loss_nonormalizer.md 2025-01-09

average loss of datasize 500, equal_weight = False

train_type	train_sp_L	uniform	exponential	linear	mixed
exponential	False	0.00812	0.00496	0.00497	0.00602
exponential	independently	0.00577	0.00308	0.00337	0.00407
exponential	together	0.00776	0.0039	0.00413	0.00526
linear	False	0.00537	0.05091	0.00383	0.02004
linear	independently	0.0046	0.0418	0.00299	0.01646
linear	together	0.0053	0.04385	0.00378	0.01764
mixed	False	0.00641	0.00619	0.00469	0.00576
mixed	independently	0.0044	0.00387	0.00328	0.00385
mixed	together	0.00576	0.00484	0.00392	0.00484
uniform	False	0.00389	0.0518	0.00416	0.01995
uniform	independently	0.00365	0.05421	0.00391	0.02059
uniform	together	0.00387	0.05159	0.00384	0.01977

average loss of datasize 1000, equal_weight = False

train_type	train_sp_L	uniform	exponential	linear	mixed
exponential	False	0.00749	0.00274	0.00307	0.00443
exponential	independently	0.00459	0.00197	0.00219	0.00292
exponential	together	0.00667	0.00212	0.00256	0.00378
linear	False	0.00343	0.03354	0.00221	0.01306
linear	independently	0.00348	0.0372	0.00194	0.01421
linear	together	0.00332	0.02825	0.00209	0.01122
mixed	False	0.00277	0.00317	0.00248	0.0028
mixed	independently	0.00227	0.00229	0.00204	0.0022
mixed	together	0.0025	0.00218	0.00192	0.0022
uniform	False	0.00211	0.03527	0.00256	0.01331
uniform	independently	0.00157	0.03447	0.00197	0.01267
uniform	together	0.00179	0.03487	0.00208	0.01291

loss nonormalizer.md 2025-01-09

average loss of datasize 1500, equal_weight = False

train_type	train_sp_L	uniform	exponential	linear	mixed
exponential	False	0.00635	0.00196	0.00215	0.00349
exponential	independently	0.00446	0.00152	0.00184	0.00261
exponential	together	0.00454	0.00165	0.00185	0.00268
linear	False	0.00284	0.01825	0.00172	0.0076
linear	independently	0.00228	0.02059	0.00148	0.00812
linear	together	0.00265	0.0233	0.00148	0.00914
mixed	False	0.00233	0.00236	0.0019	0.0022
mixed	independently	0.00178	0.00167	0.00153	0.00166
mixed	together	0.00211	0.00188	0.0016	0.00186
uniform	False	0.00157	0.03082	0.00186	0.01142
uniform	independently	0.00102	0.05475	0.0015	0.01909
uniform	together	0.00133	0.05082	0.00221	0.01812

biggest loss of datasize 500, equal_weight = False

train_type	train_sp_L	uniform	exponential	linear	mixed
exponential	False	0.13412	0.13208	0.03488	-
exponential	independently	0.08653	0.10819	0.04546	-
exponential	together	0.12825	0.1279	0.04235	-
linear	False	0.08725	0.12089	0.04632	-
linear	independently	0.0974	0.10644	0.03441	-
linear	together	0.09862	0.10497	0.0441	-
mixed	False	0.08452	0.12977	0.04003	-
mixed	independently	0.07665	0.10992	0.02331	-
mixed	together	0.07102	0.11465	0.034	-
uniform	False	0.07138	0.12156	0.01933	-
uniform	independently	0.07776	0.13838	0.02109	-
uniform	together	0.07731	0.12782	0.01387	-

loss nonormalizer.md 2025-01-09

biggest loss of datasize 1000, equal_weight = False

train_type	train_sp_L	uniform	exponential	linear	mixed
exponential	False	0.05738	0.03673	0.02836	-
exponential	independently	0.03821	0.03797	0.01856	-
exponential	together	0.03594	0.03779	0.01669	-
linear	False	0.04914	0.07033	0.01214	-
linear	independently	0.05173	0.06055	0.01117	-
linear	together	0.04441	0.06565	0.01363	-
mixed	False	0.02764	0.03336	0.01811	-
mixed	independently	0.03062	0.02914	0.01121	-
mixed	together	0.03697	0.02551	0.00898	-
uniform	False	0.01729	0.07579	0.01848	-
uniform	independently	0.01736	0.08089	0.0112	-
uniform	together	0.01348	0.07197	0.01425	-

biggest loss of datasize 1500, equal_weight = False

train_type	train_sp_L	uniform	exponential	linear	mixed
exponential	False	0.02678	0.02131	0.01078	-
exponential	independently	0.02375	0.01492	0.01216	-
exponential	together	0.02263	0.03275	0.01013	-
linear	False	0.02136	0.04857	0.0143	-
linear	independently	0.01342	0.05159	0.0146	-
linear	together	0.0186	0.04739	0.01066	-
mixed	False	0.03627	0.04211	0.0123	-
mixed	independently	0.02582	0.01851	0.01426	-
mixed	together	0.02891	0.02489	0.01081	-
uniform	False	0.01209	0.06849	0.01115	-
uniform	independently	0.00739	0.10046	0.00549	-
uniform	together	0.01544	0.09071	0.01093	-

loss nonormalizer.md 2025-01-09

average loss of datasize 500, equal_weight = True

train_type	train_sp_L	uniform	exponential	linear	mixed
mixed	False	0.00759	0.00661	0.00762	0.00727
mixed	independently	0.00681	0.00601	0.00696	0.00659
mixed	together	0.00814	0.00706	0.00802	0.00774

average loss of datasize 1000, equal_weight = True

train_type	train_sp_L	uniform	exponential	linear	mixed
mixed	False	0.00409	0.00395	0.00425	0.0041
mixed	independently	0.0037	0.00331	0.00374	0.00358
mixed	together	0.00383	0.00347	0.00399	0.00377

average loss of datasize 1500, equal_weight = True

	train_type	train_sp_L	uniform	exponential	linear	mixed
	mixed	False	0.00281	0.00312	0.00336	0.0031
-	mixed	independently	0.00237	0.00242	0.00288	0.00255
_	mixed	together	0.00263	0.00252	0.00306	0.00274

biggest loss of datasize 500, equal_weight = True

train_type	train_sp_L	uniform	exponential	linear	mixed
mixed	False	0.09214	0.15717	0.07208	-
mixed	independently	0.08295	0.14432	0.06875	-
mixed	together	0.09875	0.17162	0.09436	-

biggest loss of datasize 1000, equal_weight = True

train_type	train_sp_L	uniform	exponential	linear	mixed
mixed	False	0.03791	0.04754	0.03108	-
mixed	independently	0.0335	0.0298	0.03622	-

loss_nonormalizer.md 2025-01-09

train_type	train_sp_L	uniform	exponential	linear	mixed
mixed	together	0.03428	0.04147	0.03661	-

biggest loss of datasize 1500, equal_weight = True

train_ty	pe train_sp_L	uniform	exponential	linear	mixed
mixed	False	0.03228	0.06244	0.03568	-
mixed	independently	0.01814	0.02735	0.01956	-
mixed	together	0.02696	0.03594	0.02746	-