Trade-off on kmnist for TwoLayerDenseModel (raw) ($\tau = -0.5$