	0.741	1.217	0.983	1.193	1.151	4.419	1.091
ARI	0.606	1.133	1.042	0.727	1.076	0.855	1.268	1.232	4.45	1.111
AMI	0.614	1.149	1.059	0.75	1.204	0.969	1.213	1.173	4.215	1.1

Table 18 depicts the prediction F1-measure results when THR=0.0 and the mean-centered function is used.

TABLE 18. F1-MEASURE RESULTS WHEN THR=0.0 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (THR=0.0 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.965	0.876	0.891	0.938	0.892	0.927	0.842	0.810	0.137	0.871
MANH	0.964	0.875	0.884	0.935	0.884	0.922	0.846	0.814	0.152	0.879
EUCL	0.965	0.876	0.885	0.935	0.884	0.923	0.845	0.815	0.139	0.878
CHEB	0.965	0.876	0.885	0.936	0.884	0.923	0.843	0.814	0.130	0.877
PCC	0.962	0.875	0.882	0.935	0.890	0.925	0.836	0.810	0.164	0.871
CPCC	0.966	0.882	0.890	0.938	0.890	0.924	0.850	0.820	0.184	0.882
SPCC	0.963	0.876	0.884	0.936	0.891	0.926	0.838	0.810	0.162	0.871
COS	0.965	0.876	0.883	0.936	0.883	0.922	0.840	0.813	0.182	0.872
ACOS	0.966	0.882	0.890	0.938	0.901	0.930	0.854	0.811	0.210	0.886
SP	0.970	0.888	0.892	0.944	0.911	0.936	0.847	0.802	0.128	0.874
TAU	0.965	0.877	0.885	0.937	0.887	0.923	0.837	0.813	0.109	0.876
MSD	0.965	0.879	0.886	0.937	0.886	0.923	0.844	0.815	0.137	0.877
NSM	0.965	0.878	0.885	0.937	0.884	0.923	0.843	0.815	0.196	0.875
ARI	0.967	0.882	0.888	0.940	0.914	0.942	0.810	0.790	0.193	0.861
AMI	0.965	0.872	0.878	0.935	0.888	0.925	0.838	0.805	0.148	0.869

Table 19 depicts the prediction coverage results when THR=0.0 and the weighted sum function is used.

TABLE 19. PREDICTION COVERAGE RESULTS WHEN THR=0.0 AND THE WEIGHTED SUM FUNCTION IS USED

coverage (THR=0.0 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	40.46	59.74	19.39	74.03	23.37	44.74	11.08	53.01	53.07	96.31
MANH	40.73	61.44	21.84	74.75	29.17	49.89	11.19	58.95	61.12	96.52
EUCL	40.73	61.44	21.84	74.75	29.17	49.89	11.19	58.95	61.94	96.57
CHEB	40.73	61.44	21.84	74.75	29.17	49.89	11.19	58.95	61.94	96.57
PCC	33.32	54.13	20.66	69.39	23.39	41.61	8.45	54.19	52.13	93.23
CPCC	40.26	58.68	21.37	73.92	27.19	48.56	10.25	56.25	57.01	96.12
SPCC	33.31	54.12	20.66	69.39	23.38	41.6	8.44	54.18	52.12	93.23
COS	40.73	61.44	21.84	74.75	29.17	49.89	11.19	58.95	47.58	96.57
ACOS	30.51	48.48	19.93	67.56	19.56	37.24	6.66	49.22	52.55	94.72
SP	24.9	40.63	17.86	58.34	16.13	32.48	4.45	41.5	51.15	94.47
TAU	39.62	58.61	21.39	73.19	26.82	47.33	9.85	56.37	60.17	96.27
MSD	40.61	60.59	21.75	74.62	28.37	49.36	11.03	58.53	60.8	96.43
NSM	40.73	61.44	21.84	74.75	29.17	49.89	11.19	58.95	47.56	96.57
ARI	15.51	31.91	16.05	48.67	11.6	24.73	2.89	35.92	37.2	89.35
AMI	36.27	54.38	20.81	70.12	23.77	42.21	8.77	55.18	56.44	95.66

Table 20 depicts the prediction MAE results when THR=0.25 and the weighted sum function is used. When THR=0.25, the Jaccard Index presents a very low prediction coverage (~5.5% on average), where the rating prediction results are unreliable and hence it does not take part in the accuracy results.

TABLE 20. MAE RESULTS WHEN THR=0.25 AND THE WEIGHTED SUM FUNCTION IS USED

MAE (THR=0.25 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.239	0.658	0.636	0.438	0.602	0.491	0.826	0.875	2.64	0.746
EUCL	0.245	0.665	0.647	0.436	0.619	0.501	0.795	0.866	2.662	0.751
CHEB	0.246	0.664	0.647	0.432	0.621	0.5	0.788	0.859	2.662	0.751
PCC	0.257	0.676	0.658	0.435	0.613	0.485	0.851	0.868	3.121	0.784
CPCC	0.248	0.659	0.644	0.429	0.632	0.514	0.806	0.849	2.647	0.767
SPCC	0.254	0.671	0.65	0.43	0.603	0.475	0.848	0.868	3.076	0.783
COS	0.253	0.684	0.673	0.436	0.664	0.527	0.826	0.861	3.488	0.809
ACOS	0.237	0.643	0.621	0.415	0.559	0.447	0.81	0.871	2.616	0.734
SP	0.211	0.617	0.612	0.387	0.501	0.414	0.796	0.874	2.751	0.779
TAU	0.25	0.675	0.658	0.431	0.641	0.51	0.844	0.86	3	0.79
MSD	0.251	0.673	0.662	0.433	0.651	0.522	0.816	0.855	2.937	0.788
NSM	0.252	0.679	0.666	0.434	0.661	0.525	0.823	0.857	3.511	0.794
ARI	0.216	0.617	0.595	0.393	0.433	0.34	0.829	0.932	3.156	0.801
AMI	0.245	0.692	0.679	0.437	0.634	0.492	0.858	0.882	3.131	0.8

Table 21 depicts the prediction RMSE results when THR=0.25 and the weighted sum function is used.

TABLE 21. RMSE RESULTS WHEN THR=0.25 AND THE WEIGHTED SUM FUNCTION IS USED

RMSE (THR=0.25 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.6	1.139	1.054	0.756	1.177	0.975	1.212	1.201	4.103	1.074
EUCL	0.607	1.134	1.05	0.748	1.189	0.977	1.171	1.176	4.11	1.065
CHEB	0.608	1.129	1.041	0.741	1.189	0.972	1.159	1.163	4.105	1.054
PCC	0.626	1.144	1.055	0.749	1.185	0.965	1.229	1.176	4.297	1.082
CPCC	0.609	1.11	1.022	0.734	1.19	0.973	1.177	1.15	3.889	1.057
SPCC	0.621	1.139	1.049	0.744	1.175	0.954	1.23	1.176	4.268	1.079
COS	0.618	1.129	1.043	0.742	1.218	0.985	1.196	1.153	4.476	1.101
ACOS	0.604	1.121	1.028	0.732	1.133	0.932	1.192	1.191	3.892	1.039
SP	0.567	1.093	1.016	0.703	1.067	0.875	1.188	1.196	4.061	1.082
TAU	0.614	1.132	1.041	0.741	1.204	0.975	1.216	1.163	4.16	1.095
MSD	0.614	1.117	1.032	0.739	1.206	0.98	1.182	1.147	4.016	1.076
NSM	0.616	1.124	1.036	0.741	1.216	0.983	1.191	1.15	4.549	1.087
ARI	0.587	1.117	1.031	0.72	1.001	0.808	1.236	1.282	4.545	1.126
AMI	0.611	1.158	1.073	0.755	1.207	0.971	1.238	1.184	4.269	1.107

Table 22 depicts the prediction F1-measure results when THR=0.25 and the mean-centered function is used.

