Best performing model: Mean ROC (AUC $= 0.78 \pm 0.02$) Mean F1 score 0.82 ± 0.02

