

Autoregressive Baseline Training Convergence

Smoothed (window=50)

Cross-Entropy Loss

