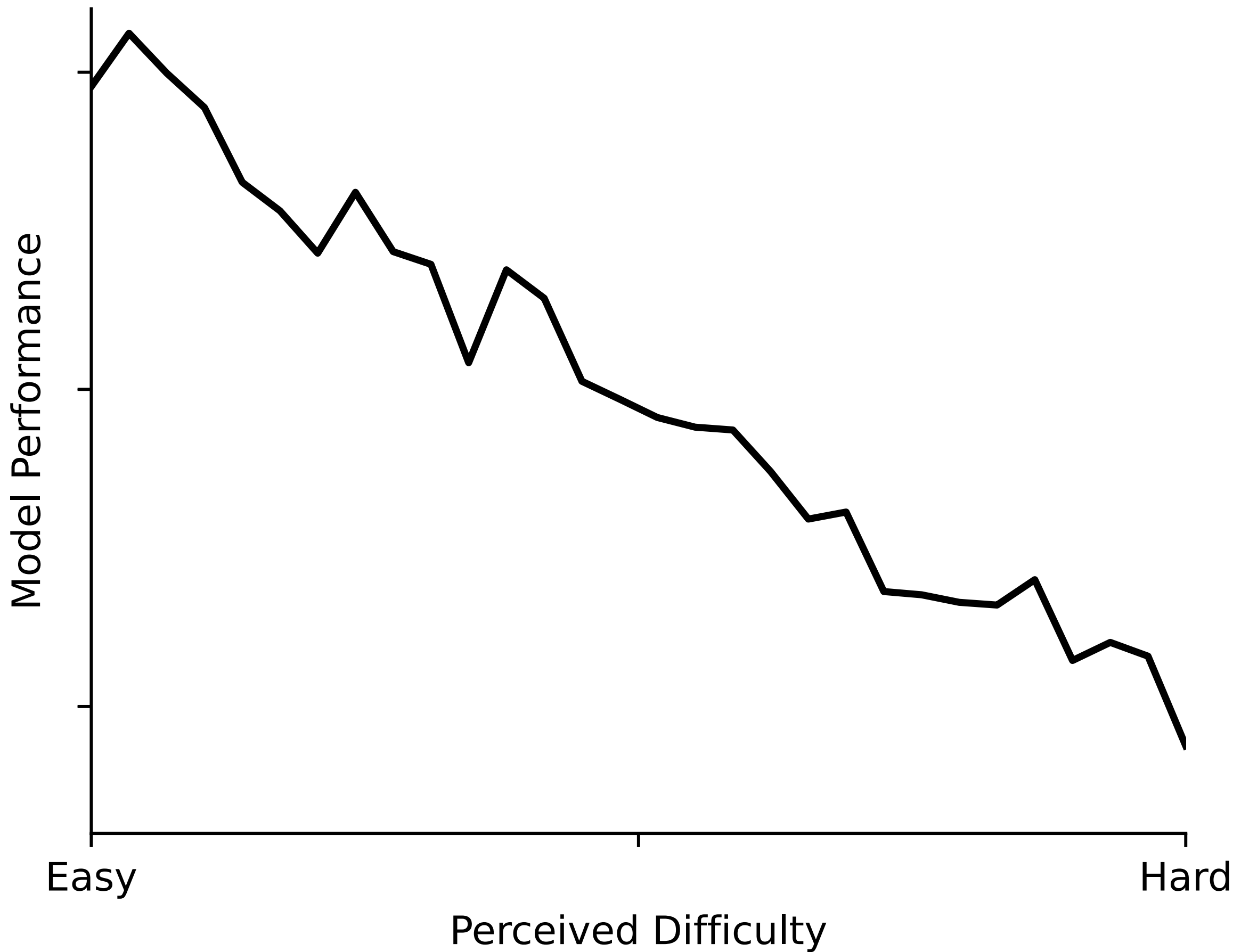


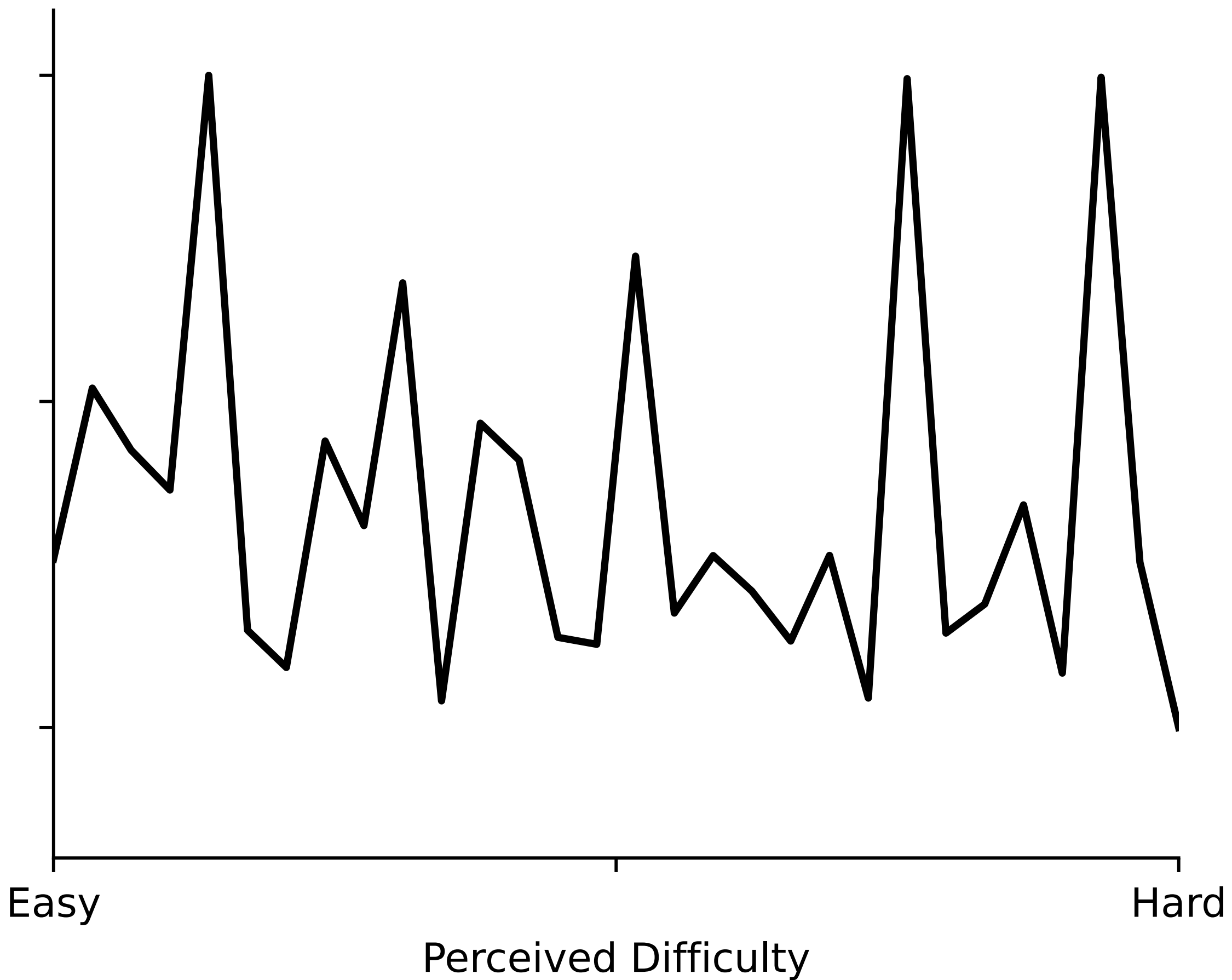
LLM's aren't bad, they're jagged

