Model Performance by Domain Across All Turns Model gemma\_2b gemma\_7b llama2\_13b 0.8 llama2\_7b Average Accuracy 0.6 0.2 0.0