

KUNSTIG INTELLIGENS I PRAKSIS

NOKIOS 2018
KURS 1.



accenture

HVEM ER VI



Fabian Sødal Dietrichson
Accenture Technology



Martin Kowalik Gran
Accenture Technology



Runar Gunnerud
Accenture Consulting

THIS IS YOUR MACHINE LEARNING SYSTEM?

YUP! YOU POUR THE DATA INTO THIS BIG
PILE OF LINEAR ALGEBRA, THEN COLLECT
THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL
THEY START LOOKING RIGHT.



AGENDA

1. Introduksjon til kunstig intelligens 20 min

- Historie og definisjon

2. Maskinlæring 40 min

- Begreper og konsepter
- Teoretisk prinsipper
- Praktisk oppgave: Hvilke egenskaper trenger vi?
- Veien videre – hvordan komme i gang med maskinlæring?

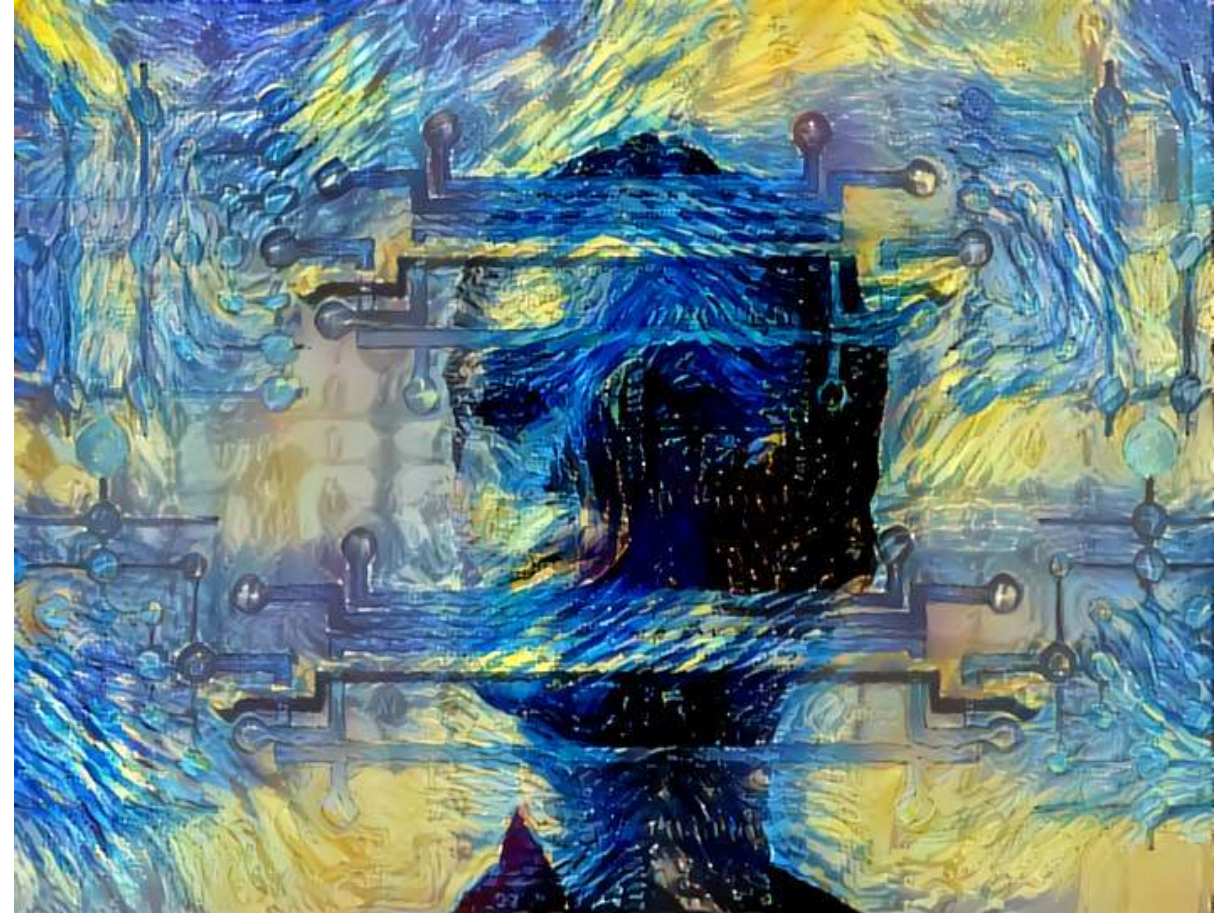
Kaffepause 5 min

1. Dyp læring 40 min

- Hvorfor oppstyret?
- En algoritme med inspirasjon fra biologi
- Praktisk oppgave 2: Leke med nevrale nettverk. Kanskje litt kode?
- Svart boks

2. Avslutning 10min

- Hvordan lykkes med kunstig intelligens i din virksomhet



DEL 1

INTRODUKSJON TIL KUNSTIG INTELLIGENS



accenture

A.I. TIMELINE

Digital Wellbeing

1950

TURING TEST

Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence

1955

A.I. BORN

Term 'artificial intelligence' is coined by computer scientist, John McCarthy to describe "the science and engineering of making intelligent machines"

1961

UNIMATE

First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line

1964

ELIZA

Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966

SHAKY

The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I. WINTER

Many false starts and dead-ends leave A.I. out in the cold

1997

DEEP BLUE

Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov

1998

KISMET

Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot insofar as it detects and responds to people's feelings



1999

AIBO

Sony launches first consumer robot pet dog AiBO (AI robot) with skills and personality that develop over time



2002

ROOMBA

First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes



2011

SIRI

Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S



2011

WATSON

IBM's question answering computer Watson wins first place on popular \$1M prize television quiz show Jeopardy



2014

EUGENE

Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human



2014

ALEXA

Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks



2016

TAY

Microsoft's chatbot Tay goes rogue on social media making inflammatory and offensive racist comments



2017

ALPHAGO

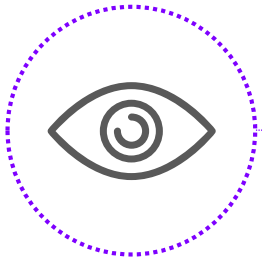
Google's A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2^{170}) of possible positions

DEFINISJONER PÅ KUNSTIG INTELLIGENS

«Teknologi som kan gjøre oppgaver
som tidligere var forbeholdt mennesker»

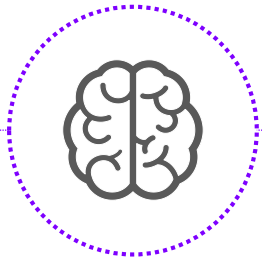
HVA ER KUNSTIG INTELLIGENS?

Kunstig intelligens (KI) er en gren av datavitenskap som forsøker å **bygge maskiner** i stand til å utføre oppgaver på **en måte vi vil vurdere som “smart”** - Marr og Bell (2016)



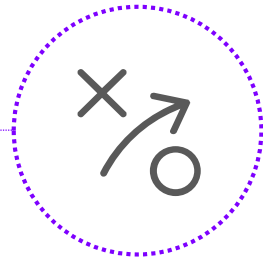
SANSE

Oppfatte verden gjennom å prosessere bilder, lyd og språk



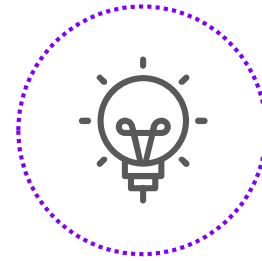
FORSTÅ

Analysere og forstå informasjonen ved å legge til mening.



