Generalization Confusion Matrices (mistral-7b-v0.2-bnb-4bit)																									
$\alpha 1_{2320} + m 1_{23}$														acc				_1							
qkvo + mlp			mlp				qkvo			qkv				qk				OV 1							
BF	0.95	0.71	0.42	BF	0.94	0.89	0.38	BF	0.94	0.82	0.41	BF	0.96	0.91	0.37	BF	0.96	0.82	0.48	BF	- 1	0.9	0.4	-0.8	
Train set MF	0.85	0.84	0.34	MF	0.93	0.88	0.37	MF	0.95	0.91	0.37	MF	0.95	0.91	0.48	MF	0.92	0.88	0.55	MF	- 0.95	0.92	0.37	-0.6 -0.4	
YL2	0.95	0.59	0.84	YL2	0.93	0.62	0.81	YL2	0.92	0.5	0.82	YL2	0.92	0.49	0.82	YL2	0.86	0.57	0.86	YL2	0.94	0.57	0.78	-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	
qkvo + mlp mlp										context_acc				~1rr				~1 <sub>z</sub>							
	q	KVO + IIIIL	)			Шр				qkvo				qkv				qk				OV		1.0	
t BF	0.9	0.58	0.57	BF	0.9	0.9	0.49	BF	0.88	0.72	0.55	BF	0.92	0.92	0.47	BF	0.92	0.9	0.78	BF	- 1	0.88	0.55	-0.8	
Train set MF	0.9	0.88	0.45	MF	0.9	0.9	0.45	MF	0.9	0.88	0.47	MF	0.9	0.88	0.69	MF	0.9	0.88	0.9	MF	- 0.92	0.9	0.45	-0.6 -0.4	
YL2	0.9	0.3	1	YL2	0.86	0.3	1	YL2	0.84	0.04	1	YL2	0.84	0.08	1	YL2	0.9	0.88	1	YL2	- 0.88	0.2	1	-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	
1 . 1										1	pr	ior_acc				~. <b>1</b> _									
	q	kvo + mlp	)			mlp				qkvo				qkv				qk				OV		1.0	
BF	1	0.84	0.27	BF	0.98	0.88	0.27	BF	1	0.92	0.27	BF	1	0.9	0.27	BF	1	0.74	0.2	BF	- 1	0.92	0.25	-0.8	
Train set MF	0.8	0.8	0.24	MF	0.96	0.86	0.29	MF	1	0.94	0.27	MF	1	0.94	0.27	MF	0.94	0.88	0.22	MF	0.98	0.94	0.29	-0.6 -0.4	
YL2	1	0.88	0.69	YL2	1	0.94	0.63	YL2	1	0.96	0.65	YL2	1	0.9	0.65	YL2	0.82	0.26	0.73	YL2	- 1	0.94	0.57	-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	