



Velkommen til maskinlæringskurs med

# netcompany

---

## FAKTA OM NETCOMPANY

---

- Veletablert tech- og innovasjonsmiljø
- Ønsker å bli Nord-Europas ledende aktør

**5 land**

Norge, Danmark, Polen,  
England og Vietnam

**+1700**

Antall ansatte globalt



# Hvem er Netcompany?

- Vi tar ansvar
- **Samfunnsnyttige** løsninger
- Tar imot mange nyutdannede
- Fokus på **samarbeid og team**





# Hvordan er det å jobbe som konsulent i Netcompany?

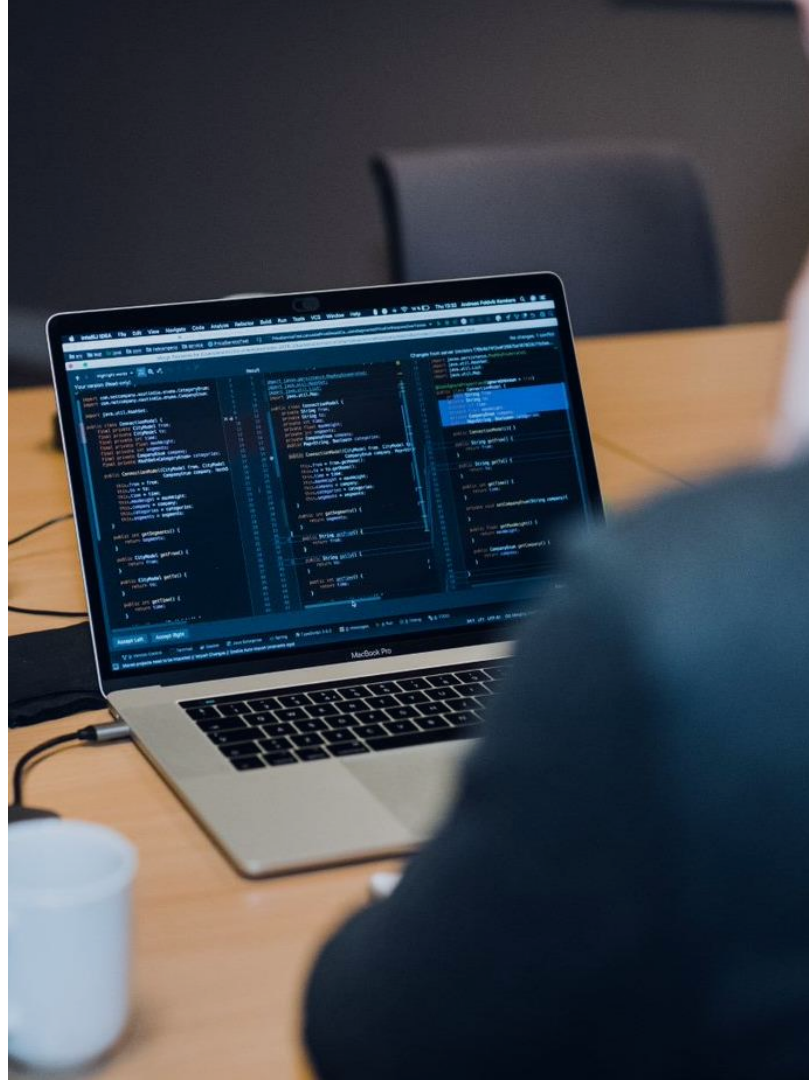
- Jobbe tett med kunden
- Skape de beste løsningene
- Vi tar ansvar for hele leveransen
- Kompetansepartner



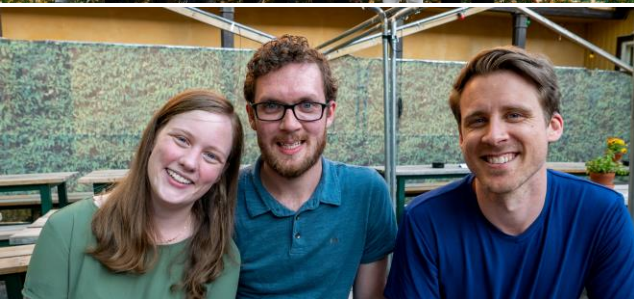
# Hvordan utvikler vi konsulentene våre?

