Category	1 (lowest)	2	3	4	5 (highest)
Technical Complexity Quality of code, architectural soundness, and advanced use of AI/ML engineering.	Non-functional or hardcoded logic rather than AI/ML implementation	hallow implementation using a single, vanilla cloud API call with no custom data processing, orchestration, or training.	Proper prompt engineering with basic contextual routing, a straightforward RAG setup, or training/evaluating a classic machine learning model using standard libraries.	Complex multi-stage AI pipelines, such as optimized RAG (hybrid search, re-ranking), multi-agent coordination, custom model fine-tuning, or highly optimized in-browser/edge inference.	Flawless model orchestration, custom fine-tuning with rigorous evaluation datasets, or highly optimized quantization and inference pipelines with minimal latency and compute overhead. Implements innovative architectures that align with the problem being solved.
Data Safety & Responsibility Rigor of privacy measures, encryption, data minimization, and mitigation of AI bias.	No input/output validation, high risk of dangerous health hallucinations, or feeding raw, unencrypted user health data directly into public third-party APIs.	Relies purely on the default safety filters of a third-party provider. No custom system prompts or guardrails to prevent harmful medical advice, prompt injection, etc.	Implements explicit system prompt constraints to restrict the model to its medical bounds. Anonymizes user data before model processing and demonstrates awareness of dataset limitations.	Utilizes dedicated verification layers or guardrail frameworks (e.g., input/output checking models). Verifies and pre-processes training data or prompts to mitigate demographic bias and prevent data leakage	Features automated hallucination detection, robust defense against adversarial attacks, and explainable AI mechanisms to show why the model reached a health conclusion. Sensitive user data is completely protected, and bias is significantly mitigated.