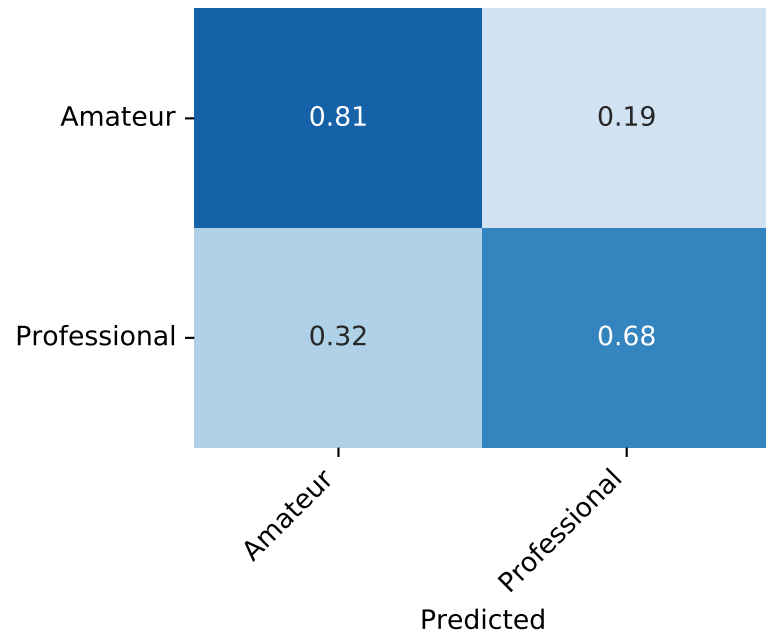
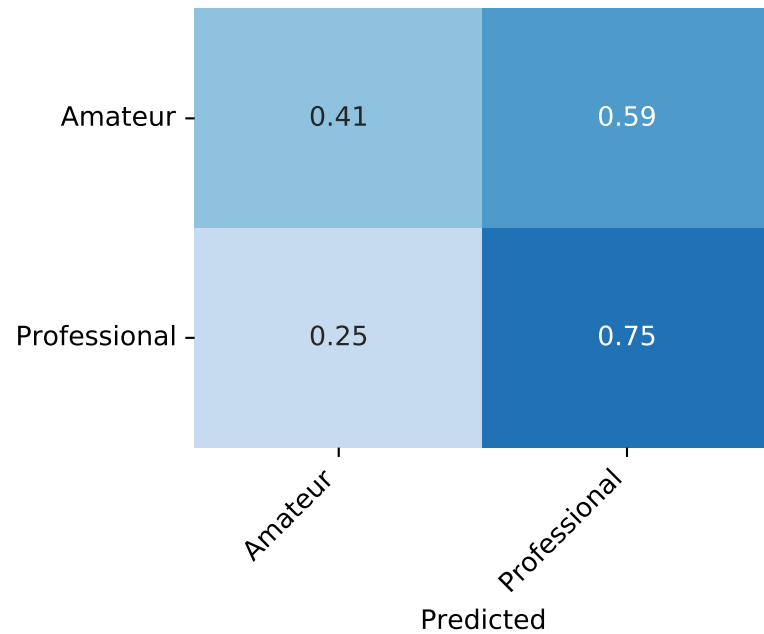


Sensor-only
F1 0.733 ± 0.029



Audio-only
F1 0.589 ± 0.030



Multimodal
F1 0.773 ± 0.021

