

Attack Success Rate by Prompt Type

