

# TEK5020/9020 Mønster-gjenkjenning Høsten 2023

Forelesning 13 – Ikke-ledet læring

Idar Dyrdal (idar.dyrdal@its.uio.no)

UiO : Institutt for teknologisystemer

3. november 2023

# Innhold i kurset

- Introduksjon til mønstergjenkjenning
- Beslutningsteori (desisjonsteori)
- Parametriske metoder
- Ikke-parametriske metoder
- Lineære og generaliserte diskriminantfunksjoner
- Evaluering av klassifikatorer
- Ikke-ledet læring
- Klyngeanalyse.

# Ikke-ledet læring

Denne og neste forelesningen tar for seg metoder som kan brukes når det man har av treningssampler ikke er merket med klassetilhørighet.

Kort fortalt:

- Ikke-ledet læring går ut på trene klassifikatorer uten slik merking av treningssamplene. I noen tilfeller lar det seg faktisk gjøre å estimere tetthetsfunksjonene til de enkelte klassene.
- Derved kan man konstruere en klassifikator, f.eks. ut fra minimum feilrateprinsippet.
- Det beslektede temaet *klyngeanalyse* (neste forelesning) dreier seg hovedsakelig om å kartlegge strukturen i et ukjent datasett.
- Målet er her å dele datasettet inn i et antall naturlige klynger ut fra innbyrdes likhet (similaritet) eller nærhet i egenskapsrommet.

Tema for denne forelesningen er ikke-ledet læring.

# Ikke-ledet læring

Trening av klassifikator vha. treningssett uten merkede sampler (dvs. ukjent klasse-tilhørighet). Behov for slike metoder bl.a. når:

- Merkede sampler ikke er tilgjengelig,
- det er for kostbart å merke sampler,
- når statistikken i problemet endres over tid, dvs. klassifikatoren må kunne oppdateres med nye data.

Maksimum-likelihood metoden kan her brukes til å estimere parametrene i en såkalt *blandingstetthet*

$$p(\mathbf{x}|\boldsymbol{\theta}) = \sum_{i=1}^c p(\mathbf{x}|\omega_i, \boldsymbol{\theta}_i)P(\omega_i),$$

der  $p(\mathbf{x}|\omega_i, \boldsymbol{\theta}_i)$  *komponenttettheter*,  $P(\omega_i)$  er *blandingsparametre* og  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^t, \dots, \boldsymbol{\theta}_c^t)^t$  er den ukjente parametervektoren.

## Maksimum-likelihood estimering av parametervektor

Metoden er som i ledet læring, men beregningene blir som oftest mer kompliserte (multimodal fordeling, mangel på suffisiente observatorer). Det er likevel mulig å komme frem til en løsning i enkelte tilfeller.

Skal estimere parametervektoren  $\theta$  i blandingstettheten, dvs. hver av subvektorene  $\theta_1, \dots, \theta_c$  skal estimeres ved hjelp av et treningssett  $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  med *ukjent* klassetilhørighet.

Dette krever at  $p(\mathbf{x}|\theta)$  er *identifiserbar*, dvs. at  $\theta$  er unik (forutsettes her). Dette er som oftest tilfelle for blandinger av vanlige (kontinuerlige) tetthetsfunksjoner, mens diskrete fordelinger ofte ikke er identifiserbare.

Antar (i første omgang):

- Antall klasser  $c$  er kjent,
- A priorisannsynlighetene  $P(\omega_i)$ ,  $i = 1, \dots, c$  er kjente,
- Tetthetsfunksjonene  $p(\mathbf{x}|\omega_i, \theta_i)$ ,  $i = 1, \dots, c$  har kjent form (funksjoner av  $\theta_i$ ).

# Maksimum-likelihood metoden

Likelihoodfunksjonen er

$$p(\mathcal{X}|\boldsymbol{\theta}) = \prod_{k=1}^n p(\mathbf{x}_k|\boldsymbol{\theta}) \quad (\text{som for ledet læring}).$$

Faktoriseringen er mulig fordi det forutsettes, som for ledet læring, at treningssamplene er innbyrdes uavhengige.

Log-likelihoodfunksjonen blir som tidligere

$$\mathcal{L}(\boldsymbol{\theta}) = \ln p(\mathcal{X}|\boldsymbol{\theta}) = \sum_{k=1}^n \ln p(\mathbf{x}_k|\boldsymbol{\theta}).$$

Denne skal maksimaliseres med hensyn på parametervektoren  $\boldsymbol{\theta}$ .

## Maksimum-likelihood metoden (forts.)

Ved innsetting for blandingstettheten blir gradienten til log-likelihoodfunksjonen

$$\begin{aligned}\nabla_{\boldsymbol{\theta}_i} \mathcal{L}(\boldsymbol{\theta}) &= \sum_{k=1}^n \nabla_{\boldsymbol{\theta}_i} \ln p(\mathbf{x}_k | \boldsymbol{\theta}) \\ &= \sum_{k=1}^n \frac{1}{p(\mathbf{x}_k | \boldsymbol{\theta})} \nabla_{\boldsymbol{\theta}_i} \left[ \sum_{j=1}^c p(\mathbf{x}_k | \omega_j, \boldsymbol{\theta}_j) P(\omega_j) \right] \\ &= \sum_{k=1}^n \frac{P(\omega_i)}{p(\mathbf{x}_k | \boldsymbol{\theta})} \nabla_{\boldsymbol{\theta}_i} p(\mathbf{x}_k | \omega_i, \boldsymbol{\theta}_i).\end{aligned}$$

Den siste overgangen er mulig siden  $\boldsymbol{\theta}_i$  og  $\boldsymbol{\theta}_j$  er funksjonelt uavhengige for  $i \neq j$ .

# Maksimum-likelihood metoden (forts.)

Innsetting av Bayes regel

$$P(\omega_i | \mathbf{x}_k, \boldsymbol{\theta}) = \frac{p(\mathbf{x}_k | \omega_i, \boldsymbol{\theta}_i) P(\omega_i)}{p(\mathbf{x}_k | \boldsymbol{\theta})}$$

gir da

$$\begin{aligned} \nabla_{\boldsymbol{\theta}_i} \mathcal{L}(\boldsymbol{\theta}) &= \sum_{k=1}^n P(\omega_i | \mathbf{x}_k, \boldsymbol{\theta}) \frac{\nabla_{\boldsymbol{\theta}_i} p(\mathbf{x}_k | \omega_i, \boldsymbol{\theta}_i)}{p(\mathbf{x}_k | \omega_i, \boldsymbol{\theta}_i)} \\ &= \sum_{k=1}^n P(\omega_i | \mathbf{x}_k, \boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}_i} \ln p(\mathbf{x}_k | \omega_i, \boldsymbol{\theta}_i). \end{aligned}$$

En nødvendig betingelse for maksimum av  $\mathcal{L}$  er da gitt ved likningssystemet

$$\sum_{k=1}^n P(\omega_i | \mathbf{x}_k, \boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}_i} \ln p(\mathbf{x}_k | \omega_i, \boldsymbol{\theta}_i) = 0, i = 1, \dots, c.$$