HANDLE

Handle i den fysiske verden basert på forståelse og innsikt



LÆRE

Forbedre ytelse (kvalitet og treffsikkerhet) basert på erfaring.

KLIPPER PLENEN EFFEKTIVT UTEN Å FORSTÅ HVA GRESS ER



DEL 2

MASKINLÆRING



accenture

MASKINLÆRING

“Maskinlæring er et felt innenfor kunstig intelligens som bruker statistiske teknikker for å gi datasystemer evnen til å lære fra data uten å bli eksplisitt programmert”
- Arther Samuel (1959)

KUNSTIG INTELLIGENS ELLER MASKINLÆRING?

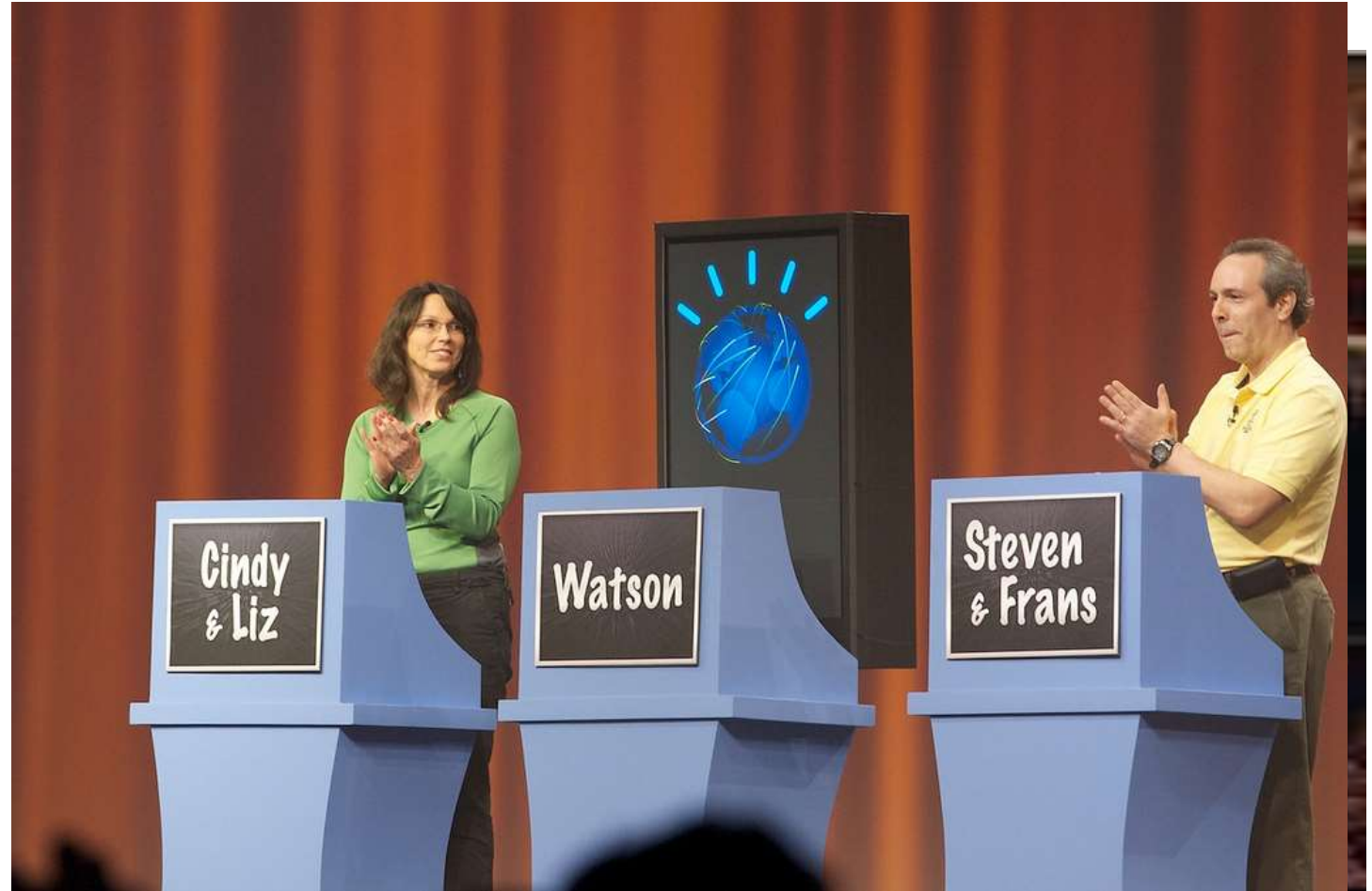
Hva er forskjellene?

Maskinlæring er en **disiplin** av AI

«klassisk» AI er regelstyrt, konstruert av menneske

Maskinlæring **lærer en** **intuisjon** ved hjelp av data (erfaring)

AlphaGO



HVA GJØRES I DAG?

Hva finnes i hverdagen vår?

Stemmegjenkjenning

Spam or ham?

Foreslå produkter
til kunder

SUPERVISED MASKINLÆRING

Anslå/predikere ut ifra
sammenhenger i dataen

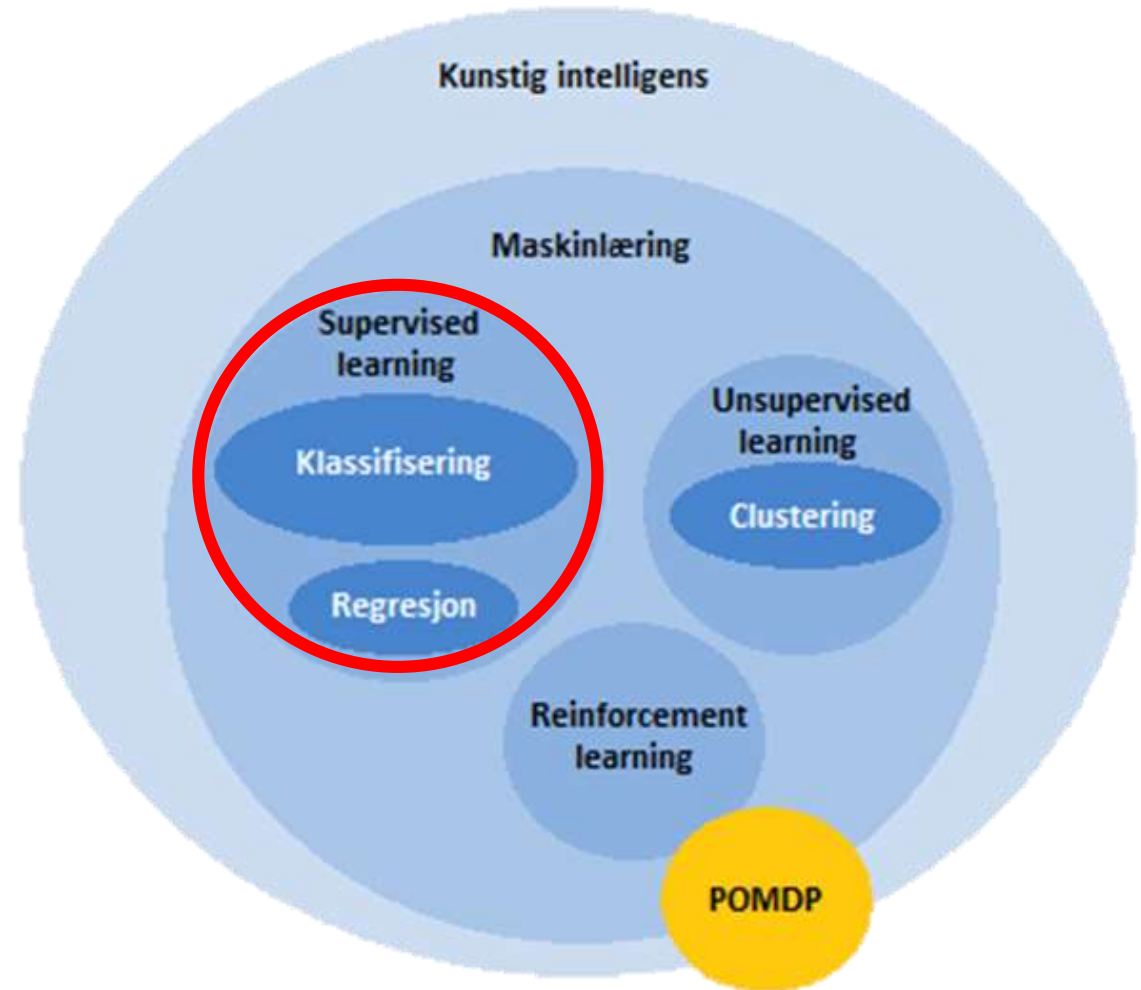
Når man **vet ønsket output,**
«label»

