

Изкуствен интелект

Тема 14: Методи за машинно обучение. Статистически методи. Смесии от Гаусови функции (GMM).



Съдържание

- ❑ Обща класификация на методите
- ❑ Класификация с невронни мрежи. Архитектури. MLPNN и PNN.
- ❑ Неврони с обратна връзка. Рекурентни и локално рекурентни невронни мрежи (TDNN, RNN, DRNN, LRPNN, GLRPNN)
- ❑ Статистически методи. Генеративни класификатори. Класификатор със смес от Гаусови функции (Gaussian Mixture Models, GMM)
- ❑ Статистически методи. Дискриминативни класификатори. Класификатор с опорни вектори (Support Vector Machine, SVM)
- ❑ Еволюционни и генетични методи за оптимизация (DE, PSO)

Класификатори

Discriminative approaches

LDA

Polynomial
classifier

TDNN
and RNN

FFNN

SVM

Decision
trees

Non-discriminative approaches

k-NN

LVQ

SOM

PNN

GMM

HMM

Generative approaches

Combined methods

GMM/SVM

HMM/ANN

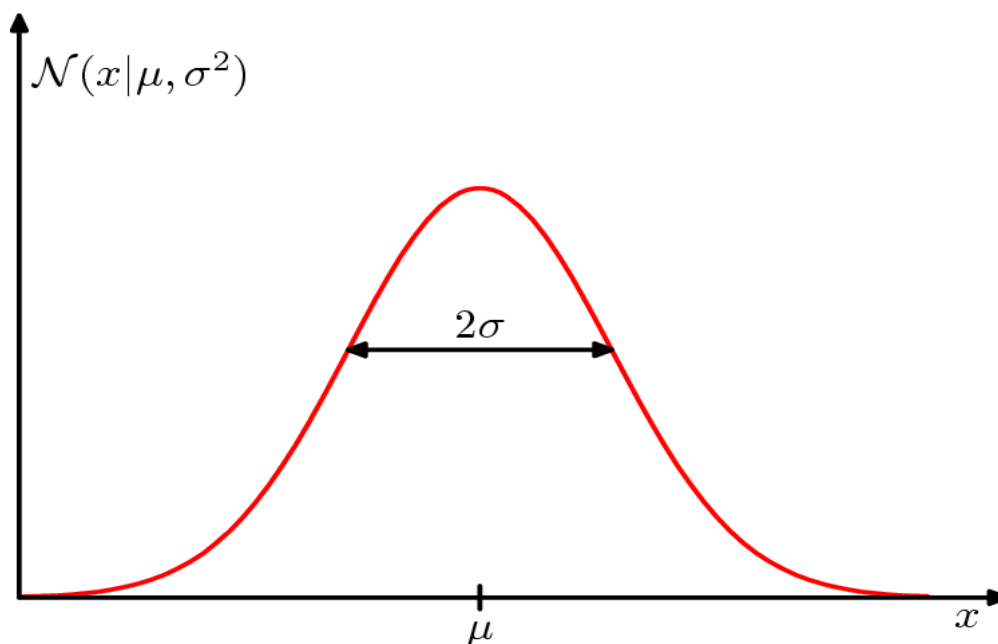
RBF

HMMs trained
discriminatively

PNN-RNN

1-D Gaussian

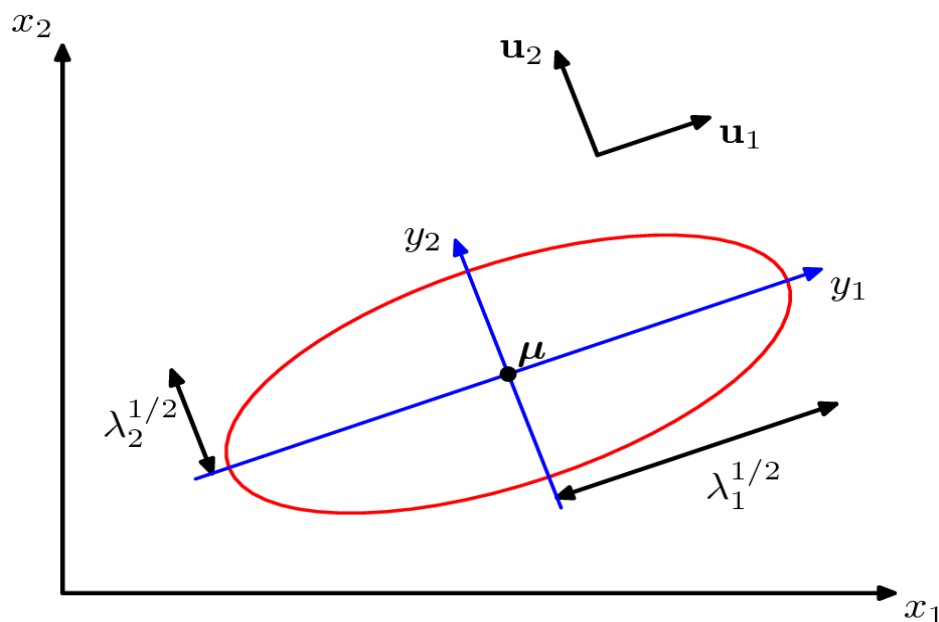
$$\text{Normal}(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\}$$



