

Maskininläarning och Deep learning

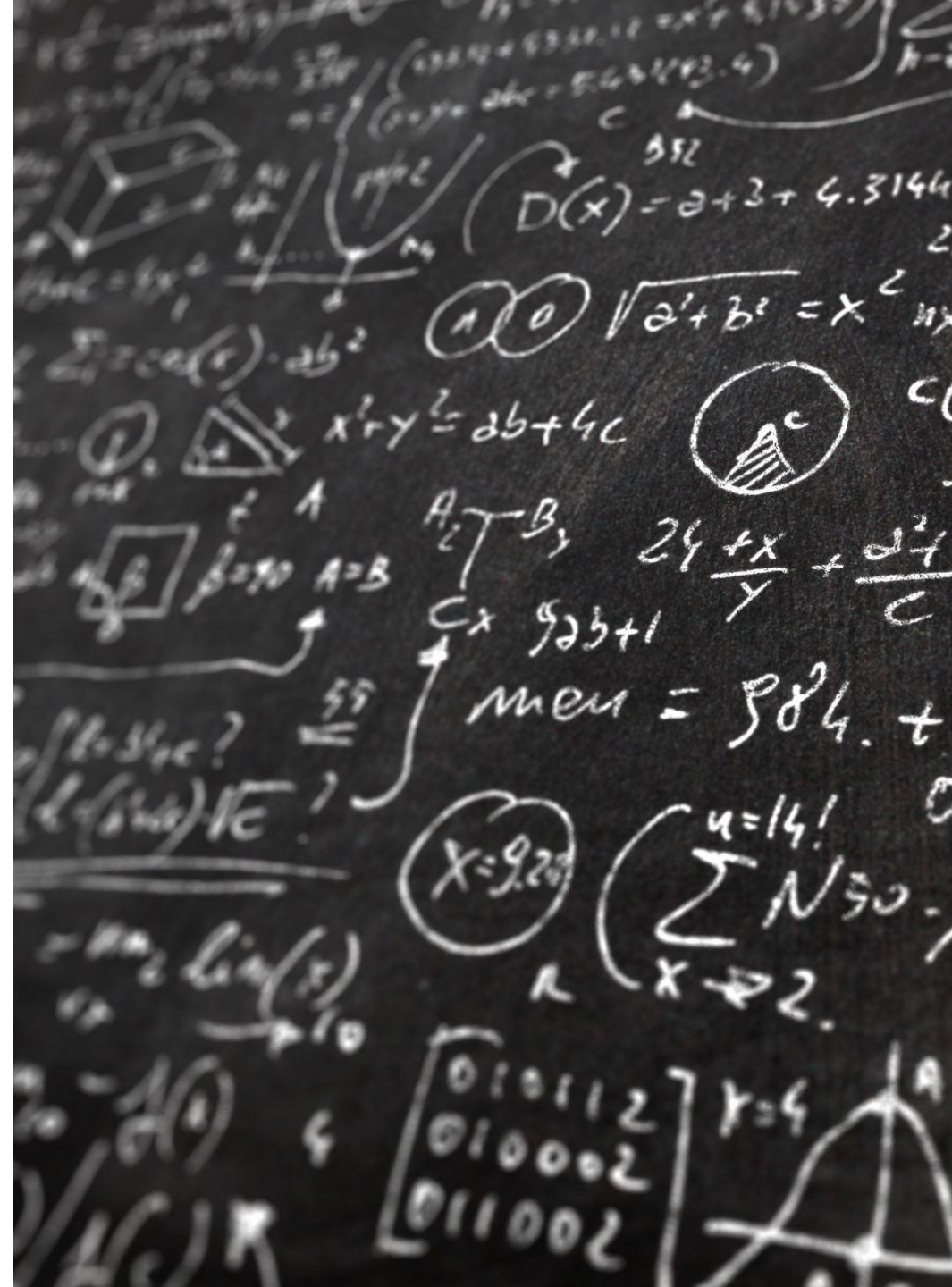
Kursintroduktion

Kursupplägg

Stora delar av kursen kommer vara "fliped classroom". Den kommer bestå mycket av videoklipp, diskussioner och quiz. Det innebär också att det inte kommer spelas in lika mycket från teams.

Under kursen kommer olika algoritmer behandlas varje vecka och det kommer vara en inlämning på varje algoritm (nästan).

Varje inlämning kommer ha en G och en VG nivå.



