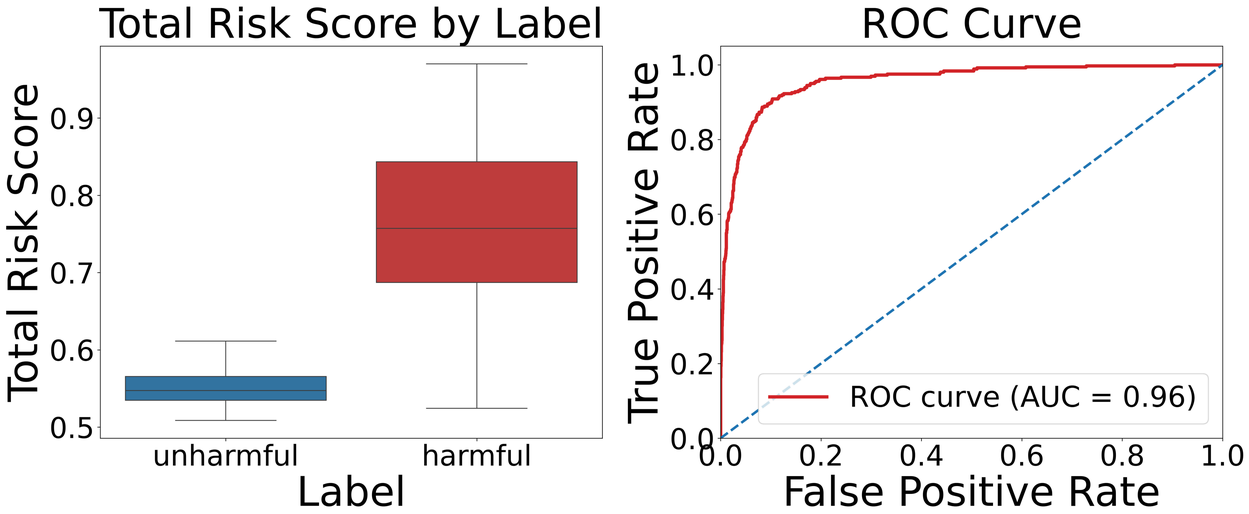
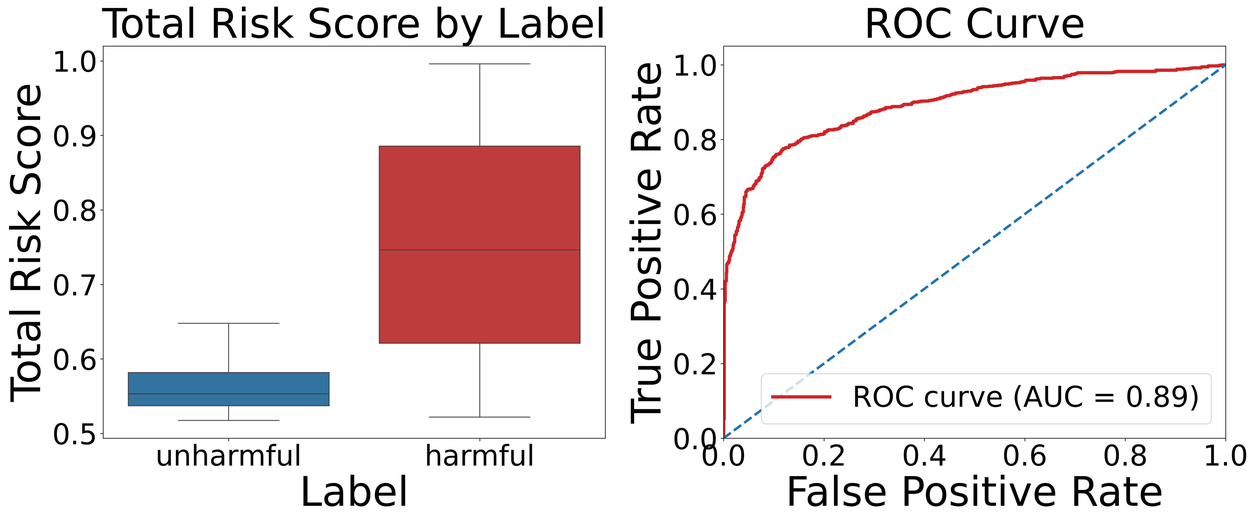


(a) ToxicChat



(b) WildGuardMix