Russian dataset for physical reasoning

Ivan Belashkin
Independent
`t2468ttt@gmail.com`

Abstract

This dataset contains 100 questions in Russian about physical reasoning.

1 Description of dataset

Dataset contains 100 manually created questions in Russian. These questions cover topics such as cooking, safety measures, basic physics, and basic computer use. Some questions are also based on Russian culture.

I used common sense, and basic knowledge of physics, avoiding complex questions. Some physical paradoxes were included. I also included cooking tasks, including ones related to Russian culture.

Some incorrect answers don not follow pattern of 'moving or changing one or two words in the correct answer', but they were included for plausibility.

2 Format

The TSV file (100 lines), which contains 5 columns: question, answer 0 (correct), answer 1 (incorrect), position of correct answer (always "0"), comma-separated tags ('ru', 'safety', 'phys', 'cooking', 'comp', or 'no' if no tags).

Description of tags:

- cooking - for questions related to recipes.

- comp - for basic computer-related questions.

- no - for questions without tags.

- phys - for physical questions (almost every such question of this type can appear in a physics textbook).

- ru - for Russian traditions, food, etc.

- safety - for safety rules.

3 Notice

The text of this PDF document was corrected with LLM-tools. Dataset created manually.