Dataset Description Paper : PIQA Thai Language

Naphat Permpredanun

1 Dataset Description

This dataset composes of the 105 examples of the PIQA data in Thai Language. The dataset includes

1. **Physical reasoning**: The example relates to physical properties of one or more objects. There are 65 examples of this type. All of them are categorized based on the following topics :

- Heat and Temperature : Heat preservation, Effect of heat

- Water and Humidity : Comparison of objects on water absorption, etc.

- Force, Motion and Stability : Balancing things, Two objects collision, Movement of object after impact

- Buoyancy and Floating : Idea of density in real life

- Everyday Objects : Usage of utensils or woodwork

- Scientific Facts : Ideas of weight and volume, Natural phenomenon

2. **Common sense**: The examples include what average person who speaks your language natively should know the answer. We include around 20 of these examples. The example targets in bringing the fact within Thailand based on the weather, season and routine. This includes the manner, actions, and the general landscape of Thailand.

For example, the fact that the temperature of Thailand is not too cold in any season compared to USA, we have the different appliances for each house, etc.

3. **Culturally specific**: The examples include culturally-relevant examples for their language(s). For example, some items may not be easily translatable into English, or may require regional and/or cultural commonsense.

For this group, which contains 20 of them, are purely characteristic of the objects (or a part of). This part includes landmarks, arts, culinary, and manner that are unique to Thailand. The difference between 2 solution should be about description of the main appearance of the objects since some other minute details of these culturally specific things could be adapted.