

Bengali Physical Commonsense Reasoning Dataset for MRL 2025 Shared Task

1. Introduction

We present a Bengali dataset for the MRL 2025 Shared Task on Multilingual Physical Commonsense Reasoning. The dataset follows the [PIQA](#)-style multiple-choice format and is designed to evaluate the ability of models to perform everyday physical reasoning in Bengali. Bengali is the seventh most spoken language in the world, yet high-quality benchmark datasets for commonsense reasoning in Bengali remain limited. This contribution addresses that gap with culturally grounded examples reflecting daily life in Bangladesh and West Bengal, India.

2. Task Scope and Alignment

The shared task defines three central requirements:

1. **Physical Reasoning:** Each example involves reasoning about the physical properties of objects or substances.
 - In our dataset, examples focus on material behaviors (e.g., metals rusting, water evaporating, fabric drying in the sun, food spoiling in humidity).
2. **Commonsense Knowledge:** The correct answer is evident to an average native Bengali speaker.
 - All prompts were designed to reflect familiar experiences such as handling kitchen utensils, seasonal weather changes, or traditional storage practices.
3. **Culturally Specific Content:** Whenever possible, prompts highlight Bengali cultural and environmental contexts.
 - Examples include *lep* (cotton quilt), *pata shukano* (drying leaves in the sun), winter fog, clay utensils, and local food preparation methods. These capture cultural nuance and make the dataset authentic.

By grounding the dataset in shared everyday knowledge, we ensured that the task remains realistic and meaningful for native speakers, while still presenting a challenge to computational models.

3. Data Collection Methodology

The dataset was entirely manually constructed by a native Bengali speaker with expertise in both language and physical commonsense. The process involved several steps:

1. **Topic Brainstorming:** A list of candidate topics was generated from everyday Bengali life such as household chores, seasonal weather, agriculture, cooking, storage, and material usage.
2. **Scenario Design:** Each example begins with a short scenario or question written in natural and conversational Bengali. These scenarios describe familiar real-world situations (e.g., drying a cotton quilt in the sun, washing a hot pan with cold water). Both single sentence and multi sentence variations were included to provide diversity.
3. **Solution Construction:** For each prompt, two solutions were created:
 - One correct, based on real physical commonsense.
 - One incorrect, written to be plausible but subtly wrong, ensuring the task is non-trivial.
4. **Manual Validation:** Each example was reviewed for linguistic clarity, correctness, and adherence to the task scope.
5. **Final Compilation:** The dataset was formatted into the required TSV structure (`prompt`, `solution0`, `solution1`, `label`).

We did not use LLMs or automatic translators to design the context; all examples were manually written and curated. This ensures cultural authenticity and reduces risks of overly generic patterns.

4. Dataset Statistics

- **Total examples:** 103
- **Average prompt length:** ~22 words
- **Longest prompt:** 48 words
- **Dialect/Script:** Standard Bengali, written in Eastern Nagari script
- **Domains covered:** Household practices, weather phenomena, agriculture, food preparation, material interactions

5. Illustrative Examples by Task Scope

To demonstrate the nature of the dataset and its alignment with the shared task scope, we present representative examples categorized into three dimensions: **Physical Reasoning**, **Commonsense Knowledge**, and **Cultural Specificity**. Each entry contains the original Bengali prompt, its English translation, candidate solutions, and the correct label.

A. Physical Reasoning

These examples require reasoning over physical properties of materials and objects.

