

Datathon@IndoML 2021

Cash prizes up
to ₹3,50,000
to be won!

Register Now*



*Mandatory to be eligible for prizes

GLUECoS-NLI-Knowledge

This is a Conversational Code-switched Natural Language Inference dataset, in English and Hindi (written in Roman script). Here, the premise is a multi-turn dialog and hypothesis is a sentence. Deviating from the traditional NLI task, the goal is to determine whether inferring the entailment label for a premise-hypothesis pair would require some sort of external knowledge or not.

URL:

<http://www.kaggle.com/c/datathon-gluecos-nli>

IndoRE

IndoRE is a relation extraction task for three low-resource Indian Languages (Bengali, Telugu, and Hindi).

URL:

<http://www.kaggle.com/c/indore-datathon-2021>

E-Manual

This is a *question-answering task*. Here the dataset is a set of manuals of electronic instruments, and the goal here is to extract an answer and/or the section from the manual where an answer is present for a given question.

URL: <http://www.kaggle.com/c/e-manual>

Multilingual Abusive Comment Identification

Multilingual Abusive Comment Identification Challenge is a challenge towards combating abusive comments (in multiple Indian regional languages) on Moj, which is one of India's largest short-video apps.

URL:

<http://www.kaggle.com/c/multilingualabusivecomment>