We appreciate the reviewers' comments. As a common response. We have performed a more comprehensive set of experiments on the public datasets along with statistical significance tests. Within these experiments we also compare with the other common lifelong learning techniques with all the backbone models. We will include these in the camera ready version.

Table 2: Evaluation on public dataset Diginetica. We compare the average performance of the continuously updated model with two common lifelong learning techniques, i.e., regularization (Reg.), replay (Rep.), and MoDSA. Statistical significance of differences of MoDSA vs. the best baseline is determined by a paired t-test, where * is for p-value ≤ 0.05, ** is for p-value ≤ 0.01, and *** is for p-value ≤ 0.001.

Methods	HR@10	HR@20	MRR@10	MRR@20	NDCG@10	NDCG@20
SR-GNN	0.278	0.292	0.126	0.214	0.169	0.182
SR-GNN+Reg.	0.295	0.299	0.138	0.231	0.174	0.184
SR-GNN+Rep.	0.310	0.328	0.159	0.239	0.172	0.190
SR-GNN+MoDSA	**0.321**	**0.340**	**0.197**	**0.273**	**0.180**	**0.192**
improv.	3.548%**	3.659%**	23.899%***	14.226%***	3.448%*	1.053%*
GRU4Rec	0.231	0.268	0.142	0.245	0.108	0.125
GRU4Rec+Reg.	0.248	0.277	0.144	0.247	0.111	0.128
GRU4Rec+Rep.	0.254	0.276	0.145	0.250	0.109	0.130
GRU4Rec+MoDSA	**0.271**	**0.309**	**0.150**	**0.256**	**0.120**	**0.138**
improv.	6.693%***	11.552%***	3.448%***	2.4%**	8.108%***	6.154%***
NARM	0.307	0.320	0.150	0.165	0.170	0.195
NARM+Reg.	0.310	0.321	0.153	0.170	0.184	0.196
NARM+Rep.	0.316	0.328	0.163	0.174	0.194	0.218
NARM+MoDSA	**0.350**	**0.366**	**0.201**	**0.230**	**0.196**	**0.224**
improv.	10.759%***	11.585%***	23.313%***	32.184%***	1.031%*	2.752%*
SASRec	0.404	0.420	0.238	0.307	0.203	0.220
SASRec+Reg.	0.420	0.428	0.250	0.316	0.216	0.229
SASRec+Rep.	0.428	0.432	0.255	0.322	0.224	0.240
SASRec+MoDSA	**0.469**	**0.479**	**0.279**	**0.354**	**0.228**	**0.244**
improv.	9.579%***	10.88%***	9.412%***	9.938%***	1.786%*	1.667%*
BERT4Rec	0.356	0.442	0.161	0.192	0.149	0.190
BERT4Rec+Reg.	0.359	0.445	0.168	0.200	0.166	0.200
BERT4Rec+Rep.	0.362	0.460	0.180	0.206	0.173	0.204
BERT4Rec+MoDSA	**0.373**	**0.489**	**0.200**	**0.217**	**0.185**	**0.213**
improv.	3.039%**	6.304%**	11.111%***	5.34%**	6.936%***	4.412%**

Table 3: Evaluation on public dataset Yoochoose. We compare the average performance of the continuously updated model with two common lifelong learning techniques, i.e., regularization (Reg.), replay (Rep.), and MoDSA. Statistical significance of differences of MoDSA vs. the best baseline is determined by a paired t-test, where * is for p-value ≤ 0.05, ** is for p-value ≤ 0.01, and *** is for p-value ≤ 0.001.

