Appendix I: Hyperparameter Ranges and Optimal Values

YANG LIU, University of Helsinki, Finland

ALAN MEDLAR, University of Helsinki, Finland

DOROTA GŁOWACKA, University of Helsinki, Finland

Table 1. Hyperparameter values for all models on three datasets.

Model	Hyperparameter	Range	MovieLens-100K	MovieLens-1M	Amazon-Books
AFM	learning rate	[5e-05, 0.0001, 0.0005]	0.0005	0.0005	5e-05
	dropout prob	[0.0, 0.1]	0.1	0.1	0.1
	attention size	[20, 30]	30	30	30
	reg weight	[2, 5]	5	5	5
AutoInt	learning rate	[0.001, 0.005]	0.005	0.005	0.001
	dropout prob	[0.0, 0.1]	0.0	0.1	0.0
	attention size	[8, 16, 32]	32	8	8
	mlp hidden size	[[64, 64, 64], [128, 128, 128], [256, 256, 256]]	[64, 64, 64]	[256, 256, 256]	[64, 64, 64]
DCN	learning rate	[0.0001, 0.0005, 0.001, 0.005, 0.006]	0.006	0.001	0.0001
	mlp hidden size	[[128, 128, 128], [256, 256, 256], [512, 512, 512], [1024, 1024, 1024]]	[128, 128, 128]	[512, 512, 512]	[1024, 1024, 1024]
	reg weight	[1, 2, 5]	5	1	2
	cross layer num	[6]	6	6	6
	dropout prob	[0.1, 0.2]	0.2	0.2	0.2
DCNV2	learning rate	[0.005, 0.001, 0.0005]	0.005	0.005	0.0005
	mlp hidden size	[[256, 256], [512, 512], [768, 768], [1024, 1024]]	[256, 256]	[256, 256]	[256, 256]
	cross layer num	[2, 3, 4]	4	2	2
	dropout prob	[0.1, 0.2]	0.2	0.2	0.2
	reg weight	[1, 2, 5]	5	2	1

Authors' addresses: Yang Liu, University of Helsinki, Helsinki, Finland, yang.liu@helsinki.fi; Alan Medlar, University of Helsinki, Helsinki, Finland, alan.j.medlar@helsinki.fi; Dorota Głowacka, University of Helsinki, Helsinki, Finland, dorota.glowacka@helsinki.fi.

Model	Hyperparameter	Range	MovieLens-100K	MovieLens-1M	Amazon-Books
DeepFM	learning rate	[0.001, 0.005, 0.01]	0.005	0.005	0.001
	dropout prob	[0.0, 0.1]	0.1	0.1	0.1
	mlp hidden size	[[128, 128, 128], [256, 256, 256]]	[256, 256, 256]	[256, 256, 256]	[128, 128, 128]
FFM	learning rate	[0.0001, 0.0005, 0.001, 0.005, 0.05]	0.0005	0.001	0.0001
FM	learning rate	[5e-05, 0.0001, 0.0002, 0.0005, 0.001, 0.005]	0.0001	0.005	5e-05
FNN	learning rate	[0.0005, 0.001, 0.003, 0.005]	0.003	0.003	0.0005
	dropout prob	[0.0, 0.1]	0.1	0.1	0.1
	mlp hidden size	[[128, 256, 128], [128, 128, 128]]	[128, 256, 128]	[128, 256, 128]	[128, 128, 128]
FwFM	learning rate	[0.0001, 0.0005, 0.001, 0.005, 0.01]	0.0005	0.001	0.0001
	dropout prob	[0.0, 0.2, 0.4]	0.4	0.0	0.4
LR	learning rate	[5e-05, 0.0001, 0.0002, 0.0005, 0.001, 0.005]	0.0005	0.005	0,005
NFM	learning rate	[5e-05, 8e-05, 0.0001, 0.0005, 0.001]	8e-05	0.001	0.001
	dropout prob	[0.1, 0.2, 0.3]	0.3	0.3	0.3
	mlp hidden size	[[20, 20, 20], [40, 40, 40], [50, 50, 50]]	[20, 20, 20]	[50, 50, 50]	[20, 20, 20]
PNN	learning rate	[0.001, 0.003, 0.005, 0.006, 0.01]	0.006	0.005	0.001
	dropout prob	[0.0, 0.1]	0.0	0.1	0.1
	mlp hidden size	[[64, 64, 64], [128, 128, 128], [256, 256, 256]]	[128, 128, 128]	[256, 256, 256]	[64, 64, 64]
	reg weight	[0.0]	0.0	0.0	0.0
WideDeep	learning rate	[0.0005, 0.001, 0.005, 0.01]	0.01	0.005	0.0005
	dropout prob	[0.0, 0.2]	0.2	0.2	0.2
	mlp hidden size	[[64, 64, 64], [128, 128, 128], [256, 256, 256]]	[64, 64, 64]	[256, 256, 256]	[64, 64, 64]
ADMMSLIM	alpha	[0.25, 0.5, 0.75, 1]	0.75	1	/
	lambda1	[0.1, 0.5, 5, 10]	0.5	5	/
	lambda2	[5, 50, 1000, 5000]	1000	5000	/