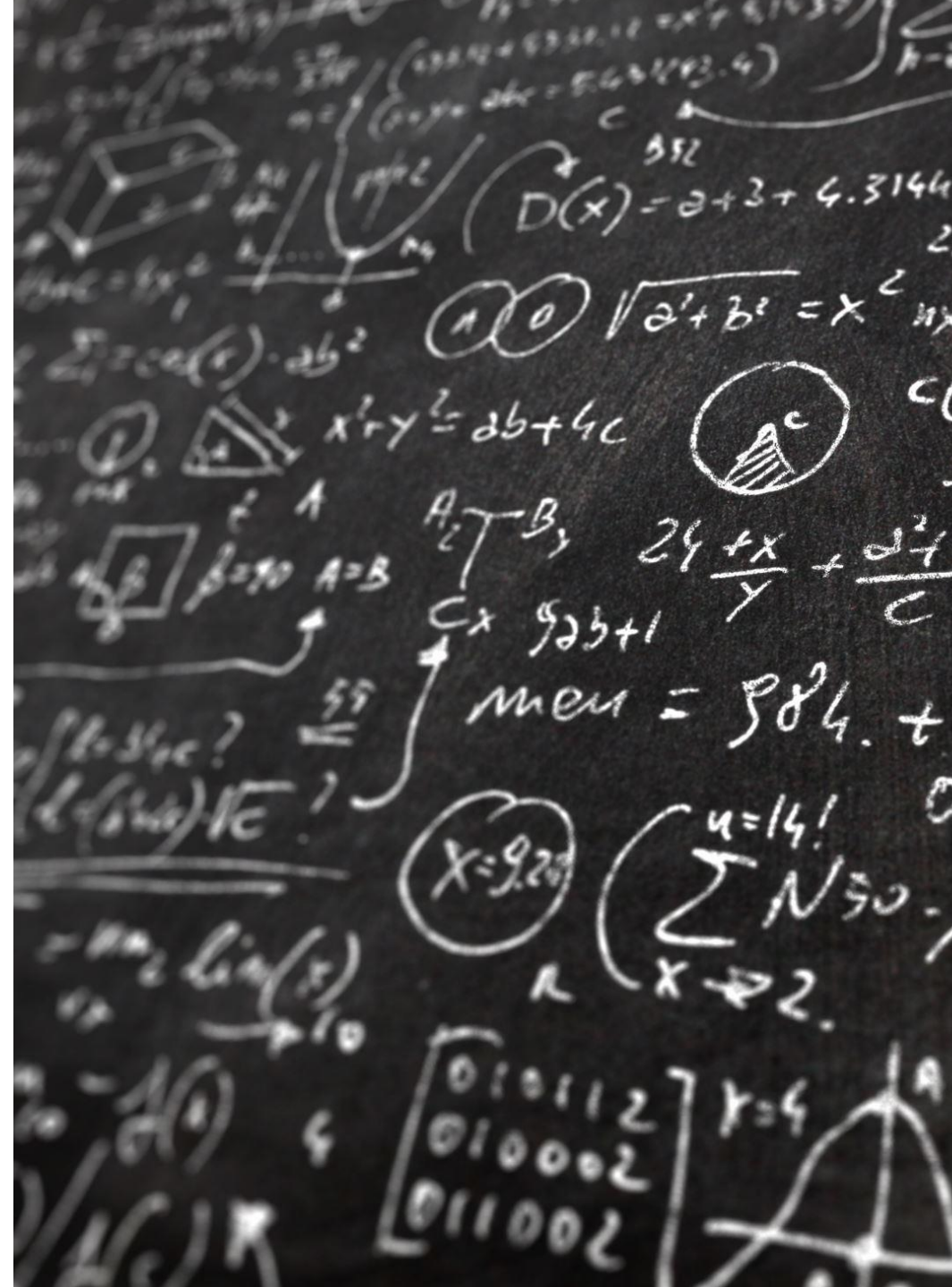
Kursupplägg

För att få godkänt på kursen ska man få minst Godkänt på alla moment, 8 stycken totalt.

För att få Väl Godkänt krävs det Väl Godkänt på 4 av 8 inlämningar.

Väl Godkänt kan bara fås på första inlämningstillfället för varje uppgift och kan ej kompletteras upp.

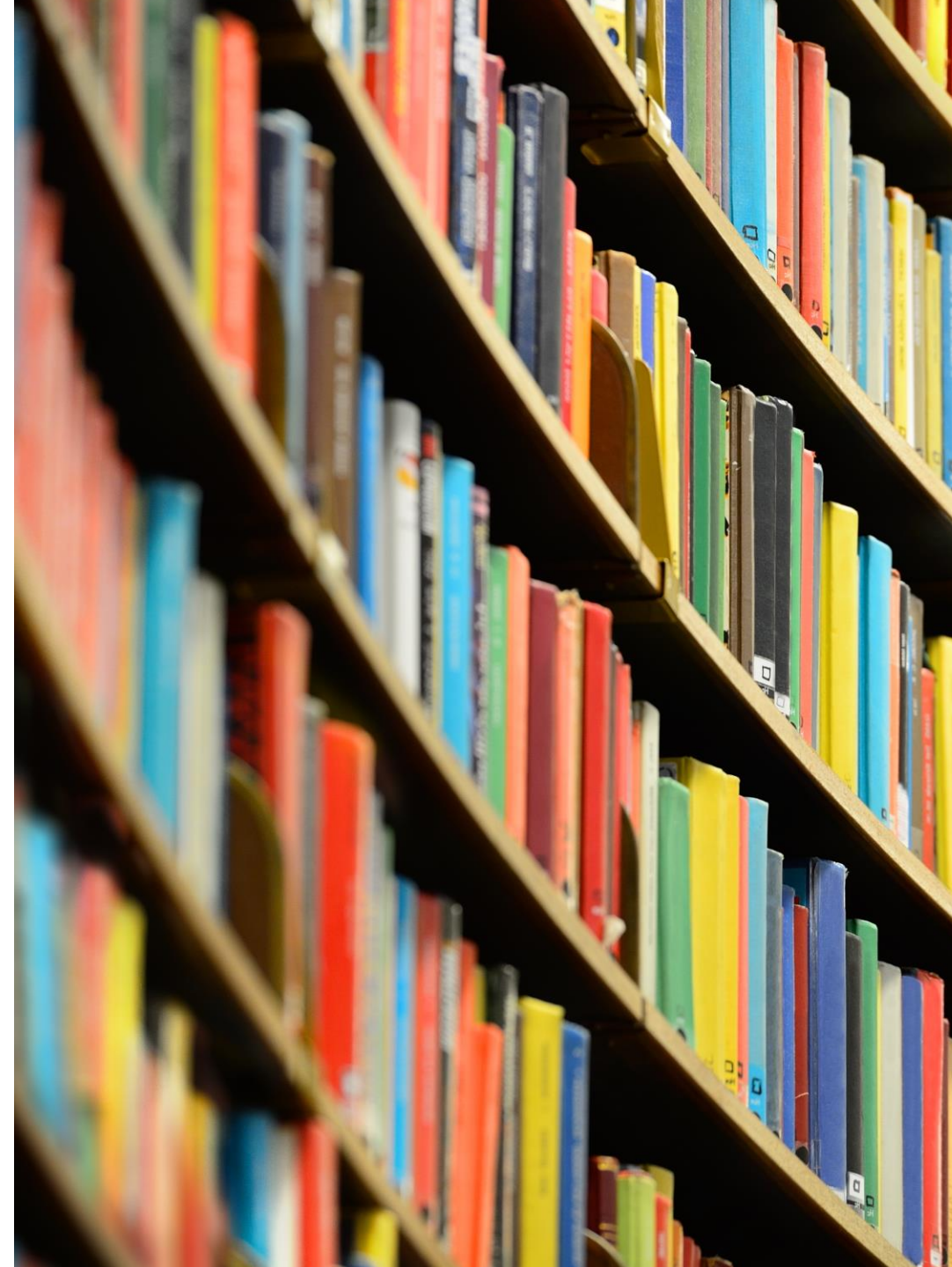
Alla inlämningar kommer kunna kompletteras till slutet av kursen.



Kursmaterial

Kursen har ingen bok utan vi kommer utgå från videoklipp och artiklar online.

Är det någon som absolut vill ha en bok, meddela mig så kan vi kolla på alternativ.



Schema

Lärrarledda dagar två dagar i veckan,
08:00-16:00.

Vilka dagar skiljer sig mycket mellan
veckorna (bland annat på grund av röda
dagar). Men för de flesta veckorna är
det Tisdag och Torsdag.

