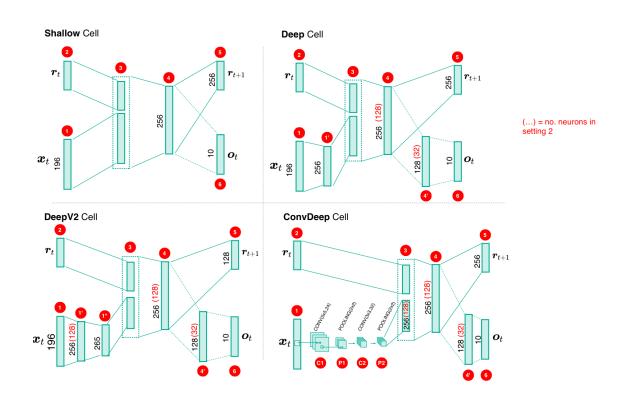
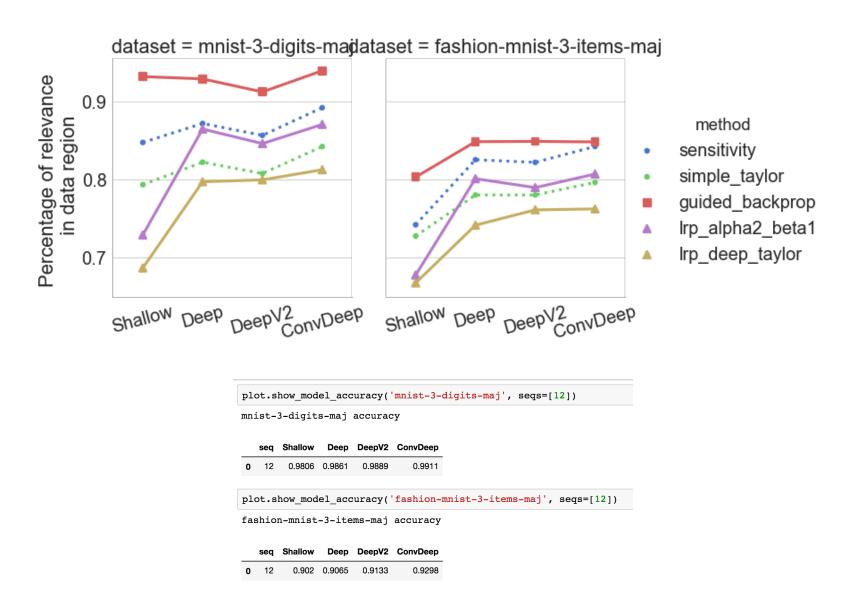
Results after improving weight initialization and add one linear layer after pooling for ConvDeep

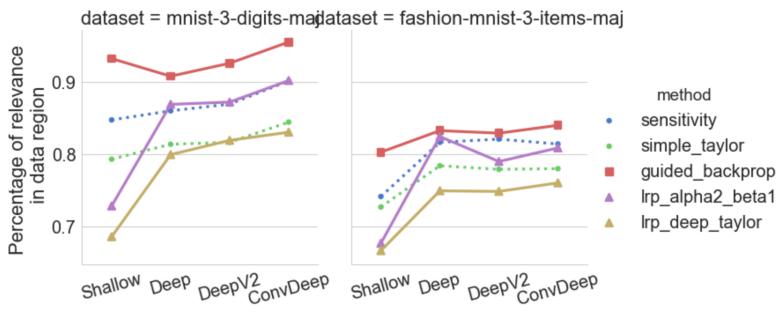


# trainable variables	Shallow	Deep	DeepV2	ConvDeep
Setting 1 (discussed yesterday)	184330	281738	347530	353818
Setting 2	184330	153578	161386	151802

Setting 1 Result



Setting 2 Result



pl	ot.s	how_mod	lel_acc	uracy(mnist-3-	-di	lig	git	s-n	naj	', s	eqs	=[1	2])		
mn.	ist-	3-digit	s-maj	accurac	у											
	seq	Shallow	Deep	DeepV2	ConvDeep	>										
0	12	0.9806	0.9839	0.9809	0.9903	3										
pl	ot.s	how_mod	lel_acc	uracy(fashion-	-mn	ni	ist	-3-	-it	ems-	maj	٠,	seq	[s=[12
fa	shio	n-mnist	-3-ite	ms-maj	accuracy	У										
	seq	Shallow	Deep	DeepV2	ConvDeep)										
0	12	0.902	0 9065	0.9071	0.9238	3										