Model	Hyperparameter	Range	MovieLens-100K	MovieLens-1M	Amazon-Books
BPR	learning rate	[5e-05, 0.0001, 0.0005, 0.0007, 0.001, 0.005, 0.007]	0.0005	0.0001	0.0005
xDeepFM	learning rate	[0.0001, 0.001, 0.005, 0.006]	0.001	0.001	0.0001
	dropout prob	[0.0, 0.1]	0.1	0.1	0.0
	mlp hidden size	[[128, 128, 128], [256, 256, 256], [512, 512, 512]]	[256, 256, 256]	[512, 512, 512]	[512, 512, 512]
	cin layer size	[[60, 60, 60], [100, 100, 100]]	[60, 60, 60]	[100, 100, 100]	[100, 100, 100]
	reg weight	[1e-05, 0.0005]	0.0005	0.0005	1e-05
CDAE	learning rate	[0.01, 0.001, 0.005, 0.0005]	0.005	0.01	0.005
	loss type	[BCE, MSE]	BCE	BCE	BCE
	corruption ratio	[0.5, 0.3, 0.1]	0.5	0.3	0.5
	reg weight 1	[0.0, 0.01]	0.0	0.01	0.01
	reg weight 2	[0.0, 0.01]	0.0	0.01	0.0
ConvNCF	learning rate	[0.005, 0.01, 0.02]	0.02	0.02	0.02
	cnn channels	[[1, 32, 32, 32, 32, 32], [1, 64, 32, 32, 32, 32]]	[1, 32, 32, 32, 32]	[1, 64, 32, 32, 32, 32]	[1, 32, 32, 32, 32]
	cnn kernels	[[2, 2, 2, 2, 2, 2], [4, 2, 2, 2, 2]]	[4, 2, 2, 2, 2]	[4, 2, 2, 2, 2]	[4, 2, 2, 2, 2]
	cnn strides	[[2, 2, 2, 2, 2, 2]', [4, 2, 2, 2, 2]]	[4, 2, 2, 2, 2]	[4, 2, 2, 2, 2]	[4, 2, 2, 2, 2]
	dropout prob	[0.0, 0.1, 0.2, 0.3]	0.2	0.3	0.2
	reg weights	[[0.0, 0.0], [0.1, 0.1]]	[0.1, 0.1]	[0.1, 0.1]	[0.1, 0.1]
DGCF	learning rate	[0.0005, 0.001, 0.005, 0.01]	0.001	0.0005	0.001
	n factors	[2, 4, 8]	4	2	2
	reg weight	[0.001, 0.01]	0.001	0.001	0.001
	cor weight	[0.001, 0.01]	0.001	0.001	0.01
DMF	learning rate	[5e-05, 0.0005, 0.0003, 0.0001, 0.005, 0.001]	0.0005	0.0005	0.0005
	user hidden size	[[64, 64], [64, 32]]	[64, 64]	[64, 64]	[64, 64]
	item hidden size	[[64, 64], [64, 32]]	[64, 64]	[64, 64]	[64, 64]
ENMF	learning rate	[0.005, 0.01, 0.05]	0.005	0.005	0.01
	dropout prob	[0.3, 0.5, 0.7]	0.7	0.3	0.3
	negative weight	[0.1, 0.2, 0.5]	0.2	0.5	0.1

