Y2: performance of single vs ensemble model on subgroup-by-race: white one-model-for-all one-model-per-subgroup 100 ensemble-on-all ensemble-on-subgroup 85.5 85.8 87.3 87.6 83.1 83.2 84.3 84.3 81.1 82.8 81.4 78.5 79.3 80.7 80.4 79 80 Score(%) 60 40 20 0 precision f1 recall accuracy