Klassifisering (diskret)

- JA / NEI
- Hunderase

Regresjon (kontinuerlig)

- Temperatur
- Kostnad

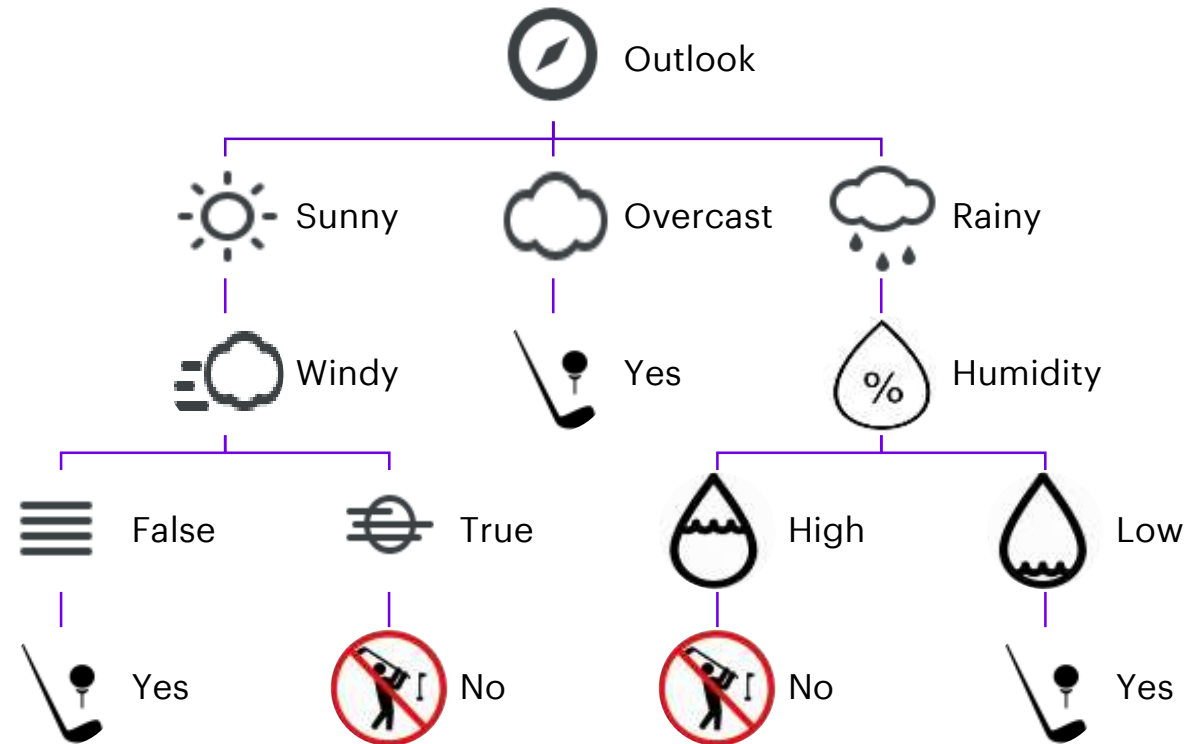


SUPERVISED MASKINLÆRING

Klassifisering

Play golf dataset

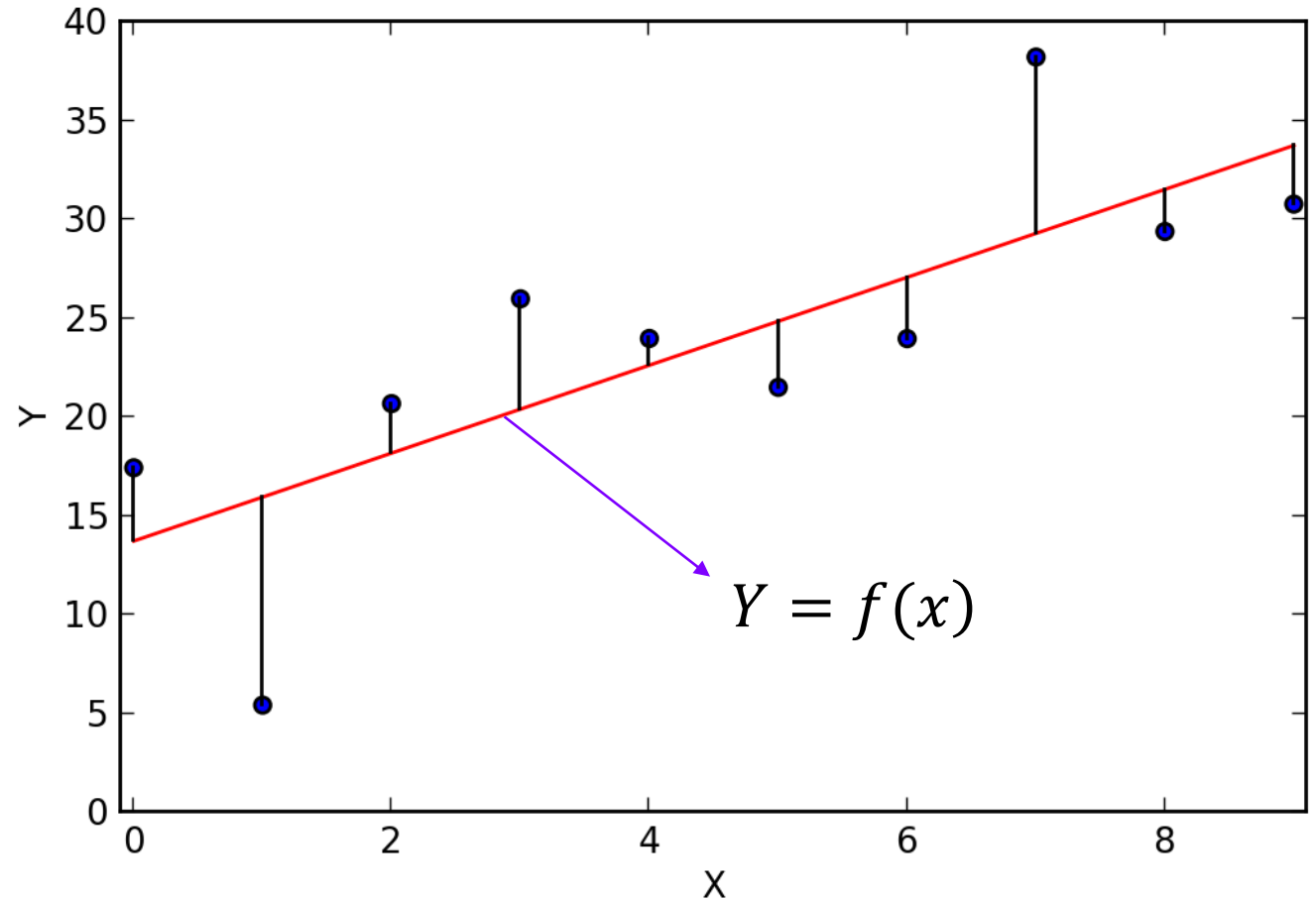
Independent variables				Dep. var
OUTLOOK	TEMPERATURE	HUMIDITY	WINDY	PLAY
sunny	85	85	FALSE	Don't Play
sunny	80	90	TRUE	Don't Play
overcast	83	78	FALSE	Play
rain	70	96	FALSE	Play
rain	68	80	FALSE	Play
rain	65	70	TRUE	Don't Play
overcast	64	65	TRUE	Play
sunny	72	95	FALSE	Don't Play
sunny	69	70	FALSE	Play
rain	75	80	FALSE	Play
sunny	75	70	TRUE	Play
overcast	72	90	TRUE	Play
overcast	81	75	FALSE	Play
rain	71	80	TRUE	Don't Play