TABLE 22. F1-MEASURE RESULTS WHEN THR=0.25 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (THR=0.25 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.966	0.880	0.886	0.935	0.896	0.925	0.854	0.814	0.259	0.887
EUCL	0.966	0.879	0.886	0.936	0.893	0.924	0.857	0.815	0.278	0.887
CHEB	0.966	0.879	0.886	0.936	0.892	0.925	0.853	0.814	0.282	0.887
PCC	0.963	0.876	0.882	0.936	0.893	0.927	0.838	0.810	0.291	0.872
CPCC	0.966	0.883	0.891	0.938	0.891	0.925	0.853	0.820	0.262	0.883
SPCC	0.964	0.877	0.884	0.936	0.895	0.929	0.840	0.809	0.291	0.871
COS	0.965	0.876	0.883	0.936	0.883	0.922	0.840	0.813	0.315	0.872
ACOS	0.966	0.884	0.891	0.939	0.905	0.932	0.859	0.811	0.298	0.887
SP	0.971	0.890	0.893	0.945	0.917	0.940	0.858	0.800	0.257	0.874
TAU	0.965	0.877	0.885	0.937	0.887	0.924	0.839	0.813	0.265	0.876
MSD	0.965	0.879	0.886	0.937	0.886	0.923	0.846	0.815	0.269	0.877
NSM	0.965	0.878	0.885	0.937	0.884	0.923	0.843	0.816	0.328	0.876
ARI	0.968	0.887	0.891	0.942	0.926	0.949	0.834	0.785	0.275	0.863
AMI	0.965	0.871	0.875	0.935	0.887	0.925	0.835	0.803	0.280	0.870

Table 23 depicts the prediction coverage results when THR=0.25 and the weighted sum function is used.

TABLE 23. PREDICTION COVERAGE RESULTS WHEN THR=0.25 AND THE WEIGHTED SUM FUNCTION IS USED

coverage (THR=0.25 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	5.28	1.16	2.24	16.48	2.61	8.43	0	0.18	0.26	37.45
MANH	38.73	47.64	18.32	68.81	20.01	40.8	6.93	45.11	30.31	90.59
EUCL	39.75	53.47	20.12	72.7	22.59	44.61	8.87	51.19	31.06	93.41
CHEB	39.84	54.76	20.67	73.48	23.17	45.54	9.37	53.44	31.16	94.56
PCC	32.51	50.73	19.89	67.54	21.1	38.38	7.4	51.42	45.02	92.39
CPCC	40.19	58.19	21.29	73.84	26.78	48.31	9.98	55.65	53.51	95.98
SPCC	32.25	49.63	19.66	66.79	20.47	37.49	6.96	50.57	44.18	92.23
COS	40.73	61.44	21.84	74.75	29.17	49.89	11.19	58.95	44.88	96.57
ACOS	29.54	44.38	18.92	64.9	17.43	33.92	5.46	44.88	46.97	93.17
SP	23.76	37.15	17.01	56.05	14.38	29.72	3.51	38.35	47.71	93.48
TAU	39.56	58.25	21.31	73.06	26.5	47.06	9.6	55.85	59.32	96.12
MSD	40.58	60.37	21.72	74.58	28.21	49.26	10.94	58.33	59.84	96.39
NSM	40.72	61.33	21.83	74.74	29.09	49.86	11.16	58.89	37.97	96.49
ARI	14.69	24.89	13.52	44.13	9	20.38	1.79	26.87	23.95	83.28
AMI	35.94	52.41	20.25	68.61	22.06	39.62	8.09	52.98	49.77	94.07

Table 24 depicts the prediction MAE results when THR=0.5 and the weighted sum function is used. When THR=0.5, the Jaccard Index presents a very low prediction coverage (~1.6% on average), where the rating prediction results are unreliable and hence it does not take part in the accuracy results.

TABLE 24. MAE RESULTS WHEN THR=0.5 AND THE WEIGHTED SUM FUNCTION IS USED

MAE (THR=0.5 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.209	0.603	0.569	0.413	0.488	0.401	0.796	0.898	2.499	0.734
EUCL	0.209	0.603	0.569	0.413	0.488	0.401	0.796	0.898	2.499	0.734
CHEB	0.209	0.603	0.569	0.413	0.488	0.401	0.796	0.898	2.499	0.734
PCC	0.251	0.672	0.653	0.434	0.592	0.467	0.842	0.853	3.049	0.781
CPCC	0.243	0.649	0.636	0.426	0.614	0.507	0.792	0.829	2.533	0.76
SPCC	0.244	0.652	0.636	0.423	0.56	0.444	0.819	0.856	2.904	0.775
COS	0.253	0.684	0.673	0.436	0.663	0.527	0.827	0.83	3.572	0.809
ACOS	0.232	0.639	0.615	0.415	0.539	0.433	0.816	0.875	2.583	0.735
SP	0.201	0.587	0.584	0.375	0.445	0.379	0.743	0.864	2.482	0.768
TAU	0.247	0.671	0.652	0.428	0.626	0.503	0.845	0.841	2.943	0.785
MSD	0.248	0.666	0.657	0.431	0.638	0.517	0.812	0.825	2.819	0.78
NSM	0.251	0.672	0.661	0.433	0.653	0.523	0.817	0.826	3.534	0.787
ARI	0.205	0.602	0.575	0.394	0.347	0.26	0.862	0.934	3.204	0.802
AMI	0.242	0.7	0.687	0.444	0.619	0.479	0.859	0.868	3.101	0.802

Table 25 depicts the prediction RMSE results when THR=0.5 and the weighted sum function is used.

