ToS on Gemini (More Assertive Prompt) - Accuracy and Misclassification Penalty per Language 0.64 Accuracy 0.62 0.61 **Penalty Score** 0.6 0.57 0.5 0.4 **Score** 0.3 0.22 0.21 0.20 0.19 0.2 0.1 0.0 1 en 90