GPT-4 3-shot accuracy on MMLU across languages
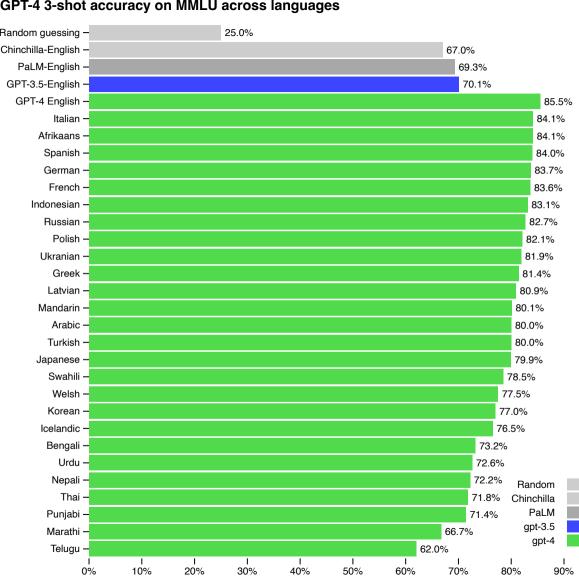
Language	Accuracy
Random guessing	25.0%
Chinchilla-English	67.0%
PaLM-English	69.3%
GPT-3.5-English	70.1%
GPT-4 English	85.5%
Italian	84.1%
Afrikaans	84.1%
Spanish	84.0%
German	83.7%
French	83.6%
Indonesian	83.1%
Russian	82.7%
Polish	82.1%
Ukranian	81.9%
Greek	81.4%
Latvian	80.9%
Mandarin	80.1%
Arabic	80.0%
Turkish	80.0%
Japanese	79.9%
Swahili	78.5%
Welsh	77.5%
Korean	77.0%
Icelandic	76.5%
Bengali	73.2%
Urdu	72.6%
Nepali	72.2%
Thai	71.8%
Punjabi	71.4%
Marathi	66.7%
Telugu	62.0%

Legend:
- Random
- Chinchilla
- PaLM
- gpt-3.5
- gpt-4

Accuracy →