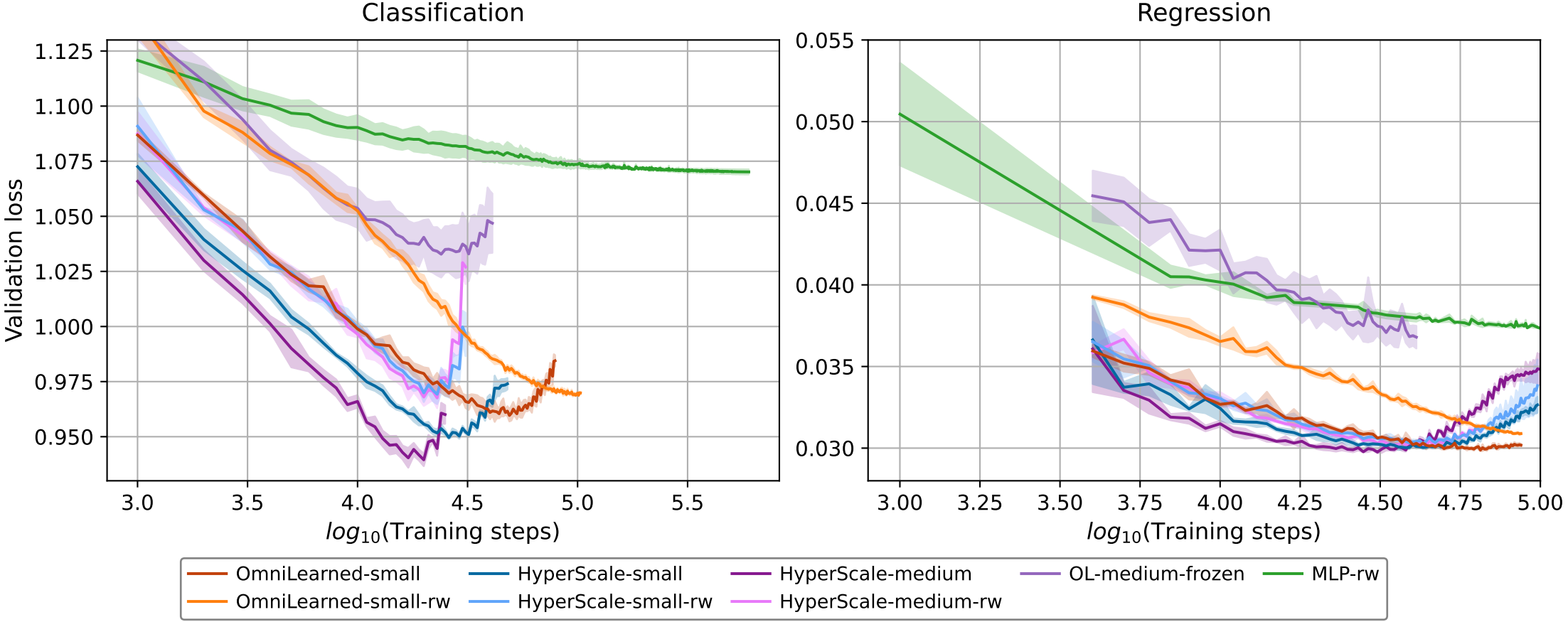

Classification
Regression
Validation loss
1.125
1.100
1.075
1.050
1.025
1.000
0.975
0.950
0.055
0.050
0.045
0.040
0.035
0.030
3.0
3.5
4.0
4.5
5.0
5.5
3.00
3.25
3.50
3.75
4.00
4.25
4.50
4.75
5.00
log10(Training steps)
log10(Training steps)
OmniLearned-small
OmniLearned-small-rw
HyperScale-small
HyperScale-small-rw
HyperScale-medium
HyperScale-medium-rw
OL-medium-frozen
MLP-rw