**Reasoning Dataset:**

LogiQA (H.Liu et al., 2023)

Common Sense QA (Talmor et al., 2018)

Social Intelligence QA (SIQA) (Sap et al., 2019)

Algebra Math with Rationale (AQuA-RAT) (Ling et al., 2017)

**LLM Models:**

1. Mistral 7B
2. Mistral 8\*7B
3. Llama3 8B
4. Llama2 13B
5. Llama2 7B

We can get the model interface from hugging face website.

The study also features the MistralAI Mistral-7B-Instruct-v0.2 model, which we compare with other prominent language models. Although Mistral-7B excels at generating text, its efficiency depends heavily on the hardware setup. We also assess the MistralAI Mixtral-8 x 7B-Instruct-v0.1, designed to optimise performance across diverse environments.

Additionally, the Meta LLaMA 3 8B model is included and it is built to generate text similar to human writing using a comprehensive training dataset. Alongside this model, the meta-llama / LLaMA-2-7b-hf and meta-llama / LLaMA-3-13b-hf offer scalable alternatives for hardware-constrained environments while maintaining strong performance and efficient resource usage.