AI MAT

MATEMATIKKEN BAG MAGIEN

Ege Rubak / Institut for matematiske fag, AAU
Malene Cramer Engebjerg / Aalborg Katedralskole

Velkommen og lidt om vores projekt

- Projektet Aalborg Intelligence, som er finansieret af Novo Nordisk Fonden, er forankret på Institut for Matematiske Fag på Aalborg Universitet (AAU), og inkluderer en repræsentant fra de fem STX-gymnasier i Aalborg.
- Projektperiode: 2020-2026

Hvad er AI-matematik?

- **Det er al den spændende matematik, som ligger bag kunstig intelligens!**
- **Noter, videoer og undervisningsforløb, som viser hvordan gymnasiematematikken bruges i en masse former for kunstig intelligens.**
- **Det handler til gengæld *ikke* om at lære, hvordan man får en kunstig intelligens til at løse sine matematikopgaver!**

Forskere fra AAU	**Gymnasielærere**
Lisbeth Fajstrup	Allan Frendrup (Nørresundby Gymnasium)
Anne Marie Svane	Nikolaj Hess-Nielsen (Aalborg Katedralskole)
Søren Højsgaard	Mette Kristensen (Hasseris Gymnasium)
Ege Rubak (projektleder)	Jan B. Sørensen (Aalborg City Gymnasium)
	Malene Cramer Engebjerg (tidl. Aalborghus Gymnasium)

Velkommen!

AI MAT
MATEMATIKKEN BAG MAGIEN

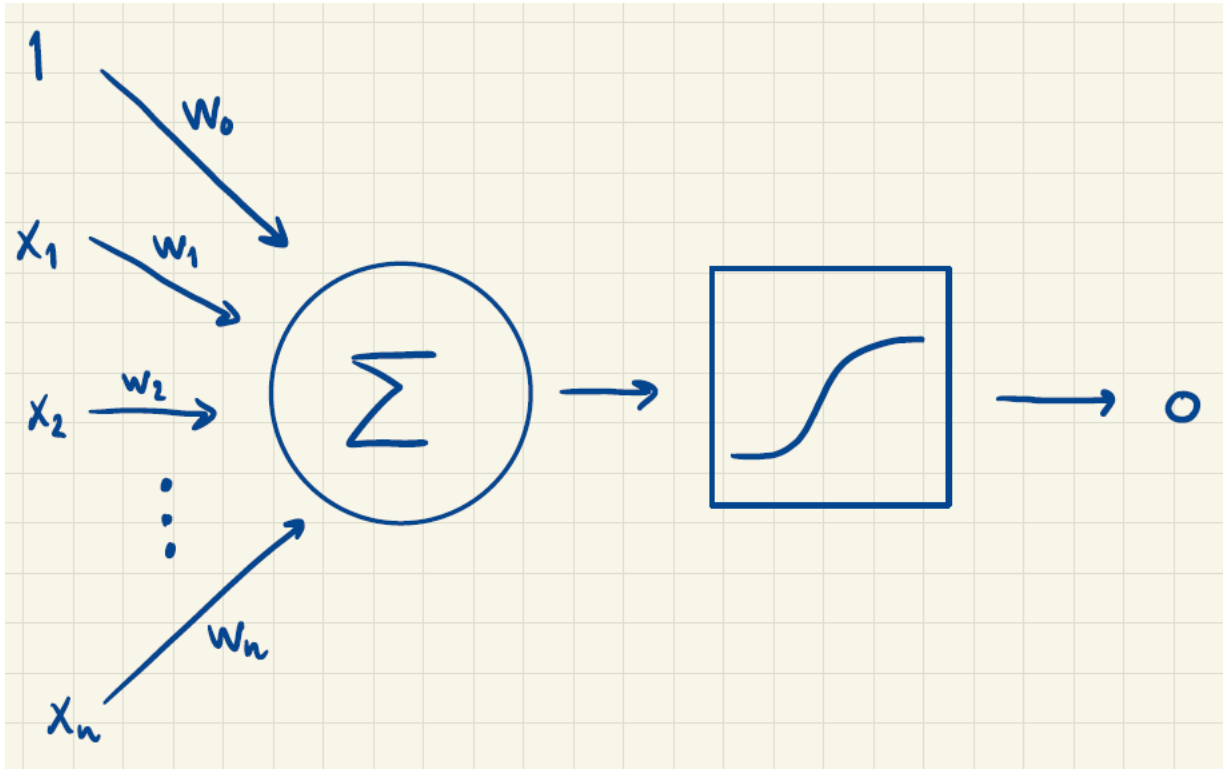
- Præsentation af hjemmesiden: AI MAT - matematikken bag magien
- Program for kurset: Program – AI MAT - matematikken bag magien

Program for denne del

- Kunstige neuroner – herunder
  - Træningsdata
  - Aktiveringsfunktioner
  - Opdateringsregler
  - Kunstig neuron app

Kunstige neuroner

- De byggesten, som et kunstigt neuralt netværk består af



Eksempel: Medlemsapp til Good Food

- Medlemsapp til kunderne i Good Food, som skal vise et begrænset udvalg af tilbud, som kunden skal aktivere for at få rabat.

Idé: En midaldrende mand med et stort forbrug, skal have vist et andet tilbud end en teenagepige, som kun sjældent handler i Good Food.



Good Food indsamler data

- x_1: kundens alder målt i år
- x_2: kundens forbrug i Good Food den seneste måned målt i kr.