Bengali Prompt	English Translation	Solution 0	Solution 1	Correct Label
একটি গরম ভাজার হাঁড়ি ধুয়ে সঙ্গে সঙ্গে ঠান্ডা পানিতে ডুবিয়ে দিলে কী হতে পারে?	What happens if a hot frying pan is immediately dipped into cold water after washing?	হাঁড়ির ধাতু ঠান্ডা হয়ে আরও বেশি চকচকে হয়ে যাবে। (The metal of the pan will cool down and become shinier.)	হাঁড়ির ধাতব পৃষ্ঠে তাপমাত্রা পরিবর্তনে বিকৃতি আসতে পারে। (The metal surface of the pan may deform due to the sudden temperature change.)	1
একটি তালা অনেকদিন খোলা হয়নি, হঠাৎ খুলতে গেলে কী ধরনের সমস্যা হতে পারে?	If a lock has not been opened for a long time, what problem might occur when trying to open it suddenly?	তালা ধাতব হলেও বাতাসে শুকিয়ে গেলে নিজে থেকেই খুলে যায়। (Even though the lock is metallic, it dries out in air and opens by itself.)	তালার ভেতরের ধাতব অংশে জং ধরে খুলতে কষ্ট হতে পারে। (The internal metal parts may rust, making it hard to open.)	1

B. Commonsense Knowledge

These examples rely on widely shared knowledge accessible to any average Bengali speaker.

Bengali Prompt	English Translation	Solution 0	Solution 1	Correct Label
শীতকালে কুয়াশায় বাইরের কাপড় ভিজে যায়, কেন?	Why do clothes left outside get wet in the fog during winter?	শীতের সময় বাতাসে আগুনের উপাদান থাকে যা কাপড়ে জমে। (In winter, the air contains fire elements that collect on clothes.)	বাতাসে জলীয় বাষ্প জমে কাপড়ে পানির ফোঁটা তৈরি করে। (Water vapor in the air condenses and forms droplets on the clothes.)	1
একটি গ্লাসে ঠান্ডা পানি ঢাললে গ্লাসের বাইরের দেয়ালে পানি জমে যায় কেন?	Why does water form on the outside of a glass when cold water is poured inside?	গ্লাস ভিজে যায় কারণ পানি বাইরে থেকে ভেতরে প্রবাহিত হয়। (The glass gets wet because water flows from outside to inside.)	বাতাসে থাকা জলীয় বাষ্প ঠান্ডা গ্লাসে জমে পানি ফোঁটা তৈরি করে। (Water vapor in the air condenses on the cold glass and forms droplets.)	1

C. Cultural Specificity

These examples reflect culturally embedded objects, foods, and practices from everyday Bengali life.

Bengali Prompt	English Translation	Solution 0	Solution 1	Correct Label
একটি মোটা তুলার লেপ রোদে কয়েক ঘন্টা রেখে দিলে কী হয়?	What happens if a thick cotton quilt (<i>lep</i>) is left in the sun for a few hours?	লেপটি রোদে নরম হয়ে গিয়ে শীতল হয়ে যায়। (The quilt becomes softer and cooler in the sun.)	লেপের ভেতরের আর্দ্রতা শুকিয়ে গিয়ে এটি আরও হালকা হয়। (The moisture inside the quilt dries out, making it lighter.)	1

গ্রামে এক বৃদ্ধা রোদে পাতা শুকাচ্ছেন। তিনি কেন এমনটা করছেন?	In a village, an elderly woman is drying leaves in the sun. Why is she doing this?	রোদে পাতাগুলো নরম হয়ে গেলে এগুলো রান্না করা যায়। (Once the leaves soften in the sun, they can be cooked.)	রোদে পাতার ভেতরের পানি শুকিয়ে গেলে সংরক্ষণ সহজ হয়। (When the moisture in the leaves dries up, they can be preserved easily.)	1
---	--	--	---	---

6. Discussion and Future Work

This dataset highlights the importance of grounding physical commonsense in culturally relevant scenarios. Many phenomena that are trivial in Bengali contexts (e.g., the behavior of a *lep* in sunlight or condensation in winter fog) may not easily translate into English. Future work could extend the dataset to include:

- Dialectal variations (e.g., Sylheti, Chittagonian).
- Examples from formal education contexts (e.g., school science problems).
- Parallel translations into English for cross-lingual evaluation.

7. Conclusion

We provide a Bengali dataset of 103 examples for physical commonsense reasoning. By carefully aligning with the shared task scope such as physical reasoning, commonsense knowledge, and cultural grounding and ensuring all examples were manually written, this dataset contributes to building a more inclusive multilingual benchmark.