Параметрите са mean (μ) и variance (σ)

n-D Gaussian

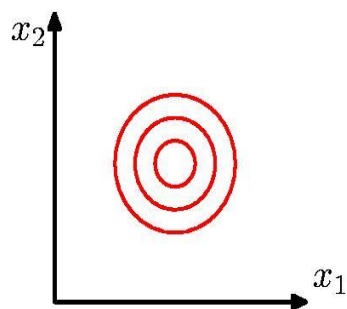
$$\text{Normal}(\mathbf{x} | \mu, \Sigma) = \frac{1}{(2\pi)^2} \frac{1}{\det(\Sigma)^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mu)^T \Sigma (\mathbf{x} - \mu) \right\}$$



При n-D дисперсията се заменя с ковариационната матрица

Ковариационна матрица на GMM

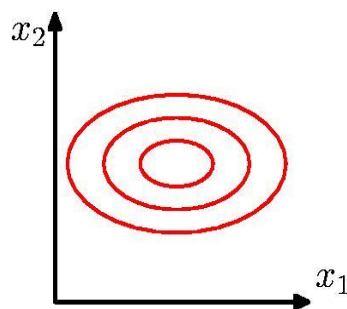
Spherical



(c)

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

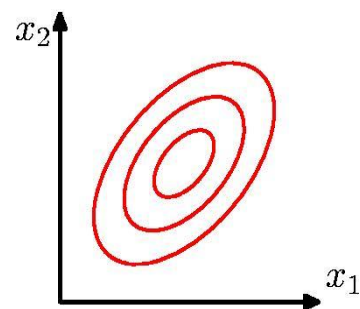
Diagonal



(b)

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

Full covariance



(a)