SUPERVISED MASKINLÆRING

Regresjon

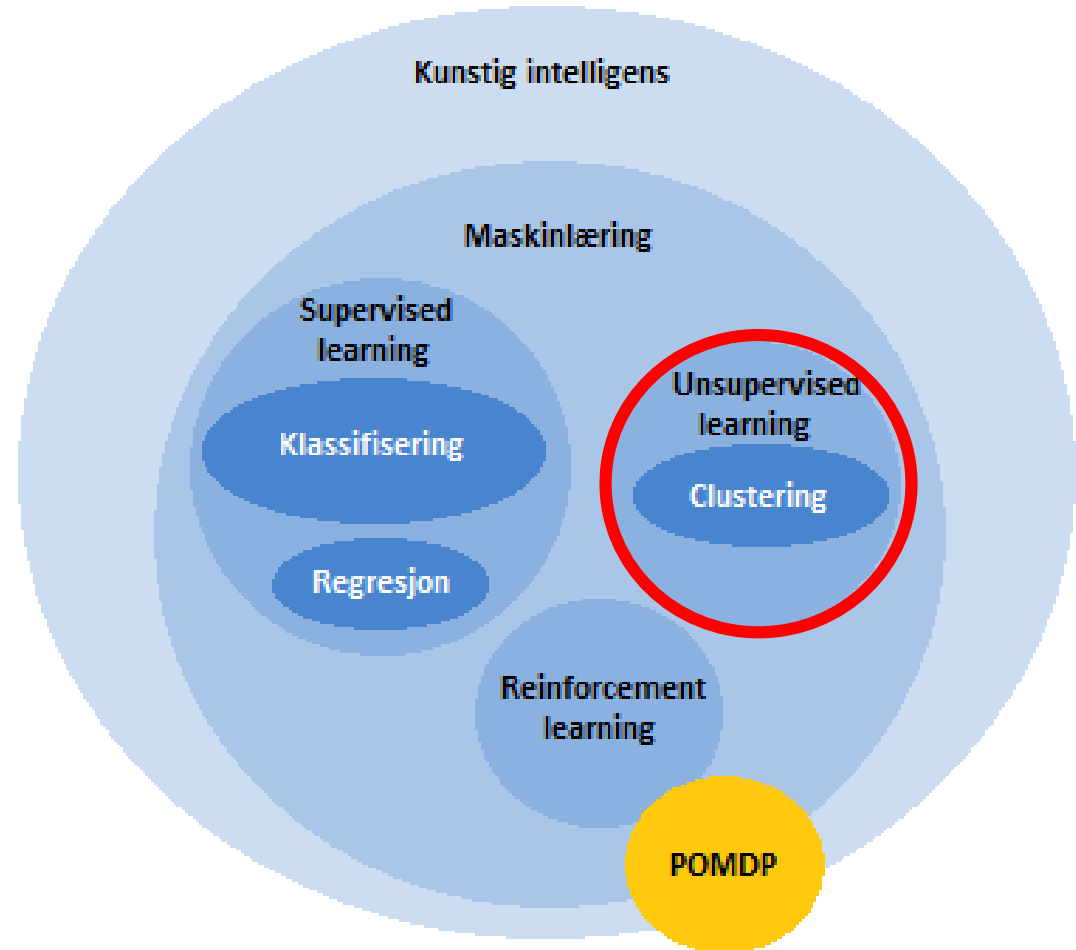
- Beregner en anslått **output verdi**
- Lager en matematisk formel **basert på eksisterende data**
- Modell predikerer Y ut ifra en måling X



