

```
prompt_template = ""
```

Objective:

You are an expert software sustainability analyst evaluating code solutions for deep learning energy-efficiency. Your task is to analyze two solutions (Human Stack Overflow vs AI ChatGPT) using the established 8 energy-efficiency patterns with numerical scoring.

Task Description:

Evaluate both code solutions against the 8 core energy-efficiency patterns using a 1-5 scoring system. Provide a concise explanation highlighting key differences and evidence.

Scoring Framework (1-5 scale):

- 1: Poor - Pattern not implemented or severely deficient
- 2: Below Average - Basic implementation with significant issues
- 3: Average - Adequate implementation meeting basic requirements
- 4: Good - Solid implementation with clear energy benefits
- 5: Excellent - Exceptional implementation demonstrating best practices

Patterns and Associated Tactics:

1. PRE-TRAINED MODEL UTILIZATION

- Measurable Tactics: Transfer Learning (T16), Knowledge Distillation (T17)
- Evidence: Pre-trained model loading, fine-tuning, teacher-student architectures

2. CHECKPOINT MANAGEMENT

- Measurable Tactics: Checkpoint Usage (T19)
- Evidence: Model saving/loading strategies, resume training capabilities

3. MODEL OPTIMIZATION STRATEGIES

- Measurable Tactics: Reduce Complexity (T8), Enhance Sparsity (T14), Energy-aware Pruning (T15)

- Evidence: Simpler architectures, regularization, pruning implementations

4. QUANTIZATION TECHNIQUES

- Measurable Tactics: Input Quantization (T4), Quantization-aware Training (T18)

- Evidence: Precision reduction, quantization APIs, model optimization

5. EFFICIENT DATA HANDLING

- Measurable Tactics: Sampling (T1), Remove Redundant Data (T2), Feature Reduction (T3), Minimize Data Referencing (T27)

- Evidence: Data pipelines, sampling techniques, feature selection, generators

6. MEMORY MANAGEMENT

- Measurable Tactics: Memory Constraints (T20), Computation Partitioning (T22)

- Evidence: Device management, memory optimization, model partitioning

7. ALGORITHM & COMPUTATION OPTIMIZATION

- Measurable Tactics: Energy-efficient Algorithms (T6), Lightweight Alternatives (T7), Dynamic Parameter Adaptation (T10), Built-in Library Functions (T11)

- Evidence: Model architecture choices, adaptive learning, optimized operations

8. MODEL MAINTENANCE & ADAPTATION

- Measurable Tactics: Graph Substitution (T13), Informed Adaptation (T28), Retrain When Needed (T29)

- Evidence: Computation graph optimization, adaptive strategies

Evaluation Guidelines:

1. Score each pattern 1-5 for both solutions independently

2. Prioritize code evidence over textual recommendations
3. For Human answers: Credit clear textual suggestions even without full code
4. For AI answers: Focus on actual code implementations and practical suggestions
5. Explicit implementations get higher scores than vague mentions
6. Identify specific tactics (T1-T30) observed in code or text
7. Calculate average scores from all 8 patterns
8. Higher average score wins; tie only for exactly 0.0 difference
9. For minimal differences (≤ 0.125), consider tie only with compelling complementary strengths
10. Provide concise analysis highlighting key differentiators and code evidence

Reporting Instructions:

Return ONLY a JSON object with this exact structure:

```
{{  
  "human_pattern_1_score": 1-5,  
  "human_pattern_2_score": 1-5,  
  "human_pattern_3_score": 1-5,  
  "human_pattern_4_score": 1-5,  
  "human_pattern_5_score": 1-5,  
  "human_pattern_6_score": 1-5,  
  "human_pattern_7_score": 1-5,  
  "human_pattern_8_score": 1-5,  
  
  "ai_pattern_1_score": 1-5,  
  "ai_pattern_2_score": 1-5,  
  "ai_pattern_3_score": 1-5,  
  "ai_pattern_4_score": 1-5,  
  "ai_pattern_5_score": 1-5,  
}}
```

```
"ai_pattern_6_score": 1-5,  
"ai_pattern_7_score": 1-5,  
"ai_pattern_8_score": 1-5,  
  
"human_average_score": "calculated average",  
"ai_average_score": "calculated average",  
"score_difference": "difference in averages",  
"winner": "Human" or "AI" or "Tie",  
  
"analysis_explanation": "Brief explanation for your analysis, highlighting key differentiating  
factors in human and AI solutions, citing relevant code snippets. In case of tie, explain  
compelling complementary strengths that justify equal evaluation despite numerical  
differences."  
}}
```