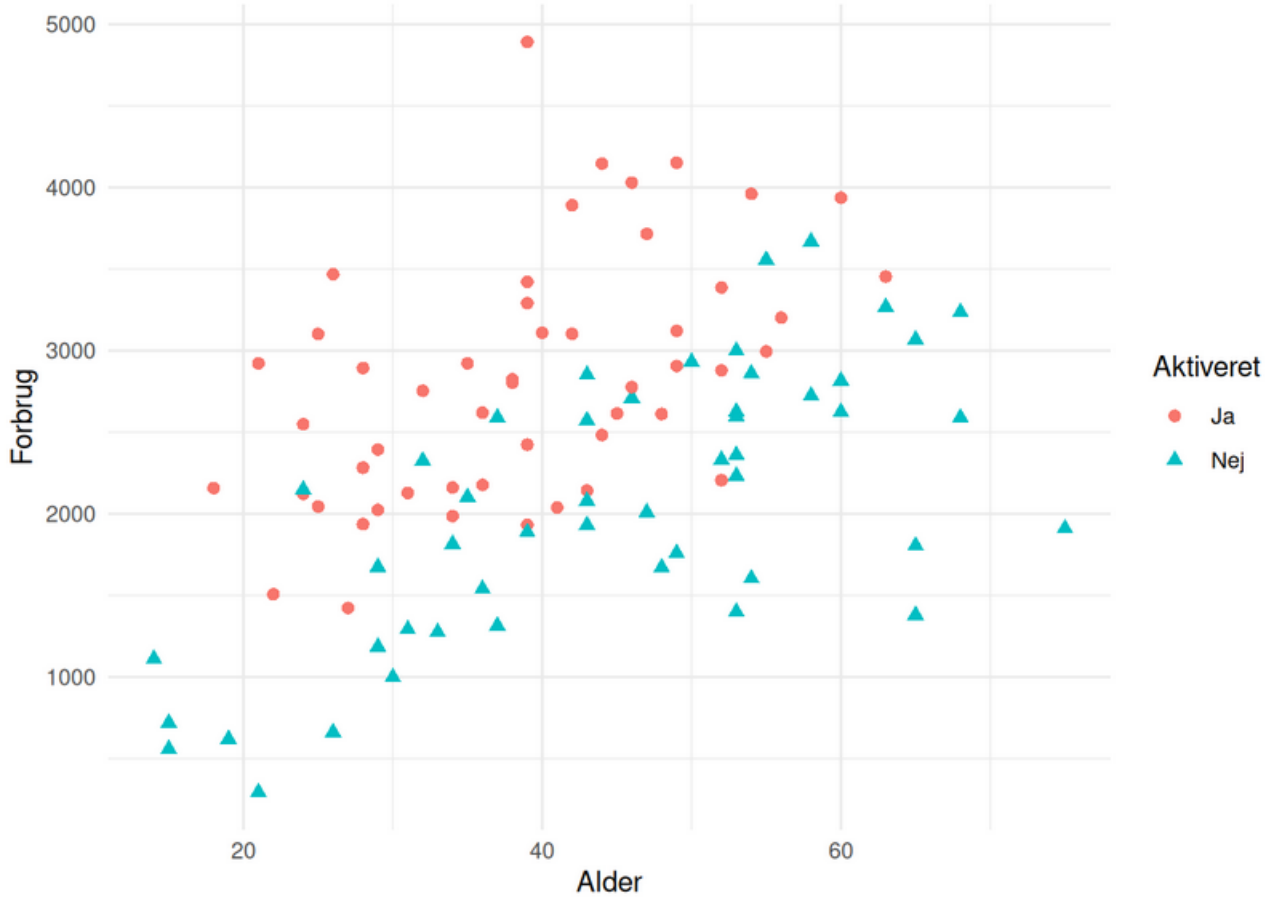
Samtidig har de i en periode for hver kunde registreret, om kunden har aktiveret et bestemt tilbud:

$$t = \begin{cases} 1 & \text{hvis tilbuddet aktiveres} \\ 0 & \text{hvis tilbuddet ikke aktiveres} \end{cases}$$

Træningsdatasæt

t kaldes for en **targetværdi**

Plot af træningsdata



Beregning af outputværdi o

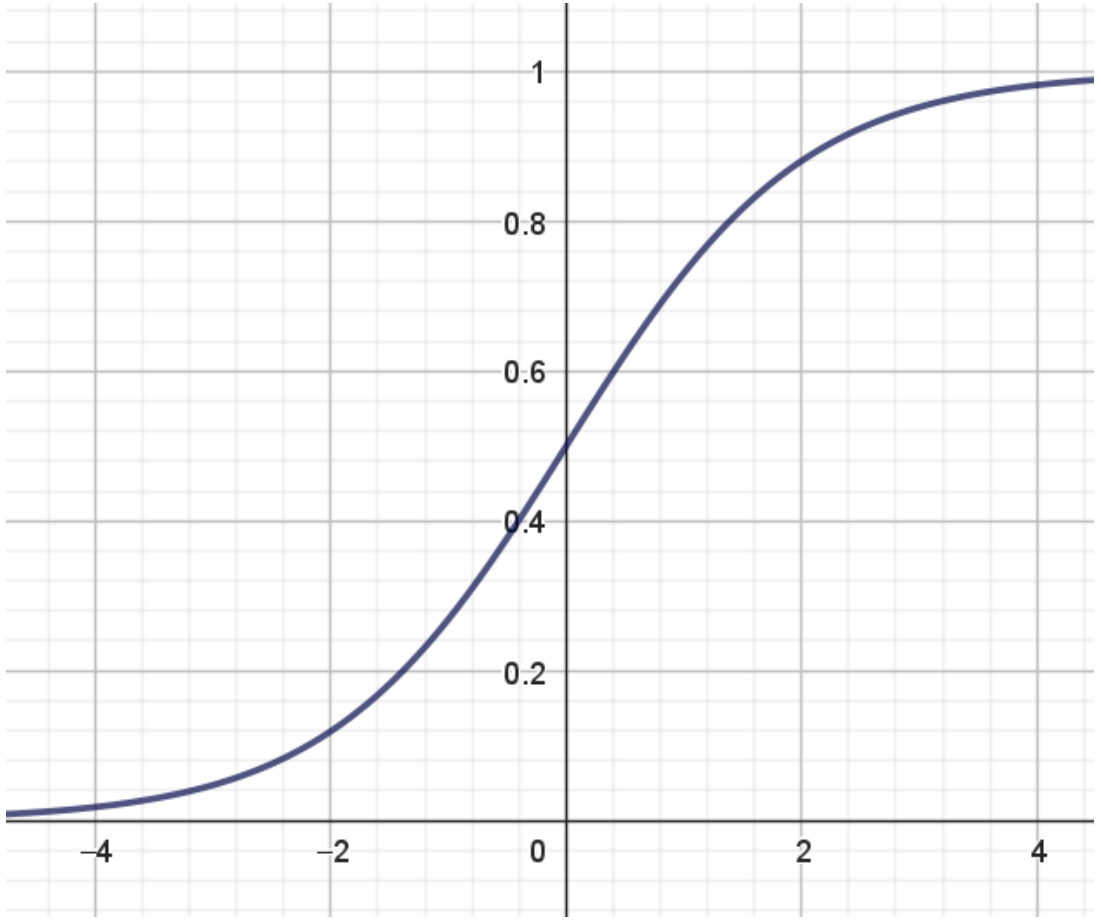
- Outputværdien o skal være et bud på sandsynligheden for, at en kunde aktiverer et tilbud