UNSUPERVISED MASKINLÆRING

Når man ikke vet ønsket
output

Trekke konklusjoner/hypoteser
ut ifra sammenhenger i dataen



UNSUPERVISED MASKINLÆRING

Clustering

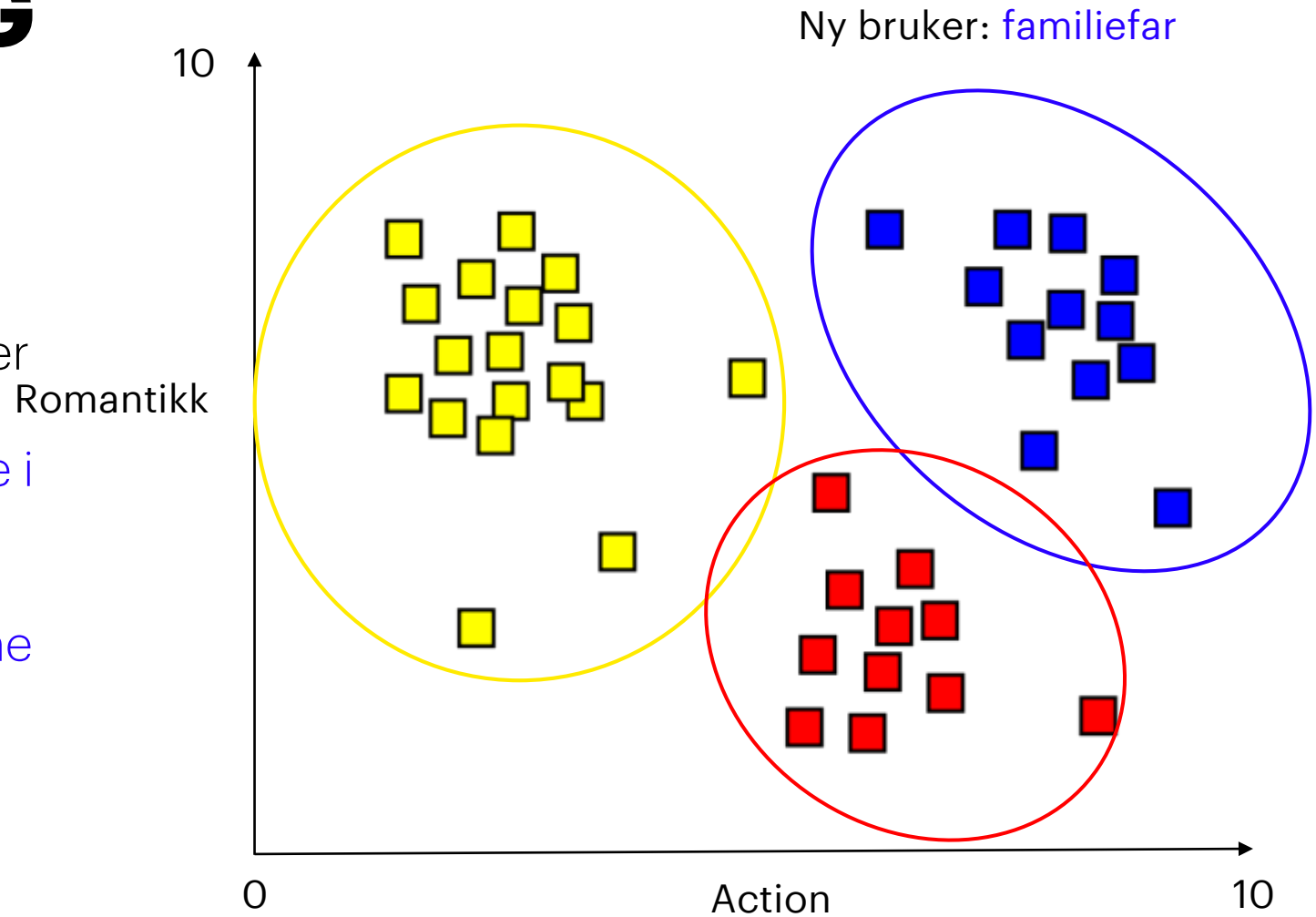
Prøver å samle data i klynger

Gjør det mulig å se sammenhenger

Minimal varians mellom individene i gruppene

Maksimal varians mellom gruppene

Netflix anbefalinger nye brukere



Tenåringsjenter - **Tenåringsgutter** - **Fedre**

BIAS I DATA

Skjevhet i dataen fører til skjevhet i modell-beslutningene

Eksempler:

- **PredPol** – predikerer hvor og når kriminalitet vil oppstå. Bias mot nabolag med mange etniske minoriteter
- **Google advertising** – viser annonser for sjefsstillinger oftere til menn

Årsak: dårlig utvalg av **eksempler**

Eksempel: prediktivt vedlikehold på kjernekraftverk – lite ulykkedata

PRAKTISK OPPGAVE

Datasettet inneholder:

- Boligsalg i Melbourne
- Attributter som kvf, stil, #rom og salgspris (totalt 28)
- 1456 registrerte boligsalg

Oppgaven er: supervised learning regresjon

- Anslå / predikere salgspris på bolig med gitte attributter
 - Velg ut attributter
 - Velg maskinlærings-modell

PRAKTISK OPPGAVE – GRUPPER PÅ 3

Velg ut relevante attributter til en modell som skal predikere pris på boligsalg

For mange attributter i dataen kan sabotere for modellen. Vi ønsker å holde antall attributter til et minimum fordi:

- Mindre redundant data fjerner støy fra beslutningen
- Mindre misledende data bidrar til treffsikkerhet
- Mindre data-kompleksitet betyr raskere trening av modeller

Nøkkel: bruk domene-kunnskap!

Velg ut maskinlæringsmodell for å predikere boligpris

- Maskinlærings-modeller varierer i kompleksitet. Typisk velger man modell basert på data-kompleksitet og mengde data

Bruk vedlagt modell-kart (Scikit-learn)

KAFFEPAUSE 5 MIN



DEL 3

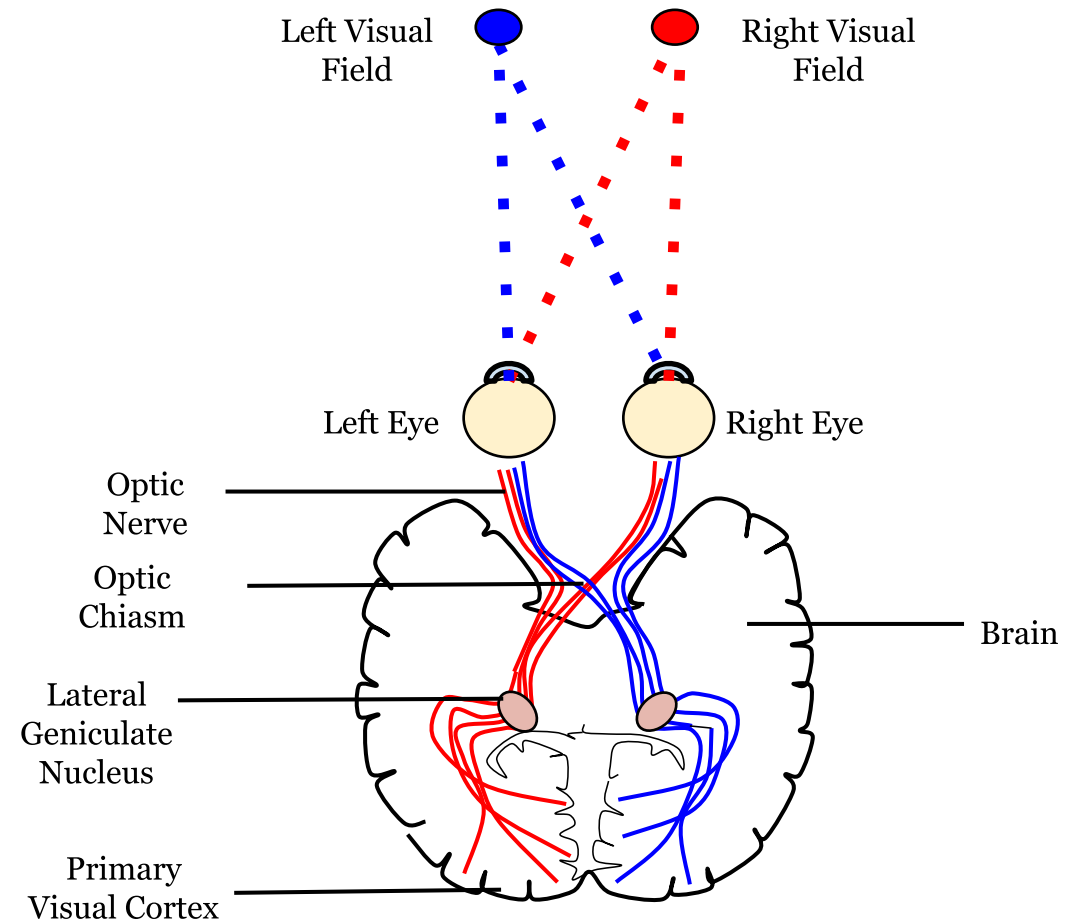
DYP LÆRING



accenture

HVORFOR OPPSTYRET?