Det som kjennetegner oss er at vi **brenner for å levere kvalitet** i arbeidet vi gjør, og i Netcompany vet vi at kvalitet henger sammen med **kompetanseutvikling**.

**DIN** utvikling er like viktig for oss som den er for deg!







# netcompany

---

Vi tar ansvar



Netcompany Norge



netcompanynorge



Netcompany

NTNU

# MASKINLÆRING

---

En introduksjon til maskinlæring

**Dato:** 24.01.2019

**Versjon:** 1.0

**Av:** Henrik Bøhler

**Kontakt:** [henrikb@netcompany.com](mailto:henrikb@netcompany.com)

**netcompany**



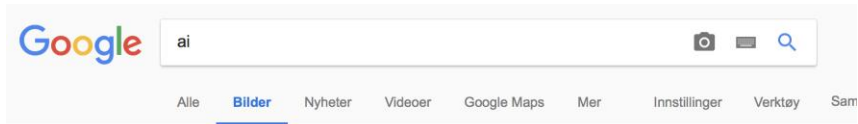
# Agenda

---

- Hva er AI og maskinlæring?
- Ulike typer maskinlæring og bruksområder
- Utviklingsprosessen
- Caser og eksempler

# HVA ER AI OG MASKINLÆRING?

---



What is AI? Everything you need to know about Ar...  
zdnet.com



7 tips for scaling your AI strategy | CIO  
cio.com



AI: Separating Artificial From Int...  
chiefexecutive.net



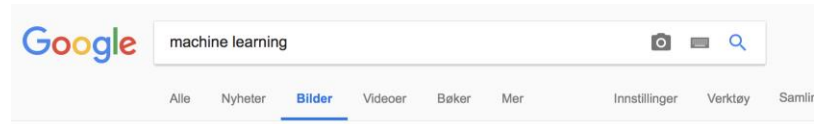
Advanced AI and machine learning help law enforcem...  
celebrite.com



3 Ways AI Can Make The Most Of Mob...  
cno.com



7 surprising companies wher...  
thenextweb.com



Machine Learning For Beginners – T...  
towardsdatascience.com



Machine Learning for dummies — exp...  
becominghuman.ai



5 questions to ask about machine learning – So...  
news.sophos.com



Machine Learning for Marketing...  
le.edu



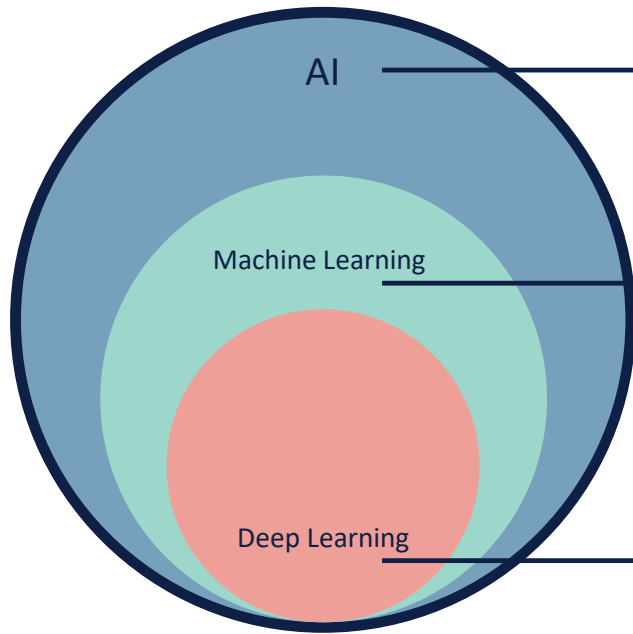
An Introduction to Machine Learning | DigitalOcean  
digitalocean.com



Machine Learning: Classification Model...  
medium.com



# AI, machine learning, deep learning...



- Kunstig intelligens – teknikker som tillater maskiner å etterligne menneskeoppførsel
- Maskinlæring - teknikker innen kunstig intelligens som gir maskiner mulighet til å lære uten at de er direkte programmert til å gjøre det
- Et fagområde innen maskinlæring