Model	Hyperparameter	Range	MovieLens-100K	MovieLens-1M	Amazon-Books
EASE	reg weight	[1.0, 10.0, 100.0, 500.0, 1000.0, 2000.0]	100.0	1000.0	/
FISM	learning rate	[0.0005, 0.0001, 0.005, 0.001]	0.0005	0.001	0.0001
	embedding size	/	64	64	64
	regs	$[[10^{-7}, 10^{-7}], [0, 0]]$	$[10^{-7}, 10^{-7}]$	$[0, 0]$	$[10^{-7}, 10^{-7}]$
	alpha	/	0.0	0.0	0.0
GCMC	accum	[stack, sum]	sum	stack	stack
	learning rate	[0.001, 0.005, 0.01]	0.005	0.001	0.001
	dropout prob	[0.3, 0.5, 0.7]	0.5	0.5	0.3
	gcn output dim	/	500	500	500
	num basis functions	/	2	2	2
ItemKNN	k	[10, 50, 100, 200, 250, 300, 400]	100	250	10
	shrink	[0.0, 0.1, 0.5, 1, 2]	0.0	1	1
LINE	learning rate	[0.0005, 0.001, 0.002]	0.001	0.001	0.002
	sample num	[1, 3, 5]	3	5	5
	second order loss	[0.3, 0.6, 1]	1	1	0.3
LightGCN	learning rate	[0.0005, 0.001, 0.002]	0.002	0.0005	0.001
	n layers	[1, 2, 3, 4]	4	1	4
	reg weight	[1e-05, 0.0001, 0.001, 0.01]	0.0001	0.01	0.01
MacridVAE	learning rate	[0.0005, 0.001, 0.005, 0.01, 0.05]	0.001	0.005	/
	kafc	[3, 5, 10, 20]	3	20	/
MultiDAE	learning rate	[5e-05, 0.0001, 0.0005, 0.0007, 0.001, 0.005, 0.007]	0.001	0.007	0.0005
MultiVAE	learning rate	[5e-05, 0.0001, 0.0005, 0.00007, 0.001, 0.005, 0.007]	0.0007	0.007	0.001
NGCF	learning rate	[0.0001, 0.0005, 0.001]	0.001	0.0005	0.001
	hidden size list	[[64, 64, 64], [128, 128, 128]]	[128, 128, 128]	[128, 128, 128]	[128, 128, 128]
	node dropout	[0.0, 0.1, 0.2]	0.1	0.0	0.2
	message dropout	[0.0, 0.1, 0.2]	0.0	0.0	0.2
	reg weight	[1e-05, 0.001, 0.1]	0.1	1e-05	0.1

Model	Hyperparameter	Range	MovieLens-100K	MovieLens-1M	Amazon-Books
NAIS	learning rate	[0.0001, 0.001, 0.01]	0.0001	0.0001	
	weight size	[64, 32]	32	32	
	reg weights	$[[10^{-7}, 10^{-7}, 10^{-5}], [0, 0, 0]]$	$[10^{-7}, 10^{-7}, 10^{-5}]$	$[10^{-7}, 10^{-7}, 10^{-5}]$	/
	alpha	[0.0]	0.0	0.0	
	beta	[0.5]	0.5	0.5	
NCEPLRec	rank	[100, 200, 450]	450	100	450
	beta	[0.8, 1.0, 1.3]	0.8	1.0	1.0
	reg weight	[0.0001, 0.01, 100.0, 15000]	15000	15000	100.0
NCL	num clusters	[100, 1000]	100	100	1000
	proto reg	[1e-06, 1e-07, 1e-08]	1e-08	1e-08	1e-08
	ssl reg	[1e-06, 1e-07]	1e-06	1e-06	1e-07
	ssl temp	[0.05, 0.07, 0.1]	0.05	0.05	0.05
NNCF	neigh num	[20, 50, 100]	100	20	20
	neigh embedding size	[64, 32]	32	64	32
	num conv kernel	[128, 64]	128	128	128
	learning rate	[5e-05, 0.0001, 0.0005]	0.0005	0.0001	0.0001
	neigh info method	[random, knn]	random	random	random
NeuMF	learning rate	[5e-07, 1e-06, 5e-06, 1e-05, 0.0001, 0.001]	0.001	1e-06	0.0001
	mlp hidden size	['[64, 32, 16]']	[64, 32, 16]	[64, 32, 16]	[64, 32, 16]
	dropout prob	[0.0, 0.1, 0.3]	0.0	0.3	0.0
RaCT	dropout prob	[0.1, 0.3, 0.5]	0.5	0.5	0.3
	anneal cap	[0.2, 0.5]	0.1	0.2	0.2
RecVAE	learning rate	[5e-05, 0.0001, 0.0005, 0.0007, 0.001, 0.005, 0.007]	0.007	0.0007	0.005
SimpleX	gamma	[0.3, 0.5, 0.7]	0.5	0.7	0.5
	margin	[0, 0.5, 0.9]	0.9	0.9	0.5
	negative weight	[1, 10, 50]	10	50	50
SpectralCF	learning rate	[0.002, 0.001, 0.0005]	0.002	0.001	0.002
	reg weight	[0.002, 0.001, 0.0005]	0.001	0.0005	0.0005
	n layers	[1, 2, 3, 4]	3	3	2