- Vi beregner en outputværdi o

$$o = \sigma(w_0 + w_1 \cdot x_1 + w_2 \cdot x_2)$$

hvor σ er sigmoid-funktionen:

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

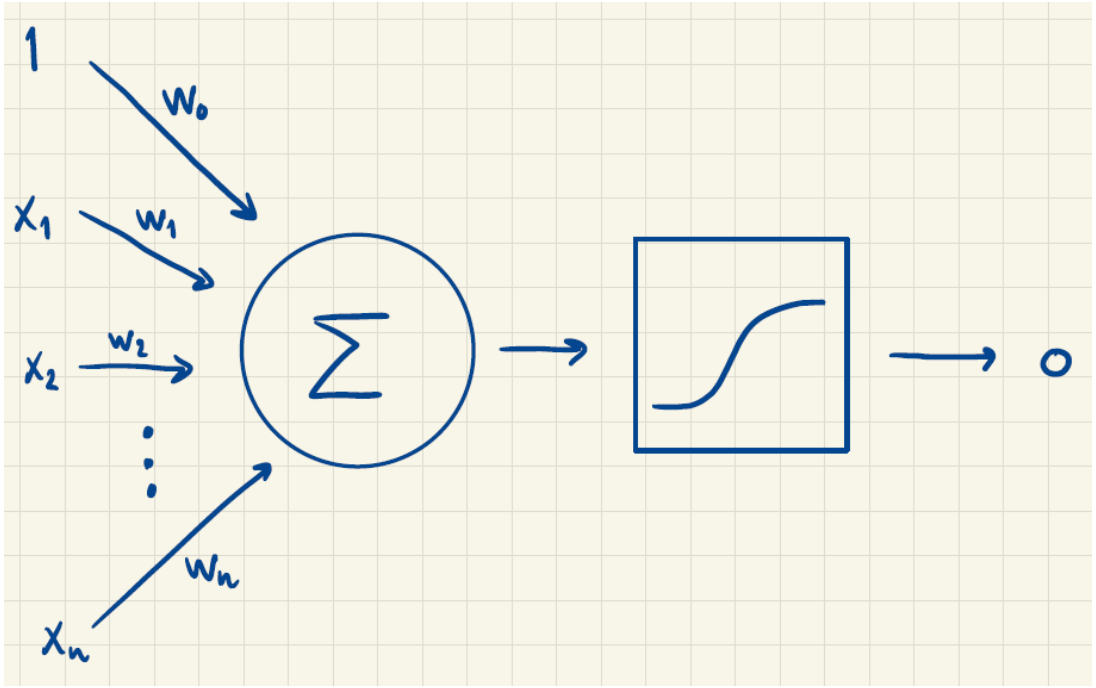


Grafisk illustration

$$o = \sigma(w_0 + w_1 \cdot x_1 + w_2 \cdot x_2)$$

Aktiveringsfunktion

(og der findes andre end sigmoid)



Tabsfunktion (*squared error*)

Vi definerer *squared error* tabsfunktionen:

$$E = \frac{1}{2}\sum (t - o)^2$$

Det vil sige

$$E = \frac{1}{2}\sum (t - \sigma(w_0 + w_1 \cdot x_1 + w_2 \cdot x_2)\,)^2$$

Vægtene w_0, w_1 og w_2 skal bestemmes, så tabsfunktionen minimeres.

Tilbage til træningsdata

- Træningseksempel 1: $\left(x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}, t^{(1)}\right)$

  ⋮

- Træningseksempel m: $\left(x_1^{(m)}, x_2^{(m)}, \dots, x_n^{(m)}, t^{(m)}\right)$

  ⋮

- Træningseksempel M: $\left(x_1^{(M)}, x_2^{(M)}, \dots, x_n^{(M)}, t^{(M)}\right)$

Tabsfunktionen generelt

$$E(w_0, w_1, \dots, w_n)$$

$$= \frac{1}{2} \sum_{m=1}^{M} \left(t^{(m)} - \sigma \left(w_0 + w_1 \cdot x_1^{(m)} + \cdots + w_n \cdot x_n^{(m)} \right) \right)^2$$

Gradientnedstigning

Vægtene opdateres ved at gå et lille stykke i den negative gradients retning:

$$w_i^{(ny)} \leftarrow w_i - \eta \cdot \frac{\partial E}{\partial w_i}$$

for $i \in \{0, 1, \ldots, n\}$.

η: learning rate

Partielle afledede ☺

Husk, at

$$E(w_0, w_1, \dots, w_n) = \frac{1}{2}\sum_{m=1}^{M}\left(t^{(m)} - \sigma\left(w_0 + w_1 \cdot x_1^{(m)} + \cdots + w_n \cdot x_n^{(m)}\right)\right)^2$$

Det vil sige, at

- Vi får brug for sum- og kædereglen
- Vi får brug for at differentiere sigmoid-funktionen

At differentiere aktiveringsfunktioner er en ting!