# Hva er da maskinlæring?

---

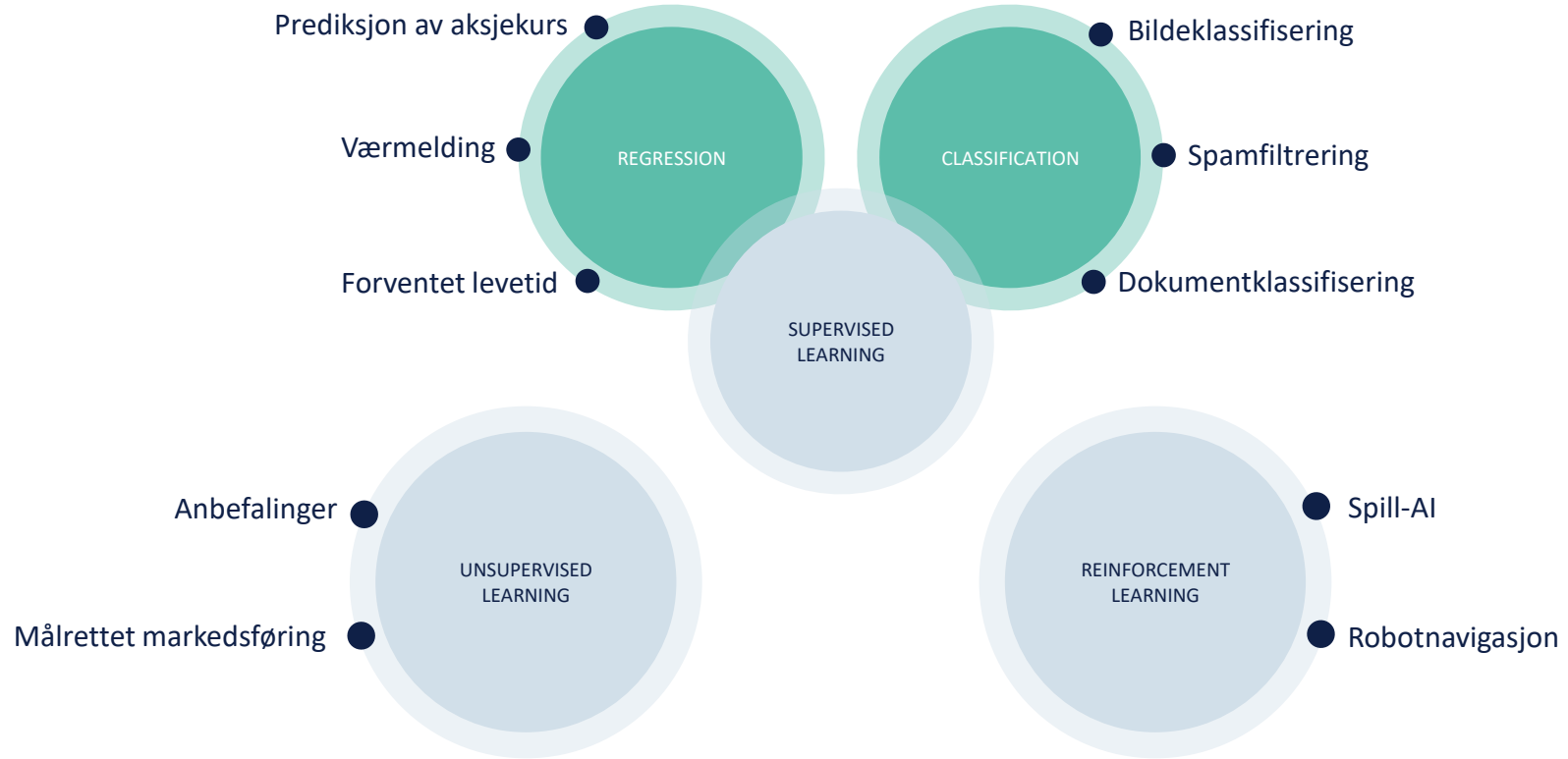
- Implementasjon av AI
- Maskiner som lærer fra data
  - Ikke gitt instruksjoner
  - Ønsker å si noe om det ukjente basert på det kjente



# ULIKE TYPER MASKINLÆRING

—  
...og hva de kan brukes til





# Maskinlæring

---

**Supervised learning:** Hvert datapunkt i datasettet har en merkelapp, dvs. et riktig svar. Målet er å finne en generell regel som tilordner rett merkelapp til gitt input.