## Eksempel – multivariate normalfordelinger med ukjent forventning

Antar nå at fordelingene er gitt ved  $N(\boldsymbol{\mu}_i, \Sigma_i)$ ,  $i = 1, \dots, c$  der  $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^t, \dots, \boldsymbol{\mu}_c^t)^t$  er ukjent. Da blir

$$\ln p(\mathbf{x}|\omega_i, \boldsymbol{\mu}_i) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^t \Sigma_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) - \ln\{(2\pi)^{d/2} |\Sigma_i|^{1/2}\}$$

slik at

$$\nabla_{\boldsymbol{\mu}_i} \ln p(\mathbf{x}|\omega_i, \boldsymbol{\mu}_i) = \Sigma_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i).$$

Innsetting i likningssystemet gir

$$\sum_{k=1}^n P(\omega_i|\mathbf{x}_k, \boldsymbol{\mu}) \Sigma_i^{-1}(\mathbf{x}_k - \boldsymbol{\mu}_i) = 0$$

og multiplikasjon med  $\Sigma_i$  på begge sider av likhetstegnet gir

$$\sum_{k=1}^n P(\omega_i|\mathbf{x}_k, \boldsymbol{\mu})(\mathbf{x}_k - \boldsymbol{\mu}_i) = 0.$$

## Eksempel (forts.)

Løsningen på likningssystemet kan da skrives som

$$\hat{\boldsymbol{\mu}}_i = \frac{\sum_{k=1}^n P(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\mu}}) \mathbf{x}_k}{\sum_{k=1}^n P(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\mu}})}, i = 1, \dots, c,$$

der

$$P(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\mu}}) = \frac{p(\mathbf{x}_k | \omega_i, \hat{\boldsymbol{\mu}}_i) P(\omega_i)}{p(\mathbf{x}_k | \hat{\boldsymbol{\mu}})}.$$

Dette er et tilfredsstillende resultat, der samplene veies med aposteriorisannsynlighetene for hver klasse, dvs. den mest sannsynlige klassen gis størst vekt. Dette er imidlertid en *implisitt* løsning for forventningsvektoren (ingen eksplisitt løsning).

## Løsning ved iterasjon

Det implisitte likningssystemet kan løses iterativt ut fra startverdier  $\hat{\mu}_i(0)$  for forventningsestimatene

$$\hat{\mu}_i(j+1) = \frac{\sum_{k=1}^n P(\omega_i | \mathbf{x}_k, \hat{\mu}(j)) \mathbf{x}_k}{\sum_{k=1}^n P(\omega_i | \mathbf{x}_k, \hat{\mu}(j))}, \quad i = 1, \dots, c,$$

der estimatene av aposteriorisannsynlighetene er

$$P(\omega_i | \mathbf{x}_k, \hat{\mu}(j)) = \frac{p(\mathbf{x}_k | \omega_i, \hat{\mu}_i(j)) P(\omega_i)}{\sum_{l=1}^c p(\mathbf{x}_k | \omega_l, \hat{\mu}_l(j)) P(\omega_l)}.$$

Her oppdateres estimatene rekursivt for  $j = 0, 1, 2, \dots$ , med rask konvergens dersom separasjonen mellom klassene er god, men man er ikke garantert et globalt maksimum, bare at gradienten er null.

## Eksempel med to klasser

To univariat normalfordelte klasser med ukjente forventningsverdier, enhetlige varianser ( $\sigma = 1$ ) og apriorisannsynlighetene  $P(\omega_1) = 1/3$  og  $P(\omega_2) = 2/3$ .

Blandingstettheten er da gitt ved

$$p(x|\mu_1, \mu_2) = \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x - \mu_1)^2\right] + \frac{2}{3\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x - \mu_2)^2\right]$$

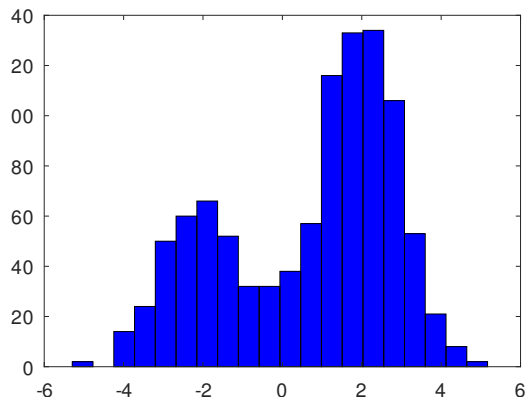
som funksjon av de ukjente parametrene  $\mu_1$  og  $\mu_2$ . Et treningssett med  $n = 1000$  sampler er trukket fra blandingstettheten med sanne parameterverdier

$$\mu_1 = -2 \text{ og } \mu_2 = 2.$$

Kan da beregne log-likelihood funksjonen som funksjon av de to (ukjente) parametrene, dvs.:

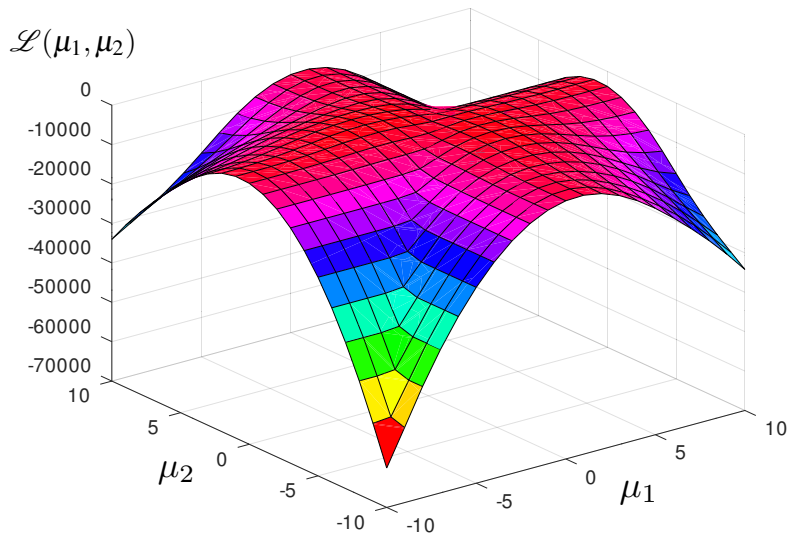
$$\mathcal{L}(\mu_1, \mu_2) = \sum_{k=1}^n \ln p(x_k|\mu_1, \mu_2).$$

# Histogram over treningssettet

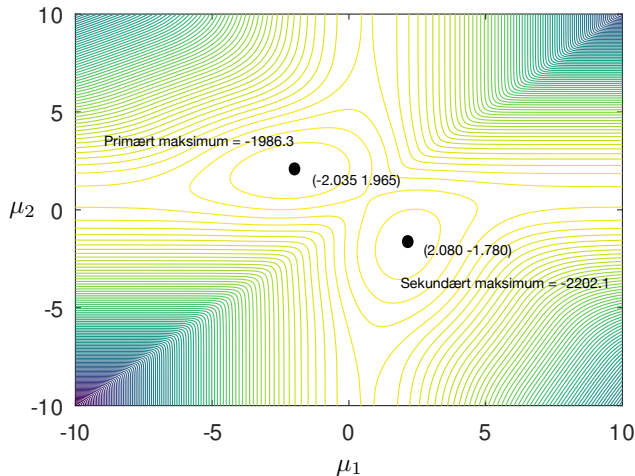


Histogrammet viser tydelig de to modene i blandingstettheten, sentrert om  $-2$  og  $2$ . Antall sampler fra hver bakenforliggende klasse er henholdsvis  $n_1 = 329$  og  $n_2 = 671$ .

