A no hype, no equation intro to the Deekseek Reasoning Language Model

Simen Eide, AI lab

1
2
Den kinesiske AI-
satsingen DeepSeek
rystet verdens
teknologiaksjer.

Why is Deepspeed a big thing?

- Comparable to openai o1 model
- Cost: $6m vs $100m training budget, 30x cheaper to run
- Giving the model and ideas away for free (Open vs closed)
- Chinese vs the West

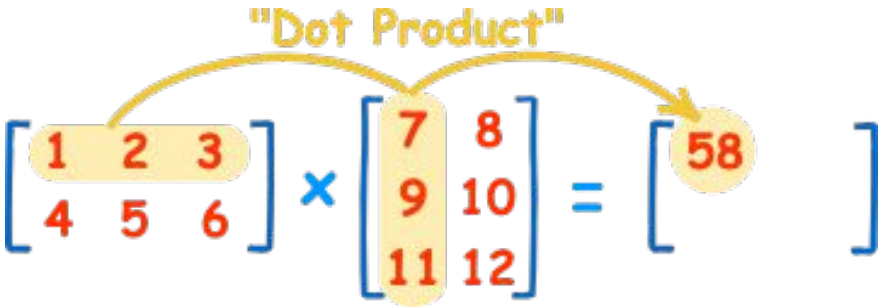
Agenda

- Product vs Model
- What are language models *really*?
- What are “reasoning language models”?
- How does language models learn?
- Discussion

Not: Stock market reaction, geopolitical effects, …

AI model vs AI product

AI model vs AI product

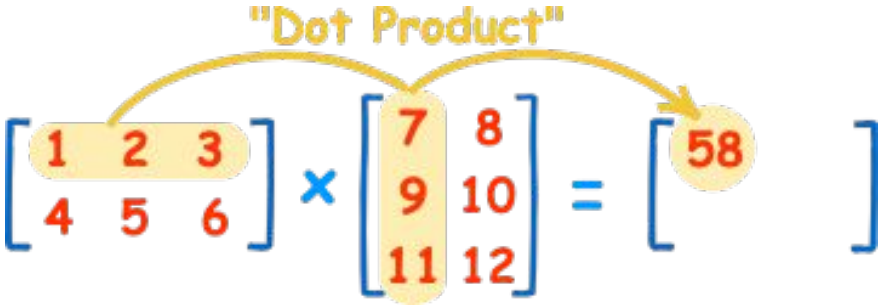


AI model:

- *Many numbers and a recipe on how to multiple these together*

- Only gives probability of next word, trained from data
- AI models can be inaccurate, hallucinate, produce harmful text

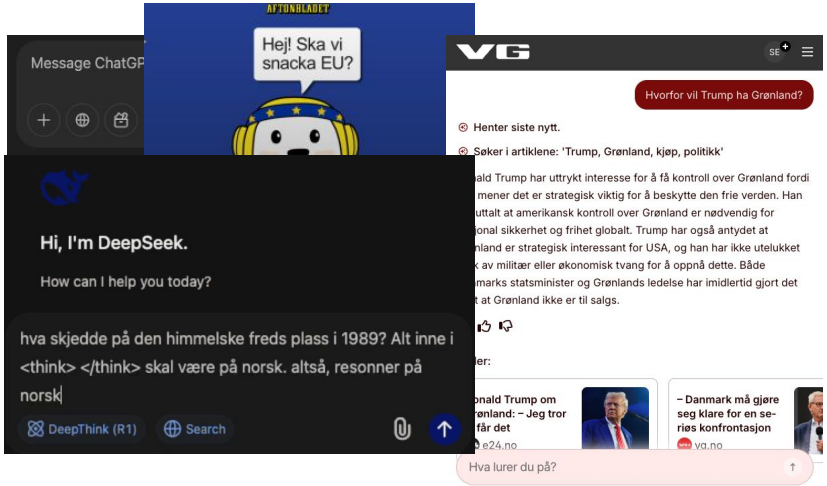
AI model vs AI product



AI model:

- *Many numbers and a recipe on how to multiple these together*

- Only gives probability of next word, trained from data
- AI models can be inaccurate, hallucinate, produce harmful text