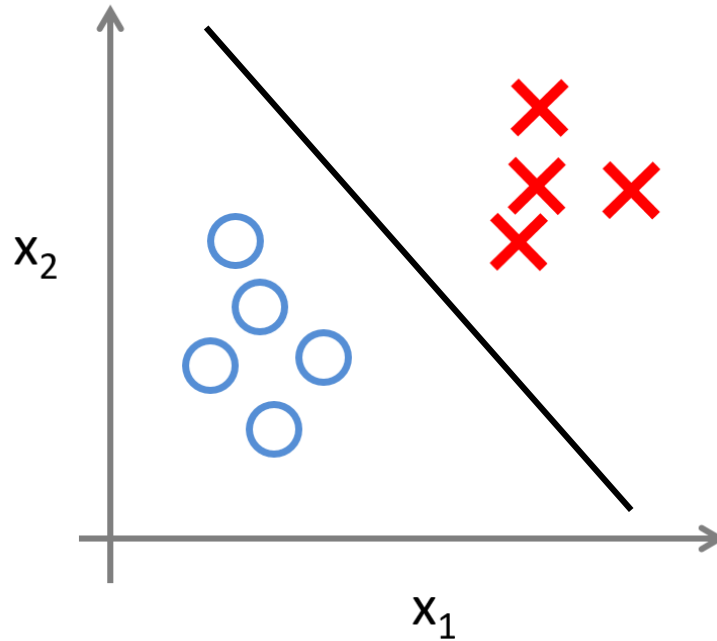
**Unsupervised learning:** Datasettene som brukes inneholder ikke en fasit. Målet er å finne en underliggende struktur til dataene som kan utnyttes.

# Spamfilter



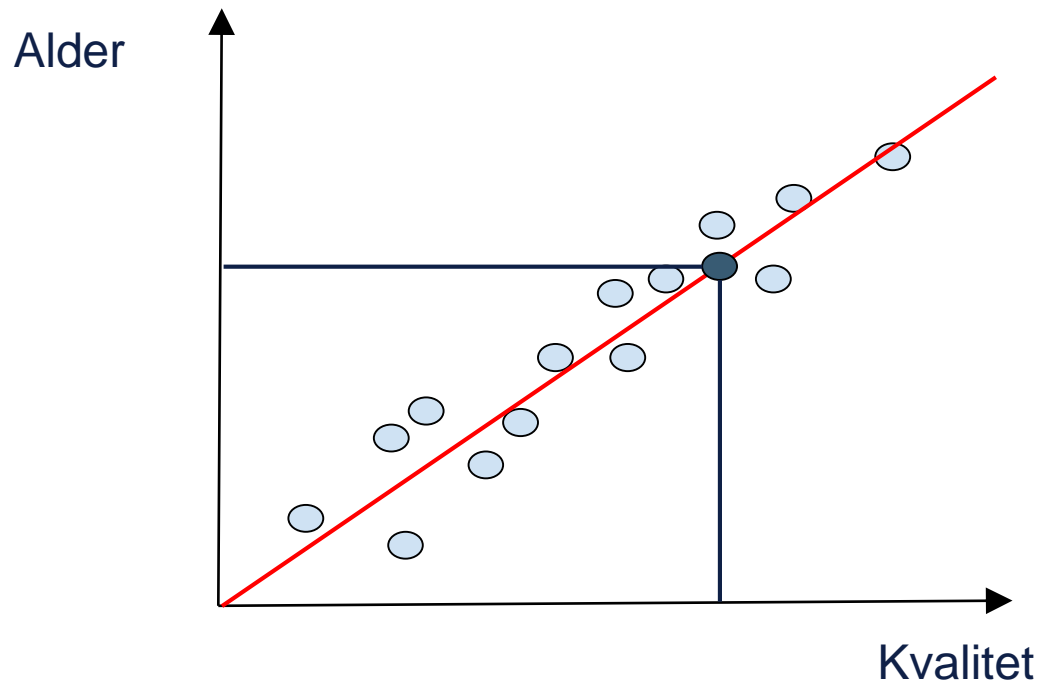
# Supervised - klassifisering

---



Hva er forholdet mellom alder og kvalitet?

