



Supplementary Figure 1: Learning rate finder figure. The learning rate is gradually increased from $10e-7$ to $10e-0$; the point where loss begins to decrease is the minimum learning rate (min LR), and the point where it begins to increase is the maximum learning rate (max LR).