Best/Worst Weight of Adaptive LR Optimizers Every 10 Epochs
e): (32. 1) Worst Param (batch_size, learning_rate): (900, 0.001) Best Param (batch_size, learning_rate): (32, 1) Weights Real Weights Weights Biases Real Bias Biases adadelta Best Param (batch size, learning rate): (32, 1) Worst Param (batch size, learning rate): (900, 0.001) Weights Real Weights Biases → Biases adagrad Worst Param (batch_size, learning_rate): (900, 0.001) Best Param (batch_size, learning_rate): (32, 0.01) Weights Real Weights Weights Biases adam Best Param (batch_size, learning_rate): (32, 0.001) Worst Param (batch size, learning rate): (900, 0.001) Weights Real Weights Weights Biases Biases rmsprop **Epoch** Epoch