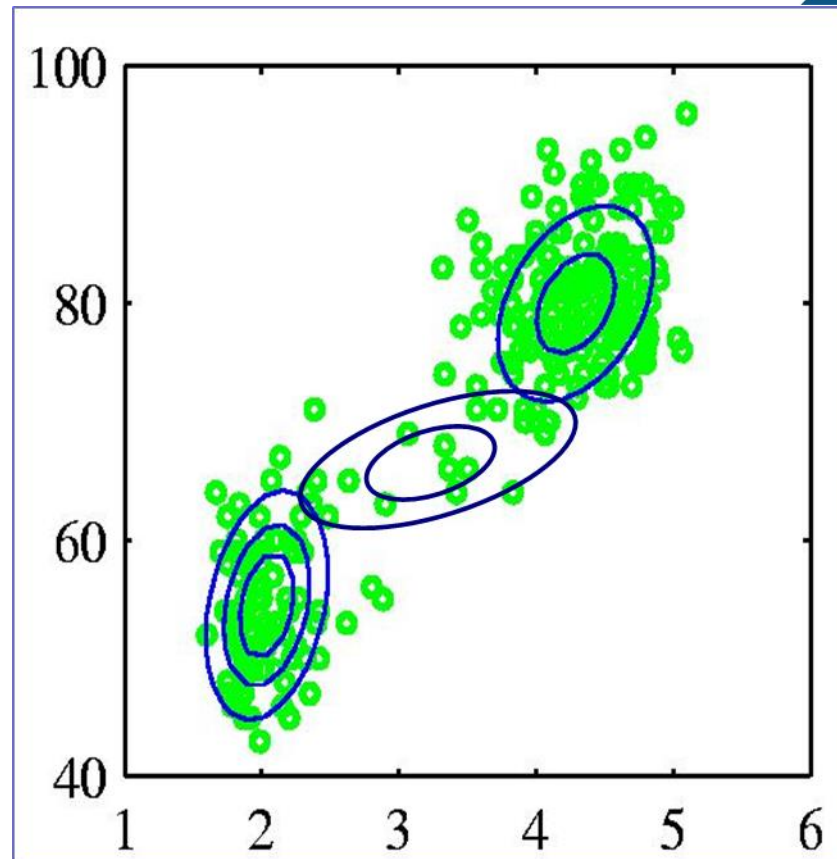
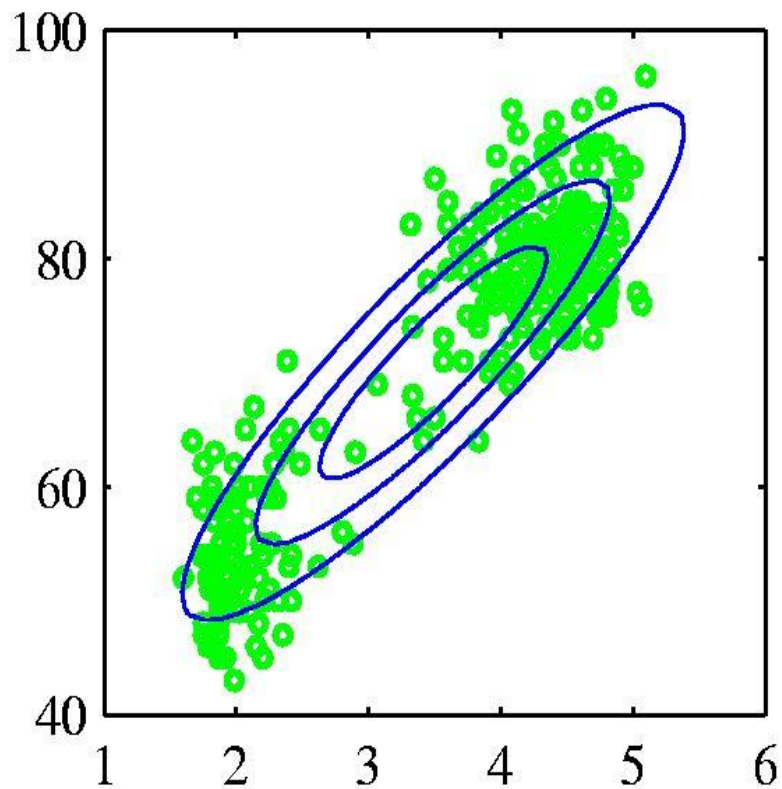
$$\Sigma = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12}^2 \\ \sigma_{12}^2 & \sigma_{22}^2 \end{pmatrix}$$

За целият набор от данни изчисляваме log-likelihood*:

$$\ln p(X) = \ln \prod_{n=1}^N \text{Normal}(\mathbf{x}_n \mid \mu, \Sigma)$$

* log-likelihood = логаритъм от правдоподобността

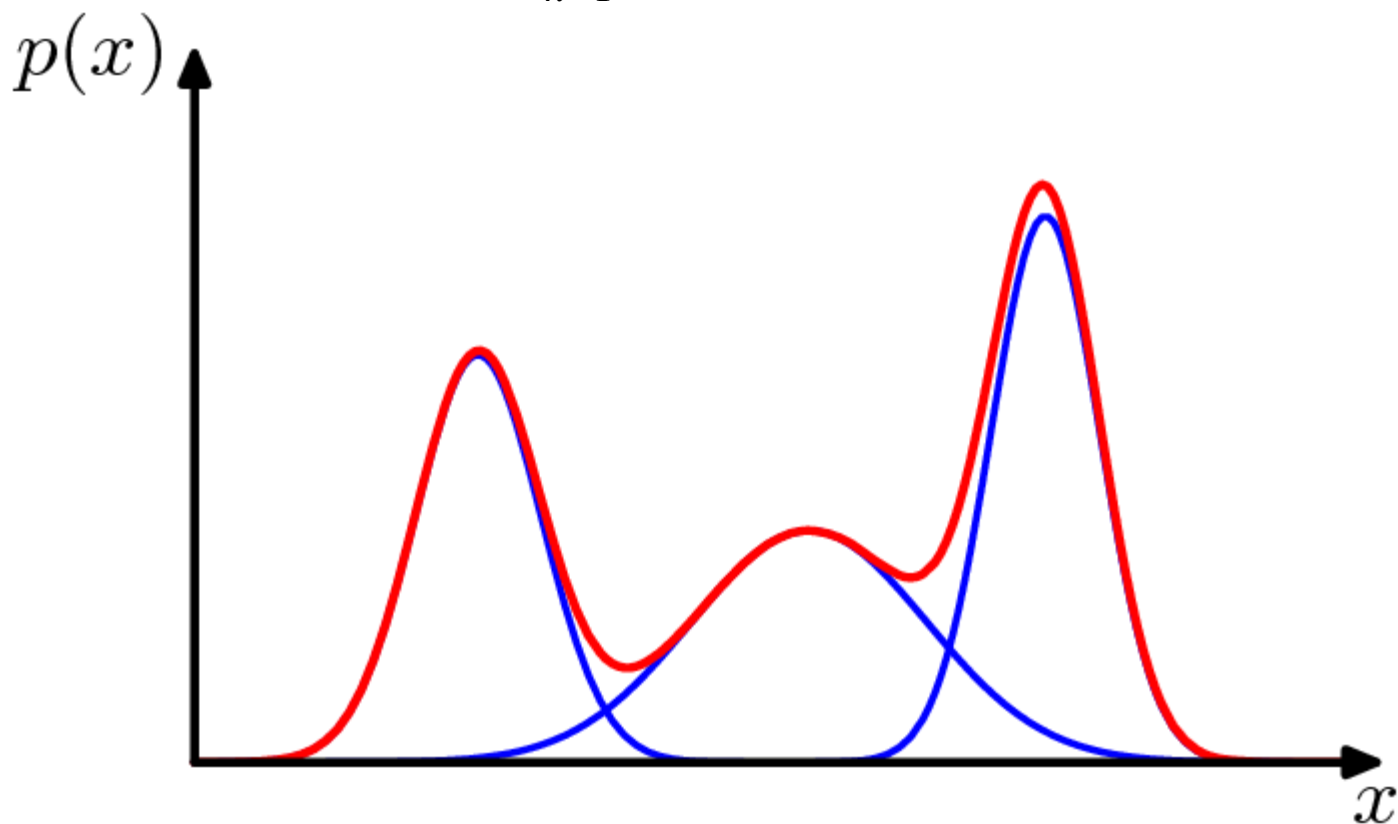
Смес от Гаусови функции



В реалния живот данните рядко могат да се опишат с една Гаусова крива, но може да се ползва комбинация от няколко...

Смес от Гаусови функции

$$p(\mathbf{x}) = \sum_{k=1}^M \pi_k \text{Normal}(\mathbf{x} \mid \mu_k, \Sigma_k)$$