TABLE 25. RMSE RESULTS WHEN THR=0.5 AND THE WEIGHTED SUM FUNCTION IS USED

RMSE (THR=0.5 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.561	1.109	1.01	0.741	1.068	0.892	1.212	1.241	4.024	1.098
EUCL	0.561	1.109	1.01	0.741	1.068	0.892	1.212	1.241	4.024	1.098
CHEB	0.561	1.109	1.01	0.741	1.068	0.892	1.212	1.241	4.024	1.098
PCC	0.618	1.149	1.065	0.753	1.168	0.95	1.226	1.154	4.298	1.088
CPCC	0.601	1.102	1.018	0.73	1.171	0.966	1.168	1.121	3.839	1.051
SPCC	0.605	1.132	1.052	0.742	1.136	0.928	1.217	1.162	4.215	1.082
COS	0.618	1.128	1.043	0.742	1.218	0.985	1.197	1.113	4.601	1.101
ACOS	0.597	1.124	1.032	0.735	1.115	0.923	1.206	1.186	3.916	1.051
SP	0.55	1.074	1.003	0.692	1.004	0.84	1.112	1.178	3.918	1.086
TAU	0.612	1.136	1.047	0.74	1.191	0.97	1.226	1.133	4.167	1.1
MSD	0.61	1.111	1.028	0.736	1.194	0.974	1.181	1.111	3.938	1.066
NSM	0.613	1.116	1.031	0.739	1.208	0.98	1.183	1.111	4.704	1.08
ARI	0.572	1.116	1.023	0.727	0.9	0.708	1.337	1.267	4.705	1.144
AMI	0.608	1.18	1.099	0.768	1.199	0.966	1.251	1.165	4.291	1.121

Table 26 depicts the prediction F1-measure results when THR=0.5 and the mean-centered function is used.

TABLE 26. F1-MEASURE RESULTS WHEN THR=0.5 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (THR=0.5 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.971	0.891	0.895	0.939	0.919	0.939	0.869	0.808	0.177	0.887
EUCL	0.971	0.891	0.895	0.939	0.919	0.939	0.869	0.808	0.177	0.887
CHEB	0.971	0.891	0.895	0.939	0.919	0.939	0.869	0.808	0.177	0.887
PCC	0.964	0.875	0.881	0.935	0.897	0.929	0.837	0.807	0.196	0.873
CPCC	0.966	0.885	0.893	0.939	0.895	0.926	0.861	0.822	0.206	0.885
SPCC	0.966	0.880	0.885	0.938	0.904	0.933	0.853	0.804	0.213	0.873
COS	0.965	0.876	0.883	0.936	0.883	0.922	0.840	0.813	0.223	0.872
ACOS	0.967	0.884	0.891	0.939	0.909	0.934	0.860	0.808	0.245	0.887
SP	0.973	0.896	0.897	0.947	0.927	0.945	0.876	0.796	0.134	0.877
TAU	0.965	0.878	0.885	0.937	0.891	0.925	0.839	0.813	0.097	0.878
MSD	0.966	0.881	0.887	0.937	0.889	0.924	0.847	0.816	0.162	0.879
NSM	0.965	0.880	0.886	0.937	0.886	0.923	0.846	0.817	0.308	0.877
ARI	0.970	0.888	0.890	0.941	0.940	0.960	0.841	0.778	0.236	0.861
AMI	0.965	0.867	0.871	0.933	0.890	0.926	0.837	0.799	0.165	0.870

Table 27 depicts the prediction coverage results when THR=0.5 and the weighted sum function is used.

TABLE 27. PREDICTION COVERAGE RESULTS WHEN THR=0.5 AND THE WEIGHTED SUM FUNCTION IS USED

coverage (THR=0.5 & weighted sum)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	3.04	0.62	1.94	2.48	2.05	7.8	0	0.01	0.11	7.51
MANH	29.41	25.25	11.58	49.04	11.02	25.14	1.8	23.44	26.88	78.95
EUCL	29.41	25.25	11.58	49.04	11.02	25.14	1.8	23.44	26.88	78.95
CHEB	29.41	25.25	11.58	49.04	11.02	25.14	1.8	23.44	26.88	78.95
PCC	31.49	45.56	18.4	64.02	18.21	34.25	5.8	46.48	37.92	90.09
CPCC	39.77	55.82	20.86	73.16	24.91	46.87	8.98	52.81	48.58	95.49
SPCC	30.5	41.78	17.47	61.14	16.41	32.08	4.57	42.68	35.32	89.24
COS	40.73	61.42	21.84	74.75	29.15	49.88	11.19	58.94	36.97	96.56
ACOS	28.47	39.31	17.33	60.81	15.37	30.73	4.36	39.9	40.29	90.59
SP	22.27	30.75	15	51.91	11.87	25.87	2.23	30.31	40.55	90.4
TAU	38.86	55.84	20.79	71.89	24.63	45.2	8.5	52.8	57.01	95.45
MSD	40.28	58.98	21.53	74.22	26.95	48.21	10.48	57.22	57.25	96.18
NSM	40.62	60.47	21.7	74.6	28.47	49.5	10.89	58.1	24.47	96.32
ARI	12.87	16.37	10.39	36.01	6.43	15.26	0.99	20.32	15.14	76.09
AMI	34.8	46.72	18.52	63.88	19.06	35.14	6.59	48.26	44.74	91.61

2.1.4 Experiments using the similarity threshold neighbor selection with the mean-centered prediction function

Table 28 depicts the prediction MAE results when THR=0.0 and the mean-centered function is used.

TABLE 28. MAE RESULTS WHEN THR=0.0 AND THE MEAN-CENTERED FUNCTION IS USED

MAE (THR=0.0 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.299	0.674	0.584	0.447	0.643	0.494	0.691	0.832	2.627	0.718
MANH	0.305	0.691	0.639	0.467	0.693	0.543	0.695	0.839	2.666	0.725
EUCL	0.304	0.687	0.636	0.464	0.692	0.543	0.693	0.833	2.652	0.722
CHEB	0.304	0.686	0.635	0.462	0.691	0.543	0.691	0.831	2.647	0.721
PCC	0.313	0.684	0.633	0.461	0.661	0.521	0.718	0.837	2.693	0.721
CPCC	0.3	0.672	0.63	0.455	0.67	0.542	0.684	0.827	2.601	0.714
SPCC	0.312	0.681	0.629	0.458	0.659	0.517	0.717	0.836	2.689	0.719
COS	0.304	0.683	0.64	0.458	0.692	0.551	0.691	0.83	2.866	0.722
ACOS	0.295	0.668	0.625	0.453	0.621	0.491	0.696	0.852	2.576	0.714
SP	0.268	0.639	0.606	0.425	0.571	0.462	0.723	0.843	2.529	0.71
TAU	0.302	0.682	0.637	0.457	0.678	0.54	0.706	0.833	2.629	0.72
MSD	0.302	0.678	0.637	0.457	0.684	0.547	0.689	0.827	2.628	0.718
NSM	0.303	0.682	0.638	0.458	0.691	0.55	0.69	0.829	2.867	0.72
ARI	0.279	0.652	0.606	0.432	0.539	0.412	0.768	0.872	2.808	0.735
AMI	0.298	0.687	0.641	0.459	0.667	0.52	0.716	0.843	2.694	0.723

Table 29 depicts the prediction RMSE results when THR=0.0 and the mean-centered function is used.