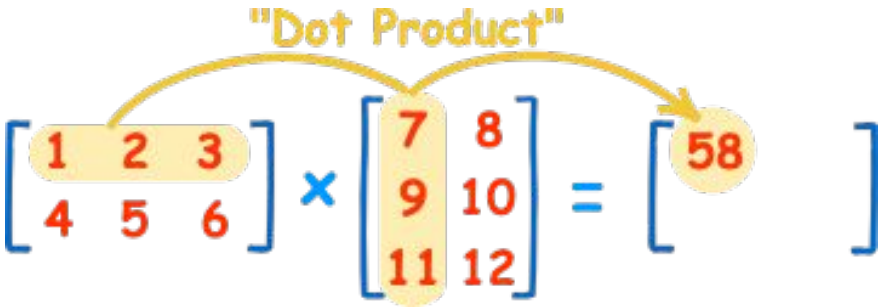


AI product:

- *Software that use an AI model*

- Contains databases, logging, web search, …
- *AI products handles your data*, and they can “steal it”

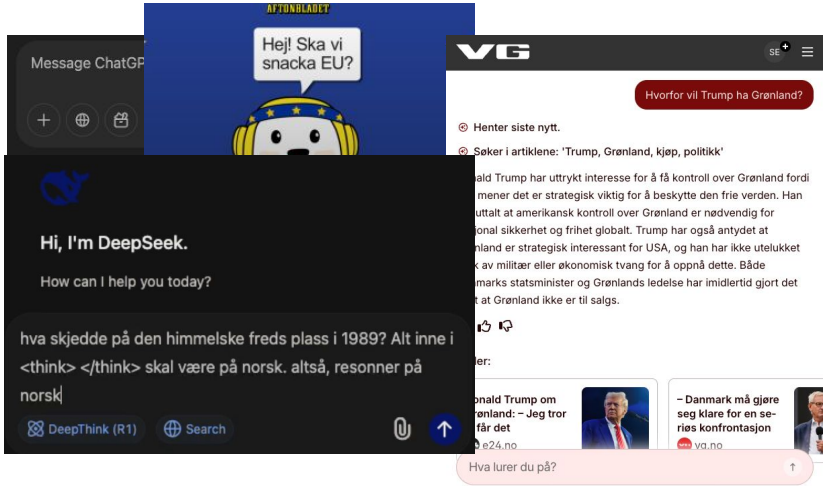
AI model vs AI product



AI model:

- *Many numbers and a recipe on how to multiple these together*

- Only gives probability of next word, trained from data
- AI models can be inaccurate, hallucinate, produce harmful text



AI product:

- *Software that use an AI model*

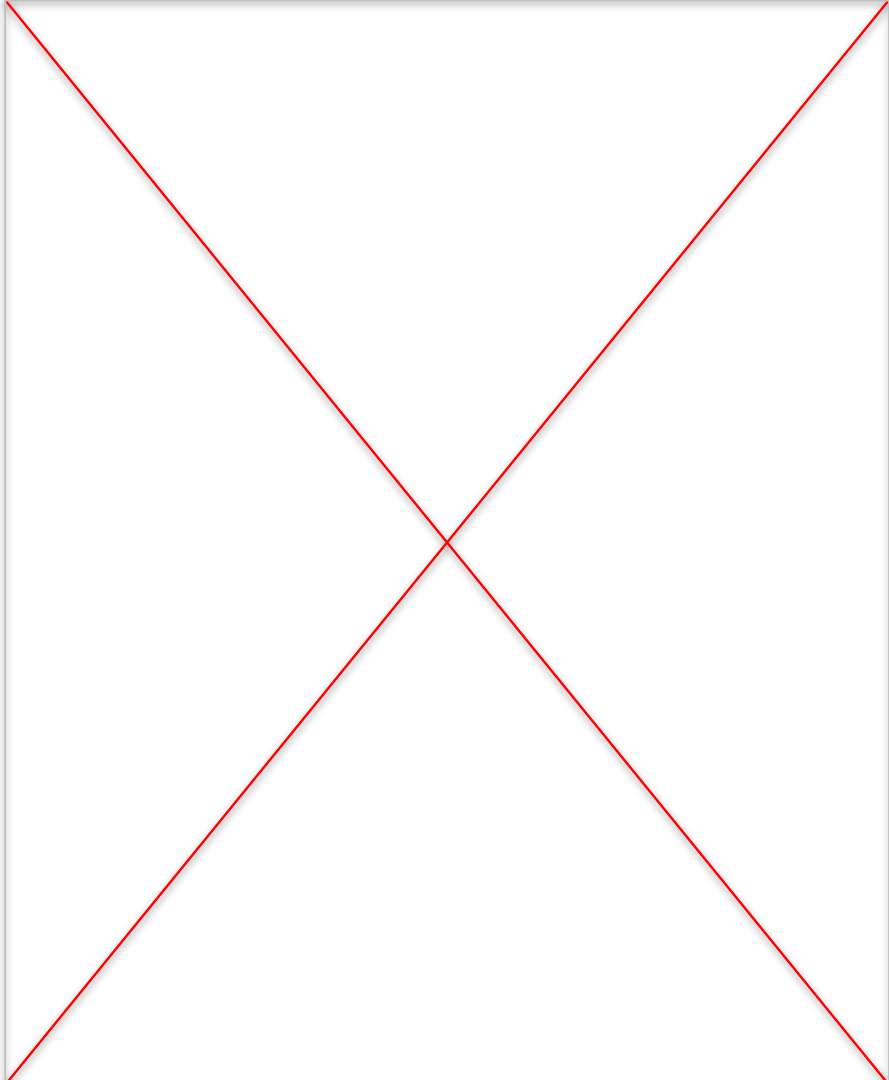
- Contains databases, logging, web search, …
- *AI products handles your data*, and they can “steal it”

Humans interact with AI products

Two layer of censorship in Deepseek

1. In the AI Product
   a. Deepseek can do “anything”
   b. Seems to be some keyword matching and classifier AI models

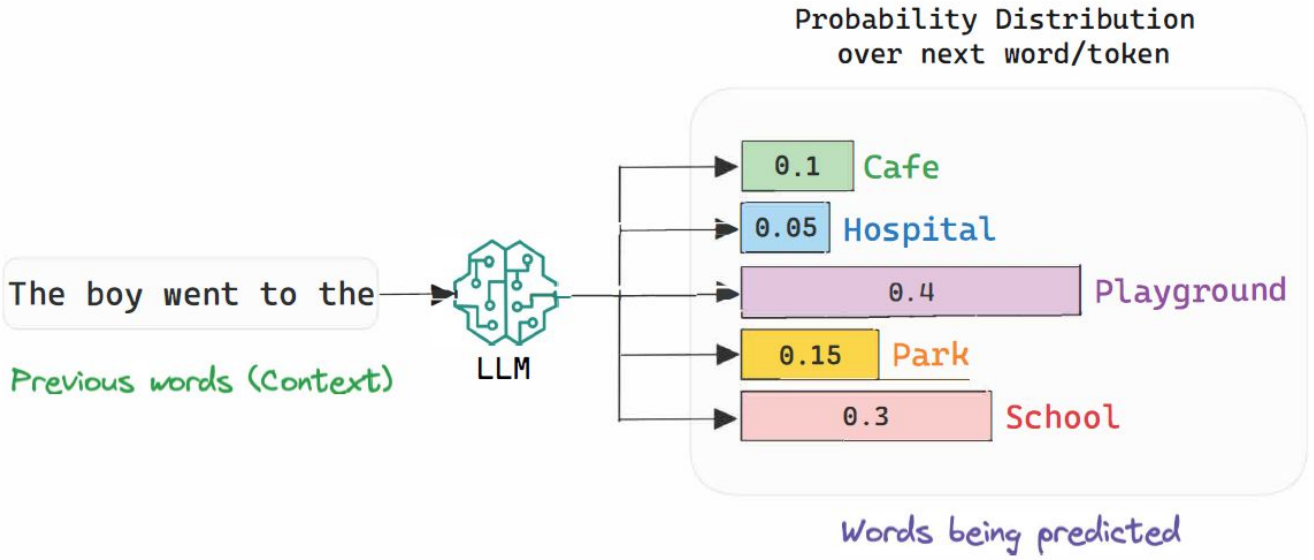
2. In the AI model
   a. This learnt by data.
   b. “Helpfulness and harmlessness”
   c. Same as openAI does, but with different glasses



What are language models?