# Log-likelihood funksjonen for blandingstettheten

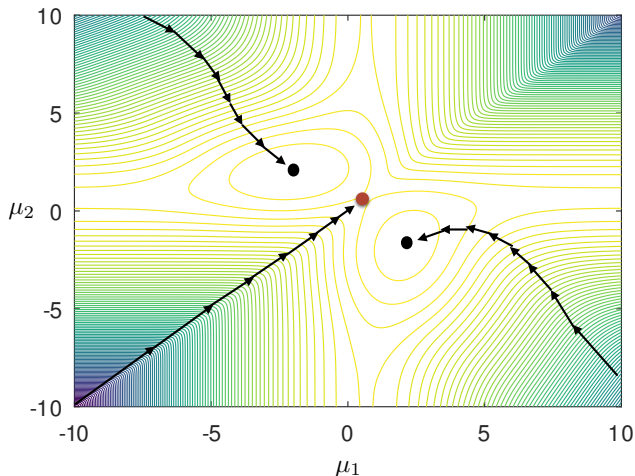


# Konturplott av log-likelihood funksjonen



- Primært maksimum i  $[\mu_1 = -2.035, \mu_2 = 1.965]$  der  $\mathcal{L} = -1986.3$ .
- Sekundært maksimum i  $[\mu_1 = 2.080, \mu_2 = -1.780]$  der  $\mathcal{L} = -2202.1$ .

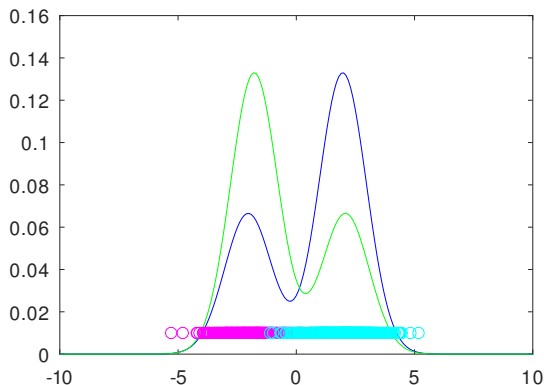
## Iterasjonsprosessen med forskjellige startpunkt



Startpunkt hvilket maksimum man ender i. Uheldig valg av startverdier kan gi konvergens mot et sadelpunktet (rødt symbol), der gradienten også er null.



## Primær og sekundær løsning



Estimerte tetthetsfunksjoner for hver løsning; primærløsningen (blå kurve) og sekundærløsningen (grønn kurve). Samplene i datasettet er plottet nederst.

## Generalisering – ukjente a priori sannsynligheter

$P(\omega_i)$ ,  $i = 1, \dots, c$  kan inkluderes blant de ukjente. Det kan vises at  $\hat{P}(\omega_i)$  og  $\hat{\theta}_i$  må tilfredsstille likningssystemet

$$\left. \begin{aligned} \hat{P}(\omega_i) &= \frac{1}{n} \sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\theta}) \\ \sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\theta}) \nabla_{\theta_i} \ln p(\mathbf{x}_k | \omega_i, \hat{\theta}_i) &= 0 \end{aligned} \right\} \quad i = 1, \dots, c,$$

der

$$\hat{P}(\omega_i | \mathbf{x}_k, \hat{\theta}) = \frac{p(\mathbf{x}_k | \omega_i, \hat{\theta}_i) \hat{P}(\omega_i)}{\sum_{j=1}^c p(\mathbf{x}_k | \omega_j, \hat{\theta}_j) \hat{P}(\omega_j)}, \quad i = 1, \dots, c.$$

Dette forutsetter at  $\mathcal{L}$  er deriverbar mhp.  $\hat{P}(\omega_i)$  og at  $\hat{P}(\omega_i) \neq 0$  for alle  $i$ .

## Eksempel – multivariate normalfordelinger (alle parametre ukjente)

For  $p(\mathbf{x}|\omega_i, \boldsymbol{\theta}_i) = N(\boldsymbol{\mu}_i, \Sigma_i)$  for alle klasser blir likningssystemet (kan vises):

$$\left. \begin{aligned} \hat{P}(\omega_i) &= \frac{1}{n} \sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}}) \\ \hat{\boldsymbol{\mu}}_i &= \frac{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}}) \mathbf{x}_k}{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}})} \\ \hat{\Sigma}_i &= \frac{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}}) (\mathbf{x}_k - \hat{\boldsymbol{\mu}}_i)(\mathbf{x}_k - \hat{\boldsymbol{\mu}}_i)^t}{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}})} \end{aligned} \right\} \quad i = 1, \dots, c.$$

Dette likningssystemet kan også løses ved iterasjon.