TABLE 29. RMSE RESULTS WHEN THR=0.0 AND THE MEAN-CENTERED FUNCTION IS USED

RMSE (THR=0.0 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.602	1.057	0.944	0.713	1.126	0.912	0.978	1.123	3.616	1.016
MANH	0.608	1.074	0.983	0.733	1.172	0.95	0.981	1.128	3.689	1.022
EUCL	0.607	1.066	0.975	0.728	1.167	0.945	0.978	1.12	3.651	1.017
CHEB	0.606	1.063	0.973	0.725	1.166	0.944	0.977	1.117	3.625	1.016
PCC	0.619	1.073	0.984	0.729	1.145	0.934	1.005	1.131	3.698	1.016
CPCC	0.599	1.051	0.969	0.716	1.144	0.94	0.969	1.116	3.599	1.007
SPCC	0.618	1.07	0.979	0.726	1.142	0.93	1.003	1.129	3.69	1.014
COS	0.604	1.056	0.976	0.72	1.162	0.946	0.977	1.114	3.842	1.019
ACOS	0.6	1.068	0.983	0.722	1.109	0.914	0.976	1.153	3.634	1.011
SP	0.566	1.033	0.961	0.69	1.055	0.868	1.009	1.144	3.534	1.007
TAU	0.602	1.06	0.976	0.72	1.153	0.938	0.993	1.122	3.577	1.016
MSD	0.602	1.053	0.973	0.719	1.155	0.944	0.975	1.112	3.583	1.013
NSM	0.604	1.055	0.973	0.72	1.161	0.945	0.975	1.113	3.859	1.016
ARI	0.592	1.059	0.977	0.705	1.033	0.826	1.075	1.181	3.912	1.035
AMI	0.602	1.075	0.989	0.728	1.15	0.931	1.002	1.133	3.708	1.024

Table 30 depicts the prediction F1-measure results when THR=0.0 and the mean-centered function is used.

TABLE 30. F1-MEASURE RESULTS WHEN THR=0.0 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (THR=0.0 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.965	0.883	0.900	0.938	0.893	0.928	0.861	0.820	0.269	0.880
MANH	0.965	0.879	0.892	0.935	0.885	0.922	0.863	0.818	0.279	0.878
EUCL	0.965	0.881	0.894	0.936	0.885	0.923	0.863	0.819	0.281	0.879
CHEB	0.965	0.881	0.894	0.936	0.885	0.922	0.862	0.820	0.279	0.879
PCC	0.962	0.880	0.892	0.936	0.891	0.925	0.855	0.817	0.292	0.877
CPCC	0.966	0.885	0.896	0.938	0.891	0.923	0.865	0.822	0.263	0.881
SPCC	0.963	0.881	0.893	0.936	0.891	0.925	0.855	0.817	0.293	0.877
COS	0.965	0.882	0.893	0.937	0.885	0.922	0.861	0.821	0.311	0.879
ACOS	0.965	0.885	0.894	0.938	0.901	0.930	0.862	0.814	0.293	0.881
SP	0.970	0.892	0.899	0.944	0.911	0.937	0.856	0.810	0.263	0.880
TAU	0.965	0.883	0.894	0.937	0.888	0.923	0.857	0.820	0.265	0.880
MSD	0.965	0.884	0.894	0.937	0.887	0.922	0.862	0.821	0.273	0.880
NSM	0.965	0.883	0.894	0.937	0.885	0.922	0.862	0.821	0.314	0.879
ARI	0.966	0.886	0.896	0.941	0.914	0.942	0.818	0.798	0.278	0.870
AMI	0.965	0.879	0.890	0.936	0.889	0.925	0.853	0.814	0.280	0.878

As far as the prediction coverage when THR=0.0 is concerned, it is the exact same, when the weighted sum prediction function was used (TABLE 19).

Table 31 depicts the prediction MAE results when THR=0.25 and the mean-centered function is used. When THR=0.25, the Jaccard Index presents a very low prediction coverage (~5.5% on average), where the rating prediction results are unreliable and hence it does not take part in the accuracy results.

TABLE 31. MAE RESULTS WHEN THR=0.25 AND THE MEAN-CENTERED FUNCTION IS USED

MAE (THR=0.25 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.294	0.677	0.634	0.468	0.641	0.519	0.686	0.862	2.587	0.729
EUCL	0.299	0.68	0.638	0.464	0.656	0.528	0.673	0.848	2.601	0.72
CHEB	0.299	0.677	0.635	0.459	0.658	0.527	0.666	0.838	2.601	0.713
PCC	0.31	0.68	0.633	0.459	0.647	0.51	0.716	0.839	2.718	0.72
CPCC	0.299	0.67	0.629	0.454	0.667	0.54	0.677	0.827	2.553	0.713
SPCC	0.306	0.676	0.627	0.455	0.638	0.501	0.714	0.839	2.704	0.718
COS	0.304	0.683	0.64	0.458	0.692	0.551	0.691	0.83	2.925	0.722
ACOS	0.292	0.662	0.621	0.45	0.601	0.476	0.693	0.859	2.537	0.716
SP	0.262	0.631	0.601	0.421	0.544	0.443	0.71	0.847	2.479	0.708
TAU	0.302	0.682	0.637	0.457	0.676	0.54	0.703	0.834	2.626	0.72
MSD	0.302	0.677	0.636	0.457	0.682	0.547	0.685	0.827	2.619	0.718
NSM	0.303	0.681	0.637	0.458	0.69	0.549	0.69	0.828	3.042	0.72
ARI	0.269	0.634	0.589	0.427	0.477	0.367	0.729	0.907	2.876	0.748
AMI	0.297	0.692	0.647	0.461	0.667	0.517	0.726	0.852	2.755	0.728

Table 32 depicts the prediction RMSE results when THR=0.25 and the mean-centered function is used.

TABLE 32. RMSE RESULTS WHEN THR=0.25 AND THE MEAN-CENTERED FUNCTION IS USED

RMSE (THR=0.25 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.593	1.077	0.998	0.737	1.127	0.937	0.972	1.168	3.746	1.029
EUCL	0.599	1.073	0.993	0.729	1.138	0.939	0.955	1.143	3.753	1.016
CHEB	0.599	1.067	0.985	0.723	1.138	0.935	0.944	1.13	3.752	1.007
PCC	0.613	1.071	0.986	0.726	1.131	0.925	0.997	1.134	3.73	1.015
CPCC	0.599	1.049	0.968	0.716	1.14	0.937	0.962	1.116	3.566	1.006
SPCC	0.609	1.066	0.981	0.722	1.122	0.915	0.995	1.134	3.721	1.012
COS	0.604	1.056	0.976	0.72	1.162	0.946	0.977	1.114	3.892	1.019
ACOS	0.595	1.063	0.981	0.718	1.088	0.897	0.973	1.163	3.614	1.014
SP	0.556	1.028	0.958	0.686	1.026	0.846	0.996	1.15	3.497	1.005
TAU	0.601	1.06	0.977	0.72	1.15	0.938	0.992	1.123	3.574	1.015
MSD	0.602	1.051	0.972	0.719	1.153	0.943	0.969	1.111	3.577	1.012
NSM	0.604	1.054	0.973	0.72	1.16	0.945	0.975	1.113	4.026	1.015
ARI	0.575	1.049	0.967	0.699	0.965	0.781	1.04	1.229	4.029	1.053
AMI	0.6	1.083	1.001	0.731	1.153	0.933	1.016	1.144	3.766	1.03

Table 33 depicts the prediction F1-measure results when THR=0.25 and the mean-centered function is used.