- Language models ONLY predict the next word in a text

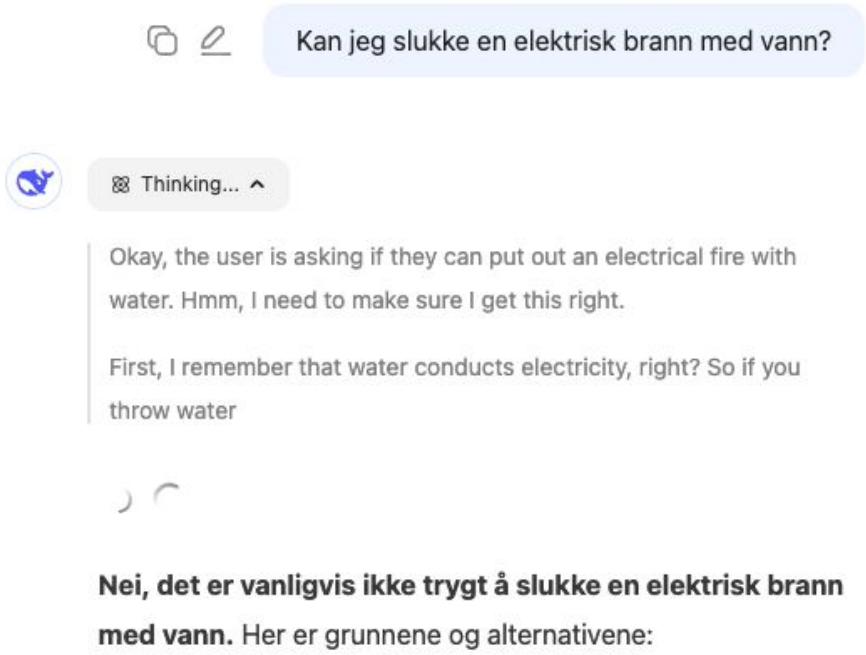
- Learns through example: Text data



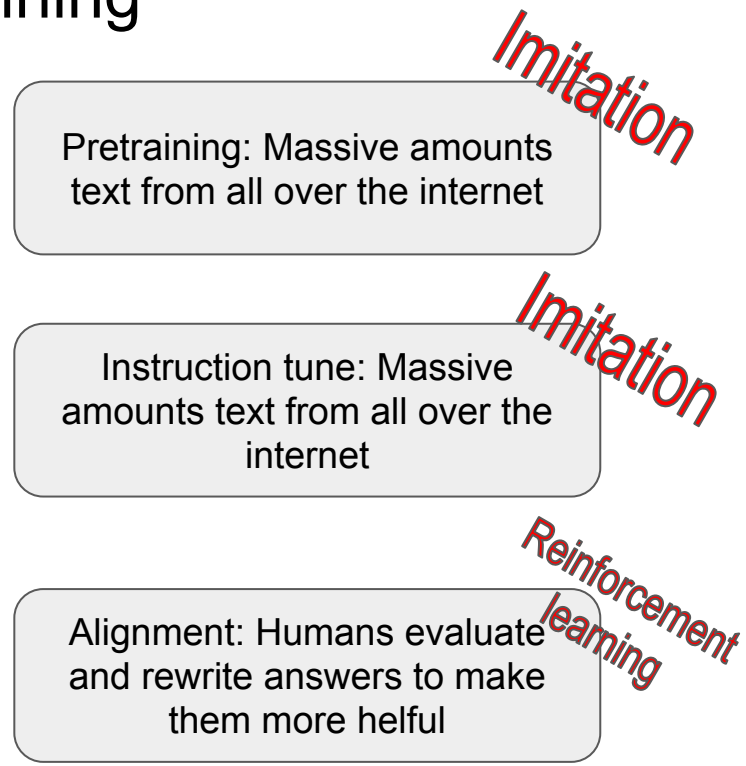
What is a “reasoning LLM”?