# Supervised - regresjon

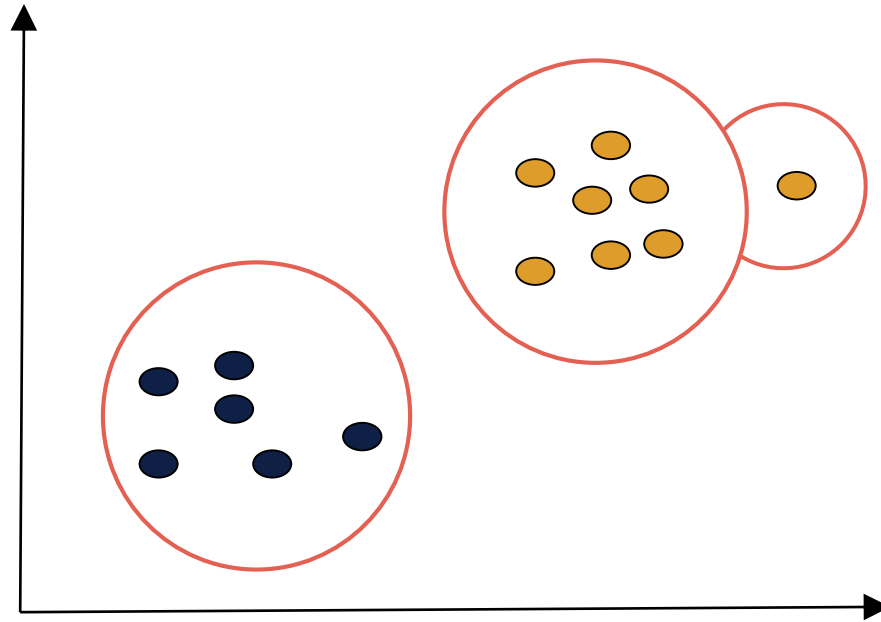


# Gruppere brukere av Spotify



# Unsupervised - clustering


---



# HVA ER FELLES FOR DE ULIKE TYPENE?

# DATA

netcompany

Blue Gene  supercomputer

**«DATA IS THE KEY  
INGREDIENT»**

**«GARBAGE IN,  
GARBAGE OUT»**

**«PAST PERFORMANCE  
IS NO GUARANTEE OF  
FUTURE RESULTS»**



# MASKINLÆRING I PRAKSIS

---

# Maskinlæring i praksis

---

- **Identifiser, klassifiser og forstå problemet**
- Innhent og analyser data
- Prosesser datasettet
- Modeller problemet
- Tren og valider
- Sett i drift



Preprosessering



Utvikling



Deployment

# Maskinlæring i praksis

---

- Identifiser, klassifiser og forstå problemet
- **Innhent og analyser data**
- Prosesser datasettet
- Modeller problemet
- Tren og valider
- Sett i drift



Preprosessering



Utvikling



Deployment

# Feature selection

- Kvadratmeter
- Byggeår
- Beliggenhet
- Hagegnomer
- Antall takstein

Feature 1 (m^2)	Feature 2 (byggeår)	Feature 3 (beliggenhet)	Resultat (over 3M)
100	1960	Moholt	False
120	2001	Singsaker	True
90	1970	Lade	False

# Maskinlæring i praksis

---

- Identifiser, klassifiser og forstå problemet
- Innhent og analyser data
- **Prosesser datasettet**
- Modeller problemet
- Tren og valider
- Sett i drift