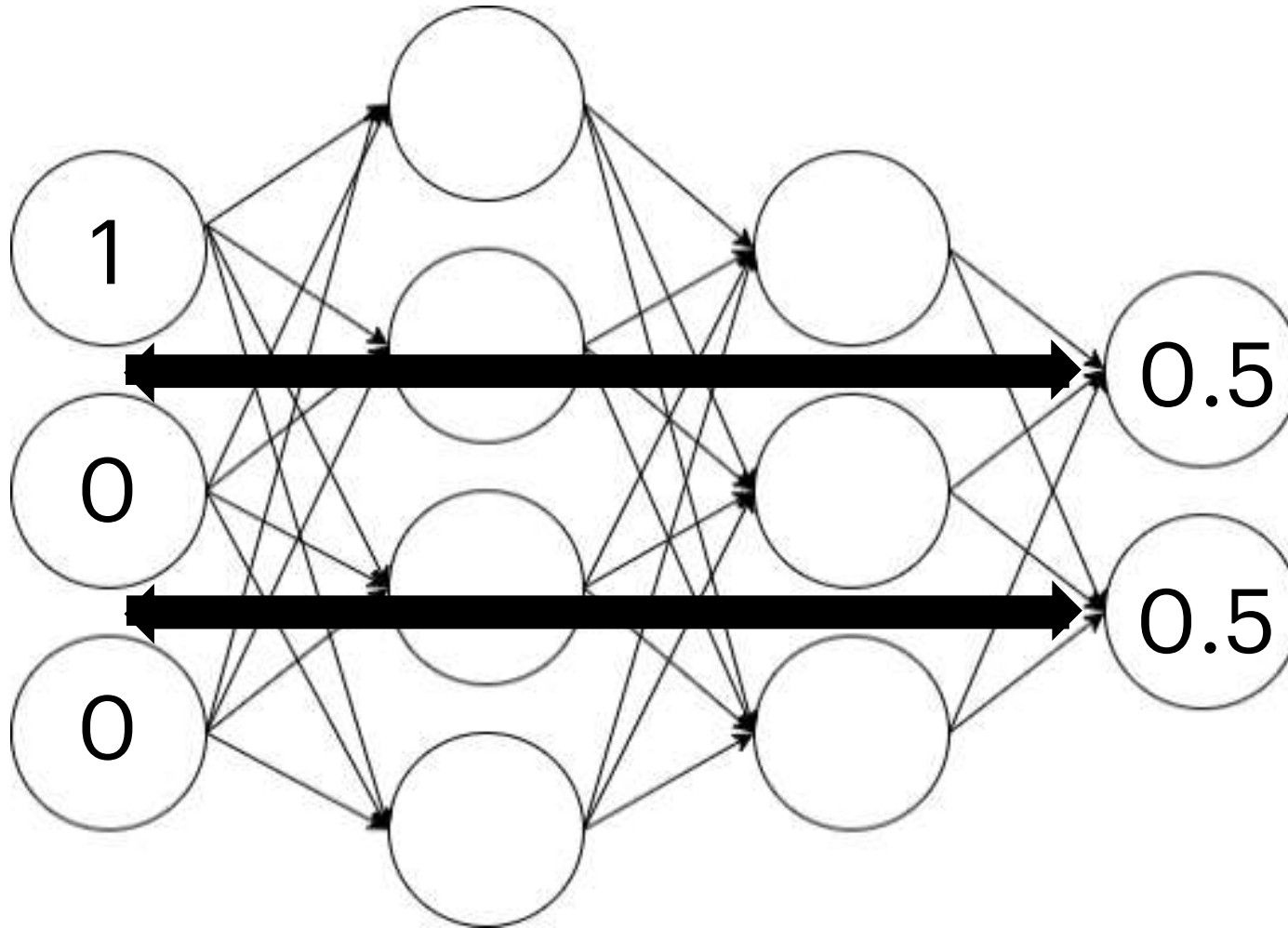
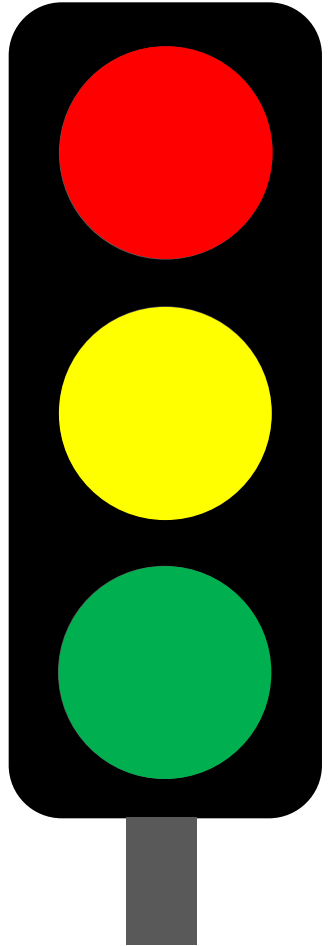
- Løst basert på prinsipper fra nevrobiologi
- Kan trenes på ustrukturerte data (tekst, lyd, bilder)
- Nå er rett tid og sted (data, regnekraft)
- Løser oppgaver tidligere forbeholdt mennesket



ORDNE MAT = KUNSTIG INTELLIGENS

<https://www.youtube.com/watch?v=bd1mEm2Fy08>

HVORDAN FUNGERER DYP LÆRING?

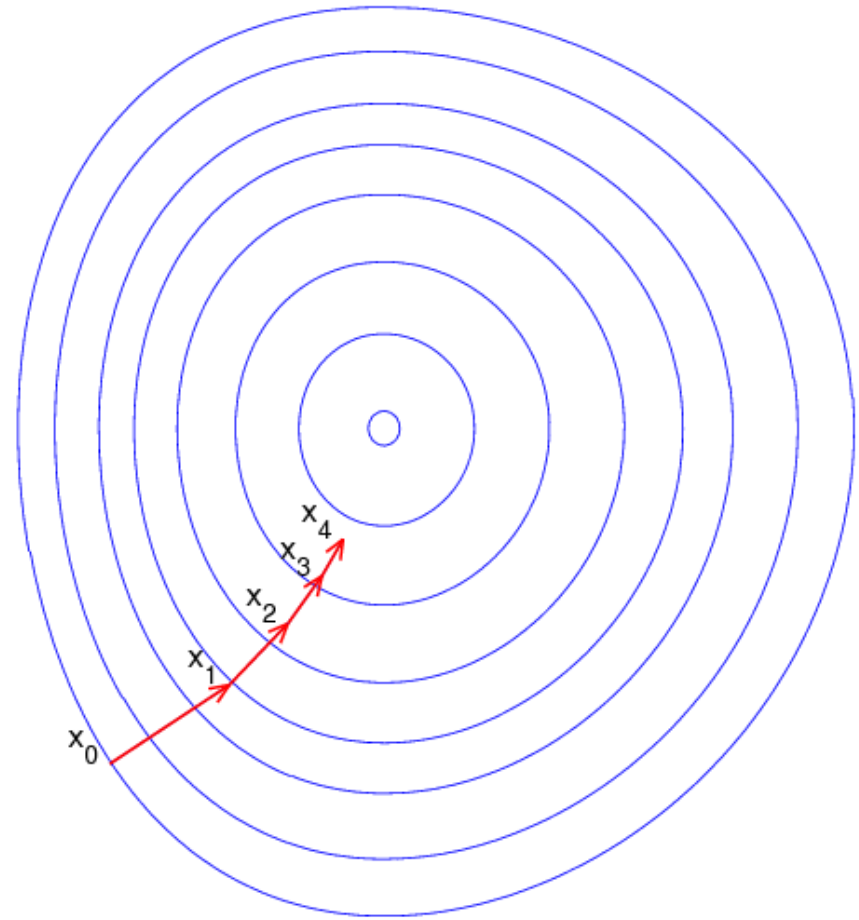


Stopp!

