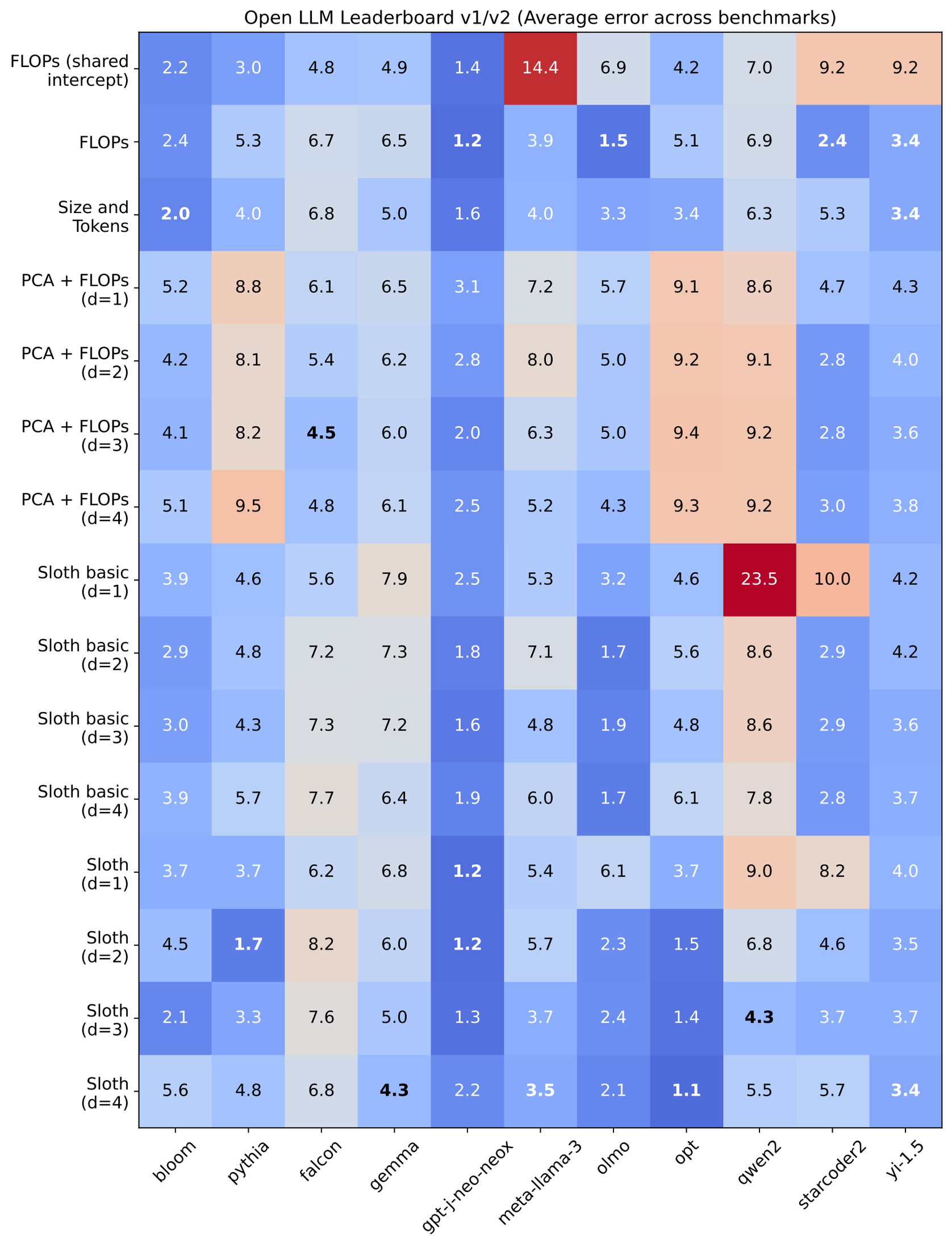
Open LLM Leaderboard v1/v2 (Average error across benchmarks)
FLOPs (shared intercept)
FLOPs
Size and Tokens
PCA + FLOPs (d=1)
PCA + FLOPs (d=2)
PCA + FLOPs (d=3)
PCA + FLOPs (d=4)
Sloth basic (d=1)
Sloth basic (d=2)
Sloth basic (d=3)
Sloth basic (d=4)
Sloth (d=1)
Sloth (d=2)
Sloth (d=3)
Sloth (d=4)
bloom
pythia
falcon
gemma
gpt-j-neo-neox
meta-llama-3
olmo
opt
qwen2
starcoder2
yi-1.5