Methods	HR@10	HR@20	MRR@10	MRR@20	NDCG@10	NDCG@20
SR-GNN	0.437	0.557	0.158	0.181	0.161	0.188
SR-GNN+Reg.	0.439	0.561	0.163	0.186	0.173	0.190
SR-GNN+Rep.	0.440	0.568	0.165	0.179	0.181	0.198
SR-GNN+MoDSA	**0.450**	**0.580**	**0.179**	**0.198**	**0.184**	**0.212**
improv.	2.273%*	2.113%*	8.485%**	10.615%***	1.657%*	7.071%***
GRU4Rec	0.409	0.420	0.131	0.162	0.233	0.266
GRU4Rec+Reg.	0.428	0.433	0.137	0.173	0.250	0.278
GRU4Rec+Rep.	0.423	0.443	0.141	0.178	0.272	0.287
GRU4Rec+MoDSA	**0.441**	**0.540**	**0.156**	**0.187**	**0.282**	**0.296**
improv.	3.037%**	21.896%***	10.638%***	5.056%**	3.676%*	3.136%*
NARM	0.440	0.536	0.165	0.181	0.295	0.343
NARM+Reg.	0.450	0.540	0.170	0.186	0.315	0.356
NARM+Rep.	0.447	0.546	0.178	0.193	0.327	0.369
NARM+MoDSA	**0.472**	**0.553**	**0.188**	**0.202**	**0.344**	**0.379**
improv.	4.889%**	1.282%*	5.618%**	4.663%**	5.199%**	2.71%*
SASRec	0.530	0.583	0.230	0.279	0.308	0.348
SASRec+Reg.	0.536	0.590	0.252	0.288	0.321	0.355
SASRec+Rep.	0.540	0.594	0.263	0.296	0.329	0.362
SASRec+MoDSA	**0.566**	**0.609**	**0.271**	**0.304**	**0.347**	**0.391**
improv.	4.815%**	2.525%**	3.042%**	2.703%*	5.471%**	8.011%***
BERT4Rec	0.670	0.741	0.247	0.271	0.321	0.361
BERT4Rec+Reg.	0.691	0.758	0.256	0.276	0.333	0.379
BERT4Rec+Rep.	0.703	0.768	0.261	0.280	0.345	0.395
BERT4Rec+MoDSA	**0.720**	**0.785**	**0.279**	**0.293**	**0.365**	**0.412**
improv.	2.418%*	2.214%**	6.897%***	4.643%***	5.797%**	4.304%**

Table 4: Evaluation on public dataset MovieLens 1M. We compare the average performance of the continuously updated model with two common lifelong learning techniques, i.e., regularization (Reg.), replay (Rep.), and MoDSA. Statistical significance of differences of MoDSA vs. the best baseline is determined by a paired t-test, where * is for p-value ≤ 0.05, ** is for p-value ≤ 0.01, and *** is for p-value ≤ 0.001.

Methods	HR@10	HR@20	MRR@10	MRR@20	NDCG@10	NDCG@20
SR-GNN	0.342	0.365	0.117	0.129	0.129	0.166
SR-GNN+Reg.	0.347	0.373	0.126	0.134	0.138	0.171
SR-GNN+Rep.	0.352	0.378	0.130	0.139	0.152	0.177
SR-GNN+MoDSA	**0.360**	**0.388**	**0.138**	**0.146**	**0.160**	**0.183**
improv.	2.273%*	2.646%*	6.154%***	5.036%**	5.263%**	3.39%**
GRU4Rec	0.507	0.530	0.121	0.144	0.151	0.165
GRU4Rec+Reg.	0.515	0.541	0.127	0.148	0.160	0.167
GRU4Rec+Rep.	0.523	0.549	0.131	0.154	0.167	0.178
GRU4Rec+MoDSA	**0.536**	**0.561**	**0.146**	**0.166**	**0.175**	**0.189**
improv.	2.486%*	2.186%*	11.45%***	7.792%***	4.79%**	6.18%***
NARM	0.534	0.605	0.137	0.166	0.187	0.207
NARM+Reg.	0.541	0.611	0.144	0.173	0.204	0.223
NARM+Rep.	0.538	0.617	0.149	0.188	0.214	0.232
NARM+MoDSA	**0.569**	**0.633**	**0.160**	**0.199**	**0.232**	**0.264**
improv.	5.176%**	2.593%*	7.383%***	5.851%**	8.411%***	13.793%***
SASRec	0.603	0.646	0.156	0.182	0.258	0.297
SASRec+Reg.	0.616	0.650	0.174	0.194	0.275	0.325
SASRec+Rep.	0.623	0.661	0.185	0.216	0.295	0.332
SASRec+MoDSA	**0.647**	**0.689**	**0.192**	**0.222**	**0.328**	**0.365**
improv.	3.852%**	4.236%**	3.784%**	2.778%*	11.186%***	9.94%***
BERT4Rec	0.657	0.722	0.146	0.172	0.371	0.390
BERT4Rec+Reg.	0.662	0.726	0.157	0.178	0.376	0.402
BERT4Rec+Rep.	0.693	0.738	0.165	0.189	0.381	0.410
BERT4Rec+MoDSA	**0.712**	**0.764**	**0.180**	**0.196**	**0.405**	**0.433**
improv.	2.742%*	3.523%**	9.091%***	3.704%**	6.299%**	5.61%**