TABLE 33. F1-MEASURE RESULTS WHEN THR=0.25 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (THR=0.25 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.966	0.884	0.892	0.936	0.897	0.926	0.870	0.816	0.224	0.879
EUCL	0.966	0.883	0.892	0.936	0.893	0.925	0.872	0.817	0.242	0.880
CHEB	0.966	0.883	0.893	0.937	0.893	0.925	0.870	0.819	0.247	0.881
PCC	0.963	0.881	0.892	0.936	0.894	0.927	0.859	0.816	0.171	0.877
CPCC	0.966	0.886	0.896	0.938	0.892	0.924	0.868	0.823	0.190	0.881
SPCC	0.963	0.882	0.893	0.937	0.896	0.929	0.860	0.816	0.172	0.877
COS	0.965	0.882	0.893	0.937	0.885	0.922	0.861	0.821	0.190	0.879
ACOS	0.966	0.887	0.895	0.939	0.905	0.933	0.865	0.813	0.222	0.881
SP	0.971	0.894	0.900	0.945	0.916	0.940	0.865	0.808	0.121	0.880
TAU	0.965	0.883	0.894	0.937	0.889	0.923	0.858	0.820	0.106	0.880
MSD	0.965	0.884	0.894	0.937	0.887	0.922	0.863	0.822	0.140	0.880
NSM	0.965	0.883	0.894	0.937	0.886	0.922	0.862	0.821	0.230	0.880
ARI	0.968	0.891	0.898	0.942	0.926	0.949	0.838	0.792	0.208	0.868
AMI	0.965	0.878	0.888	0.936	0.889	0.925	0.852	0.812	0.159	0.877

As far as the prediction coverage when THR=0.25 is concerned, it is the exact same, when the weighted sum prediction function was used (TABLE 23).

Table 34 depicts the prediction MAE results when THR=0.5 and the mean-centered function is used. When THR=0.5, the Jaccard Index presents a very low prediction coverage (~1.6% on average), where the rating prediction results are unreliable and hence it does not take part in the accuracy results.

TABLE 34. MAE RESULTS WHEN THR=0.5 AND THE MEAN-CENTERED FUNCTION IS USED

MAE (THR=0.5 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.267	0.634	0.582	0.454	0.539	0.431	0.684	0.898	2.491	0.744
EUCL	0.267	0.634	0.582	0.454	0.539	0.431	0.684	0.898	2.491	0.744
CHEB	0.267	0.634	0.582	0.454	0.539	0.431	0.684	0.898	2.491	0.744
PCC	0.304	0.679	0.633	0.459	0.628	0.493	0.709	0.853	2.722	0.725
CPCC	0.296	0.663	0.625	0.452	0.65	0.533	0.672	0.829	2.494	0.71
SPCC	0.295	0.663	0.62	0.45	0.599	0.471	0.691	0.856	2.687	0.722
COS	0.304	0.683	0.64	0.458	0.692	0.551	0.692	0.83	3.05	0.722
ACOS	0.288	0.659	0.618	0.45	0.582	0.462	0.695	0.875	2.524	0.722
SP	0.252	0.61	0.582	0.412	0.492	0.41	0.654	0.864	2.364	0.71
TAU	0.301	0.682	0.637	0.457	0.664	0.534	0.7	0.841	2.619	0.721
MSD	0.3	0.672	0.634	0.455	0.671	0.542	0.683	0.825	2.595	0.715
NSM	0.302	0.677	0.635	0.457	0.683	0.547	0.687	0.826	3.178	0.718
ARI	0.259	0.623	0.571	0.431	0.396	0.286	0.765	0.934	2.956	0.761
AMI	0.295	0.701	0.657	0.469	0.654	0.504	0.721	0.868	2.744	0.735

Table 35 depicts the prediction RMSE results when THR=0.5 and the mean-centered function is used.

TABLE 35. RMSE RESULTS WHEN THR=0.5 AND THE MEAN-CENTERED FUNCTION IS USED

RMSE (THR=0.5 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	-	-	-	-	-	-	-	-	-	-
MANH	0.558	1.052	0.96	0.725	1.029	0.858	0.964	1.241	3.674	1.057
EUCL	0.558	1.052	0.96	0.725	1.029	0.858	0.964	1.241	3.674	1.057
CHEB	0.558	1.052	0.96	0.725	1.029	0.858	0.964	1.241	3.674	1.057
PCC	0.608	1.076	0.994	0.729	1.115	0.909	0.992	1.154	3.774	1.023
CPCC	0.593	1.044	0.967	0.713	1.124	0.93	0.957	1.121	3.545	1.002
SPCC	0.596	1.063	0.984	0.72	1.087	0.891	0.984	1.162	3.774	1.021
COS	0.604	1.056	0.976	0.72	1.162	0.946	0.977	1.113	4.023	1.018
ACOS	0.59	1.066	0.984	0.72	1.072	0.888	0.989	1.186	3.635	1.023
SP	0.542	1.016	0.948	0.676	0.97	0.814	0.92	1.178	3.444	1.01
TAU	0.599	1.064	0.981	0.72	1.139	0.933	0.991	1.133	3.57	1.016
MSD	0.599	1.048	0.971	0.717	1.143	0.937	0.966	1.111	3.569	1.009
NSM	0.602	1.05	0.971	0.719	1.155	0.943	0.971	1.111	4.207	1.013
ARI	0.563	1.046	0.957	0.704	0.871	0.685	1.103	1.267	4.173	1.072
AMI	0.597	1.101	1.02	0.742	1.144	0.925	1.017	1.165	3.773	1.04

Table 36 depicts the prediction F1-measure results when THR=0.5 and the mean-centered function is used.

