

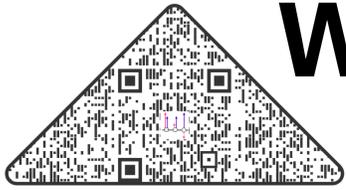
# Resolving Interference When Merging Models



THE UNIVERSITY OF NORTH CAROLINA at CHAPEL HILL

Prateek Yadav<sup>1</sup>, Derek Tam<sup>1</sup>, Leshem Choshen<sup>2</sup>, Colin Raffel<sup>1</sup>, Mohit Bansal<sup>1</sup>

Chapel Hill, <sup>2</sup>IBM Research



Model Merging - averages updates to each weight

It can be seen as a vote



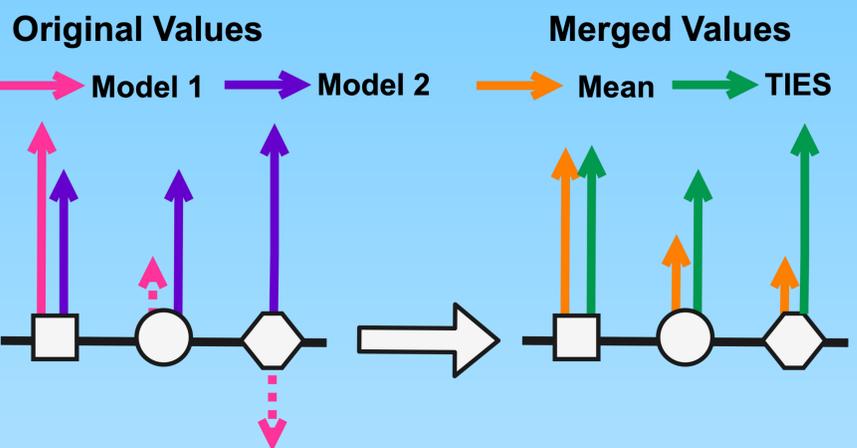
Models may *disagree*

Models may also *abstain* (not vote)

1

Two cases where updates interfere (pull towards 0):-

- Redundant changes  (almost) no change
- Sign Conflicts  Direction (+/-) disagreement
- Not interference 

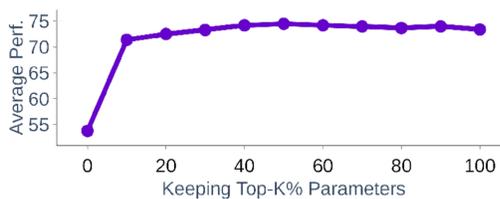


## Better Generalization

Model	T5-Base	T5-Large
Zeroshot	31.1	27.6
Simple Averaging [9, 76]	31.7	30.4
Fisher [43]	33.8	32.0
RegMean [31]	34.3	36.0
Task Arithmetic [29]	31.9	32.3
TIES-MERGING	35.3 [+1.0]	40.4 [+4.4]

3

Most gradient updates are unnecessary



4

Large gains – NLP\Vision\PEFT 

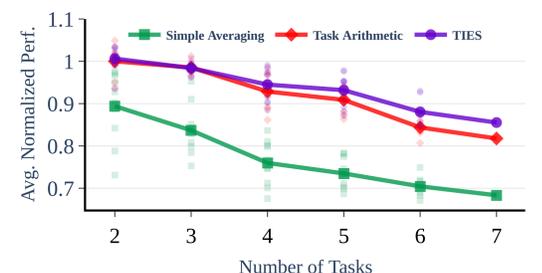
Method (↓)	Validation	PEFT		Full Finetuning		
		(IA) <sup>3</sup>	T5-Base	T5-Large	ViT-B/32	ViT-L/14
FINE-TUNED	-	71.4	82.8	88.8	90.5	94.2
MULTITASK	-	73.1	83.6	88.1	88.9	93.5
AVERAGING [76, 9]	✗	-	65.9	59.6	65.8	79.6
TASK ARITHMETIC [29]	✗	-	73.2	73.5	60.4	83.3
TIES-MERGING	✗	-	69.7 [-3.2]	74.4 [+0.9]	72.4 [+6.6]	86.0 [+2.7]
FISHER MERGING [43]	✓	62.2	68.9	64.6	68.3	82.2
REGMEAN [31]	✓	58.0	71.2	73.2	71.8	83.7
TASK ARITHMETIC [29]	✓	63.9	73.2	73.3	70.1	84.5
TIES-MERGING	✓	66.4 [+2.5]	73.9 [+0.7]	76.9 [+3.6]	73.6 [+1.8]	86.0 [+1.5]

5

Prune redundancies

6

Scales Better with #models merged



7

Sign is more important than Size

Method	Average
Fine-Tuned Multitask	73.1
Averaging [9, 76]	58.0
Task Vectors [29]	63.9
TIES-MERGING	66.4
TIES-MERGING (Oracle Sign)	72.0 [+5.6]

7

Method	T5-base	(IA) <sup>3</sup>
TIES-MERGING	74.5	70.7
- TRIM	73.0	70.6
- ELECT	73.1	69.6
- DISJOINT MEAN	72.6	67.5
- SCALE	72.0	65.5

7

Interference is Strong

