

Domain



Subjective NLP



Math

Dataset

LLM: Qwen3-32B

Judge: Qwen3-32B

Judge: DeepSeek-R1-8B

MP

CSC

Paraphrase

VariErrNLI

PRM800K

AIME

0

10000

20000

Number of Chars

0

5000

Number of Chars

0

20000

Number of Chars

