Comparison of ROC Gap and AUC across models, $q_t = 1.5, k = 3$

↓ Bias mitigation

↑ Discriminative performance

ROC Gap: 0.2, 0.1

AUC: 0.6, 0.4

- y-model (oracle)
- y-obs model
- Tested-only
- Group 0 only
- Group 1 only
- Group peer loss
- SELF
- DragonNet
- SAREM
- DCEM (ours)