# Eksempel – Fishers datasett

```
>> load fisheriris
>>
>> meas(1:10,:)

ans =

    5.1000    3.5000    1.4000    0.2000
    4.9000    3.0000    1.4000    0.2000
    4.7000    3.2000    1.3000    0.2000
    4.6000    3.1000    1.5000    0.2000
    5.0000    3.6000    1.4000    0.2000
    5.4000    3.9000    1.7000    0.4000
    4.6000    3.4000    1.4000    0.3000
    5.0000    3.4000    1.5000    0.2000
    4.4000    2.9000    1.4000    0.2000
    4.9000    3.1000    1.5000    0.1000

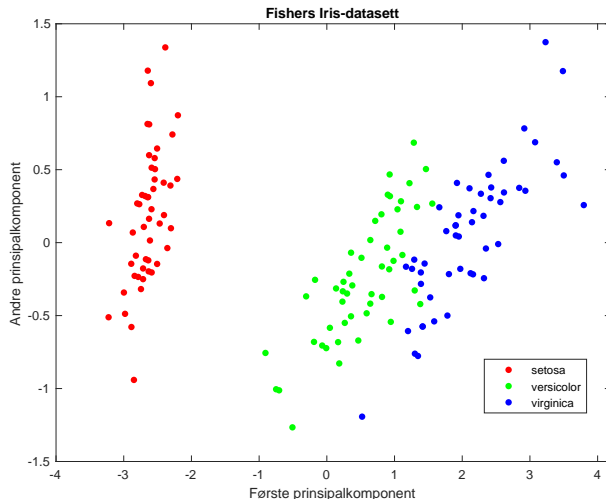
>> species(1:10)

ans =

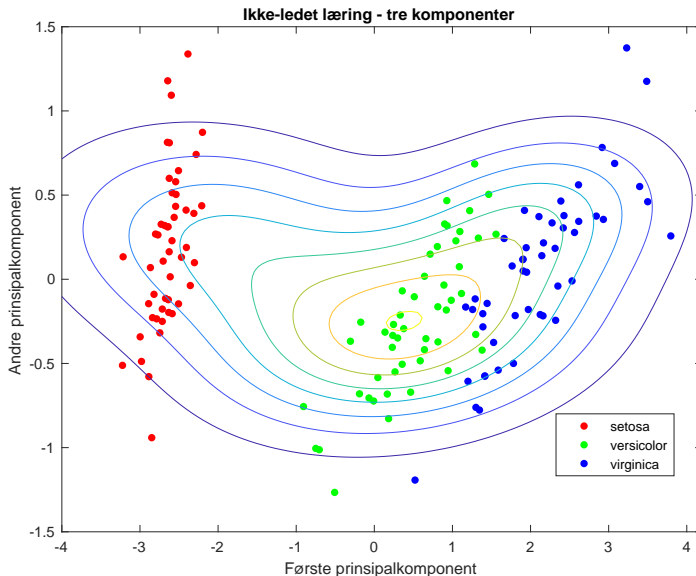
10x1 cell array

{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}

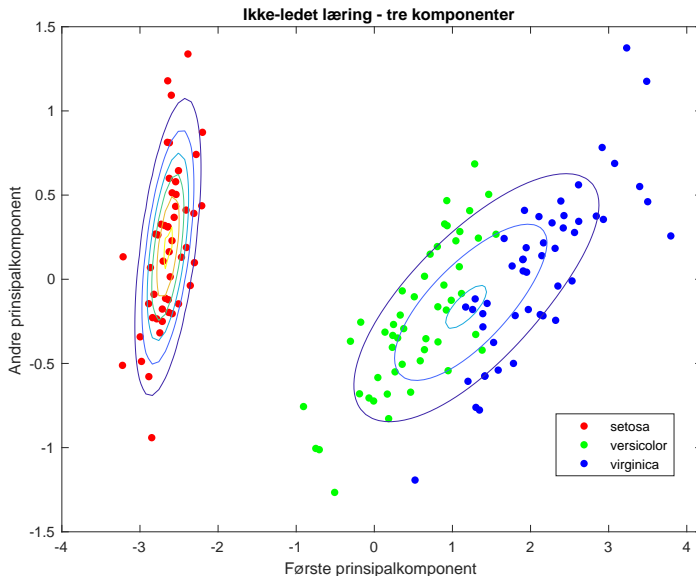
>> summary(categorical(species));
   setosa      50
 versicolor  50
  virginica   50
>> |
```



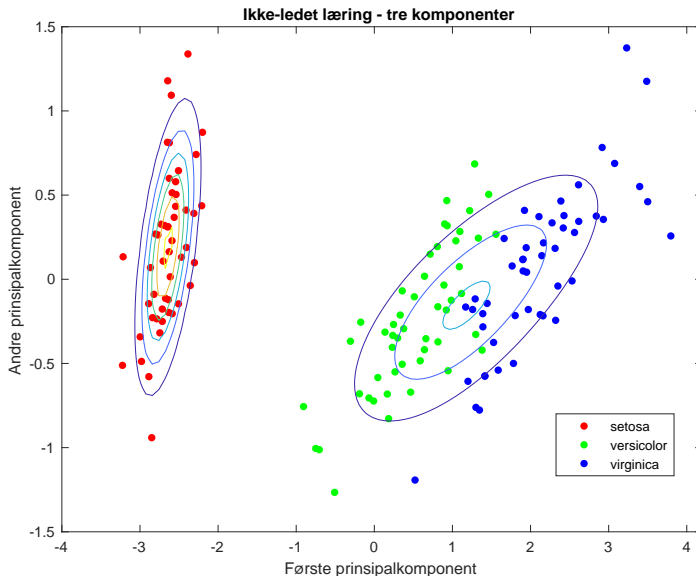
# Spredningsplott og blandingsstetthet – 1 iterasjon



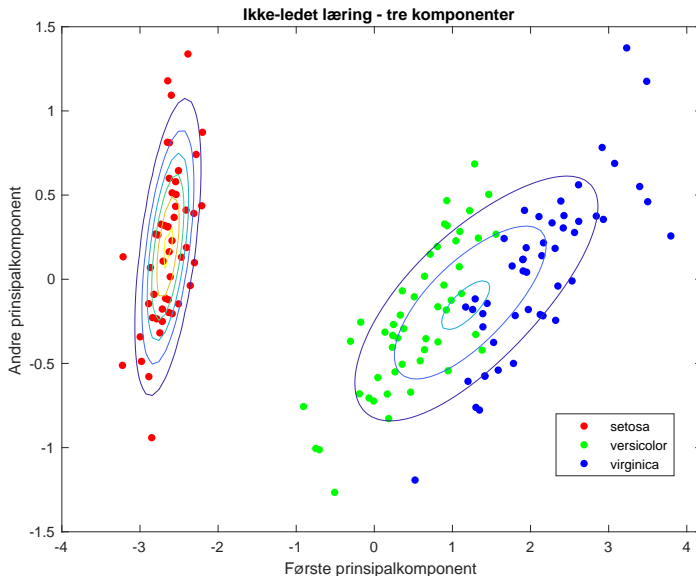
# Spredningsplott og blandingsstetthet – 11 iterasjoner



