

F1 score trends for different prompts

