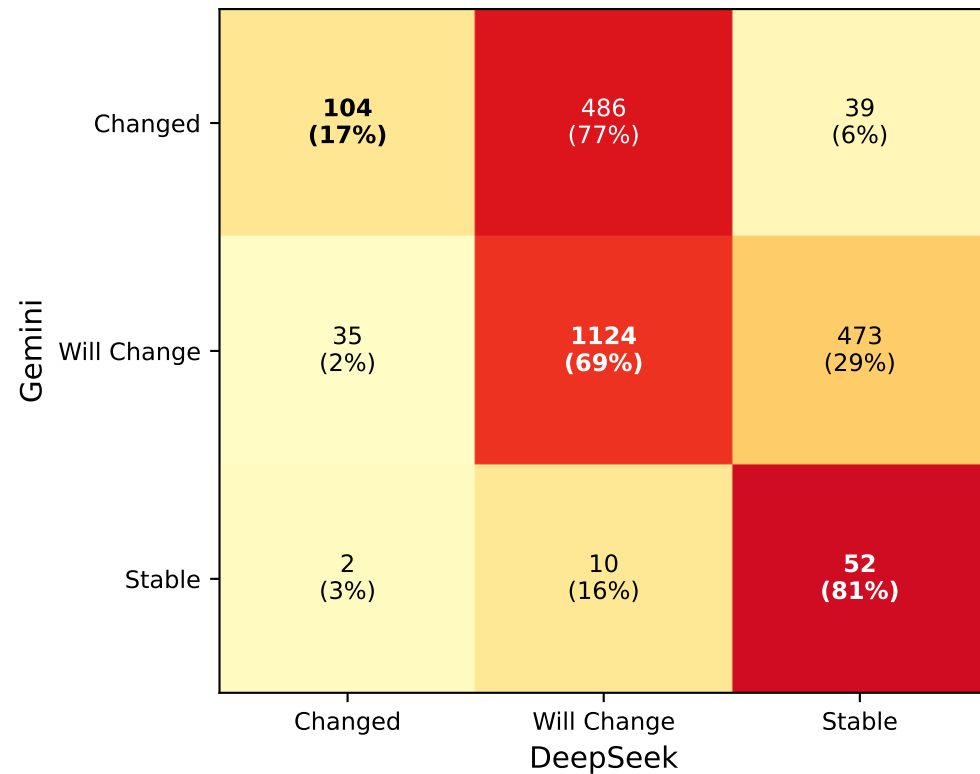
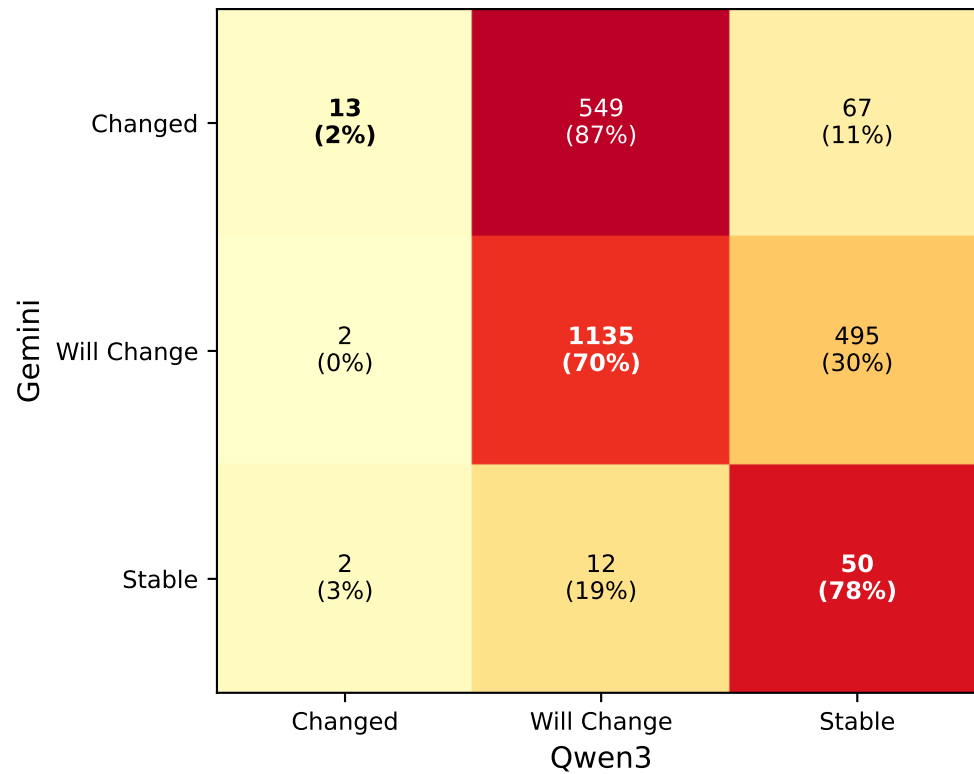


Confusion Matrices: change_status (Row-Normalized %)

Gemini-DeepSeek
 $\kappa = 0.079$



Gemini-Qwen3
 $\kappa = -0.013$



DeepSeek-Qwen3
 $\kappa = 0.448$