TABLE 36. F1-MEASURE RESULTS WHEN THR=0.5 AND THE MEAN-CENTERED FUNCTION IS USED

F1-measure (THR=0.5 & mean-centered)	Digital_music	Videogames	Movies	Kindle	Cell_phones	Musical Instruments	Ciao	Epinions	Book-crossing	Yahoo-Movies
JACC	0.971	0.895	0.900	0.940	0.918	0.940	0.881	0.812	0.205	0.877
MANH	0.971	0.895	0.900	0.940	0.918	0.940	0.881	0.812	0.205	0.877
EUCL	0.971	0.895	0.900	0.940	0.918	0.940	0.881	0.812	0.205	0.877
CHEB	0.964	0.881	0.890	0.936	0.898	0.930	0.862	0.813	0.294	0.877
PCC	0.966	0.888	0.897	0.939	0.895	0.925	0.873	0.824	0.267	0.882
CPCC	0.965	0.885	0.893	0.938	0.904	0.933	0.869	0.811	0.291	0.877
SPCC	0.965	0.882	0.893	0.937	0.885	0.922	0.861	0.821	0.330	0.879
COS	0.967	0.888	0.895	0.939	0.908	0.935	0.866	0.810	0.303	0.881
ACOS	0.973	0.899	0.902	0.947	0.927	0.945	0.883	0.804	0.250	0.880
SP	0.966	0.883	0.893	0.938	0.892	0.924	0.861	0.818	0.260	0.880
TAU	0.966	0.885	0.895	0.938	0.890	0.923	0.864	0.822	0.264	0.880
MSD	0.965	0.884	0.895	0.937	0.887	0.923	0.864	0.822	0.364	0.880
NSM	0.969	0.892	0.898	0.942	0.939	0.960	0.836	0.786	0.289	0.865
ARI	0.966	0.875	0.884	0.934	0.891	0.927	0.855	0.808	0.280	0.876
AMI	0.971	0.895	0.900	0.940	0.918	0.940	0.881	0.812	0.205	0.877

As far as the prediction coverage when THR=0.5 is concerned, it is the exact same, when the weighted sum prediction function was used (TABLE 27)

3.   Analysis of metric behaviour in relation to sparsity

In this section, we present an experiment aiming to provide insight on the behaviour of similarity metrics in relation to sparsity. To this end, we have used the Yahoo!movies and the Ciao datasets. The datasets were subsampled at subsets with size equal to 10%, 30%, 50%, 70% and 90% of the full dataset and tested the performance of the metrics in these subsets. This approach was taken due to the fact that different datasets have been found to exhibit divergent results. Therefore, in order to warrant independence from the individual dataset characteristics, we do not directly compare the results of different datasets, but rather subsets of the same dataset.

The results of the experiments, concerning the MAE, RMSE and F1 measure observed are listed in the following tables. Each tables lists the improvements of the metric (MAE, RMSE, F1) observed when the sparsity of the dataset decreases. In all cases, the baseline against which the improvement is computed is the respective performance when the subsampling ratio is 10%.

TABLE 37. MAE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE YAHOO!MOVIES DATASET, WHEN THE TOP-K NNS ARE SELECTED, USING K=250 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

MAE (Yahoo movies, Top-K, K=250)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	4.99%	4.20%	3.57%	5.44%	2.57%	3.54%	4.98%	3.67%	4.04%	4.75%	3.91%	3.12%	4.51%	3.33%	3.92%
50%	5.64%	3.97%	5.48%	7.05%	2.97%	2.82%	6.86%	3.50%	4.60%	4.63%	5.97%	2.80%	4.04%	4.22%	5.48%
70%	5.14%	3.79%	5.55%	9.24%	4.94%	5.60%	10.73%	6.87%	7.33%	6.78%	8.77%	4.09%	6.12%	4.65%	8.56%
90%	15.47%	8.91%	10.14%	13.63%	8.27%	6.65%	15.61%	10.14%	10.21%	9.92%	11.84%	6.58%	10.84%	7.96%	13.37%
100%	17.00%	10.53%	10.53%	15.38%	8.63%	7.68%	17.75%	10.45%	10.72%	11.24%	13.79%	7.31%	10.91%	9.14%	13.92%

TABLE 38. RMSE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE YAHOO!MOVIES DATASET, WHEN THE TOP-K NNS ARE SELECTED, USING K=250 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

RMSE (Yahoo movies, Top-K, K=250)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	5.65%	4.59%	4.22%	3.69%	4.25%	2.59%	4.31%	3.00%	3.30%	3.40%	4.37%	3.67%	3.31%	3.02%	3.30%
50%	5.78%	4.11%	5.37%	5.45%	3.39%	3.03%	5.80%	5.70%	3.79%	3.77%	6.52%	3.19%	5.04%	4.69%	3.87%
70%	8.94%	7.08%	6.45%	8.76%	5.73%	5.29%	8.93%	5.88%	6.87%	7.17%	8.75%	5.71%	6.51%	5.46%	6.88%
90%	15.19%	10.25%	10.79%	11.84%	9.64%	7.76%	13.82%	10.48%	8.87%	10.16%	13.42%	7.72%	10.02%	8.92%	10.38%
100%	16.82%	11.06%	11.06%	14.58%	9.87%	8.71%	16.57%	11.63%	10.71%	10.74%	14.93%	7.60%	11.43%	9.31%	11.00%

TABLE 39. F1 IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE YAHOO!MOVIES DATASET, WHEN THE TOP-K NNS ARE SELECTED, USING K=250 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

F1 (Yahoo movies, Top-K, K=250)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	1.00%	0.65%	0.72%	1.02%	1.47%	1.12%	1.15%	1.04%	1.03%	1.33%	0.92%	0.35%	1.02%	1.76%	0.58%
50%	0.81%	0.68%	0.74%	0.87%	1.60%	1.09%	1.06%	0.92%	1.01%	1.43%	0.83%	0.34%	1.14%	1.72%	0.54%
70%	1.24%	0.99%	0.92%	1.22%	2.01%	1.35%	1.33%	1.23%	1.29%	1.87%	1.15%	0.31%	1.32%	2.34%	0.83%
90%	1.77%	1.38%	1.32%	1.91%	2.88%	2.09%	2.03%	1.99%	2.02%	2.67%	1.49%	0.42%	1.96%	3.50%	1.08%
100%	3.20%	2.56%	2.57%	3.47%	5.62%	3.88%	3.94%	3.66%	3.83%	5.06%	2.95%	0.83%	3.85%	6.75%	2.01%

TABLE 40. MAE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE YAHOO!MOVIES DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0.5 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

MAE (Yahoo movies, threshold, THR=0.5)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%		4.04%	3.08%	4.47%	3.49%	5.06%	3.63%	1.77%	3.04%	4.22%	3.85%	2.79%	4.85%	0.84%	3.48%
50%		4.23%	5.06%	4.25%	4.69%	5.33%	6.46%	2.47%	4.57%	5.92%	5.46%	4.43%	5.76%	2.61%	5.86%
70%		5.45%	5.45%	6.44%	6.74%	7.54%	8.30%	4.20%	6.52%	8.24%	8.32%	6.30%	6.90%	2.98%	7.27%
90%	8.88%	8.57%	6.78%	8.36%	8.45%	9.17%	10.71%	3.77%	9.59%	10.22%	10.13%	7.17%	9.36%	3.06%	9.27%
100%	22.21%	9.93%	9.93%	12.37%	12.55%	12.45%	13.84%	5.50%	11.74%	14.04%	14.17%	9.03%	11.47%	3.18%	12.81%

TABLE 41. RMSE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE YAHOO!MOVIES DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0.5 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