Preprosessering



Utvikling



Deployment

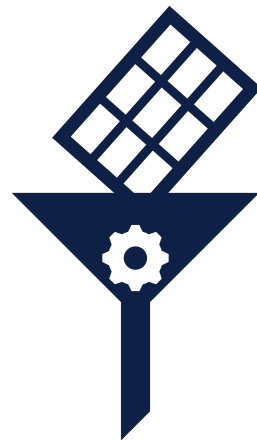
**DET MEST  
TIDKREVENDE  
ARBEIDET ER Å  
BEHANDLE (STORE  
MENGDER) DATA.**



# Preprosessering

---

- **Formaterer**
  - Er datasettet i ønsket format? (matrise/tabell)
- **Rense**
  - Mangler det features i enkelte rader?
- **Skalere**
  - Har de forskjellige featurersene ulik skala?
- **Balansere**
  - Er det for få eksempler av én klasse?



# Maskinlæring i praksis

- Identifiser, klassifiser og forstå problemet
- Innhent og analyser data
- Prosesser datasettet
- **Modeller problemet**
- Tren og valider
- Sett i drift



Preprosessering



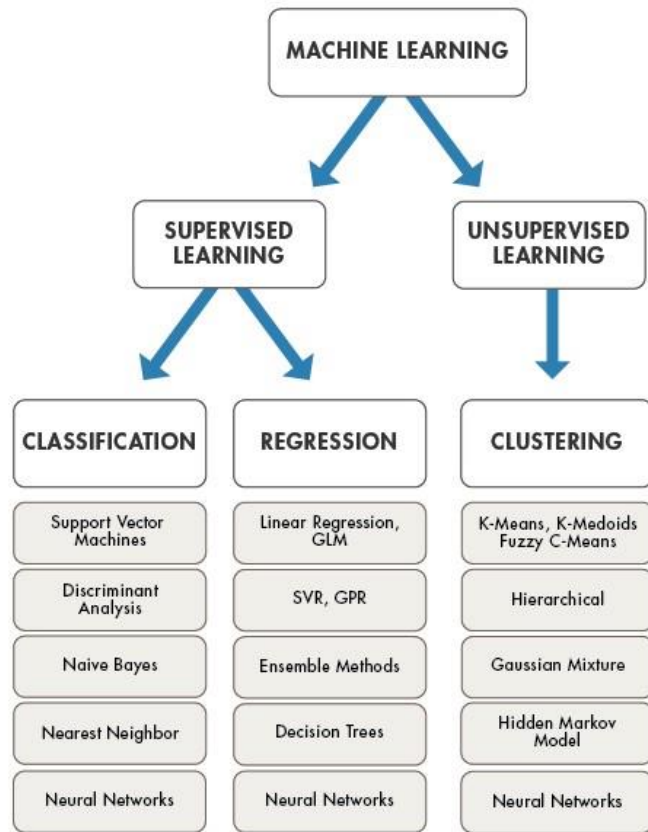
Utvikling



Deployment

# Modeller problemet

- Ha gjerne en overordnet forståelse for ML-algoritmer
- Velg en algoritme som passer problemet
- Prøv gjerne flere
- Tune gjennom parametre



# Maskinlæring i praksis

---

- Identifiser, klassifiser og forstå problemet
- Innhent og analyser data
- Prosesser datasettet
- Modeller problemet
- **Tren og valider**
- Sett i drift



