input → word

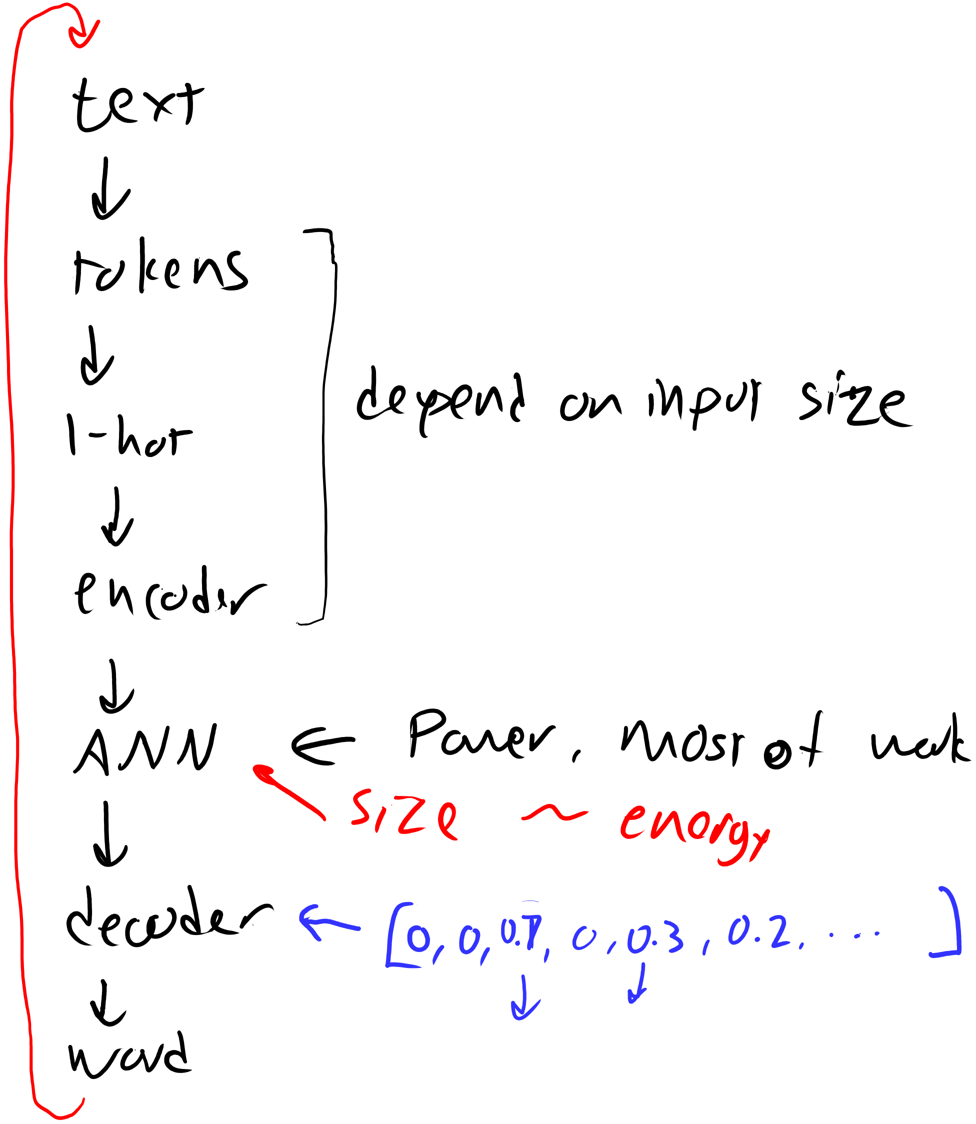
input + word → word

input word word → word

Questi

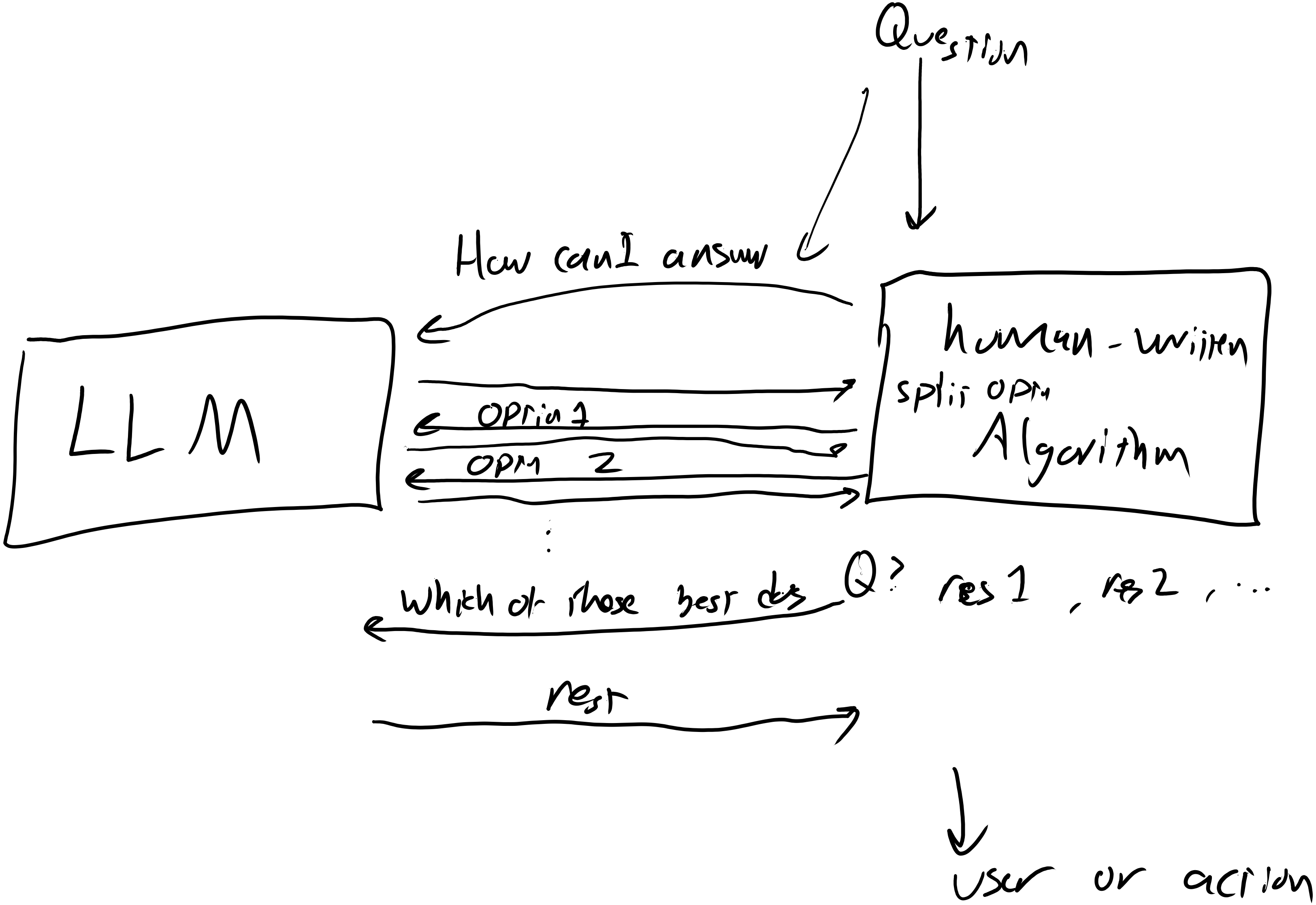
Answer

Question



Output size ~ energy

Energy used by American ⇝ 140 LLM tokens per second


Question
How can I answer
LLM
human - written
split option
Algorithm
option 1
option 2
:
Which of those best does
Q? res 1, res 2, ...
rest
user or action

Questions

Agentic vs Generative

next few years