RMSE (Yahoo movies, threshold, THR=0.5)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%		3.90%	2.93%	3.41%	5.24%	4.10%	4.99%	3.13%	2.88%	4.95%	3.86%	3.65%	4.06%	2.69%	4.27%
50%		3.64%	4.64%	5.27%	5.39%	5.59%	4.81%	2.62%	5.07%	5.60%	6.38%	5.00%	4.61%	1.61%	4.64%
70%		5.16%	5.41%	6.65%	7.02%	6.77%	8.45%	4.51%	6.62%	7.28%	8.12%	5.84%	7.82%	2.04%	5.62%
90%	12.33%	7.79%	7.22%	8.17%	9.26%	8.79%	9.55%	5.27%	8.13%	10.73%	10.39%	6.44%	9.23%	3.39%	8.96%
100%	30.84%	9.97%	9.97%	10.27%	13.52%	13.10%	13.84%	6.09%	11.35%	13.97%	14.84%	9.51%	11.99%	3.86%	10.81%

TABLE 42. F1 IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE YAHOO!MOVIES DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0.5 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

F1 (Yahoo movies, threshold, THR=0.5)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%		0.86%	0.92%	1.12%	1.87%	1.39%	1.22%	1.13%	1.40%	1.54%	1.16%	0.51%	1.28%	1.77%	0.87%
50%		0.96%	0.82%	1.01%	1.85%	1.39%	1.18%	1.20%	1.30%	1.64%	1.14%	0.54%	1.22%	1.78%	0.83%
70%		1.18%	1.24%	1.50%	2.44%	1.77%	1.60%	1.55%	1.67%	2.06%	1.53%	0.60%	1.67%	2.45%	1.04%
90%	5.61%	1.58%	1.70%	2.07%	3.64%	2.68%	2.16%	2.24%	2.64%	3.06%	2.22%	0.93%	2.55%	3.58%	1.54%
100%	14.02%	3.15%	3.15%	4.01%	6.89%	4.99%	4.25%	4.40%	4.89%	6.03%	4.30%	1.65%	4.75%	6.77%	2.86%

TABLE 43. MAE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE CIAO DATASET, WHEN THE TOP-K NNS ARE SELECTED, USING K=250 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

MAE (Ciao, threshold, Top-K, K=250)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	3.38%	2.77%	2.66%	6.25%	2.81%	3.39%	3.95%	4.92%	1.06%	2.12%	5.91%	2.71%	3.34%	4.43%	2.97%
50%	4.37%	3.83%	3.58%	6.99%	3.00%	2.94%	4.10%	7.48%	3.06%	3.76%	6.17%	2.80%	3.64%	4.76%	4.69%
70%	6.02%	4.30%	5.61%	11.92%	3.09%	3.85%	5.46%	9.29%	2.46%	4.76%	10.86%	3.65%	4.74%	6.89%	6.23%
90%	6.16%	5.24%	5.57%	14.70%	2.49%	4.68%	8.90%	12.83%	3.59%	6.06%	12.85%	5.02%	5.61%	8.67%	7.42%
100%	8.13%	7.75%	8.27%	20.16%	3.27%	7.00%	10.74%	18.27%	3.49%	6.45%	17.82%	6.63%	7.41%	12.50%	8.89%

TABLE 44. RMSE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE CIAO DATASET, WHEN THE TOP-K NNS ARE SELECTED, USING K=250 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

RMSE (Ciao, threshold, Top-K, K=250)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	2.63%	3.83%	3.55%	5.33%	2.76%	3.98%	2.96%	6.34%	3.16%	3.00%	5.81%	3.52%	1.91%	4.47%	2.11%
50%	3.25%	3.52%	4.70%	7.81%	2.90%	4.12%	4.95%	7.34%	3.20%	2.71%	8.29%	2.12%	3.38%	5.71%	4.79%
70%	5.67%	4.52%	5.24%	10.79%	4.24%	4.04%	6.70%	10.19%	3.89%	4.27%	12.23%	4.27%	4.60%	5.85%	4.34%
90%	6.99%	5.48%	6.67%	14.56%	4.19%	6.34%	8.65%	14.45%	4.10%	4.94%	14.24%	5.77%	6.44%	8.23%	6.92%
100%	7.92%	7.92%	8.64%	21.25%	4.85%	7.93%	11.27%	19.75%	6.06%	7.34%	20.48%	6.33%	7.60%	11.24%	8.40%

TABLE 45. F1 IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE CIAO DATASET, WHEN THE TOP-K NNS ARE SELECTED, USING K=250 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

F1 (Ciao, threshold, Top-K, K=250)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	1.14%	0.55%	0.61%	0.74%	0.82%	0.86%	0.23%	0.65%	0.47%	0.90%	0.87%	0.91%	0.80%	0.29%	0.22%
50%	1.02%	0.60%	0.47%	0.84%	0.72%	0.76%	0.13%	0.68%	0.57%	0.89%	0.96%	0.91%	0.68%	0.26%	0.20%
70%	1.43%	0.67%	0.62%	0.95%	0.94%	1.01%	0.31%	0.98%	0.60%	1.19%	1.21%	1.00%	0.96%	0.36%	0.21%
90%	2.14%	1.05%	0.95%	1.35%	1.48%	1.46%	0.18%	1.30%	0.81%	1.67%	1.77%	1.52%	1.53%	0.46%	0.17%
100%	3.97%	1.73%	1.76%	2.58%	2.59%	2.66%	0.34%	2.46%	1.58%	2.96%	3.41%	3.00%	2.70%	0.65%	0.09%

TABLE 46. MAE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE CIAO DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0.5 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

MAE (Ciao, threshold, threshold, THR=0.5)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	2.76%	2.46%	4.26%	2.55%	3.80%	2.96%	4.32%	2.24%	3.71%	2.39%	2.04%	2.56%	2.13%	3.53%	2.76%
50%	3.82%	3.79%	4.24%	3.03%	2.81%	3.46%	5.39%	3.35%	5.90%	2.84%	2.55%	3.28%	3.24%	4.04%	3.82%
70%	4.70%	5.99%	5.94%	3.48%	5.15%	5.53%	6.26%	4.56%	8.44%	5.88%	3.83%	5.20%	4.33%	4.86%	4.70%
90%	7.35%	7.69%	9.03%	3.93%	6.41%	7.35%	7.41%	5.48%	9.92%	7.26%	5.48%	5.39%	5.96%	7.48%	7.35%
100%	8.92%	8.92%	11.86%	4.58%	7.69%	10.26%	10.59%	6.59%	12.92%	8.50%	6.44%	6.40%	7.16%	9.08%	8.92%

TABLE 47. RMSE IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE CIAO DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0.5 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

RMSE (Ciao, threshold, threshold, THR=0.5)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	4.78%	3.02%	4.73%	3.24%	3.92%	3.32%	3.78%	1.77%	4.53%	2.35%	2.06%	2.88%	1.24%	2.21%	4.78%
50%	5.49%	5.48%	5.58%	2.83%	3.43%	4.01%	4.78%	2.19%	6.69%	5.04%	2.15%	3.48%	1.86%	2.56%	5.49%
70%	5.97%	6.23%	7.62%	4.47%	5.15%	5.96%	5.05%	3.23%	7.98%	5.66%	5.07%	4.29%	3.08%	5.24%	5.97%
90%	8.97%	8.67%	8.81%	5.70%	5.63%	7.67%	6.91%	5.67%	10.39%	6.91%	5.82%	6.21%	4.17%	4.81%	8.97%
100%	11.32%	11.32%	12.68%	5.61%	7.98%	9.97%	9.79%	5.63%	15.36%	9.33%	6.30%	6.90%	4.50%	6.53%	11.32%