Model	Hyperparameter	Range	MovieLens-100K	MovieLens-1M	Amazon-Books
SGL	type	/	ED	ED	ED
	ssl tau	[0.1, 0.2, 0.5]	0.5	0.5	0.2
	drop ratio	[0.1, 0.2, 0.4, 0.5]	0.1	0.1	0.2
	ssl weight	[0.05, 0.1, 0.5]	0.05	0.05	0.1
SLIMElastic	alpha	[0.2, 0.4, 0.6, 0.8]	0.2	0.2	0.2
	l1 ratio	[0.1, 0.05, 0.01, 0.005]	0.005	0.05	0.005
	hide item	/	True	True	True
	positive only	/	True	True	True
CFKG	learning rate	[0.01, 0.005, 0.001, 0,0005, 0,0001]	0.005	0.001	0.001
	loss function	[inner product, transe]	transe	inner product	transe
	margin	[0.5, 1.0, 2.0]	1.0	1.0	0.5
CKE	learning rate	[5e-05, 0.0001, 0.0005, 0.0007, 0.001]	0.0005	0.0005	0.001
	kg embedding size	[16, 32, 64, 128]	64	128	16
	reg weights	[[0.1, 0.1], [0.01, 0.01], [0.001, 0.001]]	[0.1, 0.1]	[0.01, 0.01]	[0.01, 0.01]
KGAT	learning rate	[0.01, 0.005, 0.001, 0.0005, 0.0001]	0.0005	0.001	0.005
	layers	[[64, 32, 16], [64, 64, 64], [128, 64, 32]]	[64, 64, 64]	[128, 64, 32]	[64, 32, 16]
	reg weight	[0.0001, 5e-05, 1e-05, 5e-06, 1e-06]	5e-06	5e-05	5e-05
	mess dropout	[0.1,0.2,0.3,0.4,0.5]]	1	0.1	0.5
KGCN	learning rate	[0.002, 0.001, 0.0005]	0.002	0.001	0.001
	n iter	[1, 2]	1	1	2
	aggregator	[sum, concat, neighbor]	sum	sum	sum
	l2 weight	[0.001, 1e-05, 1e-07]	1e-05	1e-07	0.001
	neighbor sample size	/	4	4	4
KGNNLS	learning rate	[0.002, 0.001, 0.0005]	0.0005	0.0005	0.002
	n iter	[1, 2]	1	1	2
	aggregator	/	sum	sum	sum
	l2 weight	[0.001, 1e-05]	0.001	1e-05	0.001
	neighbor sample size	/	4	4	4
	ls weight	[1, 0.5, 0.1, 0.01, 0.001]	0.5	1	0.01

Model	Hyperparameter	Range	MovieLens-100K	MovieLens-1M	Amazon-Books
KGIN	learning rate	[0.0001, 0.001, 0.005]	0.0001	0.0001	0.001
	node dropout rate	[0.1, 0.3, 0.5]	0.5	0.5	0.3
	mess dropout rate	[0.0, 0.1]	0.0	0.0	0.0
	context hops	[2, 3]	2	2	2
	n factors	[4, 8]	8	4	4
	ind	[cosine, distance]	cosine	cosine	distance
KTUP	learning rate	[0.01, 0.005, 0.001, 0.0005, 0.0001]	0.0005	0.001	0.001
	L1 flag	[True, False]	False	False	False
	use st gumbel	[True, False]	True	True	True
	train rec step	[8, 10]	8	8	8
	train kg step	[0, 1, 2, 3, 4, 5]	3	5	5
MKR	learning rate	[5e-05, 0.0001, 0.001, 0.005, 0.01]	0.001	0.001	0.005
	low layers num	[1, 2, 3]	1	1	2
	high layers num	[1, 2]	1	1	1
	l2 weight	[1e-06, 0.0001]	0.0001		1e-06
	kg embedding size	[16, 32, 64]	16	64	16