Kunstig intelligens i hjerneforskning

Esten H. Leonardsen

19.10.23

Plan for dagen

1. Historien bak kunstig intelligens (AI)
2. Terminologi og konsepter
3. Teori
   - Hva er en maskinlæringsmodell?
   - Hvordan trene en maskinlæringsmodell?
   - Intuisjon bak bildegjenkjenningsmodeller
   - Intuisjon bak språkmodeller
4. AI i hjerneforskning
   - Segmentering
   - Diagnostisk klassifikasjon
   - Hjernealder
   - Forklarbar kunstig intelligens
   - Genererte hjernebilder
5. Forskningsetiske spørsmål ved bruk av AI
6. Generelle etiske spørsmål ved bruk av AI

Historien bak kunstig intelligens

Historien bak kunstig intelligens

Turing-
testen
(1950)

Alan Turing

Historien bak kunstig intelligens

Turing-testen (1950)

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. *The Imitation Game.*

I propose to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

Computing Machinery and Intelligence, A. M. Turing, *Mind*, 1950.

Historien bak kunstig intelligens

Historien bak kunstig intelligens

Historien bak kunstig intelligens

Historien bak kunstig intelligens

ELIZA: a very basic Rogerian psychotherapist chatbot

Historien bak kunstig intelligens

MYCIN: a knowledge-based consultation program for infectious disease diagnosis, William van Melle, *International Journal of Man-Machine Studies*, 1978

Historien bak kunstig intelligens

```
1) Patient's name: (first-last)
** FRED BRAUN
2) Sex:
** M                                                    {background patient data}
3) Age:
** 55
4) Are there any cultures for Fred Braun which may be related to the present illness, and
   from which organisms have been grown successfully in the microbiology laboratory?
** y

   --------CULTURE-1--------
   5) From what site was the specimen for CULTURE-1 taken?
   ** BLOD
=BLOOD
   6) Please give the date and time when CULTURE-1 was obtained. (mo/da/yr time)
   ** JUN 20, 1977
   The first organism isolated from the blood culture of 20-JUN-77 (CULTURE-1) will be
   referred to as:

      --------ORGANISM-1--------
      7) Enter the laboratory-reported identity of ORGANISM-1:
      ** UNKNOWN                                         {typically identity is not yet known}
      8) The stain (Gram or Ziehl-Neelsen acid-fast) of ORGANISM-1:
      ** NEG                                             {preliminary lab results give some clues}
      9) Is ORGANISM-1 a rod or coccus (etc.):
      ** ROD
      10) What is the form of the individual organisms (e.g. Lancet-shaped
          for cocci, fusiform for rods, etc.)?
      ** FUSIFORM

     {...more questions follow in order to gather sufficient information
     to infer the identity and significance of the infecting organisms...}
```

MYCIN: a knowledge-based consultation program for infectious disease diagnosis, William van Melle, *International Journal of Man-Machine Studies*, 1978

Historien bak kunstig intelligens

Learning representations by back-propagating errors, Rumelhart, D., Hinton, G. & Williams, R, *Nature*, 1986

Historien bak kunstig intelligens

Learning representations by back-propagating errors, Rumelhart, D., Hinton, G. & Williams, R, *Nature*, 1986

Historien bak kunstig intelligens

Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams*

* Institute for Cognitive Science, C-015, University of California, San Diego, La Jolla, California 92093, USA
† Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Philadelphia 15213, USA

We describe a new learning procedure, back-propagation, for networks of neurone-like units. The procedure repeatedly adjusts the weights of the connections in the network so as to minimize a measure of the difference between the actual output vector of the net and the desired output vector. As a result of the weight adjustments, internal 'hidden' units which are not part of the input or output come to represent important features of the task domain, and the regularities in the task are captured by the interactions of these units. The ability to create useful new features distinguishes back-propagation from earlier, simpler methods such as the perceptron-convergence procedure[1].