TABLE 48. F1 IMPROVEMENT ACHIEVED BY THE 15 METRICS, AT DIFFERENT SUBSAMPLING RATIOS OF THE CIAO DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0.5 (BASE: METRIC PERFORMANCE FOR SUBSAMPLING RATIO=10%)

F1 (Ciao, threshold, threshold, THR=0.5)	jacc	manh	eucl	cheb	pcc	cpcc	spcc	cos	acos	sp	tau	msd	nsm	ari	ami
30%	0.42%	0.43%	0.71%	1.13%	1.48%	0.54%	0.61%	1.90%	0.37%	1.45%	0.53%	0.86%	0.41%	0.45%	0.42%
50%	0.25%	0.32%	0.72%	1.13%	1.49%	0.61%	0.48%	1.89%	0.43%	1.38%	0.51%	0.85%	0.30%	0.36%	0.25%
70%	0.50%	0.44%	1.04%	1.41%	2.12%	0.80%	0.77%	2.31%	0.45%	1.94%	0.64%	0.94%	0.42%	0.47%	0.50%
90%	0.54%	0.56%	1.33%	1.97%	3.05%	1.13%	1.12%	3.64%	0.71%	2.83%	0.94%	1.52%	0.53%	0.65%	0.54%
100%	0.98%	0.98%	2.51%	3.83%	5.79%	2.06%	1.85%	6.93%	1.28%	5.40%	1.75%	2.82%	1.02%	1.01%	0.98%

4. Implicit feedback

In this subsection we report on experiments aiming to provide insight on the behaviour of the metrics when ratings are not directly provided by users, but rather computed on the basis of implicit feedback. In particular, the following experiments were conucted: the Last.FM dataset was obtained, and song listening frequencies were used as traits for the estimation of the ratings. Subsequently, the dataset was subsampled to increase sparsity. Finally, rating prediction experiments were conducted on the subsampled dataset, to assess the performance of each similarity metric in this context. The results obtained for the different experiments are depicted in the following tables.

TABLE 49. MAE, RMSE AND F1 ACHIEVED BY THE METRICS, USING THE PROCESSED AND SUBSAMPLED LAST.FM DATASET, WHEN THE TOP-K METHOD IS USED FOR NN SELECTION, USING K=250

	MAE	RMSE	F1
JACC	0.339	0.703	0.395
EUCL	0.347	0.72	0.319
CHEB	0.345	0.713	0.34
PCC	0.342	0.707	0.368
CPCC	0.346	0.713	0.344
SPCC	0.342	0.707	0.368
COS	0.346	0.713	0.343
SP	0.346	0.642	0.451
TAU	0.334	0.649	0.378
MSD	0.346	0.714	0.342
NSM	0.344	0.711	0.351
ARI	0.334	0.649	0.395
AMI	0.347	0.729	0.417

TABLE 50. MAE, RMSE AND F1 ACHIEVED BY THE METRICS, USING THE PROCESSED AND SUBSAMPLED LAST.FM DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0

	MAE	RMSE	F1
JACC	0.339	0.703	0.395
EUCL	0.347	0.72	0.319
CHEB	0.345	0.713	0.34
PCC	0.342	0.707	0.368
CPCC	0.346	0.713	0.344
SPCC	0.342	0.707	0.368
COS	0.346	0.713	0.343
SP	0.346	0.642	0.451
TAU	0.334	0.649	0.378
MSD	0.346	0.714	0.342
NSM	0.344	0.711	0.351
ARI	0.334	0.649	0.395
AMI	0.347	0.729	0.417

TABLE 51. MAE, RMSE AND F1 ACHIEVED BY THE METRICS, USING THE PROCESSED AND SUBSAMPLED LAST.FM DATASET, WHEN THE THRESHOLD METHOD IS USED FOR NN SELECTION, USING THR=0.5

	MAE	RMSE	F1
JACC	0.339	0.703	0.395
EUCL	0.347	0.72	0.319
CHEB	0.345	0.713	0.34
PCC	0.342	0.707	0.368
CPCC	0.346	0.713	0.344
SPCC	0.342	0.707	0.368
COS	0.346	0.713	0.343
SP	0.346	0.642	0.451
TAU	0.334	0.649	0.378
MSD	0.346	0.714	0.342
NSM	0.344	0.711	0.351
ARI	0.334	0.649	0.395
AMI	0.347	0.729	0.417

5. Conclusions

In this report we have presented the experimental findings from applying 15 user similarity metrics in 10 sparse CF datasets. Furthermore, we used 2 NNs selection methods and 2 rating prediction formulation equations.

6. References

[1] Fkih, F. (2022). Similarity measures for Collaborative Filtering-based Recommender Systems: Review and experimental comparison. Journal of King Saud University-Computer and Information Sciences, 34(9), 7645-7669.

[2] Al-Shamri, M. (2022). Similarity Measures for Recommender Systems: Drawbacks and Neighbors Formation.

[3] Jain, G., Mahara, T., & Tripathi, K. N. (2020). A survey of similarity measures for collaborative filtering-based recommender system. In Soft Computing: Theories and Applications: Proceedings of SoCTA 2018 (pp. 343-352). Springer Singapore.

[4] Stephen, S. C., Xie, H., & Rai, S. (2017, July). Measures of similarity in memory-based collaborative filtering recommender system: A comparison. In Proceedings of the 4th multidisciplinary international social networks conference (pp. 1-8).

[5] Shardanand, U., & Maes, P. (1995, May). Social information filtering: Algorithms for automating "word of mouth". In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 210-217).

[6] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994, October). Grouplens: An open architecture for collaborative filtering of netnews. In Proceedings of the 1994 ACM conference on Computer supported cooperative work (pp. 175-186).

[7] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001, April). Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th international conference on World Wide Web (pp. 285-295).

[8] Aggarwal, C. C. (2016). Recommender systems (Vol. 1). Cham: Springer International Publishing.

[9] R4 - Yahoo! Movies User Ratings and Descriptive Content Information, v.1.0. . Available online: https://webscope.sandbox.yahoo.com/catalog.php?datatype=r (accessed on 11 December 2023).

[10] Book-Crossing Dataset. Available online: https://www.kaggle.com/datasets/somnambwl/bookcrossing-dataset (accessed on 11 December 2023).

[11] Amazon Review Data (2018). Available online: https://nijianmo.github.io/amazon/index.html (accessed on 11 December 2023).

[12] Epinions (trust network). Available online: https://www.kaggle.com/datasets/masoud3/epinions-trust-network (accessed on 11 December 2023).

[13] CiaoDVD dataset. Available online: https://guoguibing.github.io/librec/datasets.html (accessed on 11 December 2023).