# Spredningsplott og blandingsstetthet – 21 iterasjoner

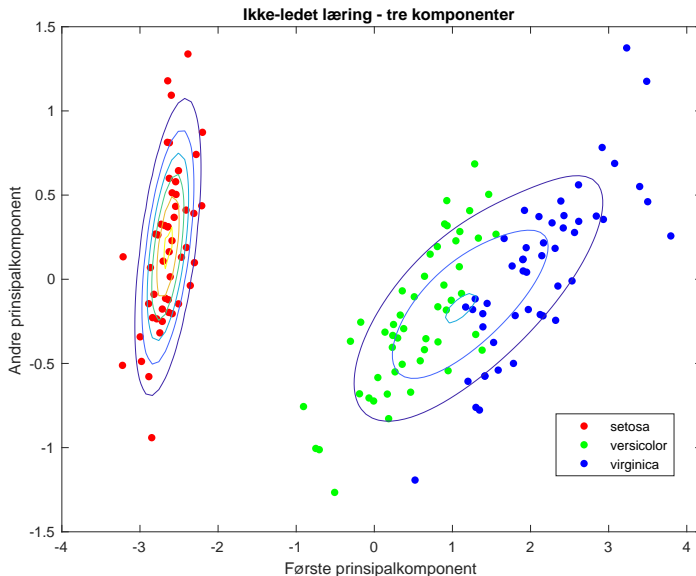


# Spredningsplott og blandingsstetthet – 31 iterasjoner

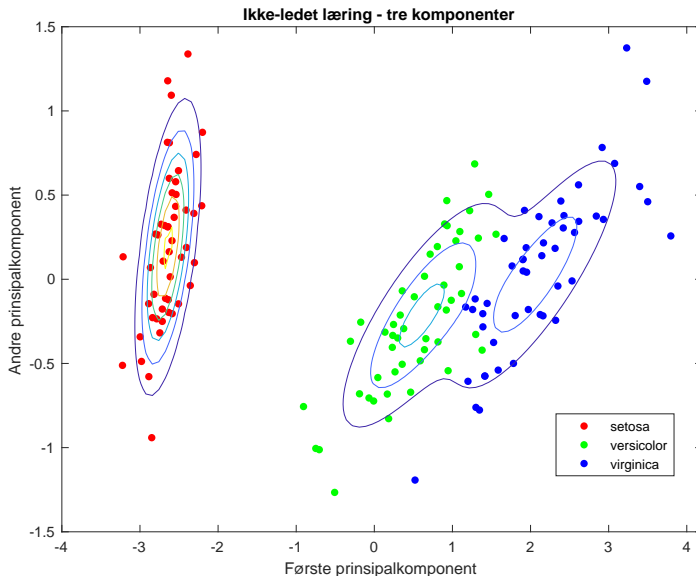




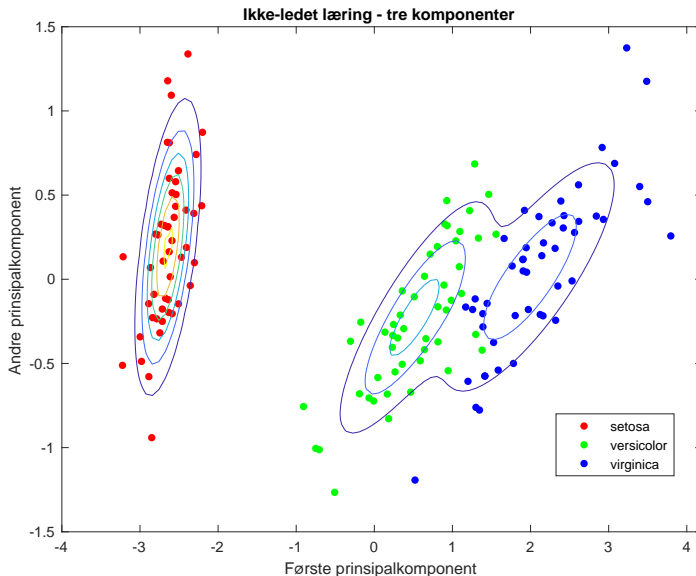
# Spredningsplott og blandingsstetthet – 41 iterasjoner



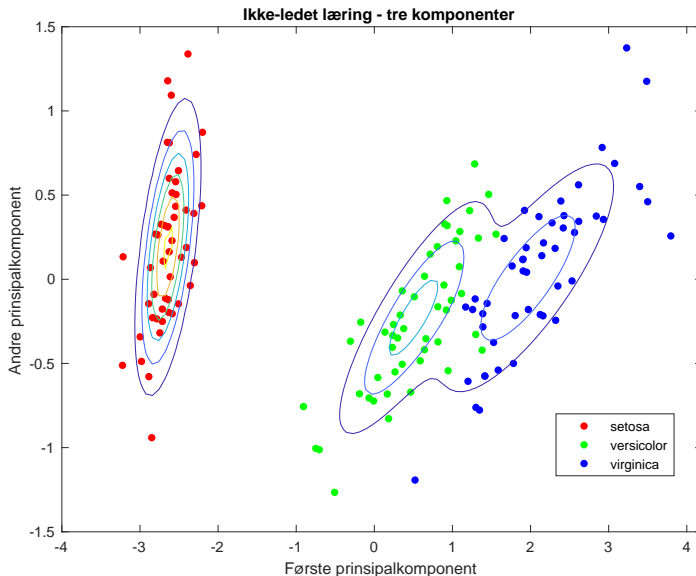
# Spredningsplott og blandingsstetthet – 51 iterasjoner



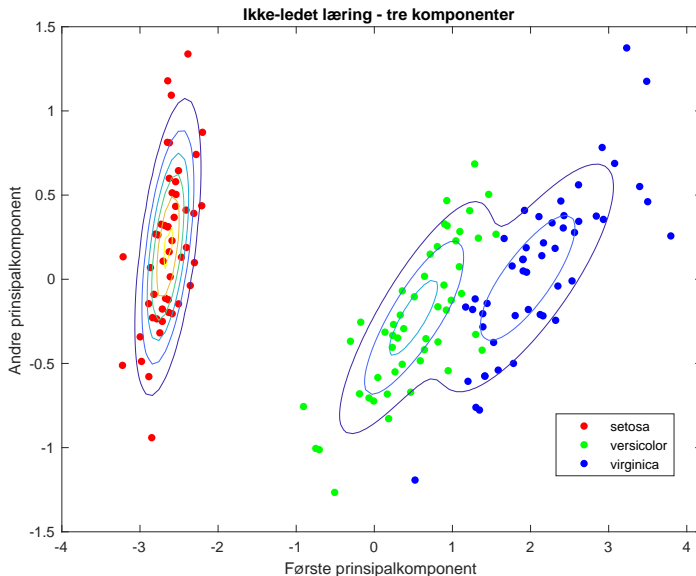
# Spredningsplott og blandingsstetthet – 61 iterasjoner



# Spredningsplott og blandingstetthet – 71 iterasjoner



# Spredningsplott og blandingsstetthet – 91 iterasjoner



# Parameterestimatene

Gaussisk blandingsfordeling med tre komponenter i to dimensjoner

Klasse 1 (versicolor):

Blandingsparameter: 0.291250

Middel: 0.4798 -0.2291

Sigma 1:

0.3528 0.2203

0.2203 0.1924

Klasse 2 (virginica):

Blandingsparameter: 0.375417

Middel: 1.9740 0.0083

Sigma 2:

0.5964 0.2971

0.2971 0.2303

Klasse 3 (setosa):

Blandingsparameter: 0.333333

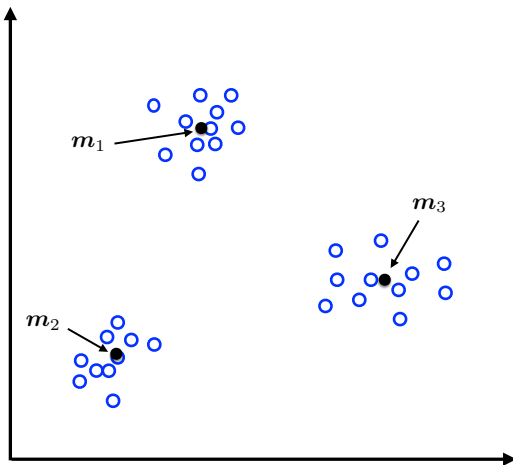
Middel: -2.6424 0.1909

Sigma 3:

0.0480 0.0549

0.0549 0.2133

## Forenkling – datasett med tette klynger omkring sampelmidlene



Anta her at samplene danner tette, adskilte klynger, slik at

$$\hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}}) \approx \begin{cases} 1 & \mathbf{x}_k \in \omega_i \\ 0 & \text{ellers.} \end{cases}$$

## Tilnærmet resultat – tette klynger

Likningssystemet reduseres da til

$$\begin{aligned}\hat{P}(\omega_i) &= \frac{1}{n} \sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}}) \approx \frac{n_i}{n} \\ \hat{\boldsymbol{\mu}}_i &= \frac{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}}) \mathbf{x}_k}{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}})} \approx \frac{1}{n_i} \sum_{\mathbf{x}_k \in \mathcal{X}_i} \mathbf{x}_k = \mathbf{m}_i \\ \hat{\boldsymbol{\Sigma}}_i &= \frac{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}}) (\mathbf{x}_k - \hat{\boldsymbol{\mu}}_i)(\mathbf{x}_k - \hat{\boldsymbol{\mu}}_i)^t}{\sum_{k=1}^n \hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}})} \approx \frac{1}{n_i} \sum_{\mathbf{x}_k \in \mathcal{X}_i} (\mathbf{x}_k - \mathbf{m}_i)(\mathbf{x}_k - \mathbf{m}_i)^t\end{aligned}$$

der  $n_i$  være antall sampler i klasse  $\omega_i$ . Dette er et tilfredsstillende (intuitivt riktig) resultat.



## Isodata-algoritmen (K-Means-Clustering)

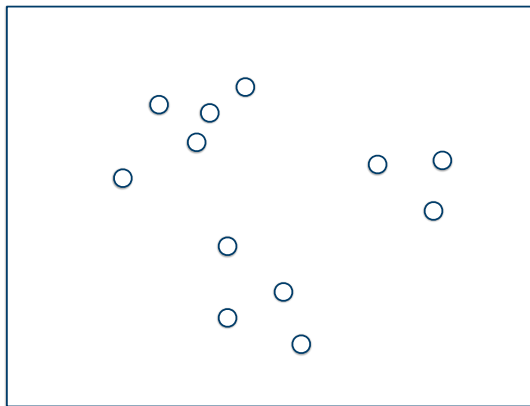
For multivariate normalfordelinger er  $\hat{P}(\omega_i | \mathbf{x}_k, \hat{\boldsymbol{\theta}})$  stor når Mahalanobisavstanden  $r_i^2 = (\mathbf{x}_k - \boldsymbol{\mu}_i)^t \hat{\boldsymbol{\Sigma}}_i^{-1} (\mathbf{x}_k - \boldsymbol{\mu}_i)$  er liten.

Dersom  $r_i^2$  erstattes med Euclidsk avstand  $\|\mathbf{x}_k - \hat{\boldsymbol{\mu}}_i\|^2$  fra hvert klassemiddel, vil det foregående resultatet antyde følgende enkle iterasjonsprosess:

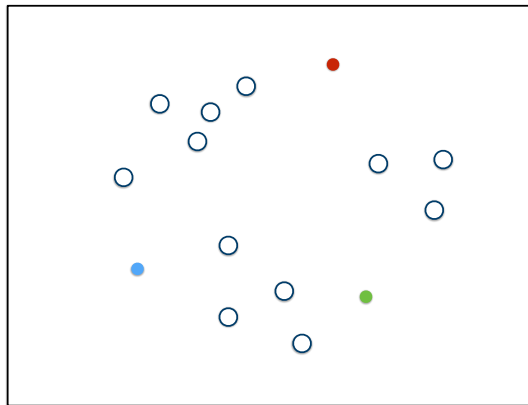
- Initialiser  $\hat{\boldsymbol{\mu}}_1, \dots, \hat{\boldsymbol{\mu}}_c$
- Gjenta inntil ferdig:
  - Klassifiser  $\mathbf{x}_k, k = 1, \dots, n$  til nærmeste middel
  - Oppdatér  $\hat{\boldsymbol{\mu}}_1, \dots, \hat{\boldsymbol{\mu}}_c$
  - Hvis ingen endring  $\rightarrow$  ferdig.

Dette er den grunnleggende *Isodata-algoritmen* – et eksempel på en *klyngeanalysemetode*.

## Eksempel – Isodata-algoritmen (1)



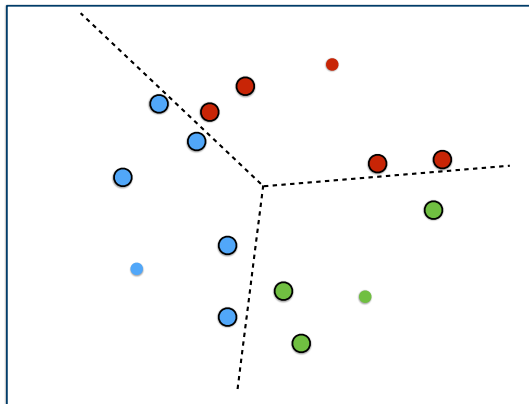
Umerket datasett



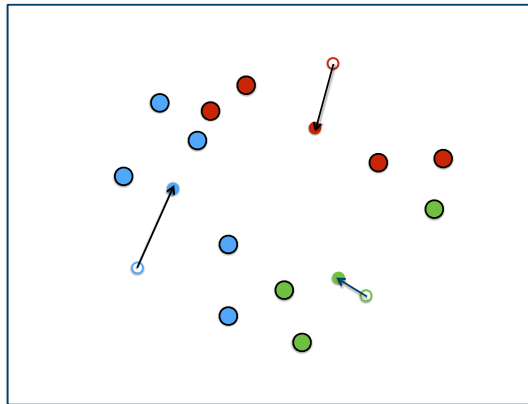
$\hat{\mu}_1$  (rød) ,  $\hat{\mu}_2$  (blå) og  $\hat{\mu}_3$  (grønn)

*Initialisering*

## Eksempel – Isodata-algoritmen (2)



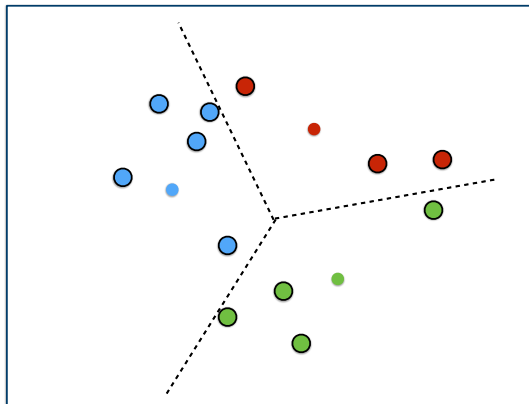
Klassifisering til nærmeste middel



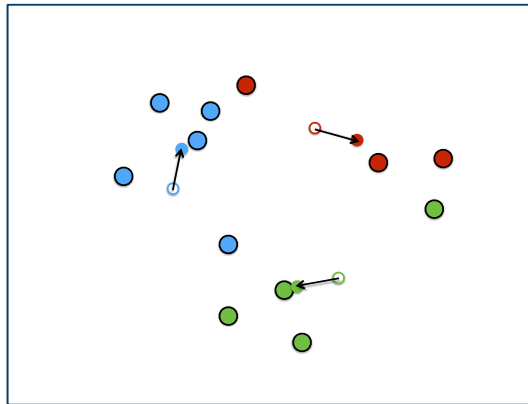
Oppdatering av  $\hat{\mu}_1$ ,  $\hat{\mu}_2$  og  $\hat{\mu}_3$