Resterande dagar kan jag nås på mail
eller teams (har bättre koll på teams)



Schema

Idag

- ML/AI genom historien
- Era tankar om maskininlärning
- Träning och testning
- Linjär regression



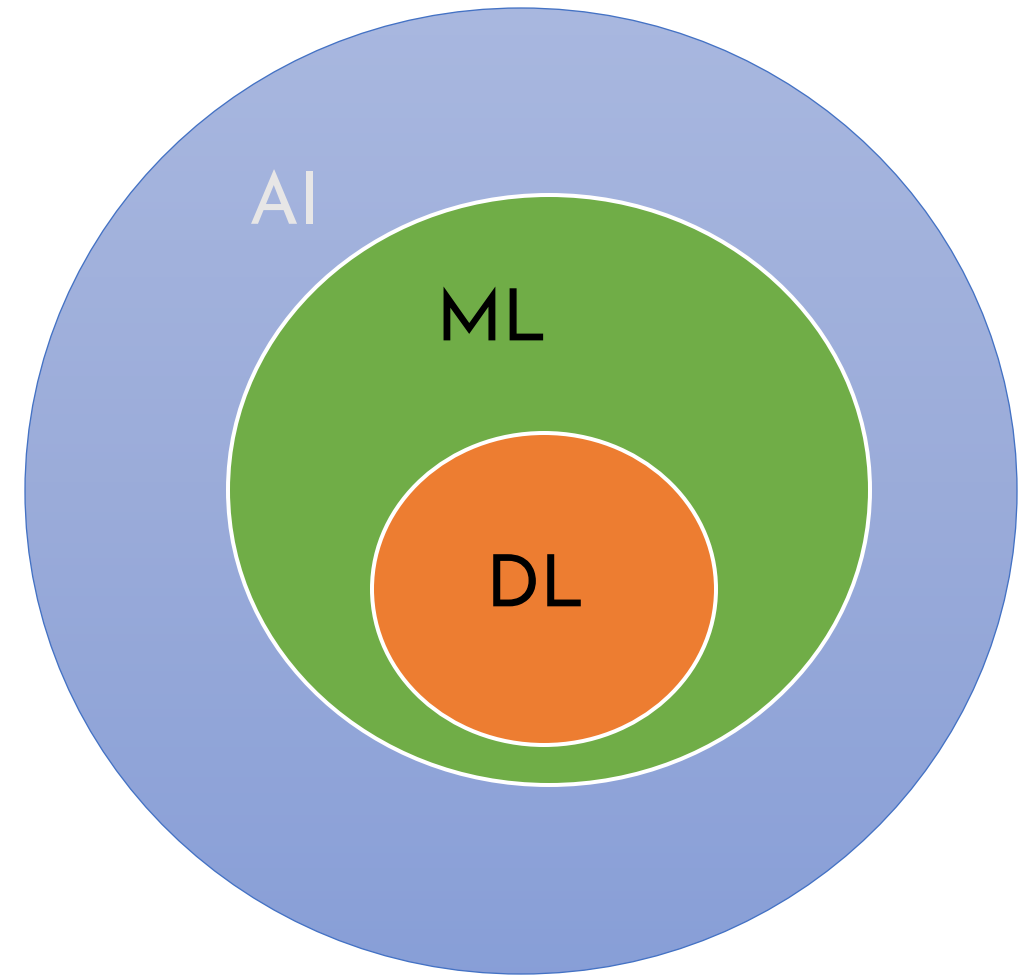
Schema

- Logistisk regression
- Decision tree
- Random Forest
- State vector machines
- K Nearest Neighbour
- Klustering
- Tidsserier
- XGBoost
- Neurala nätverk (i olika format)



