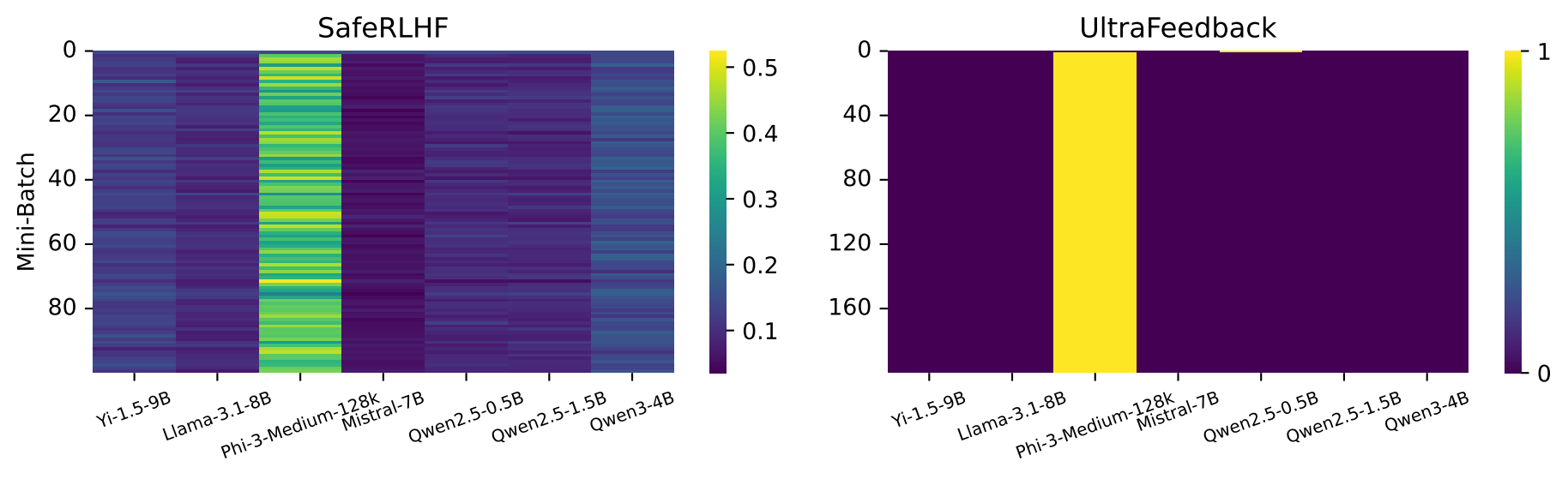
SafeRLHF
Mini-Batch
0
20
40
60
80
0.5
0.4
0.3
0.2
0.1
Yi-1.5-9B
Llama-3.1-8B
Phi-3-Medium-128k
Mistral-7B
Qwen2.5-0.5B
Qwen2.5-1.5B
Qwen3-4B
UltraFeedback
0
40
80
120
160
1
0
Yi-1.5-9B
Llama-3.1-8B
Phi-3-Medium-128k
Mistral-7B
Qwen2.5-0.5B
Qwen2.5-1.5B
Qwen3-4B