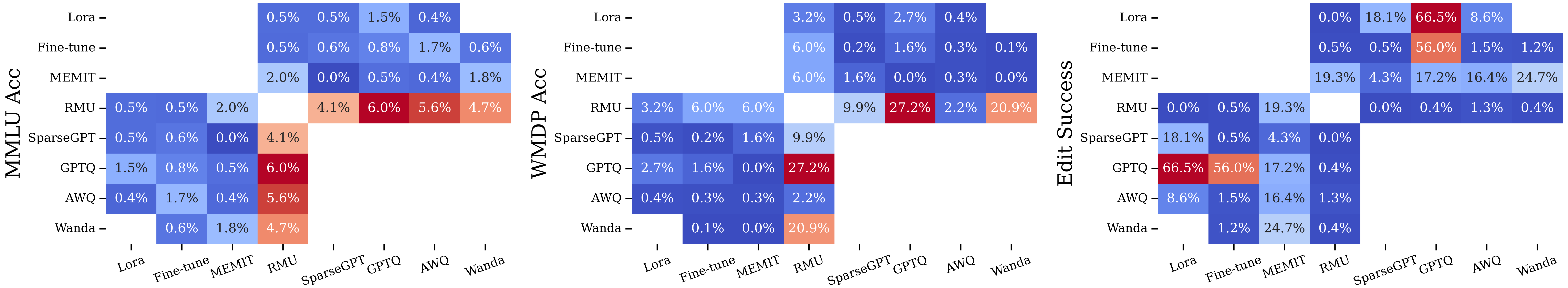 MMLU Acc
Lora
Fine-tune
MEMIT
RMU
SparseGPT
GPTQ
AWQ
Wanda
0.5% 0.5% 1.5% 0.4%
0.5% 0.6% 0.8% 1.7% 0.6%
2.0% 0.0% 0.5% 0.4% 1.8%
0.5% 0.5% 2.0% 4.1% 6.0% 5.6% 4.7%
0.5% 0.6% 0.0% 4.1%
1.5% 0.8% 0.5% 6.0%
0.4% 1.7% 0.4% 5.6%
0.6% 1.8% 4.7%
Lora
Fine-tune
MEMIT
RMU
SparseGPT
GPTQ
AWQ
Wanda
WMDP Acc
Lora
Fine-tune
MEMIT
RMU
SparseGPT
GPTQ
AWQ
Wanda
3.2% 0.5% 2.7% 0.4%
6.0% 0.2% 1.6% 0.3% 0.1%
6.0% 1.6% 0.0% 0.3% 0.0%
3.2% 6.0% 6.0% 9.9% 27.2% 2.2% 20.9%
0.5% 0.2% 1.6% 9.9%
2.7% 1.6% 0.0% 27.2%
0.4% 0.3% 0.3% 2.2%
0.1% 0.0% 20.9%
Lora
Fine-tune
MEMIT
RMU
SparseGPT
GPTQ
AWQ
Wanda
Edit Success
Lora
Fine-tune
MEMIT
RMU
SparseGPT
GPTQ
AWQ
Wanda
0.0% 18.1% 66.5% 8.6%
0.5% 0.5% 56.0% 1.5% 1.2%
19.3% 4.3% 17.2% 16.4% 24.7%
0.0% 0.5% 19.3% 0.0% 0.4% 1.3% 0.4%
18.1% 0.5% 4.3% 0.0%
66.5% 56.0% 17.2% 0.4%
8.6% 1.5% 16.4% 1.3%
1.2% 24.7% 0.4%
Lora
Fine-tune
MEMIT
RMU
SparseGPT
GPTQ
AWQ
Wanda