Generalization Confusion Matrices (llama-2-7b-chat-bnb-4bit)																								
_1											acc				1									
	qkvo + mlp				mlp			qkvo			qkv					qk				OV1.0				
BF	0.9	0.78	0.61	BF	0.89	0.57	0.56	BF	0.94	0.74	0.66	BF	0.98	0.76	0.47	BF	0.8	0.38	0.36	BF	0.98	0.93	0.6	-0.8
Train set MF	0.91	0.82	0.62	MF	0.88	0.82	0.64	MF	0.94	0.84	0.6	MF	0.9	0.79	0.6	MF	0.61	0.45	0.32	MF	0.94	0.8	0.64	-0.6 -0.4
YL2	0.9	0.73	0.78	YL2	0.89	0.71	0.8	YL2	0.85	0.75	0.78	YL2	0.91	0.76	0.76	YL2	0.41	0.35	0.33	YL2	0.89	0.73	0.75	-0.2
	BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2	0.0
context_acc																								
	qr	kvo + mlp)			mlp				qkvo				qkv				qk				OV		1.0
BH -	0.82	0.74	0.94	BF	0.84	0.74	0.96	BF	0.9	0.76	0.98	BF	0.98	0.92	0.69	BF	0.84	0.76	0.65	BF	0.96	0.94	0.92	-0.8
Train set MF	0.84	0.74	0.98	MF	0.86	0.82	1	MF	0.88	0.8	1	MF	0.82	0.72	0.96	MF	0.8	0.84	0.53	MF	0.88	0.74	1	-0.6 -0.4
YL2	0.82	0.58	1	YL2	0.82	0.62	1	YL2	0.72	0.68	1	YL2	0.84	0.66	0.98	YL2	0.64	0.64	0.47	YL2	0.78	0.64	1	-0.2
	BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YĹ2		BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YĹ2	0.0
					mln								or_acc				a1z				$\mathbf{O}\mathbf{V}$			
	qkvo + mlp					mlp				qkvo				qkv				qk				OV		1.0
BF	0.98	0.82	0.29	BF	0.94	0.4	0.18	BF	0.98	0.72	0.35	BF	0.98	0.6	0.25	BF	0.76	0	0.078	BF	1	0.92	0.29	-0.8
Train set MF	0.98	0.9	0.27	MF	0.9	0.82	0.29	MF	1	0.88	0.22	MF	0.98	0.86	0.25	MF	0.42	0.06	0.12	MF	1	0.86	0.29	-0.6 -0.4
YL2	0.98	0.88	0.57	YL2	0.96	0.8	0.61	YL2	0.98	0.82	0.57	YL2	0.98	0.86	0.55	YL2	0.18	0.06	0.2	YL2	1	0.82	0.51	-0.2
	BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2		BF	MF Test set	YL2	0.0