



støttet af

It-vest

samarbejdende universiteter

INTRODUKTION TIL MATEMATIK OG OPTIMERING (IMO)

— nu også med egen chatbot

Niels Lauritzen, Institut for matematik, AU

21 januar, 2025

Deep learning/Gen AI revolutionen meget kort

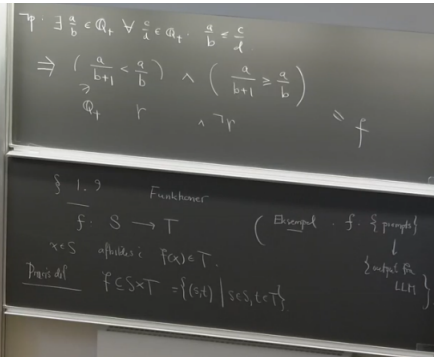
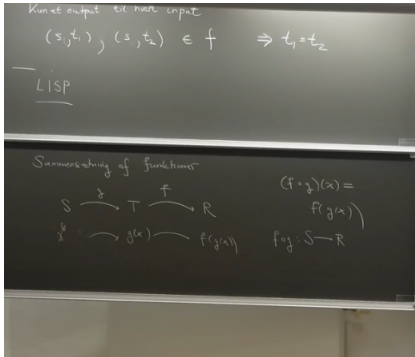
- ◇ 2012: **AlexNet** (Krizhevsky, Hinton og Sutskever), image recognition via (C)NN (og GPU)
- ◇ 2013: Word embeddings, Google, **word2vec**
- ◇ 2017: Transformer model (*Attention is all you need*)
- ◇ 2018: **GPT-1** (OpenAI)
- ◇ 2022: ChatGPT (*The Age of AI has begun*, Gates Notes)

Deep learning royalty



*The thing I find most surprising is that it actually works
(Sutskever, 2023, [YouTube](#))*

Lavpraktisk Gen AI i IMO



<https://aistudio.google.com/prompts/1jJ4XU1aQfQugcDQpw2UavcGhxAT4HV98>

Where was this picture taken?



I can't identify the exact location of the photo, but based on the background, it appears to be a waterfront area with a bridge and urban skyline in the distance. The architecture suggests it might be a modern city. Could this be in a place like Seoul, South Korea, near the Han River?

Tokens!

'Introduktion til matematik og optimering' er et introducerende matematikkursus (fra 2019) på første semester af datalogi og datavidenskab på Aarhus Universitet. Kurset bygger bro fra matematik A på gymnasiet gennem et første kapitel om matematikkens sprog (og fra efteråret 2024 med elementer af prompting) og behandler matricer, lineær optimering, euklidiske vektorrum, gradienter etc. med overvejende fokus på machine learning

<https://gptforwork.com/tools/tokenizer>

Den næste token

To date, the cleverest thinker of all time was _____



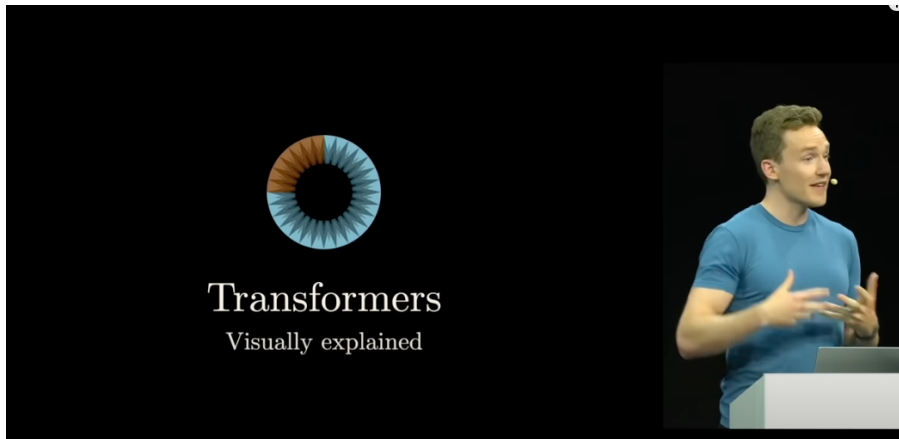
the	8%
probably	4%
John	4%
Sir	3%
Albert	3%
Ber	3%
a	3%
Isaac	2%
undoubtedly	1%
arguably	1%
⋮	

$$P(t_{n+1} | t_1, t_2, \dots, t_n)$$

Attention!