- Man kan vise (og det er sjovt ☺), at

$$\sigma'(x) = \sigma(x) \cdot \big(1 - \sigma(x)\big) = o \cdot (1 - o)$$

- Det er en eftertragtet egenskab ved aktiveringsfunktioner: at deres afledede kan udtrykkes simpelt ved funktionsværdien selv (det er en beregningsmæssig kæmpe fordel)

Partielle afledede ☺

$$E(w_0, w_1, \dots, w_n) = \frac{1}{2}\sum_{m=1}^{M}\left(t^{(m)} - \sigma\left(w_0 + w_1 \cdot x_1^{(m)} + \cdots + w_n \cdot x_n^{(m)}\right)\right)^2$$

Derfor er

$$\frac{\partial E}{\partial w_i} = -\sum_{m=1}^{M}\left(t^{(m)} - \sigma(\cdots)\right) \cdot \sigma'(\ \cdots) \cdot x_i^{(m)}$$

$$\frac{\partial E}{\partial w_i} = -\sum_{m=1}^{M}\left(t^{(m)} - o^{(m)}\right) \cdot o^{(m)} \cdot (1 - o^{(m)}) \cdot x_i^{(m)}$$

hvor $o^{(m)} = \sigma\left(w_0 + w_1 \cdot x_1^{(m)} + \cdots + w_n \cdot x_n^{(m)}\right)$ for $i \in \{1, \dots, n\}$

Partielle afledede ☺

$$E(w_0, w_1, \dots, w_n) = \frac{1}{2}\sum_{m=1}^{M}\left(t^{(m)} - \sigma\left(w_0 + w_1 \cdot x_1^{(m)} + \cdots + w_n \cdot x_n^{(m)}\right)\right)^2$$

Det vil sige, at

$$\frac{\partial E}{\partial w_0} = -\sum_{m=1}^{M}\left(t^{(m)} - o^{(m)}\right) \cdot o^{(m)} \cdot \left(1 - o^{(m)}\right)$$

og

$$\frac{\partial E}{\partial w_i} = -\sum_{m=1}^{M}\left(t^{(m)} - o^{(m)}\right) \cdot o^{(m)} \cdot \left(1 - o^{(m)}\right) \cdot x_i^{(m)}$$

for $i \in \{1, \dots, n\}$

Opdateringsregler for kunstige neuroner

$$w_0^{(ny)} \leftarrow w_0 + \eta \cdot \sum_{m=1}^{M} \left(t^{(m)} - o^{(m)}\right) \cdot o^{(m)} \cdot \left(1 - o^{(m)}\right)$$

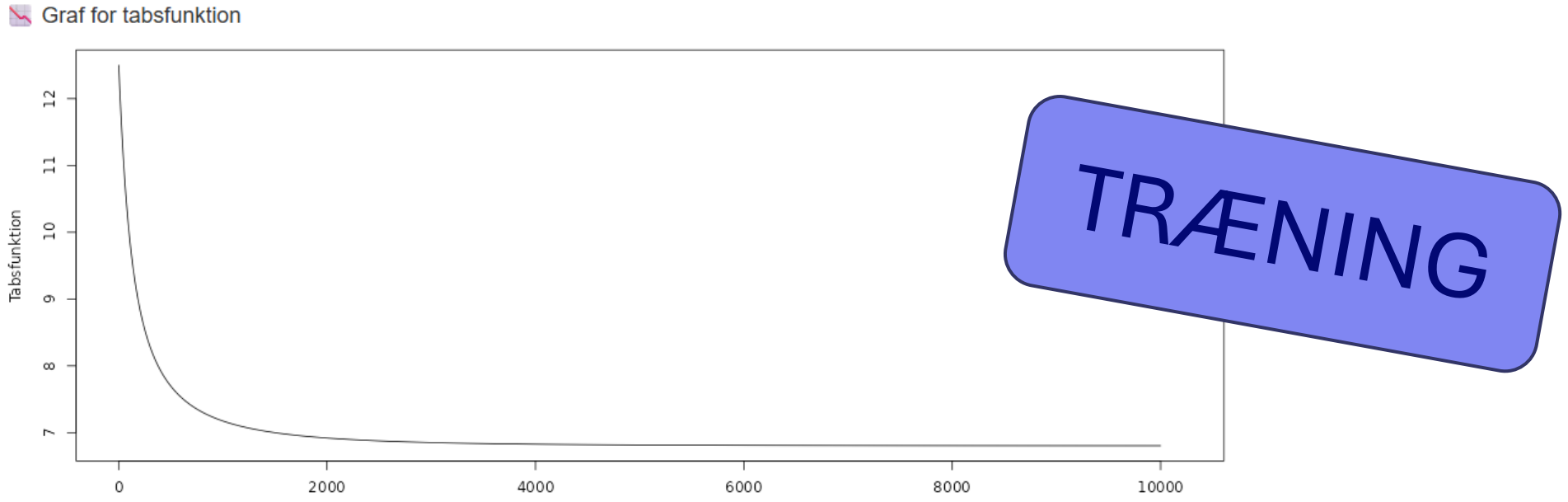
og

$$w_i^{(ny)} \leftarrow w_i + \eta \cdot \sum_{m=1}^{M} \left(t^{(m)} - o^{(m)}\right) \cdot o^{(m)} \cdot \left(1 - o^{(m)}\right) \cdot x_i^{(m)}$$

for $i \in \{1, \ldots, n\}$.

Træning af kunstig neuron

- Alle vægtene opdateres, indtil værdien af tabsfunktionen (næsten) ikke ændrer sig.



- Herefter kan AI modellen anvendes!

Prædiktion

For en ny kunde kan vi lave følgende prædiktion:

$$\text{Kunden aktiverer tilbuddet: } \begin{cases} \text{Ja} & o \geq 0.5 \\ \text{Nej} & o < 0.5 \end{cases}$$

hvor

$$o = \sigma(w_0 + w_1 \cdot x_1 + w_2 \cdot x_2)$$

Bemærk! Vægtene er nu fastlagte, fordi vi har trænet neuronen.

Skillelinje (eller plan/hyperplan)

