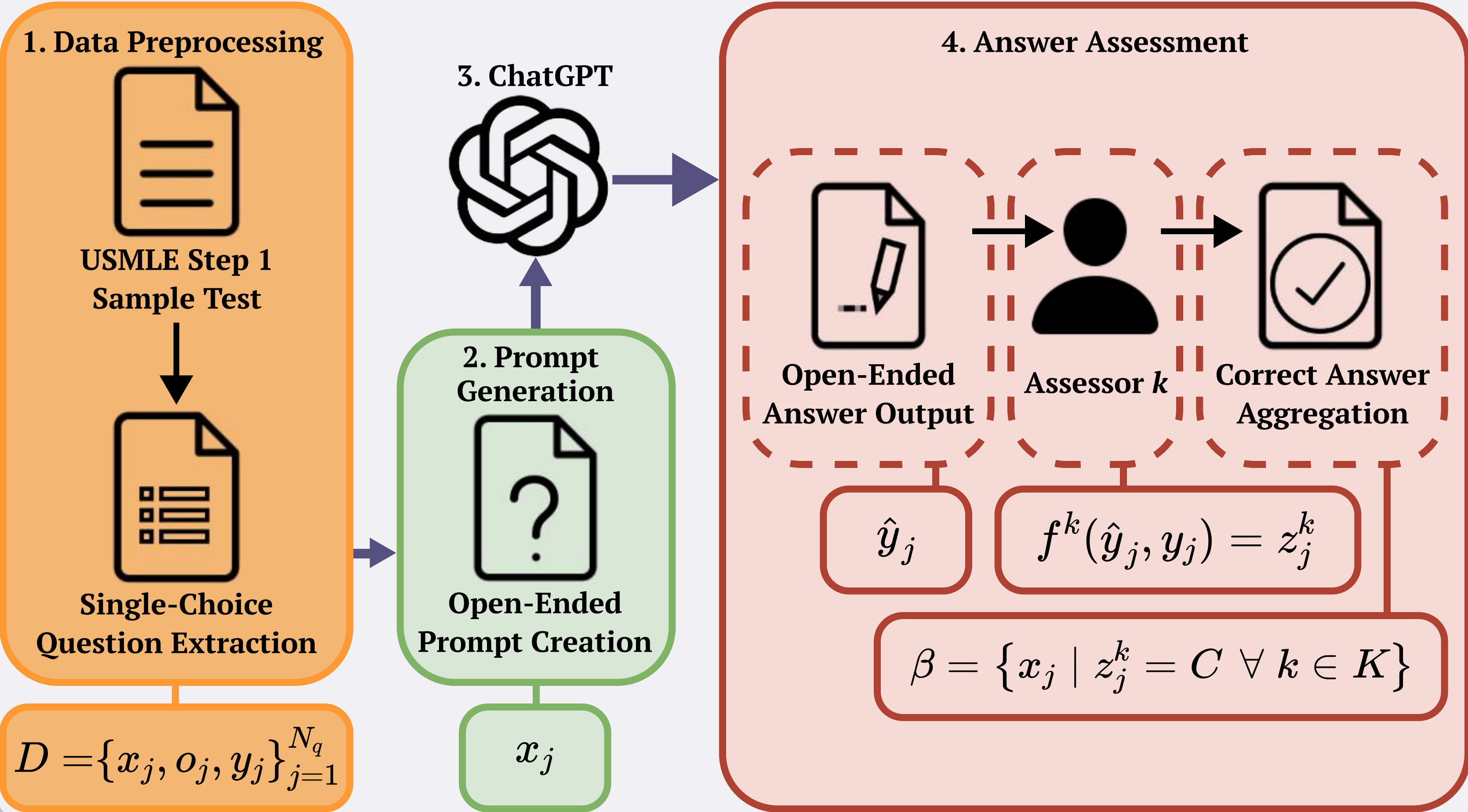


Experiment 1: ChatGPT Responses on USMLE Step 1 Open-Ended Prompts



Experiment 2: ChatGPT Robustness and Ablation Study Analysis