Learning representations by back-propagating errors, Rumelhart, D., Hinton, G. & Williams, R, *Nature*, 1986

Historien bak kunstig intelligens

Learning representations by back-propagating errors, Rumelhart, D., Hinton, G. & Williams, R, *Nature*, 1986

Historien bak kunstig intelligens

DALL-E: "A robot playing chess"

- IBMs Deep Blue slår Garry Kasparov og blir den første datamaskinen til å slå den sittende verdensmesteren i sjakk.
- Deep blue vant med 3½ poeng mot Kasparovs 2½ over 6 spill.
- Kasparov uttalte at "Deep Blue was intelligent the way your programmable alarm clock is intelligent."
- Kombinasjon av maskinlæring og hardkodet kunnskap fra sjakkeksperter.

Historien bak kunstig intelligens

Katt

Historien bak kunstig intelligens

Solsikke | Marihøne | Katt | Fly | Hvithai

ImageNet: ∼14m bilder, ∼22k kategorier

Historien bak kunstig intelligens

Historien bak kunstig intelligens

Historien bak kunstig intelligens

Historien bak kunstig intelligens

Historien bak kunstig intelligens

E

Kan du skrive et kjapt sammendrag av historien om kunstig intelligens til elever i andre klasse på vgs?

Selvfølgelig! Her er et enkelt og kort sammendrag av historien om kunstig intelligens (KI) egnet for elever i andre klasse på videregående skole:

Historien om Kunstig Intelligens (KI)

1. **Begynnelsen (1940-1950-tallet):** Alt startet da noen smarte folk tenkte: "Kan vi få maskiner til å tenke som mennesker?" De utviklet grunnleggende teorier for hvordan datamaskiner kunne løse problemer ved hjelp av matematikk.

Terminologi og konsepter

Terminologi og konsepter: Begreper

Kunstig intelligens:
Maskiner som løser oppgaver som krever intelligens

Terminologi og konsepter: Begreper

Kunstig intelligens:
Maskiner som løser oppgaver som krever intelligens

Terminologi og konsepter: Begreper

Kunstig intelligens:
Maskiner som løser oppgaver som krever intelligens

Maskinlæring:
Maskiner som løser problemer gjennom å lære seg å finne mønster i data

Terminologi og konsepter: Begreper

Kunstig intelligens:
Maskiner som løser oppgaver som krever intelligens

Maskinlæring:
Maskiner som løser problemer gjennom å lære seg å finne mønster i data

Terminologi og konsepter: Begreper

Kunstig intelligens:
Maskiner som løser oppgaver som krever intelligens

Maskinlæring:
Maskiner som løser problemer gjennom å lære seg å finne mønster i data

Dyplæring:
Maskinlæringsmodeller som er organisert i hierarkier ($\approx$ dype nevrale nettverk) inspirert av hjernen

Terminologi og konsepter: Begreper

Kunstig intelligens:
Maskiner som løser oppgaver som krever intelligens

Maskinlæring:
Maskiner som løser problemer gjennom å lære seg å finne mønster i data

Dyplæring:
Maskinlæringsmodeller som er organisert i hierarkier ($\approx$ dype nevrale nettverk) inspirert av hjernen

Konvolusjonelle nevrale nettverk:
Nevrale nettverk for bildedata

Store språkmodeller:
Nevrale nettverk for språkdata (ChatGPT)

Terminologi og konsepter: Veiledning

Veiledet læring
(supervised learning)

→ Katt

→ Hund

→ Katt

→ Hund

Ikke-veiledet læring
(unsupervised learning)

Terminology: Smal og generell AI

Smal (svak) Generell (sterk)

I stand til å løse et bredere spekter med problemer innenfor flere domener

Mer spesifikk Mer generell

Terminology: Strong and weak AI

Smal (svak)

Generell (sterk)

I stand til å løse et bredere spekter med problemer innenfor flere domener

Mer spesifikk

Mer generell

Terminology: Strong and weak AI

Smal (svak)

Generell (sterk)

I stand til å løse et bredere spekter med problemer innenfor flere domener

Bilde-diagnostikk

Forsikrings-prising

Dokument-lesing

Mer spesifikk

Mer generell

Terminology: Strong and weak AI

Smal (svak) — Generell (sterk)

I stand til å løse et bredere spekter med problemer innenfor flere domener

Bilde-diagnostikk

Forsikrings-prising

Dokument-lesing

Tesla

ChatGPT

Mer spesifikk — Mer generell

Teori

Teori: Maskinlæringsmodeller

Teori: Maskinlæringsmodeller

Teori: Maskinlæringsmodeller

Teori: Maskinlæringsmodeller

m^2	Pris
72	5.127.379
50	4.552.170
45	4.486.654
62	5.709.276
53	4.634.912
81	8.388.570
44	4.828.170
78	7.557.770
37	4.016.520
73	6.572.351

