Model Sensitivity to Bias (Fluctuation in Accuracy)

Y-axis: Standard Deviation of Accuracy (%)
X-axis: Models (Mistral, LLama, Mixtral, GPT)