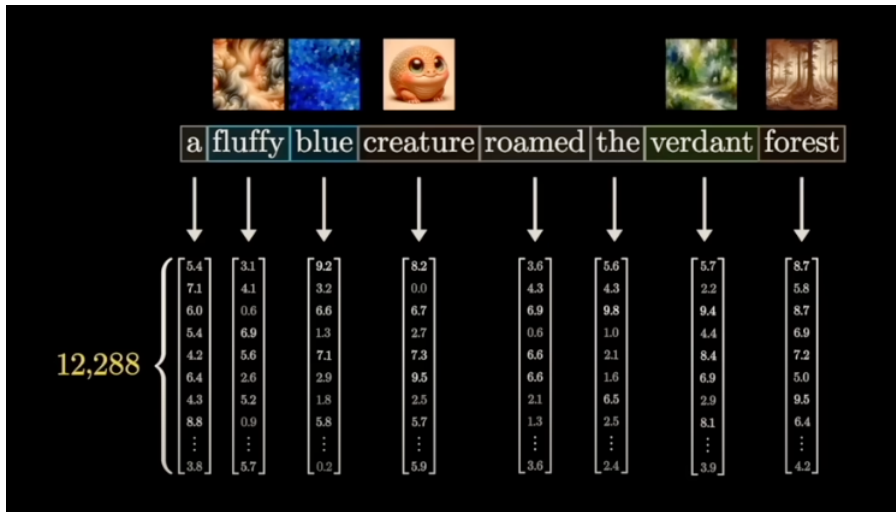


Formidling i verdensklasse



<https://youtu.be/KJtZARu03JY>

Det er altså sammen Matematik A*



Embedding layer GPT3

Matematik i deep learning

- ◇ Plus og gange med mange tal (matrixregning)
- ◇ Sammensætning af funktioner
- ◇ $\text{ReLU}(x) = \max(0, x)$ (rectified linear unit)
- ◇ Eksponentialfunktionen e^x (softmax)
- ◇ Differentialregning, udregning af f' , gradient descent

IMO kurset kort



Mantra

For de fleste studerende er det nemmere at "tale" med en lærebog end læse i den. Optimalt er det hvis lærebogen også kan være en venlig tutor.

nielsGPT – velforberedt og venlig

Hey Niels. Jeg har funktionen $f(x,y,z,t) = (1 - x - y - 2z)^2 + (1 - x - 2y - 3z)^2 + (t - y - 2z)^2$: Jeg skal finde de globale minima for f og så se om f er streng konveks. Jeg ved nu, at der for hvert t er uendelig mange løsninger og at $f(x, y, z, t) \geq 0$ for alle x, y, z, t . Men hvorfor dette? Kan du forklare mig hvorfor der er uendelig mange løsninger? Og kan du forklare hvorfor $f(x,y,z,t) \geq 0$ for alle x,y,z,t ?



Hej! Det er godt, at du allerede har gjort nogle observationer om funktionen $f(x, y, z, t)$. Lad os se nærmere på dine spørgsmål.

1. Hvorfor er der uendelig mange løsninger?

Funktionen $f(x, y, z, t)$ er en sum af kvadrater. Hver del af summen, som for eksempel $(1 - x - y - 2z)^2$, er lig med 0, når udtrykket inden for parenteser er lig med 0. Det betyder, at for at finde de globale minima, skal vi løse ligningerne:

En kursuschatbot på Azure

IMO Assistant

Hi! How can I help you?



Summarize the main concepts from Chapter 3.



Help me understand definition 5.1.



Give me a hint for exercise 4.36.



What is induction?

Microsoft Azure

Search resources, services, and docs (G+/)

Copilot

au82967@uni.au.dk
AARHUS UNIVERSITY | AARHUS...

Azure services

- Create a resource
- Power BI Embedded
- Key vaults
- Tenant properties
- Cost analysis
- Container registries
- Azure AI Studio
- Azure Machine Learning
- Azure OpenAI
- More services

Samarbejde med MGMT/BSS



- ◇ Carsten Bergenholtz og Oana Vuculescu
- ◇ Prototype (Phil) for Videnskabsteori, F24

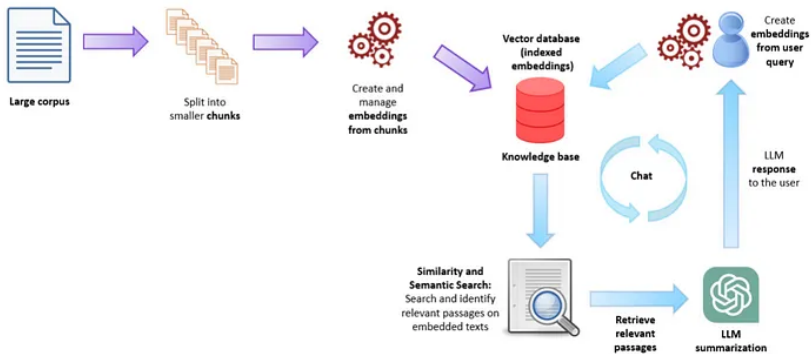
Støttet af IT-vest og USM, AU.

