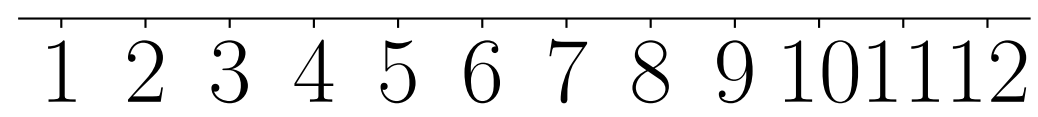
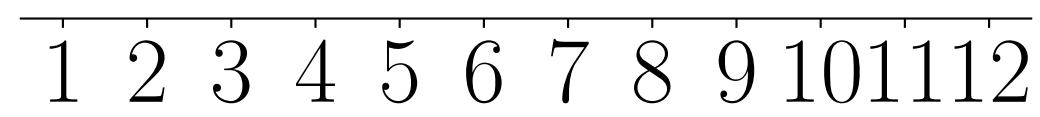


# Vicuna

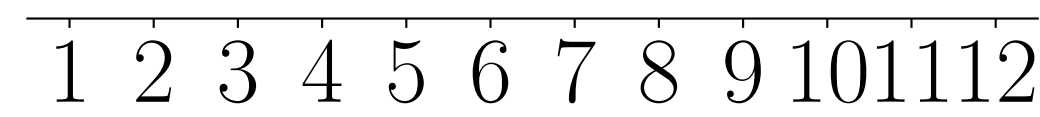
BASELINE



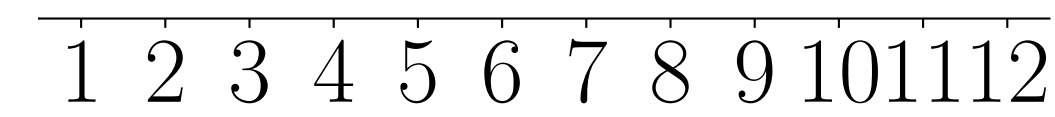
LLM GPT4



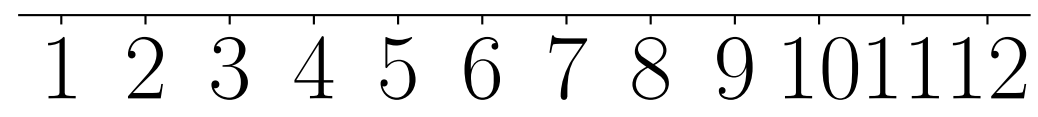
LLM CL3



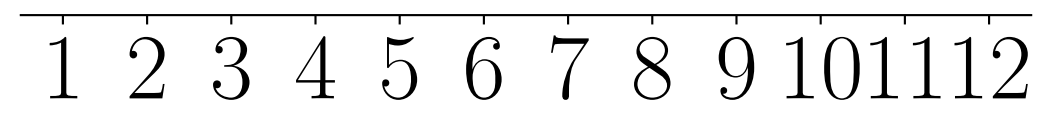
LLM GPT3.5



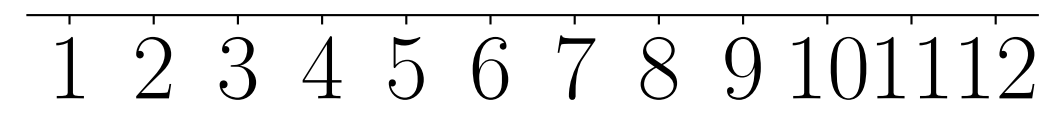
HUMAN ONLY



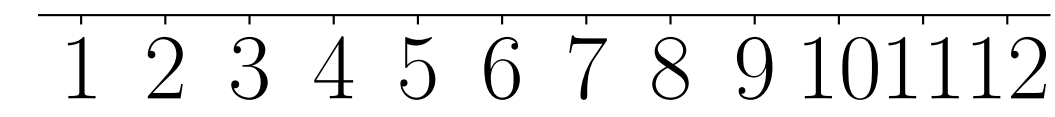
PPR GPT4



PPR CL3



PPR GPT3.5



Empirical probability

Ranking position

Ranking position

Ranking position

Ranking position