MIA threshold Results (Hybrid-Output) 0.7 -**Generation Model** minGPT minGPT DP ( $\epsilon$ =1) 0.6 -0.5 -Score - 4.0 0.3 -0.2 -0.1 -0.0 AÚC F1 Score Attack Advantage Recall Accuracy Precision Metric