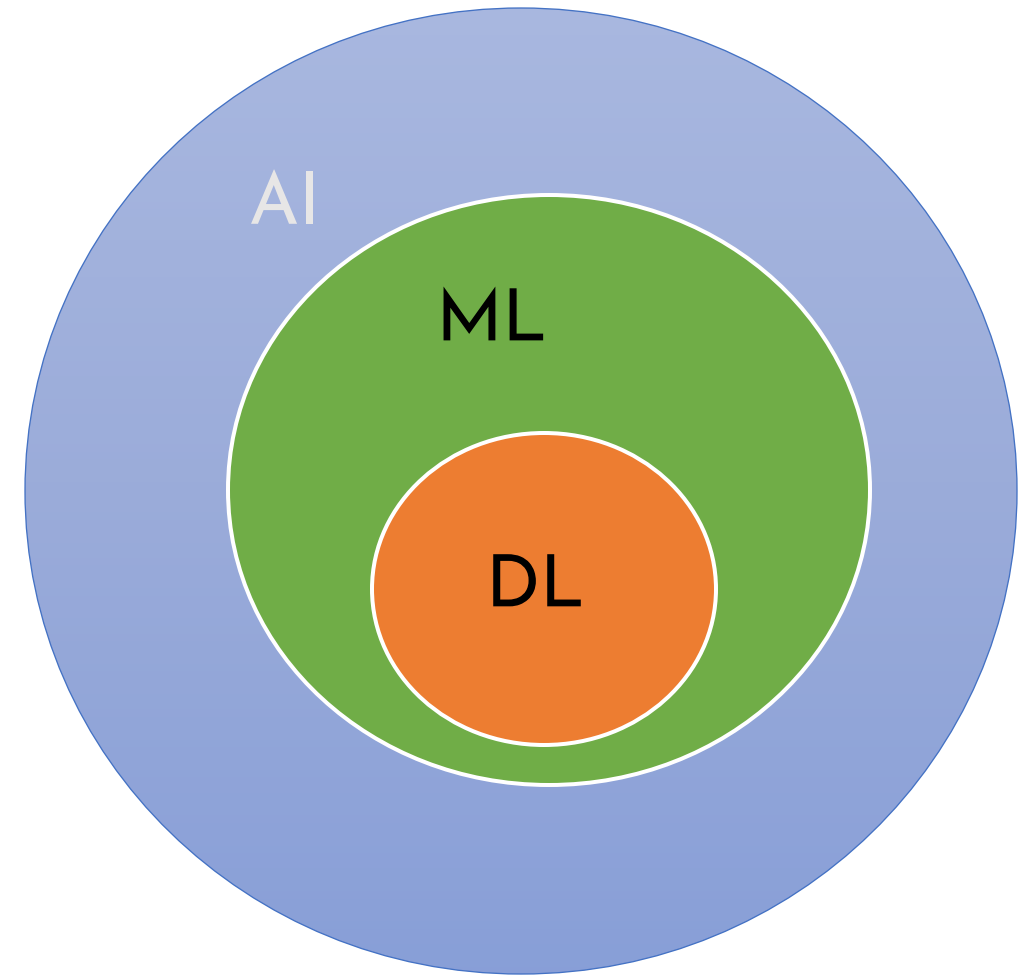
Maskininlärning

- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Reinforcement learning

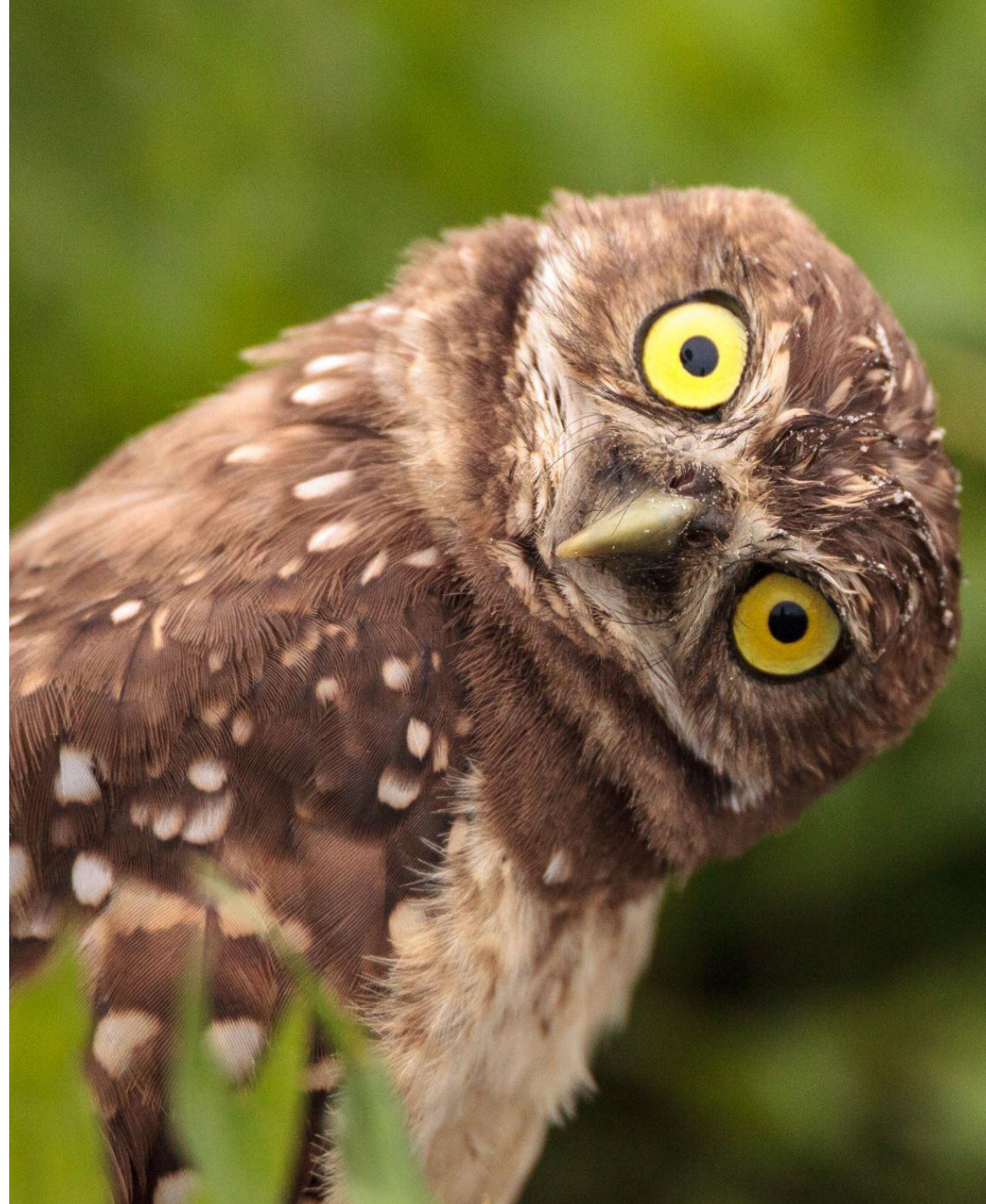


Maskininlärning

- Klassificering
- Regression
- Klustering



Frågor?



Maskininlärning och Deep learning

Maskininlärning genom historien

Maskininlärning genom historien

1959

Arthur Samuel
introducerade
termen Machine Learning

Maskininlärning genom historien

40-talet

McCulloch och Pitt
skapade mycket av
grunden för det som nu
är AI

1959

Arthur Samuel
introducerade
termen Machine Learning

Maskininlärning genom historien

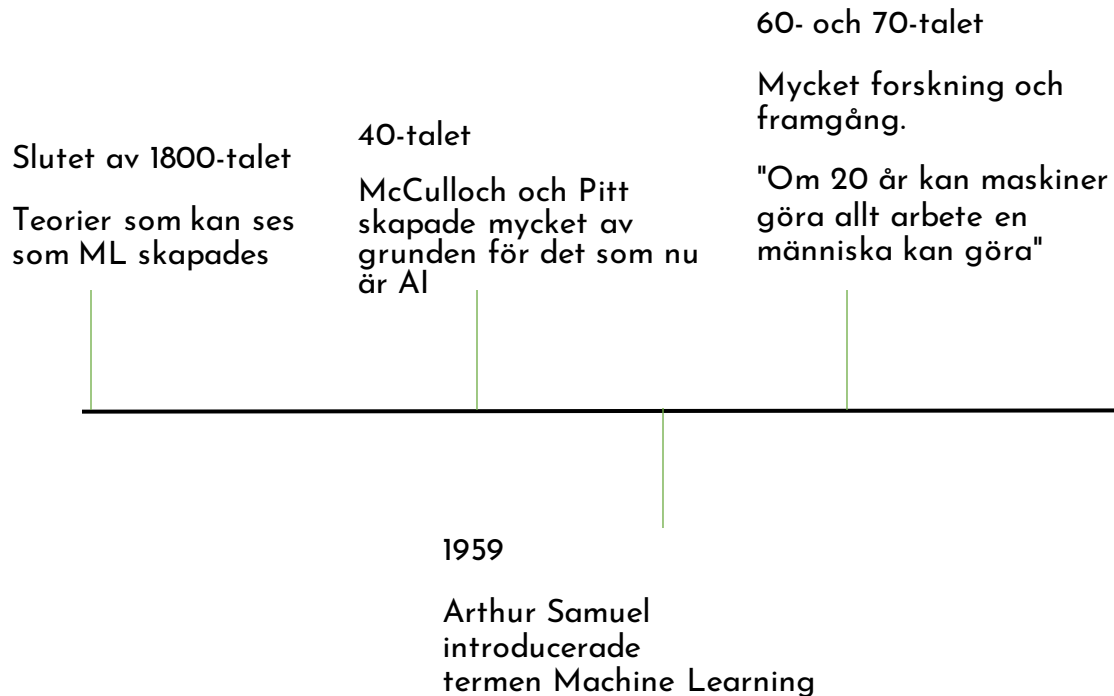
Slutet av 1800-talet
Teorier som kan ses
som ML skapades

