Quality Evaluation - judged by LLAMA4