Bemærk, at

$$\sigma(w_0 + w_1 \cdot x_1 + \cdots + w_n \cdot x_n) = 0.5$$

svarer til

$$\frac{1}{1 + e^{-(w_0 + w_1 \cdot x_1 + \cdots + w_n \cdot x_n)}} = 0.5$$

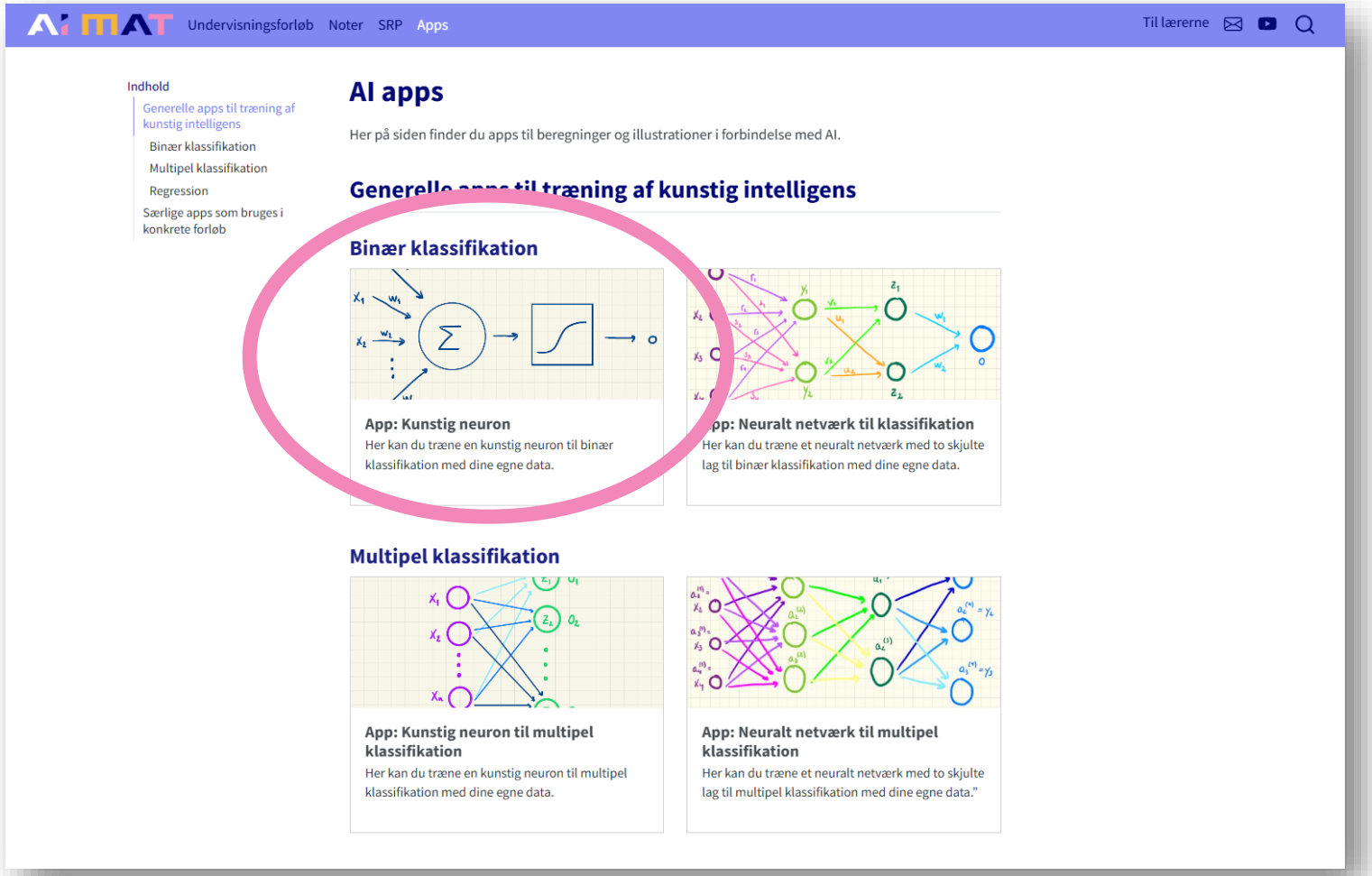
Det vil sige, at

$$e^{-(w_0 + w_1 \cdot x_1 + \cdots + w_n \cdot x_n)} = 1$$

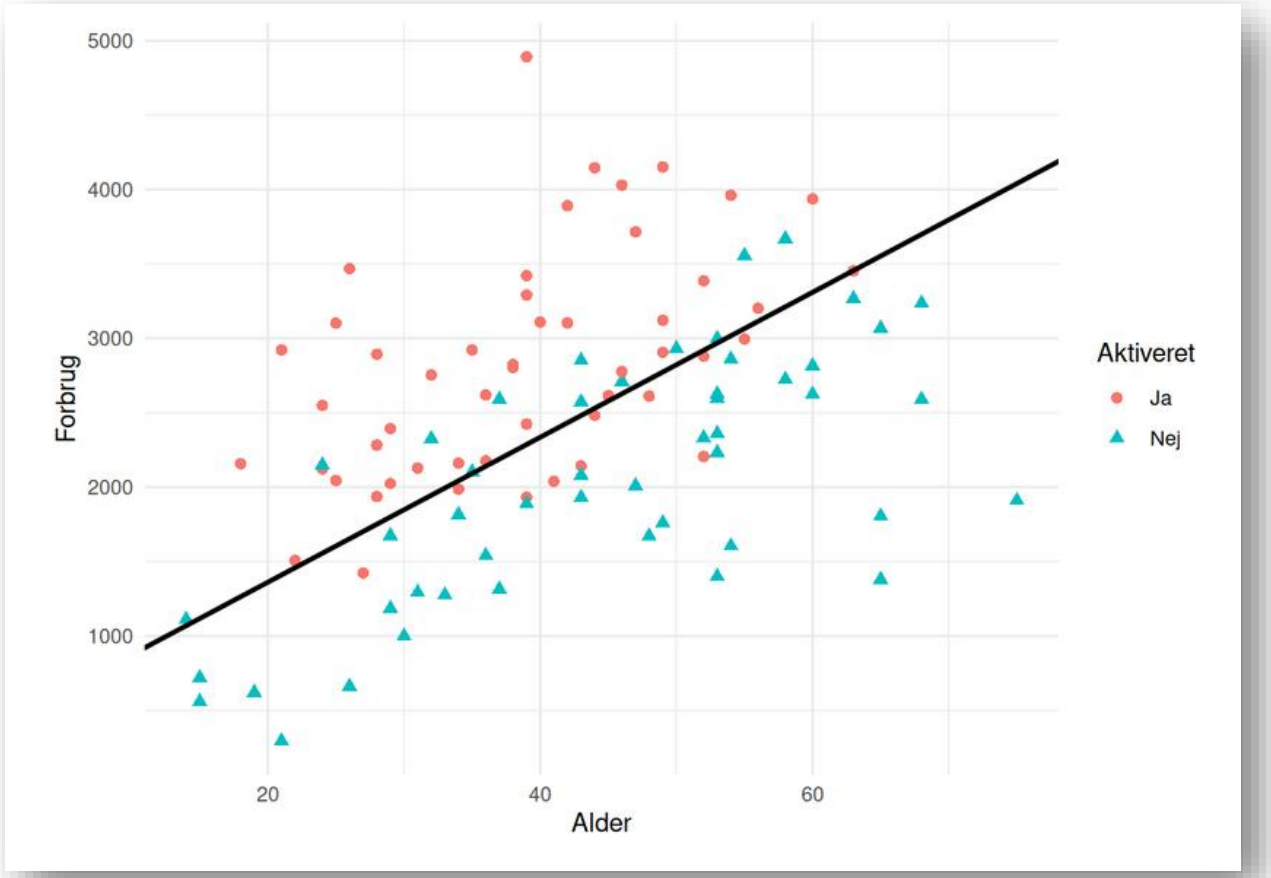
Altså

$$w_0 + w_1 \cdot x_1 + \cdots + w_n \cdot x_n = 0$$

Træning af kunstig neuron



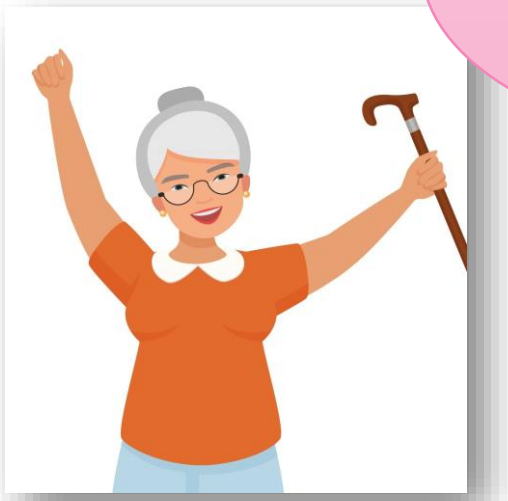