40-talet
McCulloch och Pitt
skapade mycket av
grunden för det som nu
är AI

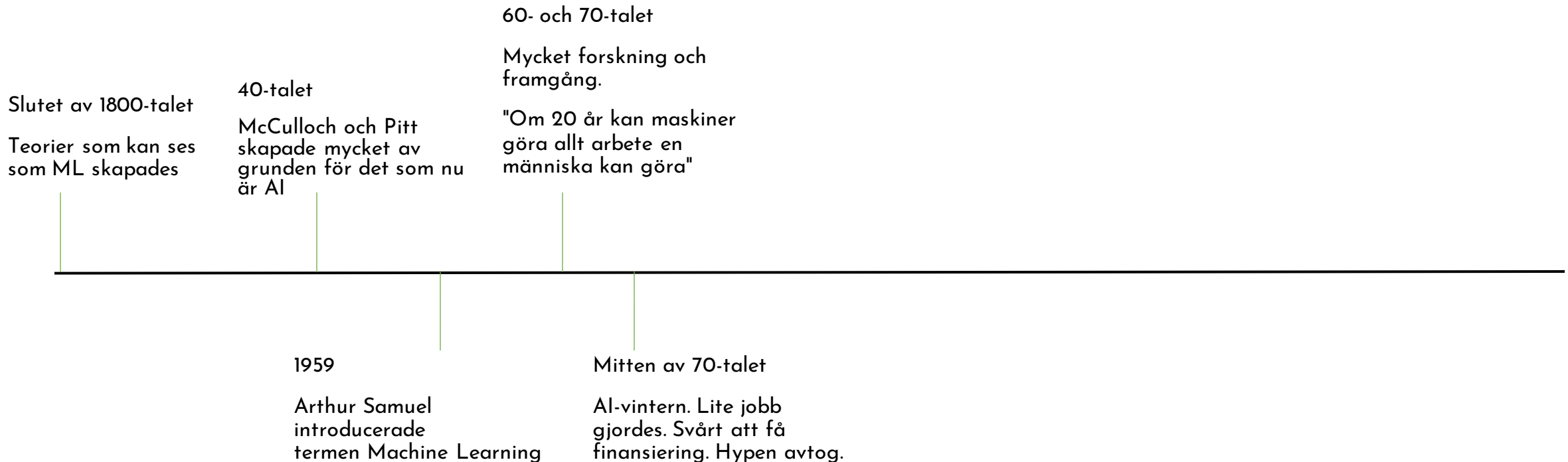
1959

Arthur Samuel
introducerade
termen Machine Learning

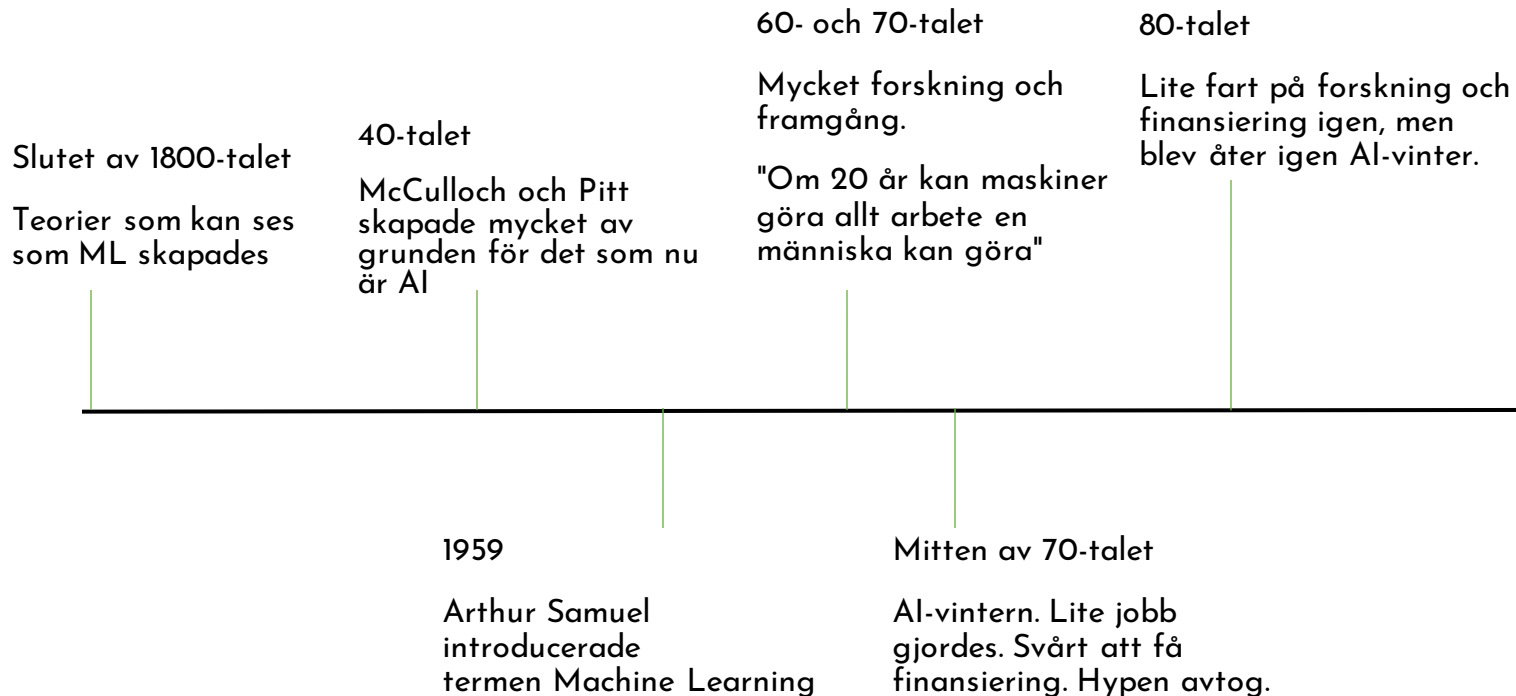
Maskininlärning genom historien



Maskininlärning genom historien



Maskininlärning genom historien



Maskininlärning genom historien



Maskininlärning genom historien



Maskininlärning genom historien



Maskininlärning genom historien



Maskininlärning genom historien



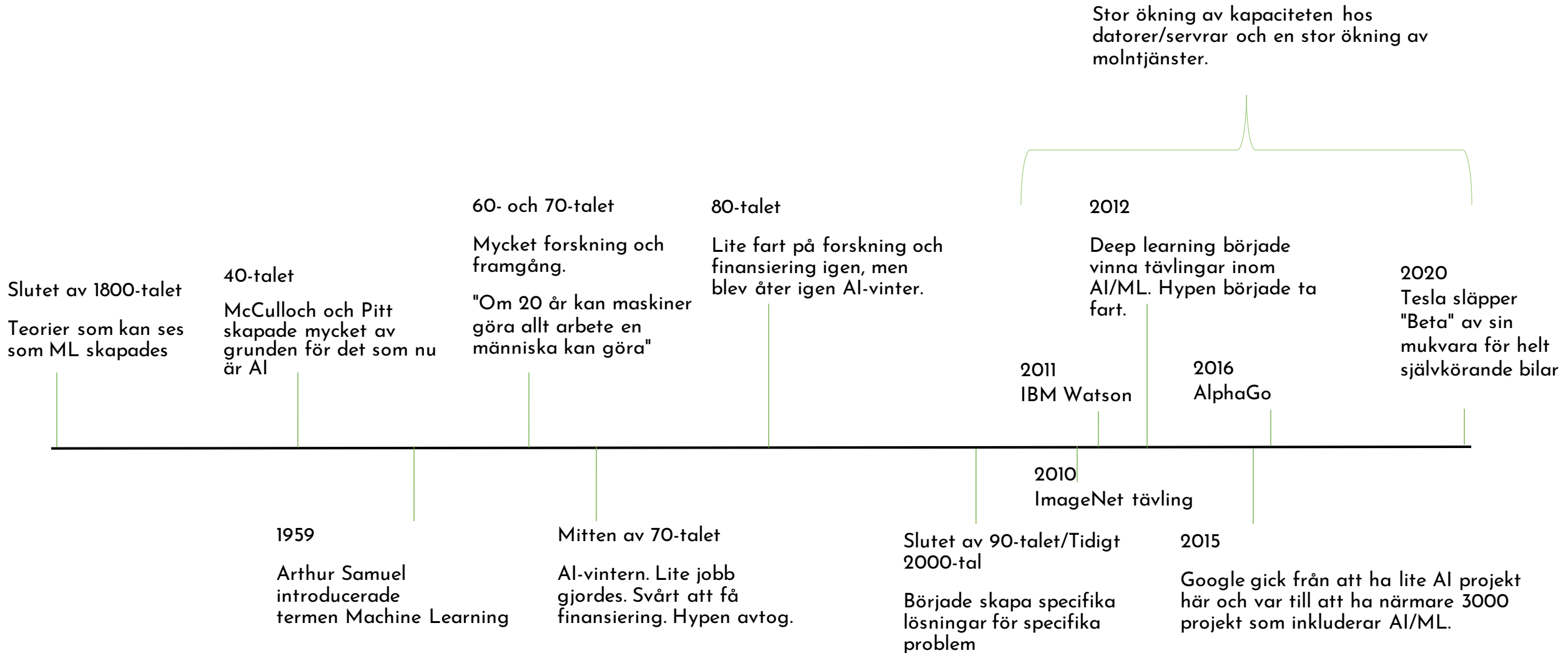
Maskininlärning genom historien



Maskininlärning genom historien



Maskininlärning genom historien



Diskussion

- Varför är du intresserad av maskininläring?
- Kom på 5 saker ni använder som utnyttjar maskininläring.
