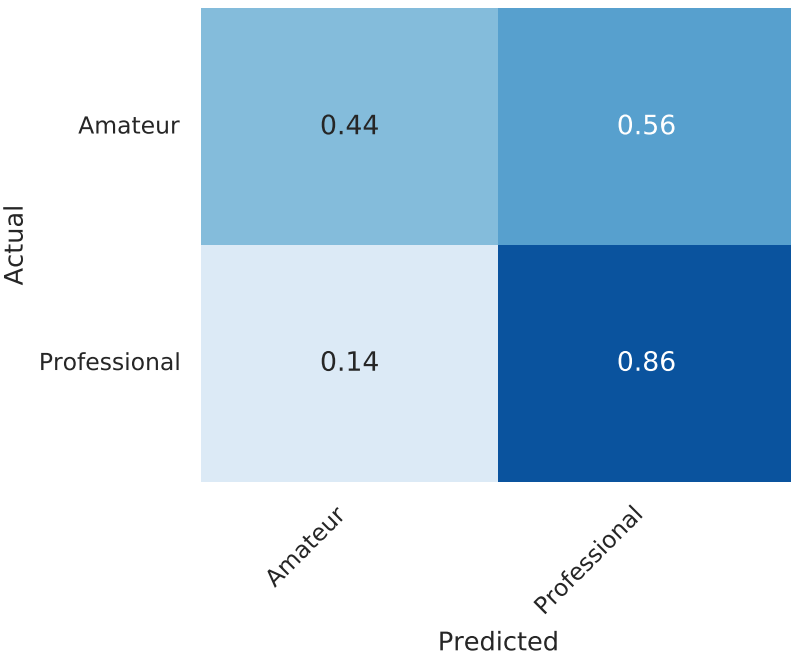
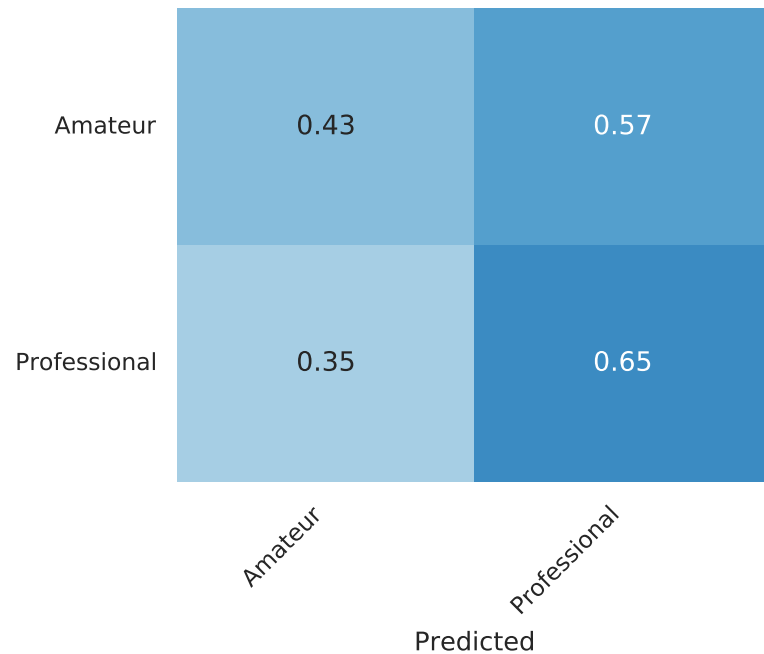


Sensor-only
F1 0.634 ± 0.150



Audio-only
F1 0.507 ± 0.058



Multimodal
F1 0.626 ± 0.141