Træning baseret på Good Food datasæt



$$w_0 = -0.7376$$
$$w_1 = -0.142$$
$$w_2 = 0.002806$$

Sig hej til Hans og Gerda

Gerda
Alder: 65 år
Forbrug: 2800 kr.



Hans
Alder: 50 år
Forbrug: 3500 kr.



Skal vi vise tilbuddet til Hans og/eller Gerda?

Hans

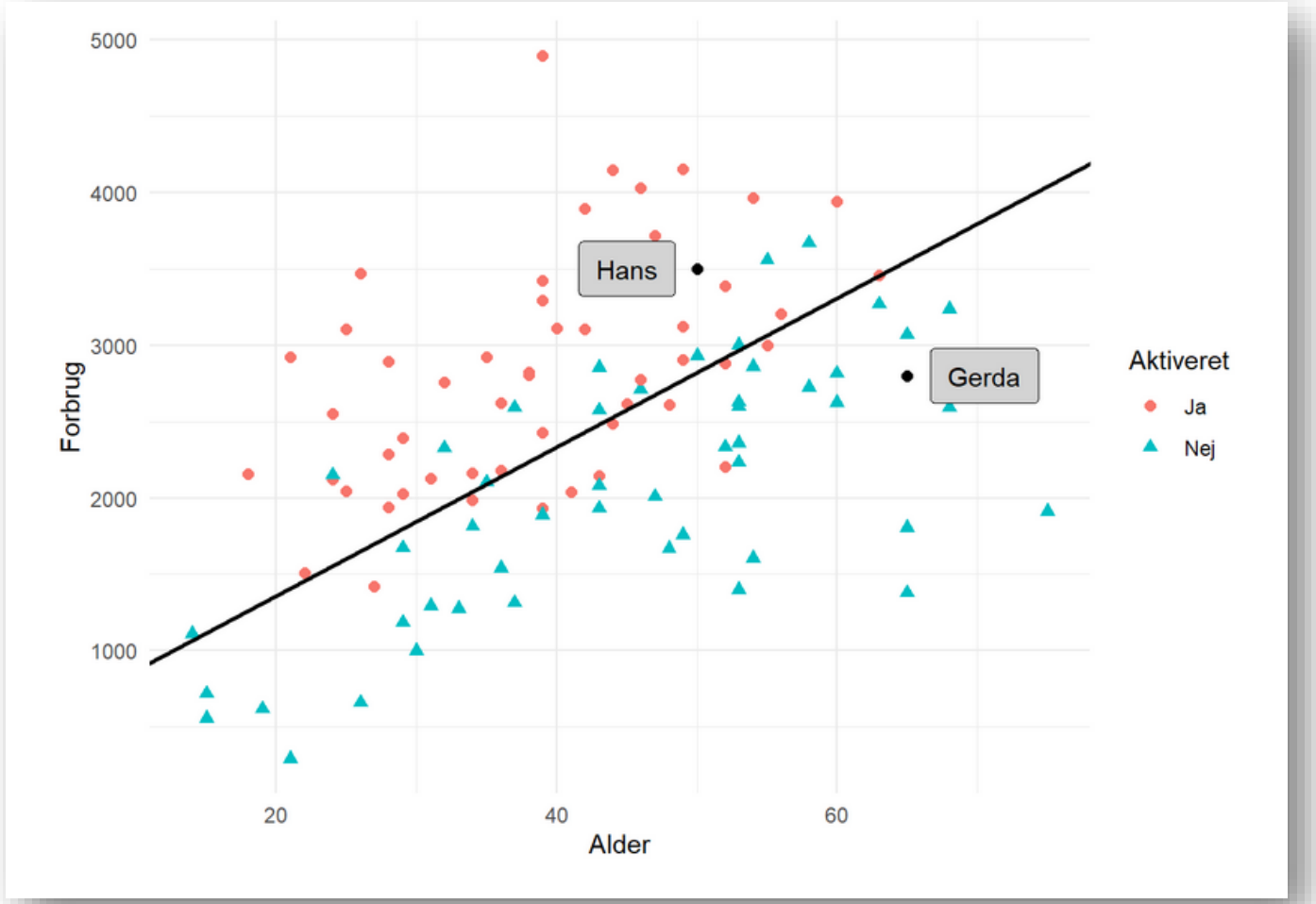
$$o = \sigma(-0.7376 - 0.142 \cdot 50 + 0.002806 \cdot 3500) = 0.879$$

Gerda

$$o = \sigma(-0.7376 - 0.142 \cdot 65 + 0.002806 \cdot 2800) = 0.108$$

Vis tilbuddet til Hans, men ikke til Gerda!

