

Logic-LM Task 5: Testing and Results

For Task 5, I reimplemented a simplified but fully compliant version of the Logic-LM reasoning pipeline. The implementation integrates a real large language model (LLM) from Together AI (Llama-3-8b-chat-hf) to translate natural language questions into symbolic logic.

I converted the CIFAR-10 image labels (like "cat", "truck", etc.) into symbolic facts and wrote simple logic rules to define which labels are animals and which are vehicles. Using a backward-chaining symbolic solver, I was able to answer questions such as "Is cat an animal?" or "Is truck a vehicle?" by checking these facts and rules.

A self-refinement loop is included to iteratively improve the symbolic logic formulation based on solver feedback, aligning with the core Logic-LM paper workflow and instructor feedback.

When I ran the tests in Google Colab, the solver correctly answered "Yes" for cases like "cat" being an animal and "truck" being a vehicle, and "No" when the label didn't fit the rule (like "automobile" as an animal). Some queries produced negated symbolic forms, which the solver currently treats as negative answers.

GitHub Repo: <https://github.com/suvalavala/llm-logic-kb>