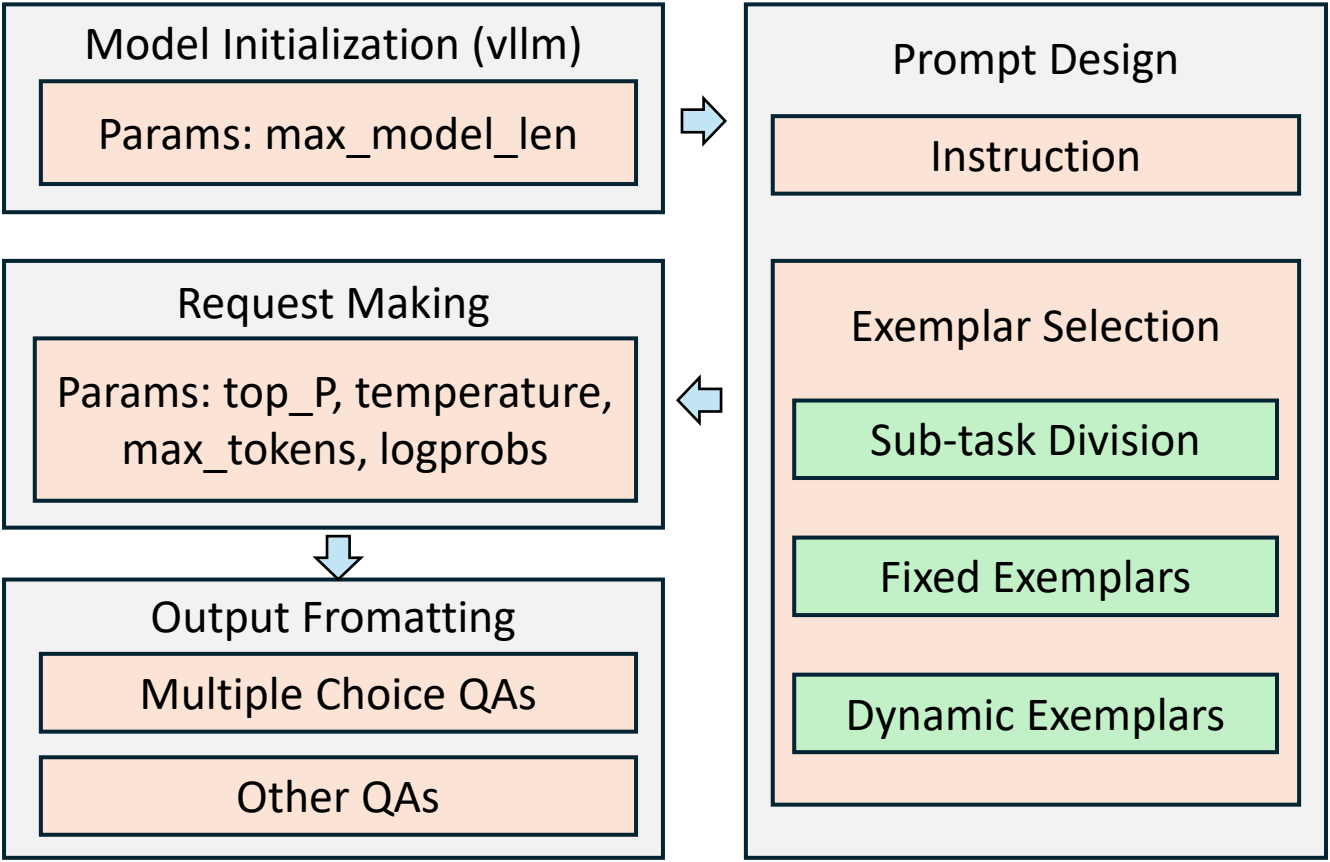
Model Initialization (vllm)
Params: max_model_len
Prompt Design
Instruction
Exemplar Selection
Sub-task Division
Fixed Exemplars
Dynamic Exemplars
Request Making
Params: top_P, temperature, max_tokens, logprobs
Output Fromatting
Multiple Choice QAs
Other QAs