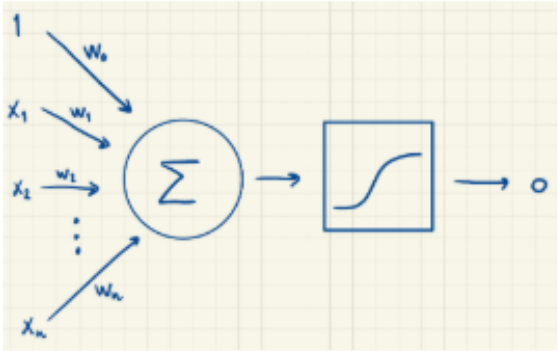
Hans og Gerda i koordinatsystemet



Det er sådan cirka indholdet af noten

Kunstige neuroner

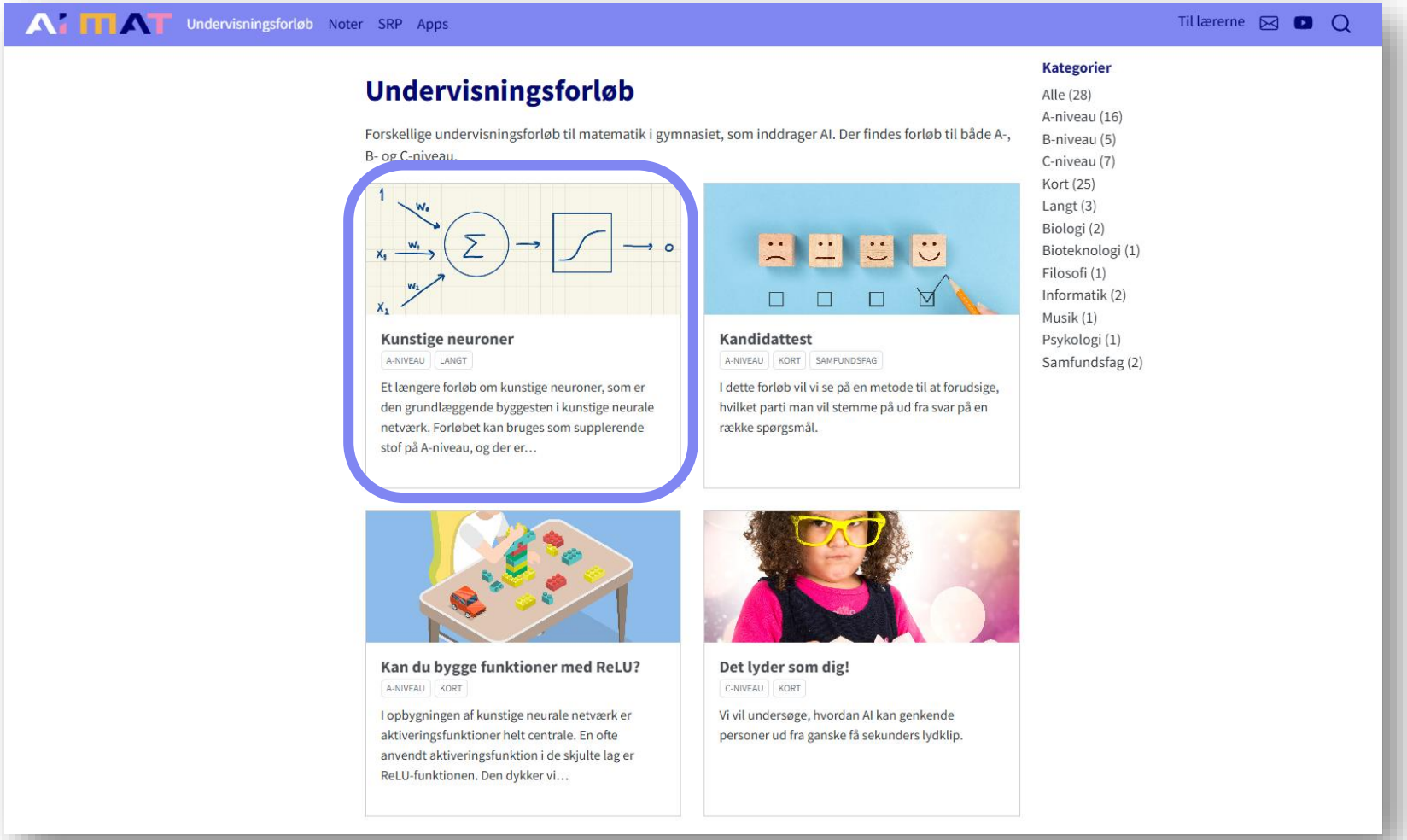
Noten behandler byggestenene til de mere generelle kunstige neurale netværk nemlig de *kunstige neuroner*. De kunstige neuroner bruges – som de beskrives her – til binær klassifikation. Opdateringsreglerne til justering af vægtene udledes ved at minimere *squared error* tabsfunktionen ved hjælp af gradientnedstigning. Vi bruger her **sigmoid**-funktionen som aktiveringsfunktion.



