#Train models
8+
7
6
5
4
3
2
<15
<5
0.07B
(X16B)
(X35B)
(X43B)
(X0B)
(X3B)
1.4B
2.8B
6.9B
(X1)
est model parameters (scale up p