

Baseline vs. Shuffled: Eval Cross-Entropy