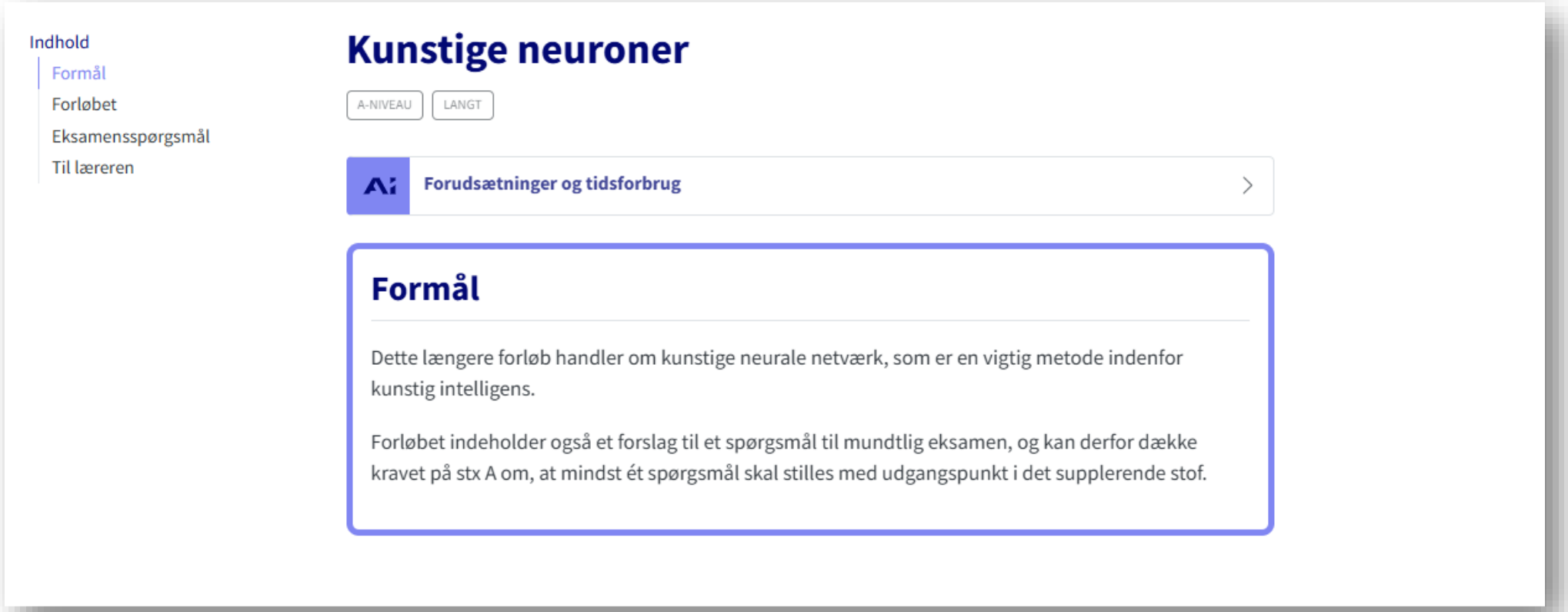
Type	Aktiveringsfunktion	Tabsfunktion
Binær klassifikation	Sigmoid	*Squared error*

Sværhedsgrad: **

Langt forløb baseret på noten



Langt forløb om kunstige neuroner



Langt forløb om kunstige neuroner

Forløbet

Forløbet består af følgende dele hvoraf de sidste tre er valgfrie og uafhængige af hinanden:

Hovedforløb

- Del 1: Aktiveringsfunktioner
- Del 2: Gradientnedstigning
- Del 3: Tabsfunktioner og opdateringsregler
- Del 4: Træning af kunstig neuron

Valgfrie dele

- Del 5: Krydsvalidering og modeludvælgelse
- Del 6: Andre tabsfunktioner
- Del 7: Vurdering af værdien af tabsfunktionen (skibidi toilet)

Eksamensspørgsmål

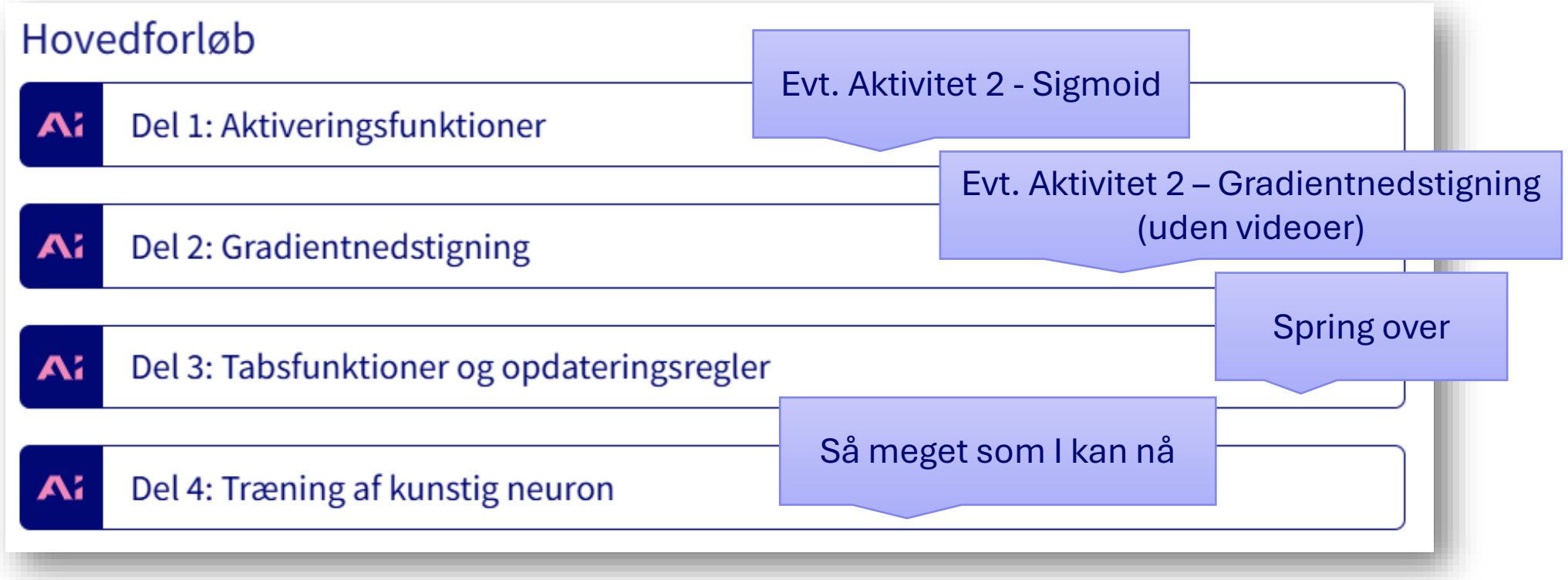
- Eksamensspørgsmål

Til læreren

Herunder findes diverse fif til læreren.

- Til læreren

Hands-on


Hovedforløb
Del 1: Aktiveringsfunktioner
Del 2: Gradientnedstigning
Del 3: Tabsfunktioner og opdateringsregler
Del 4: Træning af kunstig neuron
Evt. Aktivitet 2 - Sigmoid
Evt. Aktivitet 2 – Gradientnedstigning (uden videoer)
Spring over
Så meget som I kan nå