“According to deepseek”

- The model is allowed to “reason” by generating more words before giving the answer
- Only train the answer, not the reasoning
- Learn with reinforcement learning



Traditional Language model training



Reasoning Language model training

Traditional Language model training

Pretraining: Massive amounts text from all over the internet

Imitation

Instruction tune: Massive amounts text from all over the internet

Imitation

Alignment: Humans evaluate and rewrite answers to make them more helful

Reinforcement learning

Reasoning Language model training

Pretraining: Massive amounts text from all over the internet

Imitation

Reason training: Work on hard problems where we can evaluate the correct answer

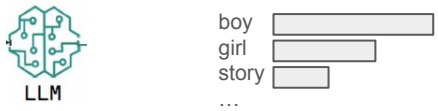
Reinforcement learning

Imitation (supervised learning)

The

Imitation (supervised learning)

The **boy**

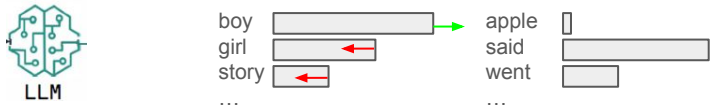


Imitation (supervised learning)

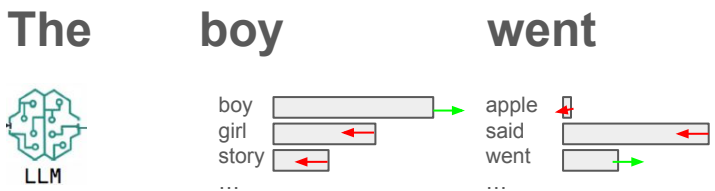
The **boy**

Imitation (supervised learning)

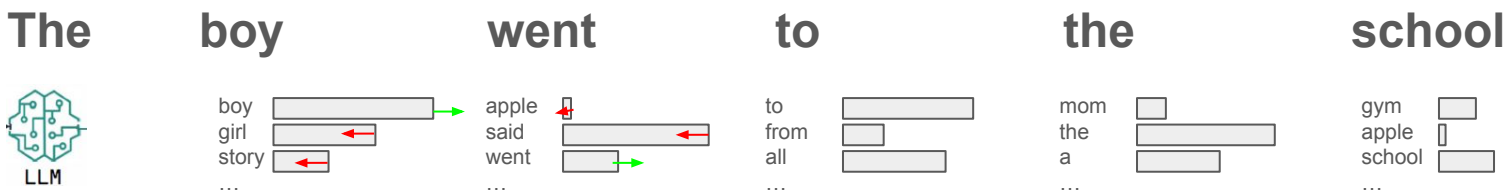
The **boy**



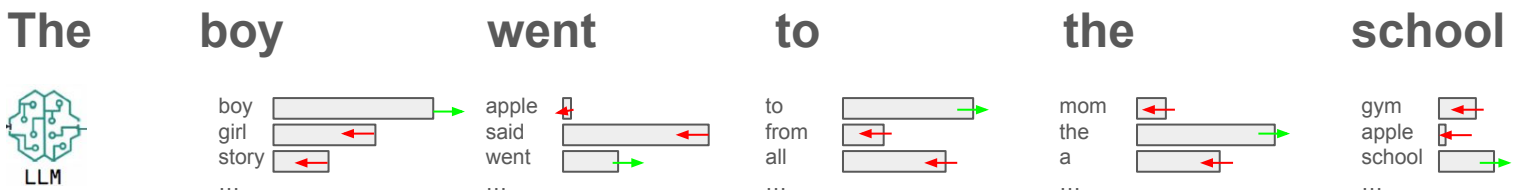
Imitation (supervised learning)



Imitation (supervised learning)

