

MRL Shared Task: Gujarati

Shaan Shah, Kaustubh Ponkshe, Raghav Singhal, Nitish Ponkshe

Our submission to the MRL shared task comprises 100 question-answer pairs in Gujarati, each designed to assess physical reasoning grounded in common sense. The dataset emphasizes cultural relevance, drawing upon familiar contexts and scenarios specific to Gujarati daily life.

The questions span a diverse set of domains, including household activities, local festivals, food, school settings, kitchen tools, farm life, animals, seasons, games, common objects, physics and geography; all reflective of Gujarati customs and environments. Each example is crafted to be answerable by a native Gujarati speaker through everyday reasoning, without requiring factual recall or specialized knowledge.

We adopted a human in the loop approach that combined creative guidance from a native Gujarati speaker with assistance from a frontier large language model (ChatGPT-5 Instant). The dataset was not directly generated by the LLM; rather, the model was used selectively as a co-creative tool under strict human supervision.

The process followed these steps:

- A human annotator (a native Gujarati speaker) selected culturally meaningful settings, such as traditional practices, local objects, or common physical activities and provided them to the LLM along with the task guidelines.
- The LLM was prompted to generate a set of question and answer pairs in English with Gujarati translations, ensuring alignment with the physical reasoning and common sense criteria.
- The human annotator curated the LLM's outputs, selecting the most suitable examples and refining them as necessary for clarity, fluency, and cultural relevance.
- Each example underwent manual verification to ensure linguistic accuracy and quality.

All examples in the final dataset were manually checked for quality and adherence to the task requirements.