- Vad skulle ni vilja jobba med inom maskininläring?
 - Sektor (sjukvård, bilindustrin, energisystem, etc)
 - Teknik (bildanalys, textanalys, etc)

Maskininlärning och Deep learning

Träning och testning

Träning och testning

När vi jobbar med data där vi har ett svar (supervised learning) behöver vi först träna vår modell och sedan testa den.

Detta gör vi genom att dela upp datan i två set, träningset och testset.

Träning

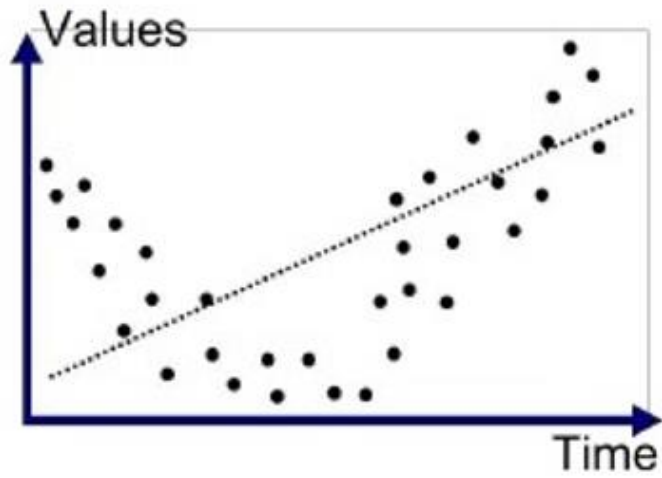
Under träningsdelen skapar vi en modell som generaliseras väl till okänd data (data modellen inte sett tidigare).

Om modellen lär sig träningsdatan för bra kallas det "overfitting".

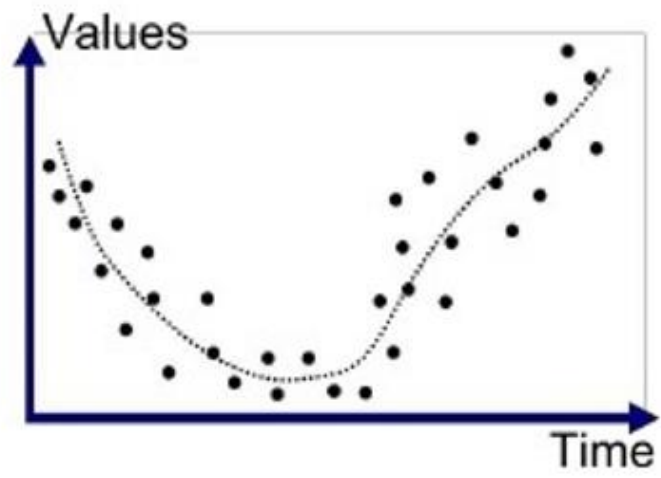
Modellen fungerar då väldigt bra på träningsdata, men den kan inte hantera okänd data.