LLMs aren't bad,

Smooth Capability Curve



Jagged Capability Curve



Coding



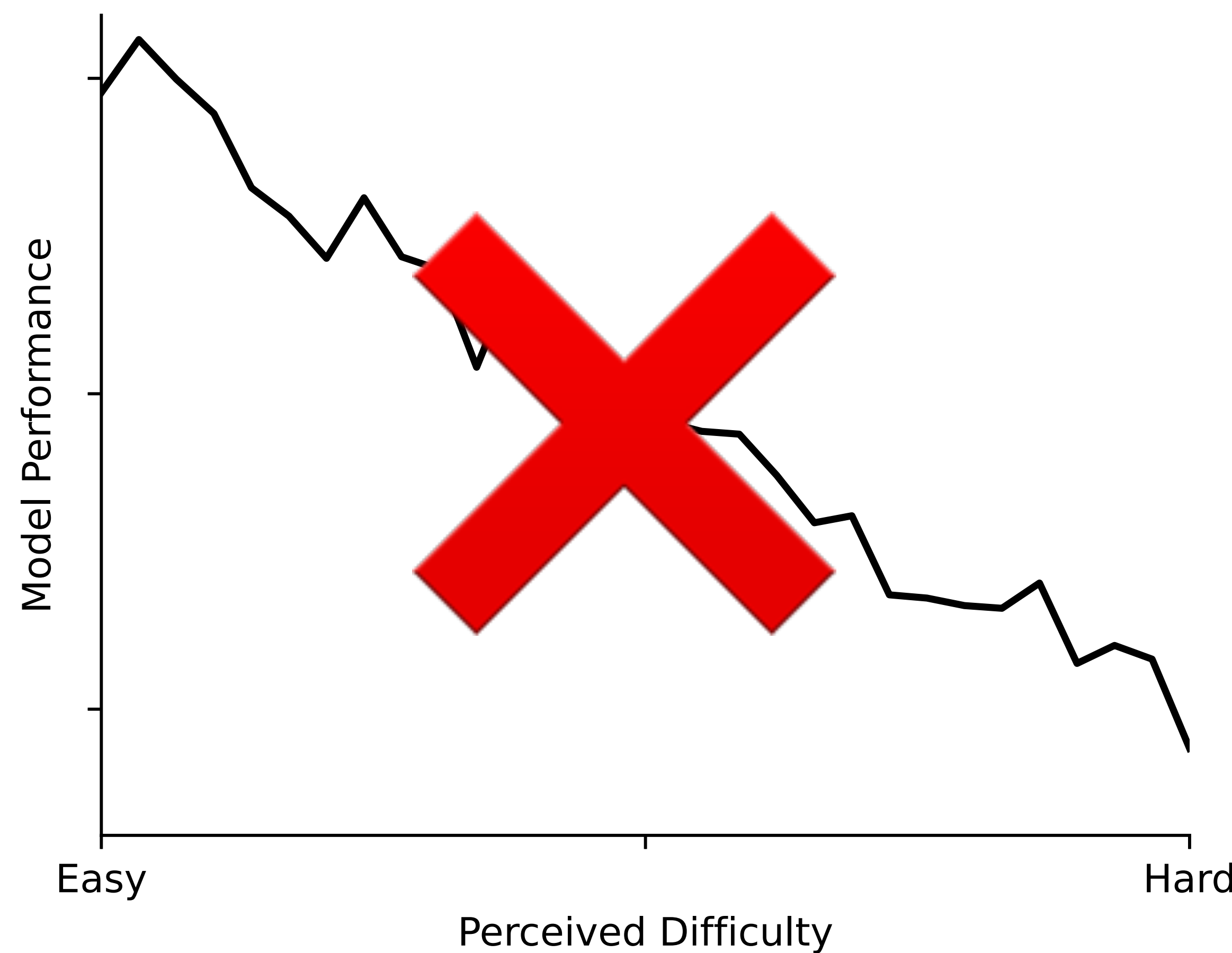


The simplest data tasks

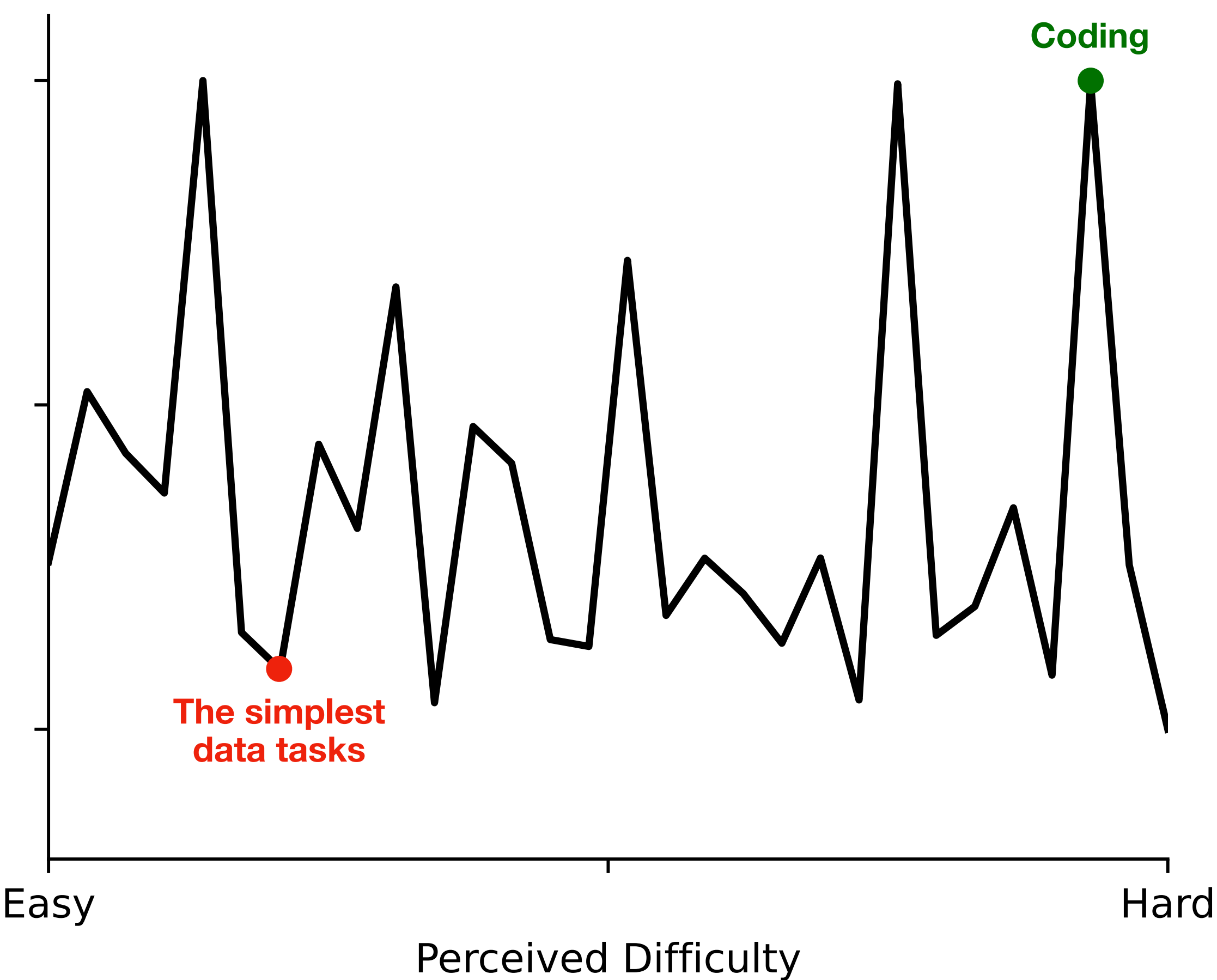


LLMs aren't bad, they're jagged

Smooth Capability Curve



Jagged Capability Curve



How bad are LLMs with data?

Example: len()