Preprosessering



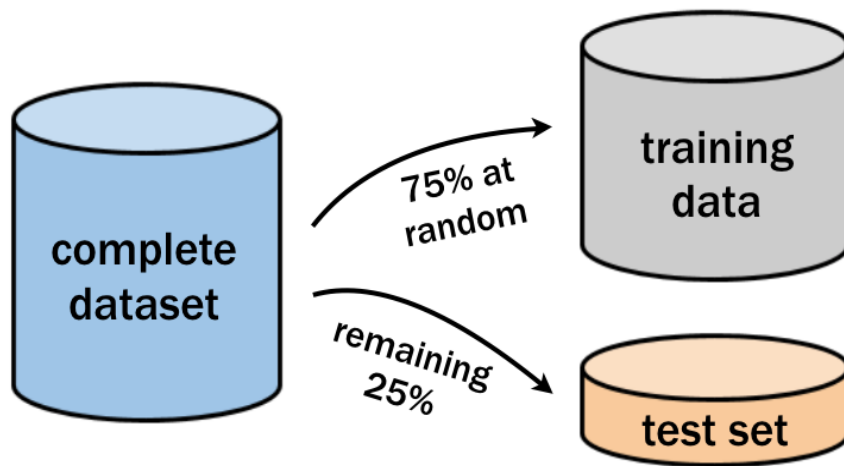
Utvikling



Deployment

# Trening og test

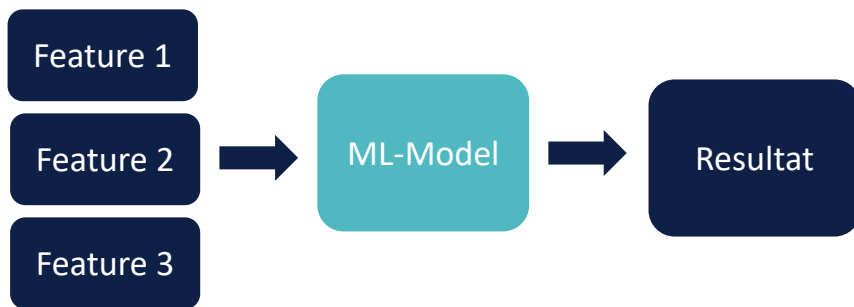
---



# Trening

---

- Resultat av trening – ML modell

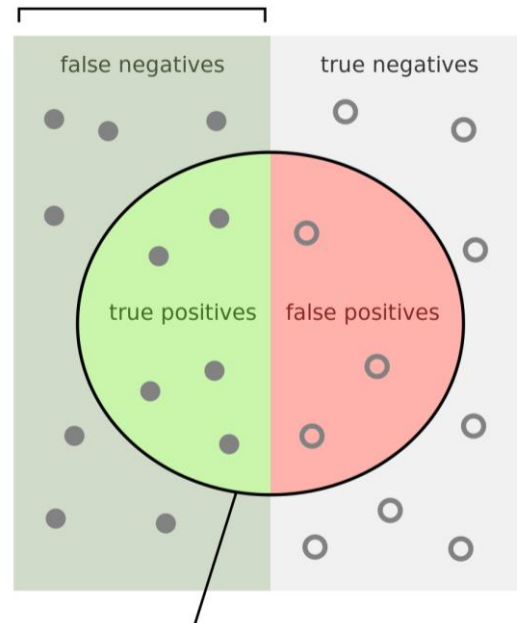




# Test

- Precision:  $\frac{\text{true positives}}{\text{true positives} + \text{false positives}}$ 
  - antall relevante elementer blant alle elementer som var valgt
- Recall:  $\frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$ 
  - antall relevante elementer som var valgt blant alle de relevante dokumentene

Relevante elementer



Valgte elementer

# Confusion Matrix

		Predikerte klasser	
		JA	NEI
Faktiske klasser	JA	True Positives	False Positives
	NEI	False Negatives	True Negatives

# Maskinlæring i praksis

---

- Identifiser, klassifiser og forstå problemet
- Innhent og analyser data
- Prosesser datasettet
- Modeller problemet
- Tren og valider
- **Sett i drift**



Preprosessering



Utvikling



Deployment

# EKSEMPEL

—

Klassifisering av Iris-blomster



# Hvordan hente ut koden

---

- Logg inn i Google
- Gå inn på <https://tinyurl.com/NCML101> og velg github i toppmenyen.
- Skriv inn *netcompanyno* og trykk Enter for å søke
- Velg ønsket notebook
  - Iris-eksempelet hvis du vil
  - Titanic, fin begynnelse
  - Games
  - Used cars
- Klikk på "Copy to drive" øverst til venstre

# VI TAR ANSVAR

[www.netcompany.com](http://www.netcompany.com)

Henrik Bøhler  
[henrikb@netcompany.com](mailto:henrikb@netcompany.com)

**netcompany**

© Copyright 2019 Netcompany. All rights reserved.

Electronic, mechanical, photographic or other reproduction, translation or copying of this document or parts thereof, for whatever purpose, is not permitted without the prior written permission of Netcompany