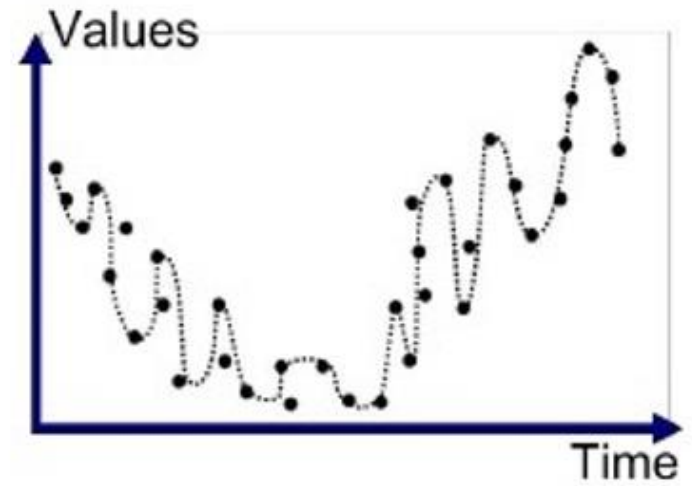
Träning



Underfitted



Good Fit/Robust

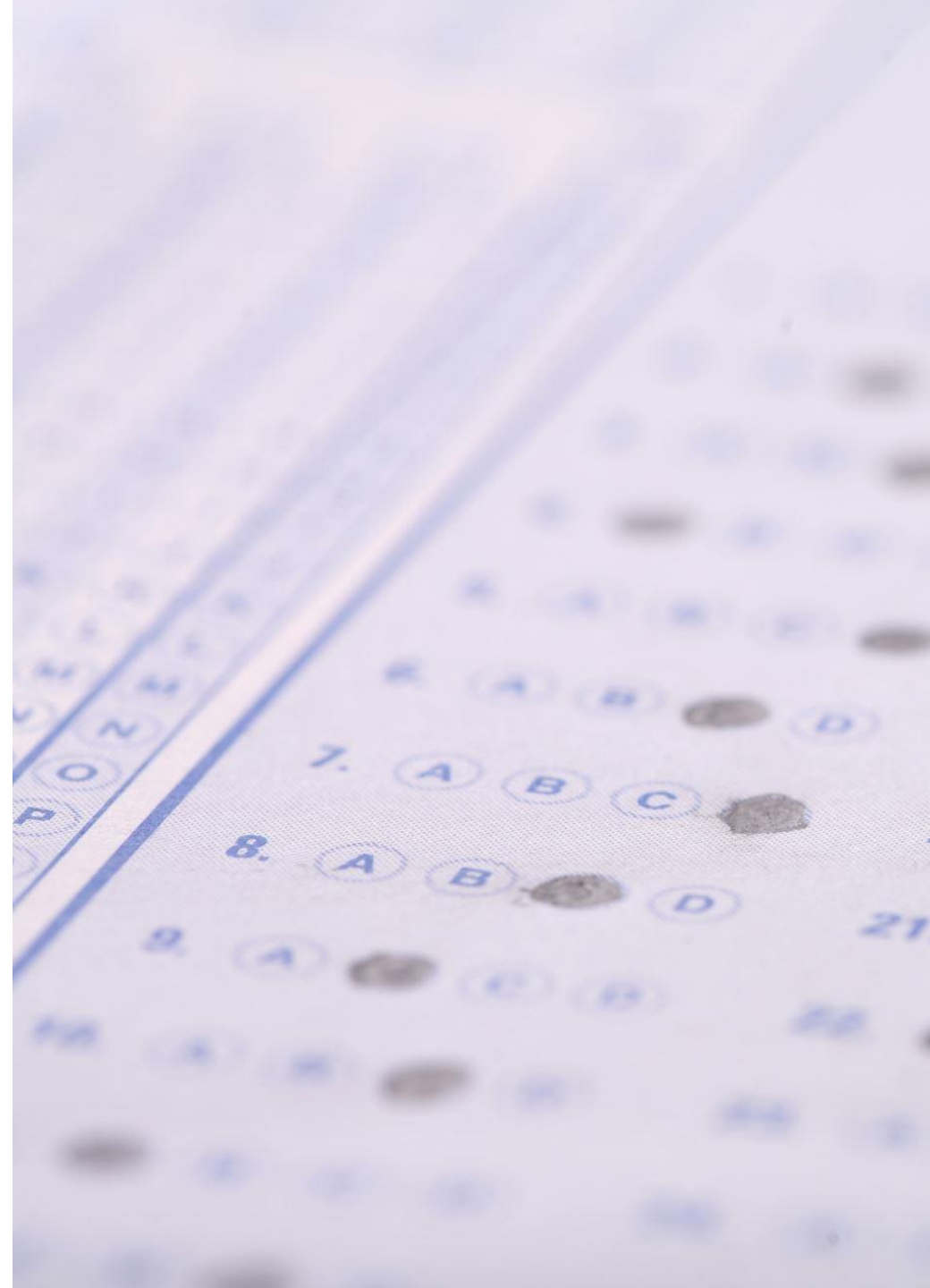


Overfitted

Testning

När vi vill test hur bra en modell faktiskt presterat använder vi oss av ett testset.

