halu_dialogue 16 models Logistic Regression 1.00 SVM XGBoost baseline 0.95 0.90 accuracy 58.0 0.80 0.75 0.4 0.6 0.5 0.8 0.3 0.7 0.9 1.0

uncertainty threshold