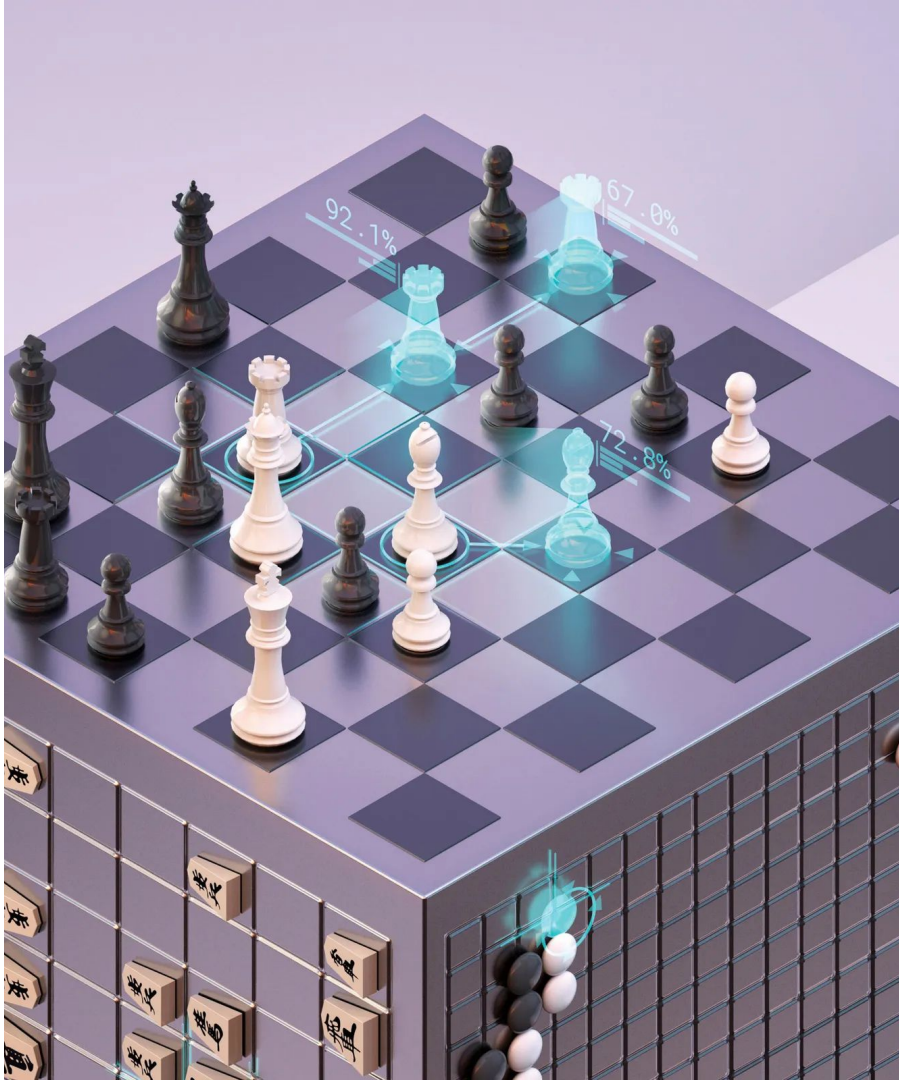


Imitation (supervised learning)



Reinforcement learning

- Alpha-zero:
  - beat humans in go, chess and shogi
- Only learn from win or lose
- If win,make all chess moves before more likely
- “Each chess move is a reasoning word”



Reinforcement learning

The boy went to the
<think> I need to figure out where boys usually go. The bar? No that is for older people, but maybe a school? Yes, that sounds right </think>
school.

Reinforcement learning

The boy went to the

<think> I need to figure out where boys usually go. The bar? No that is for older people, but maybe a school? Yes, that sounds right </think>

school.

Reinforcement learning

The boy went to the
<think> I need to figure out where boys usually go. The bar? Yes, that sounds good. </think>
bar.

Reinforcement learning

The boy went to the

<think> I need to figure out where boys usually go. The bar? Yes, that sounds good. </think>

bar.

Conclusions

- Deepseek is an impressive and efficient model
- Reinforcement learning a much more important part of training
- They closed the head start openAI had in language models
- They published their new research and model weights
- Hard to know exactly why this was so good, replication efforts ongoing