Table 1: Test performance AlphaRec vs TFCEMLP (mean ± std) over 5 seeds. All differences between models are significant (p-value < 0.05). Statistical significance was evaluated based on the recall metric.

Model	Book			Game			Movie		
	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG
AlphaRec	0.0992 ± 0.0002	0.4183 ± 0.0012	0.0832 ± 0.0002	0.1511 ± 0.0006	0.3188 ± 0.0012	0.0888 ± 0.0004	0.1221 ± 0.0002	0.5567 ± 0.0011	0.1138 ± 0.0003
TFCEMLP (default)	0.1008 ± 0.0003	0.4217 ± 0.0016	0.0838 ± 0.0005	0.1557 ± 0.0001	0.3273 ± 0.0010	0.0912 ± 0.0001	0.1231 ± 0.0003	0.5606 ± 0.0013	0.1151 ± 0.0008
TFCEMLP (tuned)	0.1049 ± 0.0003	0.4324 ± 0.0010	0.0875 ± 0.0002	0.1581 ± 0.0005	0.3315 ± 0.0012	0.0928 ± 0.0003	0.1295 ± 0.0001	0.5791 ± 0.0007	0.1227 ± 0.0007

Table 2: Test performance classical CF methods vs TFCEMLP (mean ± std) over 5 seeds. All improvements TFCE over other models are significant (p-value < 0.05). Statistical significance was evaluated based on the recall metric.

Model	Book			Game			Movie		
	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG
MF	0.0436 ± 0.0003	0.2485 ± 0.0013	0.0389 ± 0.0002	0.0327 ± 0.0005	0.0883 ± 0.0011	0.0200 ± 0.0002	0.0565 ± 0.0008	0.3350 ± 0.0037	0.0515 ± 0.0007
MultVAE	0.0721 ± 0.0004	0.3415 ± 0.0012	0.0598 ± 0.0004	0.0908 ± 0.0006	0.2094 ± 0.0011	0.0530 ± 0.0003	0.0864 ± 0.0007	0.4462 ± 0.0025	0.0784 ± 0.0009
LightGCN	0.0726 ± 0.0003	0.3507 ± 0.0015	0.0609 ± 0.0002	0.1025 ± 0.0004	0.2327 ± 0.0009	0.0595 ± 0.0005	0.0870 ± 0.0009	0.4472 ± 0.0032	0.0766 ± 0.0008
TFCE	0.0743 ± 0.0000	0.4217 ± 0.0000	0.0627 ± 0.0000	0.1319 ± 0.0000	0.2828 ± 0.0000	0.0769 ± 0.0000	0.0880 ± 0.0000	0.4397 ± 0.0000	0.0814 ± 0.0000

Table 3: Test performance AlphaRec vs AlphaRec with two-towers MLP. Two towers do not improve the performance.

Model	Book			Game			Movie		
	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG
AlphaRec	0.0992	0.4183	0.0832	0.1511	0.3188	0.0888	0.1221	0.5567	0.1138
AlphaRec (two towers)	0.0988	0.4181	0.0836	0.1501	0.3175	0.0881	0.1207	0.5534	0.1133

Table 4: Test performance TFCEMLP tuned vs TFCEMLP tuned with user ID-based learnable embedding. Enhancing with ID-based embeddings did not improve performance.

Model	Book			Game			Movie		
	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG	Recall	Hit Ratio	NDCG
TFCE	0.1038	0.4301	0.0862	0.1569	0.3296	0.0928	0.1279	0.5727	0.1199
TFCEMLP (tuned)	0.1049	0.4324	0.0875	0.1581	0.3315	0.0928	0.1295	0.5791	0.1227