Examples:

Example output of a tie:

```
{{  
  "human_pattern_1_score": 5,  
  "human_pattern_2_score": 3,  
  "human_pattern_3_score": 4,  
  "human_pattern_4_score": 2,  
  "human_pattern_5_score": 4,  
  "human_pattern_6_score": 3,  
  "human_pattern_7_score": 4,  
  "human_pattern_8_score": 3,  
  
  "ai_pattern_1_score": 3,  
  "ai_pattern_2_score": 4,  
  "ai_pattern_3_score": 3,
```

```
"ai_pattern_4_score": 5,
```

```
"ai_pattern_5_score": 3,
```

```
"ai_pattern_6_score": 4,
```

```
"ai_pattern_7_score": 3,
```

```
"ai_pattern_8_score": 5,
```

```
"human_average_score": 3.5,
```

```
"ai_average_score": 3.625,
```

```
"score_difference": 0.125,
```

```
"winner": "Tie",
```

```
"analysis_explanation": "Despite AI's slight numerical advantage (0.125), both solutions demonstrate truly complementary and equally valuable approaches. Human excels in Pre-trained Model Utilization (P1: score 5) with expert-level transfer learning strategies and practical fine-tuning guidance that significantly reduces training energy. AI counters with exceptional Quantization Techniques (P4: score 5) and Model Maintenance (P8: score 5) through advanced optimization implementations. The Human solution provides deep domain expertise that could save substantial computational resources in model development, while AI offers robust technical implementations for runtime efficiency. Given that both approaches address fundamentally different but equally critical aspects of the energy-efficiency pipeline, this evaluation results in a tie."
```

```
}}
```

Example output with clear winner:

```
{{
```

```
"human_pattern_1_score": 4,
```

```
"human_pattern_2_score": 3,
```

```
"human_pattern_3_score": 2,
```

```
"human_pattern_4_score": 1,
```

```
"human_pattern_5_score": 4,
```

```
"human_pattern_6_score": 3,
```

```
"human_pattern_7_score": 3,
```

"human_pattern_8_score": 2,

"ai_pattern_1_score": 3,

"ai_pattern_2_score": 4,

"ai_pattern_3_score": 3,

"ai_pattern_4_score": 4,

"ai_pattern_5_score": 3,

"ai_pattern_6_score": 4,

"ai_pattern_7_score": 4,

"ai_pattern_8_score": 3,

"human_average_score": 2.75,

"ai_average_score": 3.5,

"score_difference": 0.75,

"winner": "AI",

"analysis_explanation": "Both solutions effectively address the core `sess.run()` bottleneck by fetching entire weight matrices at once, improving Memory Management (P6) by reducing overhead and data transfers (Human: `h1_val = sess.run(weights['h1'])`, AI: `weight_values = sess.run(tensor)`). The Human solution provides a solid Algorithm & Computation Optimization (P7) by fixing the data extraction, but still uses manual loops for writing. The AI solution excels here by implementing `np.savetxt(fp, weight_values)`, leveraging highly optimized built-in libraries for superior computational efficiency and streamlined I/O. AI's more complete and robust approach, including support for multiple layers, provides a clearer best practice for energy-efficient weight serialization."

}}

Now evaluate the following:

Question:

{question}

Human (Stack Overflow) Answer:

{human_answer}

AI (ChatGPT) Answer:

{ai_answer}

""""