

MRL 2025 dataset description: Norwegian

Lilja Øvrelid, University of Oslo

The MRL 2025 dataset for Norwegian covers the Bokmål written variant of Norwegian and was manually created from scratch by a native speaker of Norwegian. The dataset consists of 100 questions requiring physical commonsense reasoning along with two possible solutions. The dataset creation focuses on including a combination of Norwegian-specific activities, such as the preparation of traditional food dishes or the use of traditional objects and reasoning over more generic, everyday activities that require physical commonsense. The more generic examples were constructed using the Norwegian Wordnet resource, focusing on physical, man-made objects and questioning aspects of their construction or use (e.g. how to screw in a light bulb or how to empty a vacuum cleaner). The Solutions represent minimal pairs that represent slight variations over the same answer (typically obtained by replacing one or two central words in the correct solution).