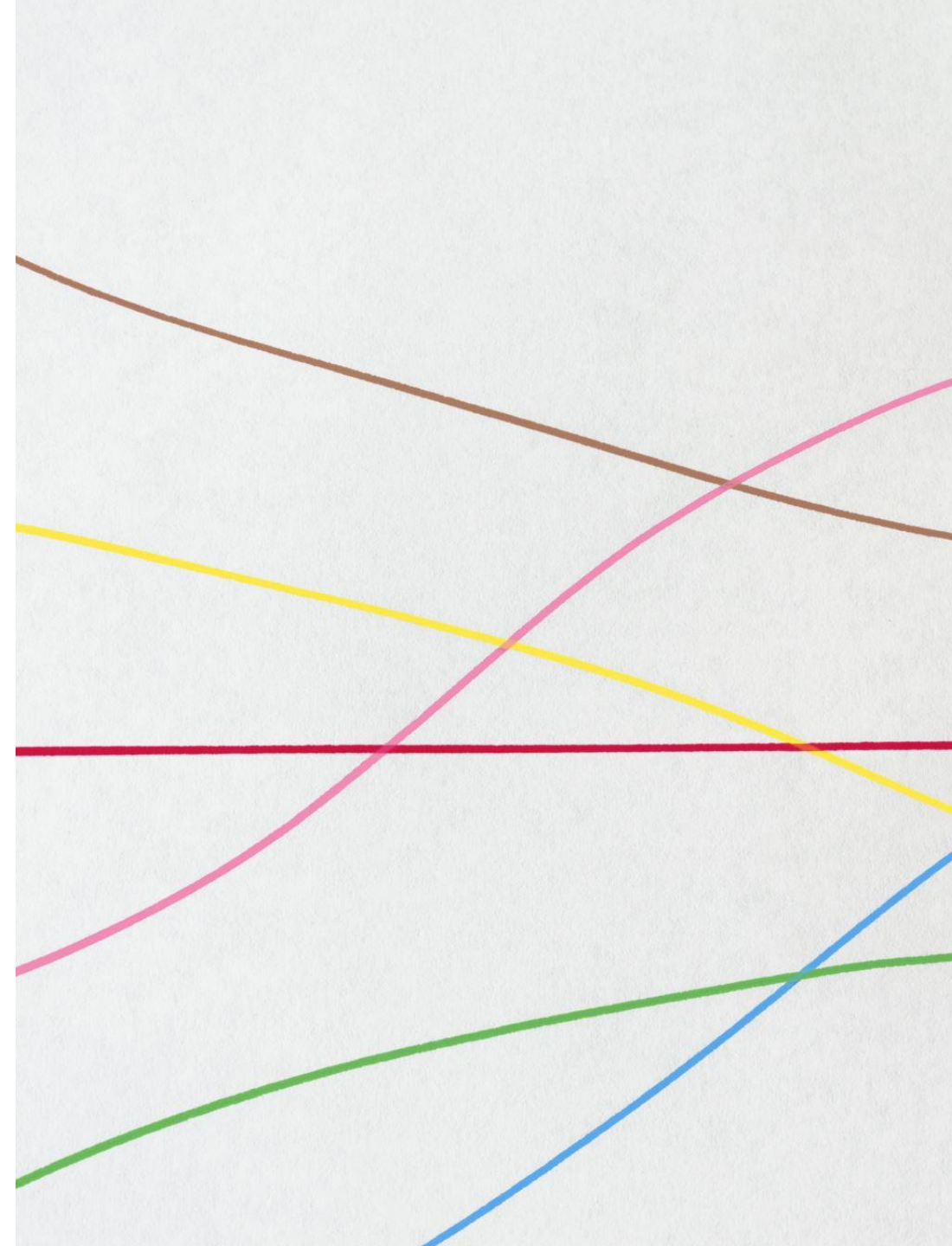
Testsetet innehåller data som modellen inte sett tidigare, men datan har samma distrubution (nästan) som träningsdatan. Har vi hälften katter och hälften hundar i vår träningsdata kan vi inte bara ha hunder i vår testdata.



Valideringsset

Valideringsdata kan användas för att vrida på parametrar i modeller, men långt ifrån alltid som man skapar ett valideringsdataset.

Valideringsdata blandas ofta ihop med testdata.

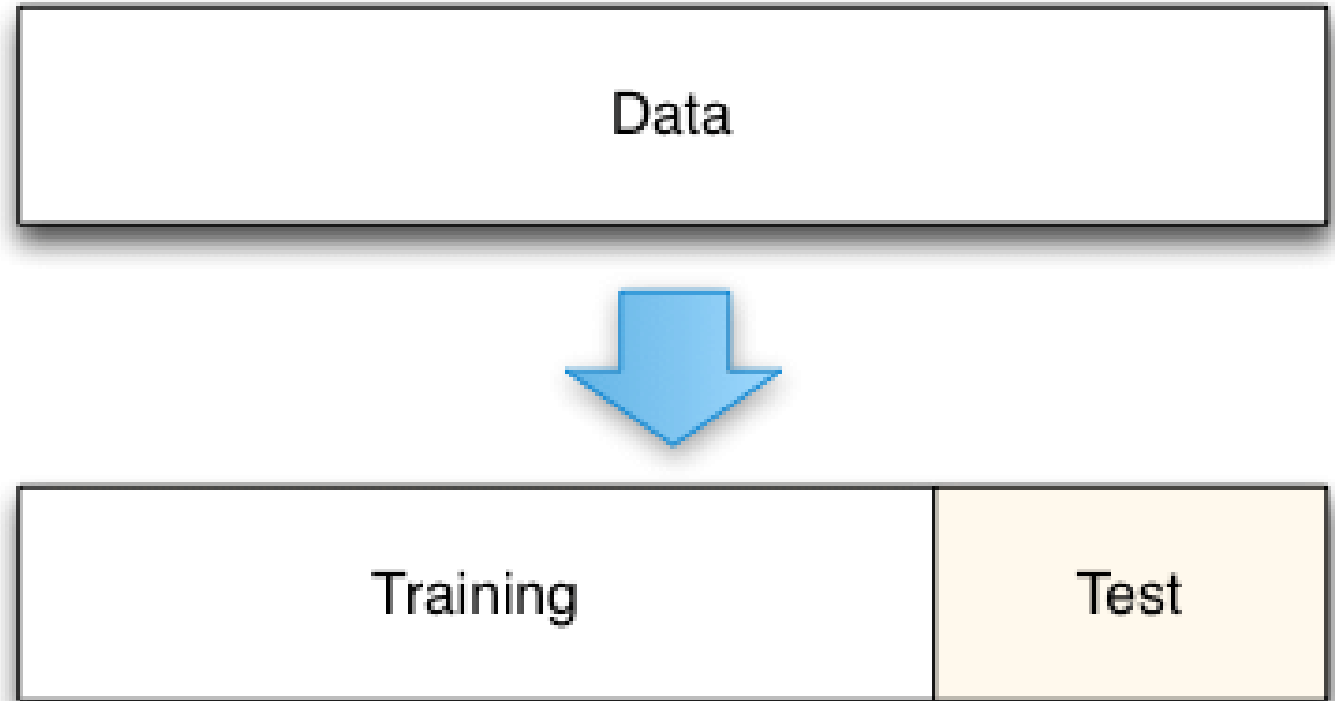


Hur ska vi dela upp datan?

Hur vi ska dela upp är olika från fall till fall och beror ofta på hur mycket data vi har.

Två vanliga uppdelningar är

- 70% träning, 30% test
- 80% träning, 20% test



Vad ska vi använda för mätvärden?

Hur prestandan på modellen skall testat skiljer sig från fall till fall.

- Procent rätt/fel (klassifiering)
 - Vilken typ av fel? (false negative, false positive)
- Hur mycket fel? (regression)
 - Straffa stora fel
- Precision
- Recall
- ROC & AUC
- Logloss
- ...

False negative, False positive

	Test tells you don't have it	Test tells you have it
You don't have it	True negative	False positive
You got it	False negative	True positive