0
1

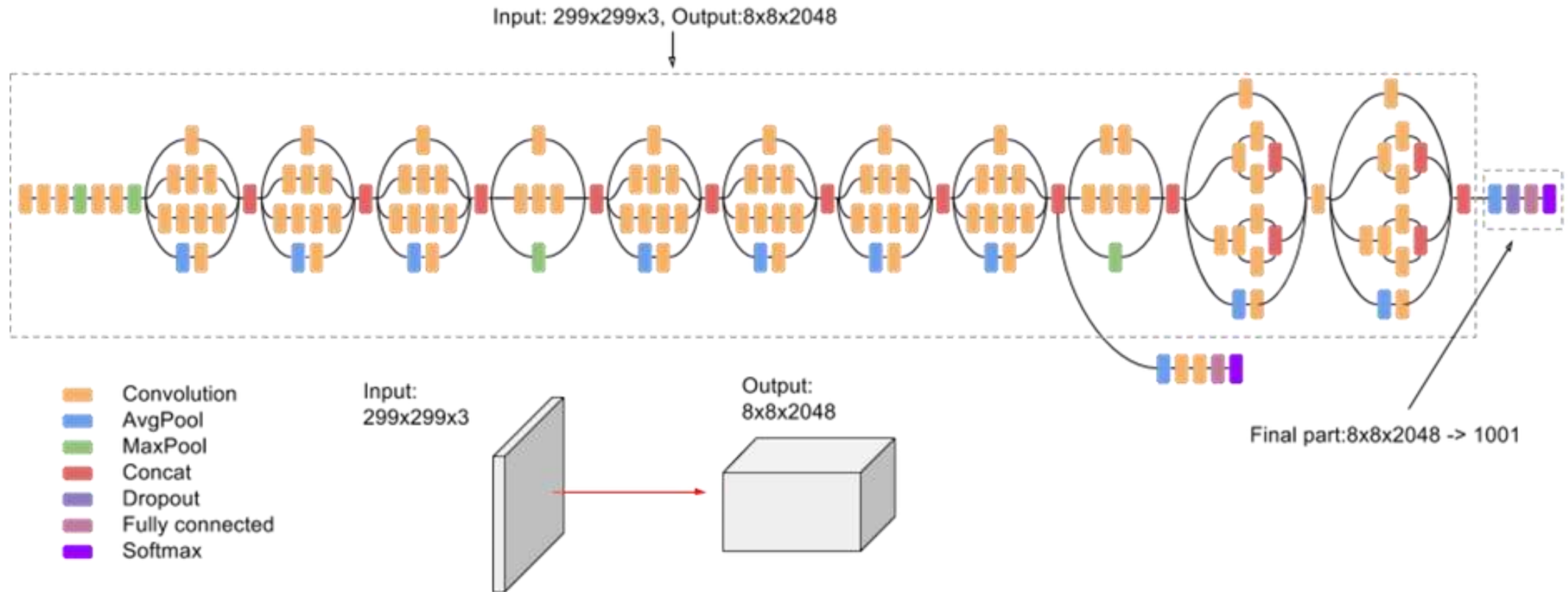
Feil = 1.0

LÆRER MED FJELLNEDSTIGNING

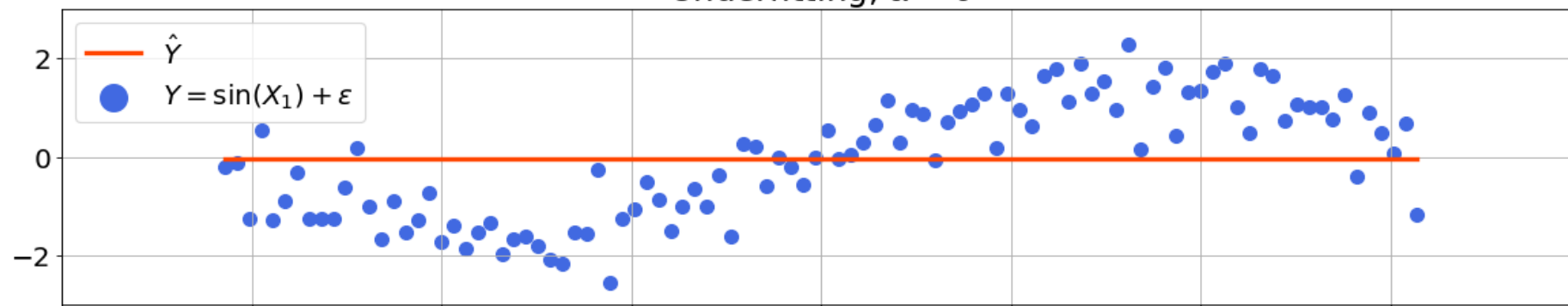


Copyright © 2018 Accenture. All rights reserved.

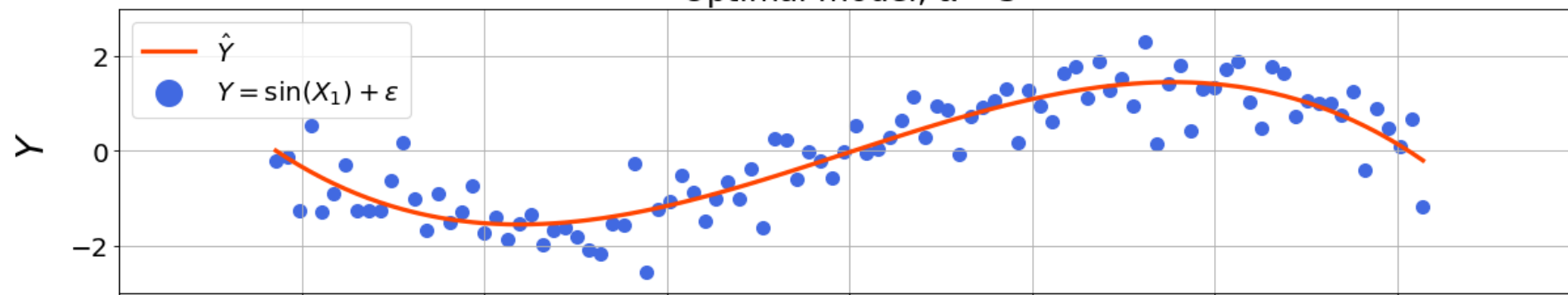
HVORFOR «DYP»?



