

Urdu Physical Commonsense Reasoning Dataset for MRL 2025 Shared Task

Maryam Sikander

Abstract

This document presents the Urdu Physical Commonsense Reasoning Dataset, developed for the MRL 2025 Shared Task. The dataset consists of 115 samples focusing on physical commonsense, culturally specific knowledge, and everyday tasks in Standard Pakistani Urdu. Each sample includes a prompt and two candidate solutions (one correct, one plausible incorrect), generated with the assistance of large language models and rigorously validated by two native Urdu speakers for cultural relevance, accuracy, and linguistic naturalness.

1 Introduction

The MRL 2025 Shared Task on Multilingual Physical Commonsense Reasoning evaluates models’ ability to apply everyday physical reasoning across different languages. This description paper introduces the Urdu dataset prepared for the task. Urdu is the national language of Pakistan and is spoken or understood by more than 231 million people worldwide [1]. Despite its wide usage, Urdu remains underrepresented in commonsense reasoning benchmarks.

The dataset is modeled after PIQA (Bisk et al., 2020), adapted to the Pakistani cultural context. It captures reasoning about physical properties, actions, and consequences that reflect daily life in Pakistan, including food preparation, household practices, and social customs.

2 Dataset Characteristics

The dataset comprises 115 distinct samples, each consisting of a prompt, ‘solution0’, ‘solution1’, and a binary ‘label’ (0 if ‘solution0’ is correct, 1 if ‘solution1’ is correct).

Content Distribution: The examples are categorized as follows:

Category	Number of Questions
Cooking	36
Everyday	54
Custom	14
Culture	11
Total	115

Table 1: Distribution of dataset examples across categories.

Linguistic and Contextual Specifics: The dataset is constructed in ****Standard Pakistani Urdu****. It heavily features culturally specific scenarios, including:

- **Local Cuisine:** Practical advice on preparing dishes like **Gajar ka Halwa**, **Nihari**, **Biryani**, and traditional sweets (**Gulab Jamun**), along with tea-making techniques.
- **Household Practices:** Methods for preserving food (e.g., milk, meat), cleaning, and maintaining everyday objects in a Pakistani home context.
- **Social Customs:** Etiquette for guest reception, practices around Eid and **Chaand Raat**, and naming ceremonies.
- **Traditional Crafts:** Descriptions of regional crafts such as Saraiki **Khes** weaving and Chitrali Topi adornments.

These examples often require an understanding of cultural context beyond mere literal translation, thereby enriching the common sense reasoning challenge.

3 Conclusion

This dataset provides culturally grounded examples of physical commonsense reasoning in Urdu. By combining LLM-assisted generation with human validation, it is a resource aligned with the goals of the MRL 2025 Shared Task and contributes to the representation of Urdu in multilingual evaluation.