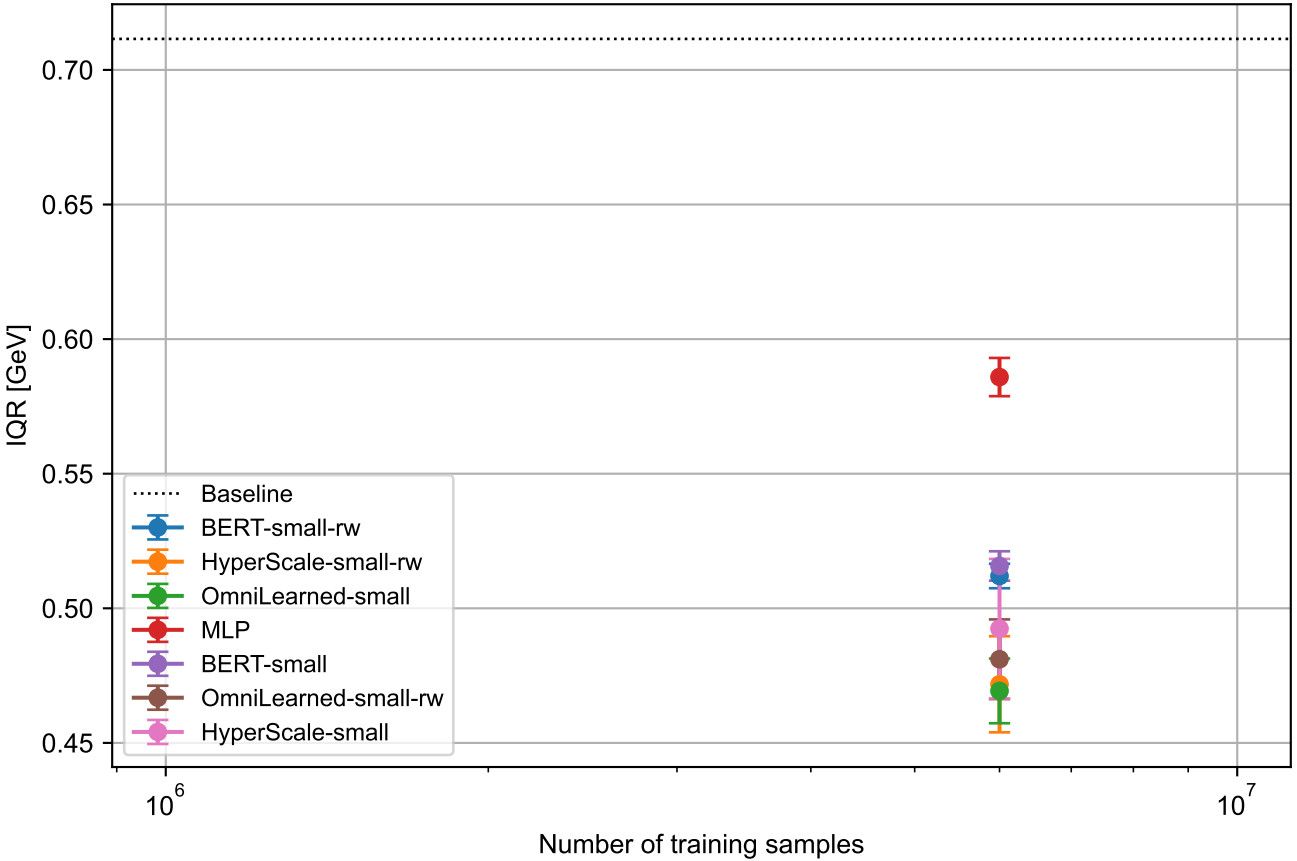

IQR vs Training Samples (q_3 bin 2, mid = 0.90 GeV)
IQR [GeV]
0.70
0.65
0.60
0.55
0.50
0.45
10^6
10^7
Number of training samples
Baseline
BERT-small-rw
HyperScale-small-rw
OmniLearned-small
MLP
BERT-small
OmniLearned-small-rw
HyperScale-small