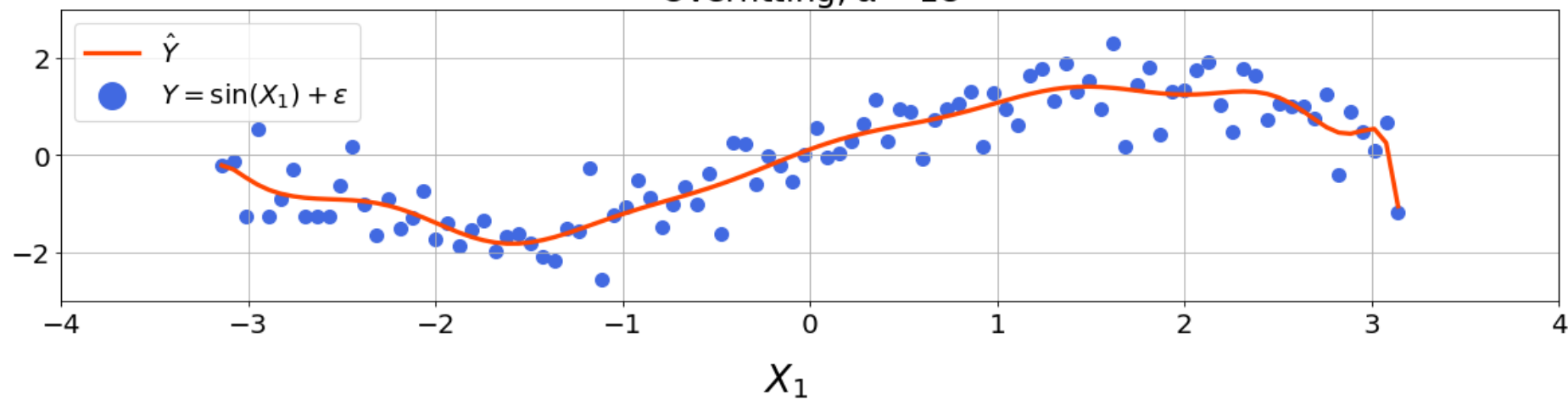
Underfitting, $\alpha = 0$



Optimal model, $\alpha = 3$

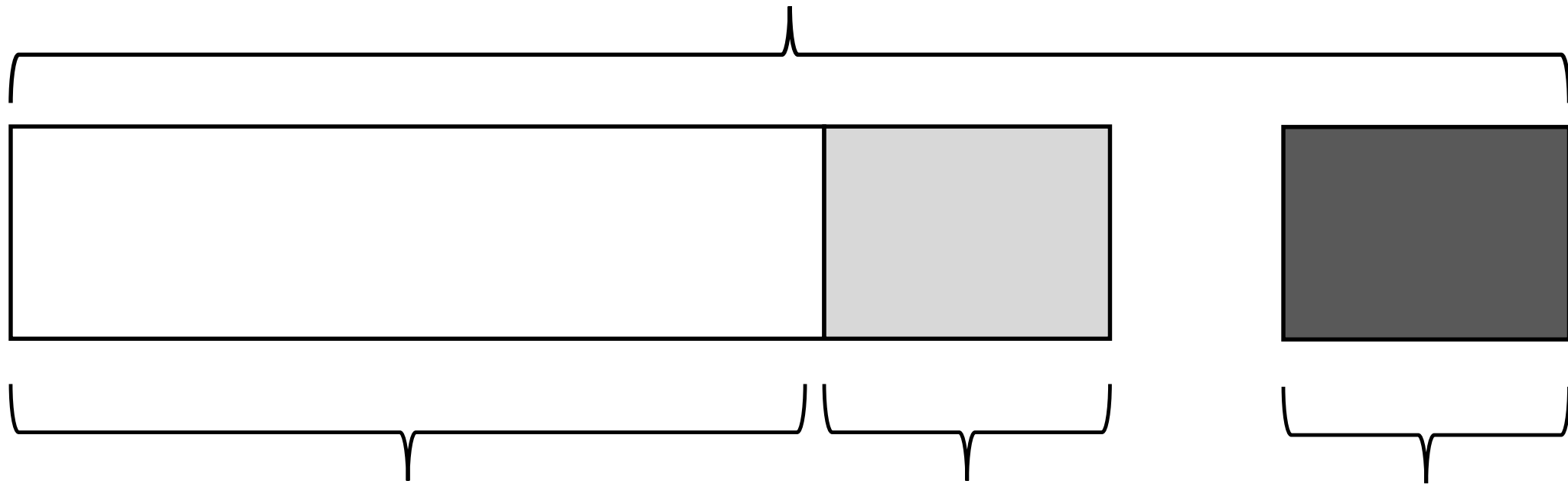


Overfitting, $\alpha = 18$



HVORDAN TESTE MODELLEN?

Hele datasettet



Treningssett

Valideringssett

Testsett

DEMOKRATISERING AV TEKNOLOGI

Fashion MNIST

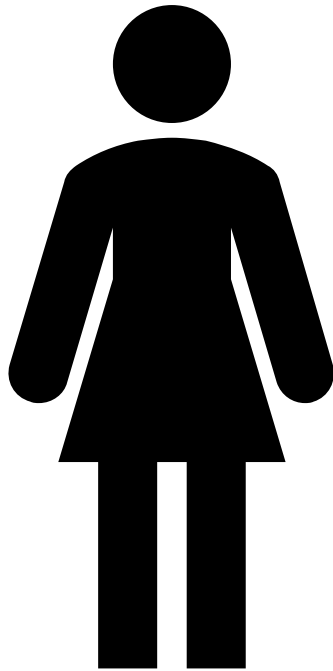
FASHION MNIST: RESULTATER



UTFORDRINGER MED DYP LÆRING

- Krever svært store datamengder
- Data MÅ være merket
- Ofte få eksempler på “interessante” klasser
- Uforutsigbart i produksjon
- Tilbøyelig for “overtilpasning”
- “Svart boks”

EGET (KATASTROFISKE) EKSEMPEL



dishwasher, dish washer, dishwashing machine (score = 0.51436)
wall clock (score = 0.09558)
spatula (score = 0.03166)
refrigerator, icebox (score = 0.02556)
cleaver, meat cleaver, chopper (score = 0.02061)

OPPGAVE: FINN EGET BRUKSOMRÅDE

1. Sett dere i grupper på 3
2. Finn et bruksområde der dyp læring kan brukes
 - a. Tekstforståelse
 - b. Bildegjenkjenning
3. Vektlegg følgende momenter:
 - a. Hva slags data trenger vi?
 - b. Hvordan skaffer vi slik data?
 - c. Drøft bruksområdet opp mot “utfordringer med dyp læring”

DEL 4

OPPSUMMERING



accenture

“SHORTLIST” FOR KUNSTIG INTELLIGENS

- Fokuser på problemet, ikke teknologi. Er problemet løsbart ved hjelp av KI?
- Har vi data for å løse problemet?
- Kan vi samle data for å løse problemet?
- Er dataene våre gode?
- Eksperimentering > bygging
- Få avklaringer tidlig: Hva må til for å produksjonssette en prototype?
 - Juridiske avklaringer
 - Integrasjon med eksisterende systemer
 - Sikkerhet
- Sørg for tverrfaglige team fra dag 1.

NYTTIGE RESSURSER

Introduksjon til Python: Udacity - Python foundation

<https://in.udacity.com/course/python-foundation-nanodegree--nd002-inpy>

Introduksjon til maskinlæring: Udacity – Intro to machine learning

<https://in.udacity.com/course/intro-to-machine-learning--ud120-india>

Rettferdighet innen maskinlæring: Google – Fairness in ML

<https://developers.google.com/machine-learning/crash-course/fairness/video-lecture>

Data science og ledelse: Coursera - Executive Data Science Specialization

<https://www.coursera.org/specializations/executive-data-science>

Ordbok for maskinlæring: Google - Machine Learning Glossary

<https://developers.google.com/machine-learning/glossary/>

BRENNER DU INNE MED SPØRSMÅL?

Runar.Gunnerud@Accenture.com