Datalogistuderende på holdet



Felix, Kasper, Timotej*, Daniel og Oliver (IMO alumner 23, 22)

RAG* baseret chatbot



ud fra mine noter <https://edtech.dk/IM024>.

Teknologistak

- ◇ Python 3.12 med fastapi og uvicorn
- ◇ Web UI bygget af CS group fra grunden i next.js
- ◇ AI API hosted af Azure (servere i EU)
- ◇ GPT-4o mini
 - ◇ USD 0.15/1M (input tokens)
 - ◇ USD 0.6/1M (output tokens)
- ◇ text-embedding-3-large (dim 3072)
- ◇ Azure Cosmos DB (vektordatabase)

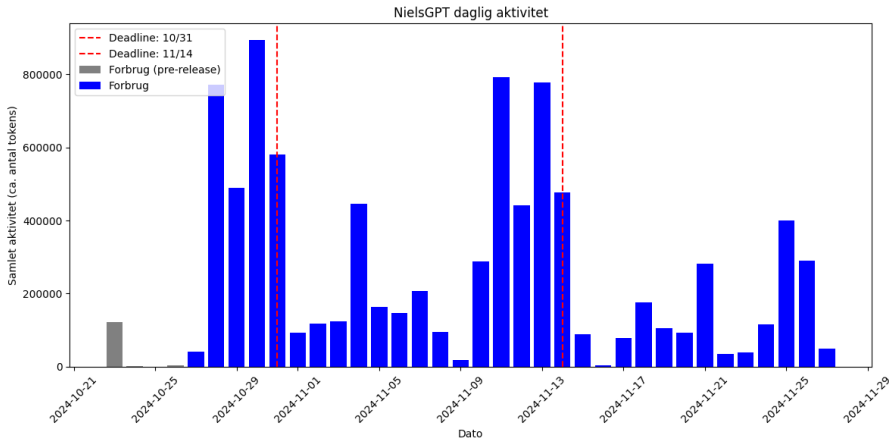
IMO chatbot statistik

En måned (november) med kun to obligatoriske opgaver:

- ◇ 928 studenter chats
- ◇ 376 chats omkring vink til opgaver
- ◇ 3614 studenter prompts (omkring 4 per chat)
- ◇ 258 sekunder (lige over 4 minutter) median chat attention span

Jeg blev positivt overrasket over median chat attention span på over 4 min

Activity timeline



Chat eksempel mm

Understanding the Hessian

System prompt og live demo (fingers crossed!).

Evaluering af kursuschatbot

<https://edtech.dk/IMO24boteval.pdf>



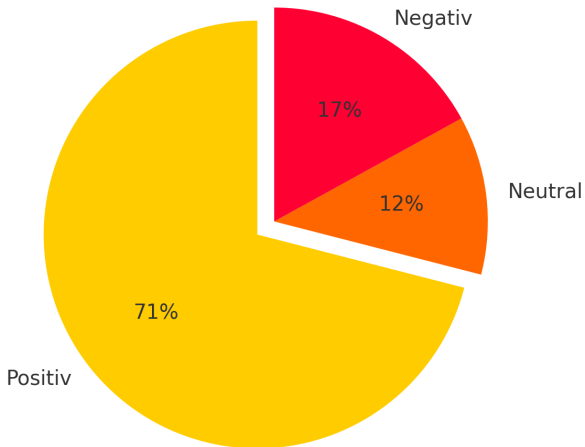
Hugging Face

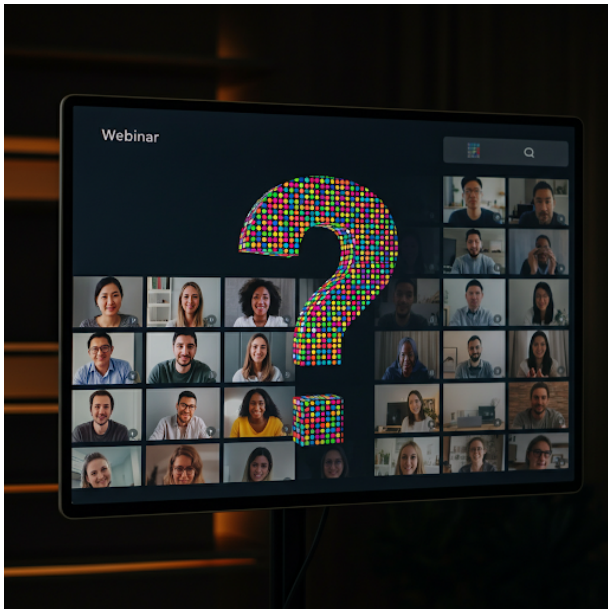
Search models, datasets, users...

cardiffnlp/twitter-xlm-roberta-base-sentiment

$n = 78$

arXiv





Courtesy of **Google imagegen 3**