Teori: Maskinlæringsmodeller

Teori: Maskinlæringsmodeller

$$\hat{y} = f(57)$$

Teori: Maskinlæringsmodeller

$$\hat{y} = 30000x + 3500000$$

Teori: Maskinlæringsmodeller

$\hat{y} = 0x + 6000000$ $\hat{y} = 30000x + 3500000$ $\hat{y} = 82031x + 706495$

Teori: Kostfunksjon

$\hat{y} = 30000x + 3500000$

Teori: Kostfunksjon

$$\hat{y} = 30000x + 3500000$$

Teori: Kostfunksjon

$\hat{y} = 30000x + 3500000$

Teori: Kostfunksjon

$$\hat{y} = 30000x + 3500000$$
$$\ell = \sum (y - \hat{y})^2$$

Teori: Kostfunksjon

$$\hat{y} = 30000x + 3500000$$
$$\ell = 1.10 \times 10^{13}$$

Teori: Kostfunksjon

$\hat{y} = 0x + 6000000$
$\ell = 2.08 \times 10^{13}$

$\hat{y} = 30000x + 3500000$
$\ell = 1.10 \times 10^{13}$

$\hat{y} = 82031x + 706495$
$\ell = 4.09 \times 10^{12}$

Teori: Trening

$$\hat{y} = wx + b$$
$$\ell = \sum (y - \hat{y})^2$$

Teori: Trening

$$\boxed{\hat{y}} = wx + b$$
$$\ell = \sum (y - \boxed{\hat{y}})^2$$

Teori: Trening

$$\ell = \sum (y - (wx + b))^2$$

Teori: Trening

$$1.10 \times 10^{13} = \sum (y - (30000x + 3500000))^2$$

Teori: Trening

$$7.24 \times 10^{12} = \sum (y - (45000x + 3000000))^2$$

Teori: Nevrale nettverk

$$\hat{y} = wx + b$$

Teori: Nevrale nettverk

$$
\begin{aligned}
\hat{y} = max(0, & w^2_{0,0} * max(0, w^1_{0,0} * max(0,\ w^0_{0,0} * x_0 + w^0_{1,0} * x_1 + b_{0,0}) + \\
& w^1_{1,0} * max(0,\ w^0_{0,1} * x_0 + w^0_{1,1} * x_1 + b_{0,1}) + \\
& w^1_{2,0} * max(0,\ w^0_{0,2} * x_0 + w^0_{1,2} * x_1 + b_{0,2}) + \\
& b_{1,0}) + \\
& w^2_{1,0} * max(0, w^1_{0,1} * max(0,\ w^0_{0,0} * x_0 + w^0_{1,0} * x_1 + b_{0,0}) + \\
& w^1_{1,1} * max(0,\ w^0_{0,1} * x_0 + w^0_{1,1} * x_1 + b_{0,1}) + \\
& w^1_{2,1} * max(0,\ w^0_{0,2} * x_0 + w^0_{1,2} * x_1 + b_{0,2}) + \\
& b_{1,1}) + \\
& w^2_{2,0} * max(0, w^1_{0,2} * max(0,\ w^0_{0,0} * x_0 + w^0_{1,0} * x_1 + b_{0,0}) + \\
& w^1_{1,2} * max(0,\ w^0_{0,1} * x_0 + w^0_{1,1} * x_1 + b_{0,1}) + \\
& w^1_{2,2} * max(0,\ w^0_{0,2} * x_0 + w^0_{1,2} * x_1 + b_{0,2}) + \\
& b_{1,2}) + \\
& b_2)
\end{aligned}
$$

Teori: Konvolusjonelle nevrale nettverk

Teori: Konvolusjonelle nevrale nettverk

AI i hjerneforskning

AI i hjerneforskning: Oversikt

Artikler om dyplæring i medisin per år

AI i hjerneforskning: Segmentering

Segmentering av hjernesvulster (glioblastoma) basert på MR-bilder

AI i hjerneforskning: Hjernealder