*Iterasjon 1*

## Eksempel – Isodata-algoritmen (3)



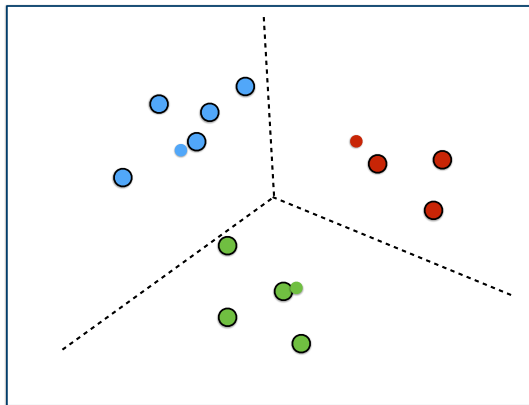
Reklassifisering til nærmeste middel



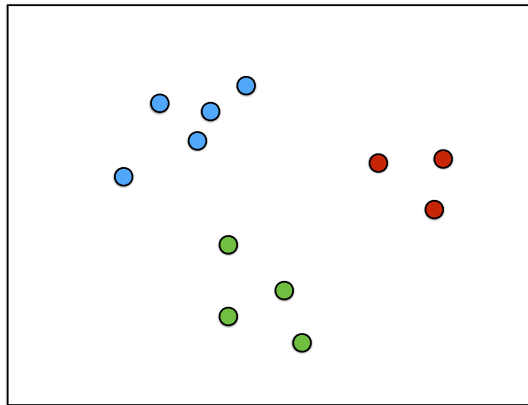
Oppdatering av  $\hat{\mu}_1, \hat{\mu}_2$  og  $\hat{\mu}_3$

*Iterasjon 2*

## Eksempel – Isodata-algoritmen (4)



Reklassifisering til nærmeste middel



Endelig kassetilordning (klyngeinndeling)

*Sluttresultat*

# Eksempel – Fishers datasett

```
>> load fisheriris
>>
>> meas(1:10,:)

ans =

    5.1000    3.5000    1.4000    0.2000
    4.9000    3.0000    1.4000    0.2000
    4.7000    3.2000    1.3000    0.2000
    4.6000    3.1000    1.5000    0.2000
    5.0000    3.6000    1.4000    0.2000
    5.4000    3.9000    1.7000    0.4000
    4.6000    3.4000    1.4000    0.3000
    5.0000    3.4000    1.5000    0.2000
    4.4000    2.9000    1.4000    0.2000
    4.9000    3.1000    1.5000    0.1000

>> species(1:10)

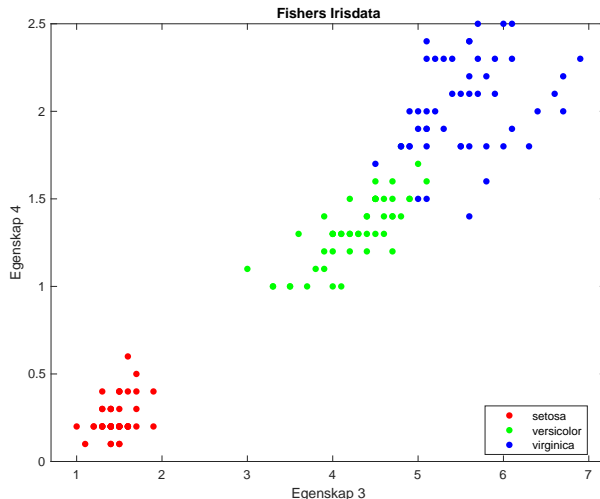
ans =

10x1 cell array

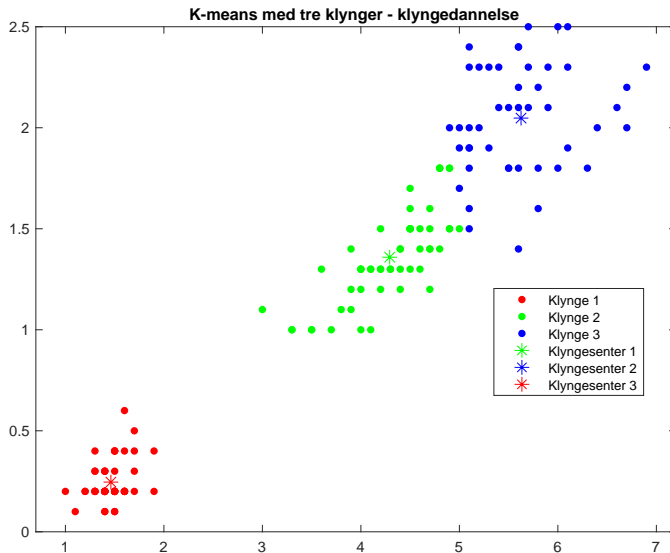
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}
{'setosa'}

>> summary(categorical(species));
   setosa      50
 versicolor  50
  virginica   50

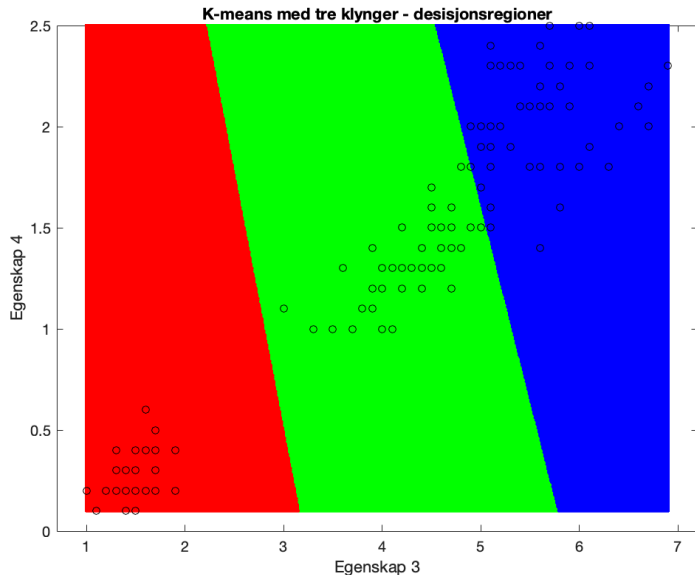
>> |
```



