

Introduction

The present article provides a dataset for the Ukrainian Physical Commonsense Task. The dataset was inspired by existing Ukrainian websites and blogs, as well as by my own knowledge. Each example includes a prompt along with two solutions, labeled as correct and incorrect.

Methodology

The solutions to the prompts pertain to the physical properties of one or more objects and are culturally specific, particularly to the Ukrainian culture. The two provided solutions differ by only one or two words. I have endeavored to balance examples of varying lengths—short, medium, and long—and to avoid initiating solutions in the same manner to enhance diversity. Furthermore, the examples were developed with consideration of the common understanding of an average individual.

In total, there are:

- **Short** solutions: < 10 words
- **Medium** solutions: around 11-20 words
- **Long** solutions: more than 21 > words

At first, I tried to equilibrate the number of short, medium, and long solutions, but, unfortunately, it was not possible since cutting or adding some words would compromise the quality and understanding of the solution. In the end, I had 40 short, 32 medium, and 28 long solutions.

The final dataset comprised 101 examples categorized into the following sections: Ukrainian cuisine (food), miscellaneous, everyday objects, and Ukrainian traditions. The miscellaneous section encompasses topics such as locations, puzzles, and life hacks, whereas Ukrainian traditions include superstitions and local Ukrainian festivities. The examples were created by me, a 31-year-old native Ukrainian speaker, who selected the prompts and solutions. Verification was subsequently conducted by a 59-year-old native Ukrainian speaker from Western Ukraine. Both of us were born in the same city in Western Ukraine and currently reside in Portugal.

Each prompt contains two solutions — solution0 and solution1. The correct assignment of the solution was chosen randomly, e.g., for one prompt, either solution0 or solution1 was correct. The examples were written in Standard Ukrainian in the Cyrillic alphabet. Each sentence for the prompts and solution was written as either referring to the second person plural¹, third person plural, or as an impersonal style/general suggestions, with omission of the subject. No translation to English was provided. However, to exemplify the content of the prompts below, please find the examples of two prompts and their respective correct

¹ In Ukrainian, the second person plural can refer to many people or to one person in a formal way (you – *Bu*).

and incorrect solutions translated from Ukrainian into American English. Please note that the following examples are culturally specific, so their translation might not make much sense in English, even though I tried my best to provide the most accurate translation of the prompts and solutions. The word in bold is the only one that changes between the correct and incorrect solutions.

Example 1.

```
{"prompt": "If you bake pancakes using kefir",  
"solution0": "they will be full.",  
"solution1": "they will be plain.",  
"label": 0}
```

Example 2.

```
{"prompt": "A woman who breaks plates or cups on purpose",  
"solution0": "tries to harm people who live in the house, or to place everything in order, or  
to put a bad eye.",  
"solution1": " tries to harm people who live in the house, or to ruin everything, or to put a  
bad eye.",  
"label": 1}
```

[Note: the highlighted words in Ukrainian make up one word only].

The final dataset includes nine columns:

1. Prompt
2. Solution0
3. Solution1
4. Label (correct or incorrect)
5. Word count of the solutions
6. Notes (a column to add additional notes about the prompt or solutions)
7. Topic
8. Length (describes the length of the solution: short, medium, or long)
9. Source (link to the webpage that inspired the prompts/solutions)

There was also no specific rule for writing numbers inside the prompts or solutions, so sometimes they appeared as numerals (e.g., 5) and other times in written form (e.g., five).