# K-means algoritmen – tre klynger

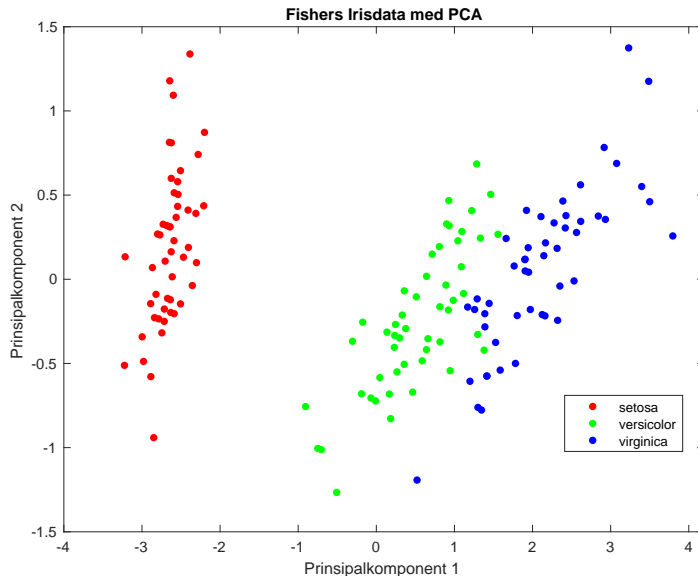


# Desisjonsregioner

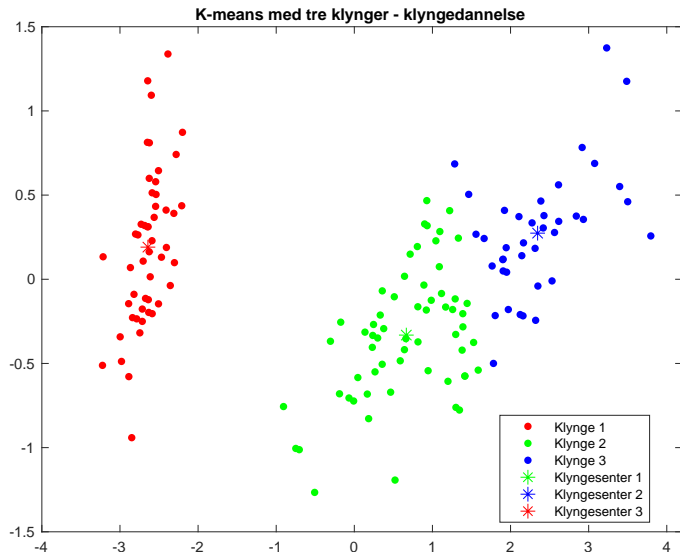




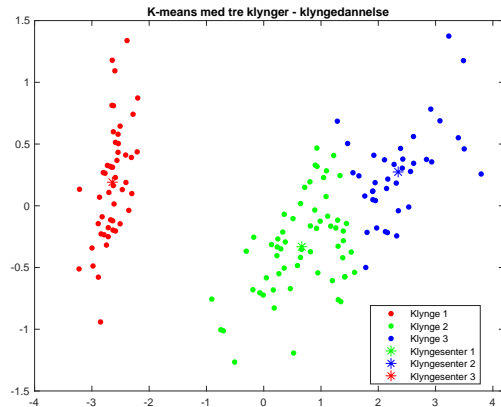
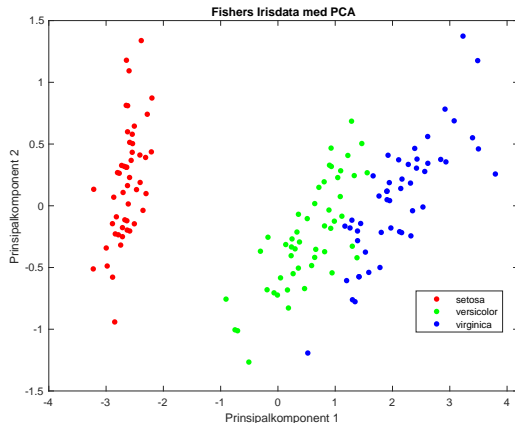
# Fishers Irisdata transformert til de to første prinsipalkomponentene



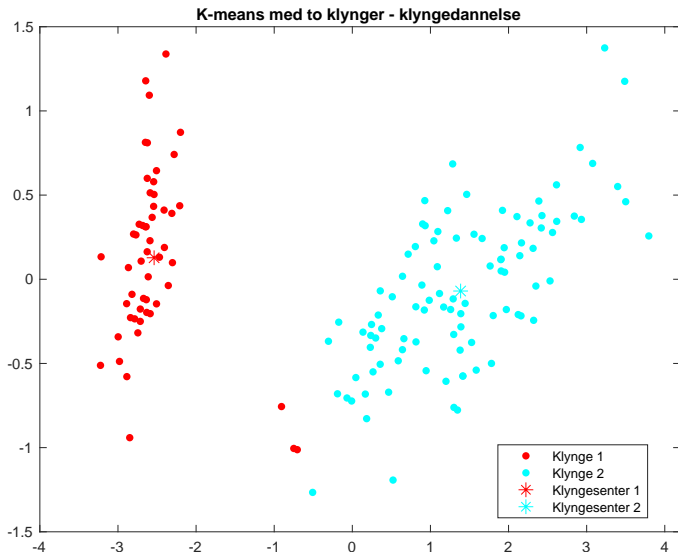
# K-means algoritmen – tre klynger (med PCA)



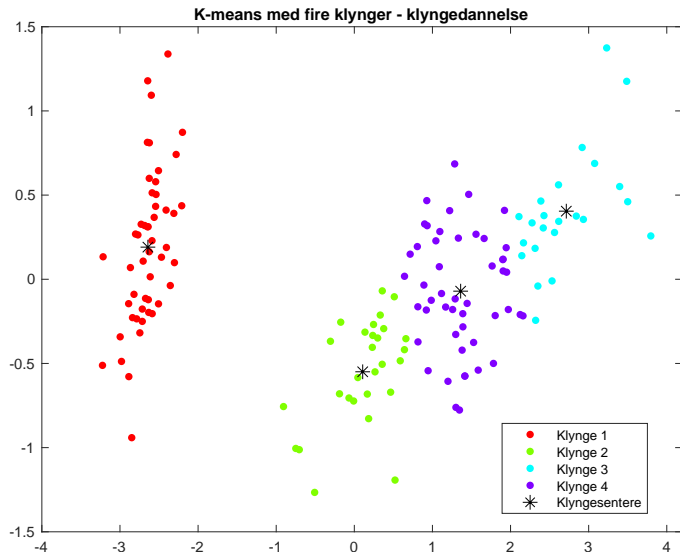
# K-means – fasit og klyngedannelse – tre klynger



# K-means algoritmen – to klynger (med PCA)



# K-means algoritmen – fire klynger (med PCA)



# Innhold i kurset

- Introduksjon til mønstergjenkjenning
- Beslutningsteori
- Parametriske metoder
- Ikke-parametriske metoder
- Lineære og generaliserte diskriminantfunksjoner
- Evaluering av klassifikatorer
- Ikke-ledet læring
- Klyngeanalyse (neste gang).