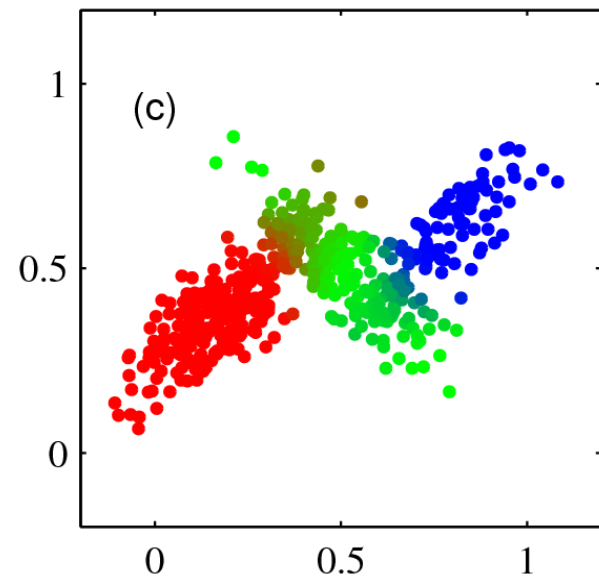
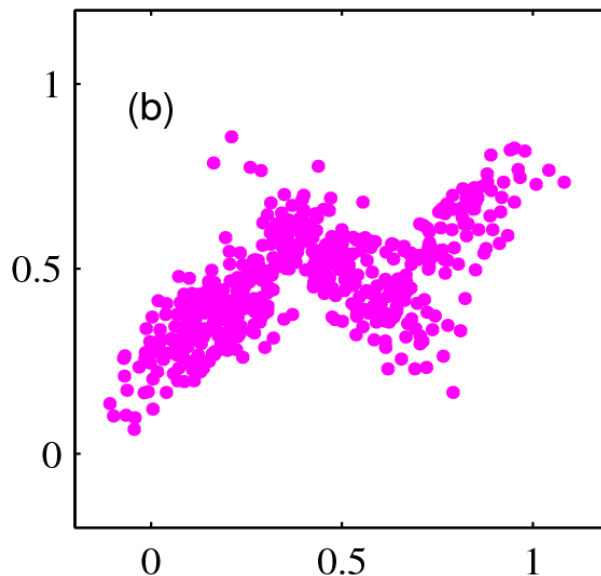
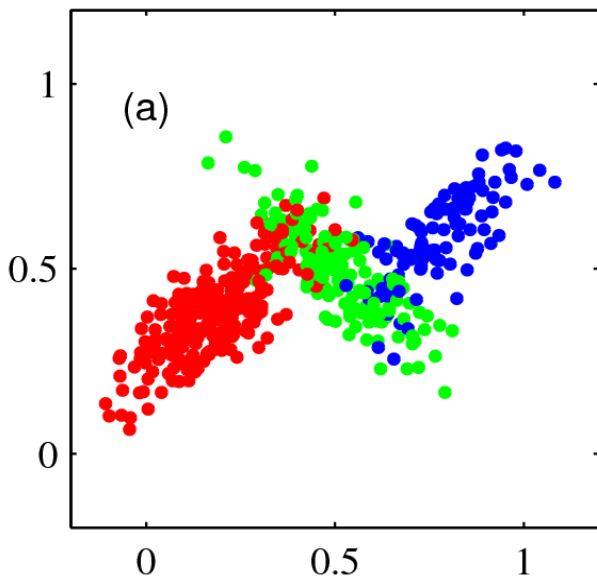
Смес от Гаусови функции

$$p(\mathbf{x}) = \sum_{k=1}^M \pi_k \text{Normal}(\mathbf{x} \mid \mu_k, \Sigma_k)$$

- В допълнение към средната стойност и ковариационните параметри (M пъти), имаме и тегловните коефициенти π_k .

$$\sum_{k=1}^M \pi_k = 1 \quad 0 \leq \pi_k \leq 1$$

Смес от Гаусови функции



$$\gamma_k(\mathbf{x}) = p(k | \mathbf{x})$$

$$p(\mathbf{x}) = \sum_{k=1}^M p(k) p(\mathbf{x} | k)$$

$$= \frac{p(\mathbf{x}) p(\mathbf{x} | k)}{\sum_l p(l) p(\mathbf{x} | l)}$$

$$= \frac{\pi_k \text{Normal}(\mathbf{x} | \mu_k, \Sigma_k)}{\sum_l \pi_l \text{Normal}(\mathbf{x} | \mu_l, \Sigma_l)}$$

Expectation Maximization (EM)

■ Цел:

Да се максимизира log-likelihood за целия набор от данни

$$\ln p(\mathbf{X} | \Pi, \mu, \Sigma) = \sum_{n=1}^N \ln \left\{ \sum_{k=1}^M \pi_k \text{Normal}(\mathbf{x}_n | \mu_k, \Sigma_k) \right\}$$

Като изчисляваме апроксимацията на класовете, можем да максимизираме индивидуално:

- ❑ средните стойности,
- ❑ ковариационните матрици,
- ❑ тегловните коефициенти за отделните компоненти

Expectation Maximization (EM)

Оценка на средна стойност:

$$\mu_k = \frac{1}{N_k} \sum_{n=1}^N \gamma_k(\mathbf{x}_n) \mathbf{x}_n \quad N_k = \sum_{n=1}^N \gamma_k(\mathbf{x}_n)$$

Оценка на ковариационните матрици

$$\Sigma_k = \frac{1}{N_k} \sum_{n=1}^N \gamma_k(\mathbf{x}_n) (\mathbf{x}_n - \mu_k)(\mathbf{x}_n - \mu_k)^T$$

Оценка на тегловните коефициенти

$$\pi_k = \frac{N_k}{N}$$

$$\gamma_k(\mathbf{x}) = p(k | \mathbf{x})$$

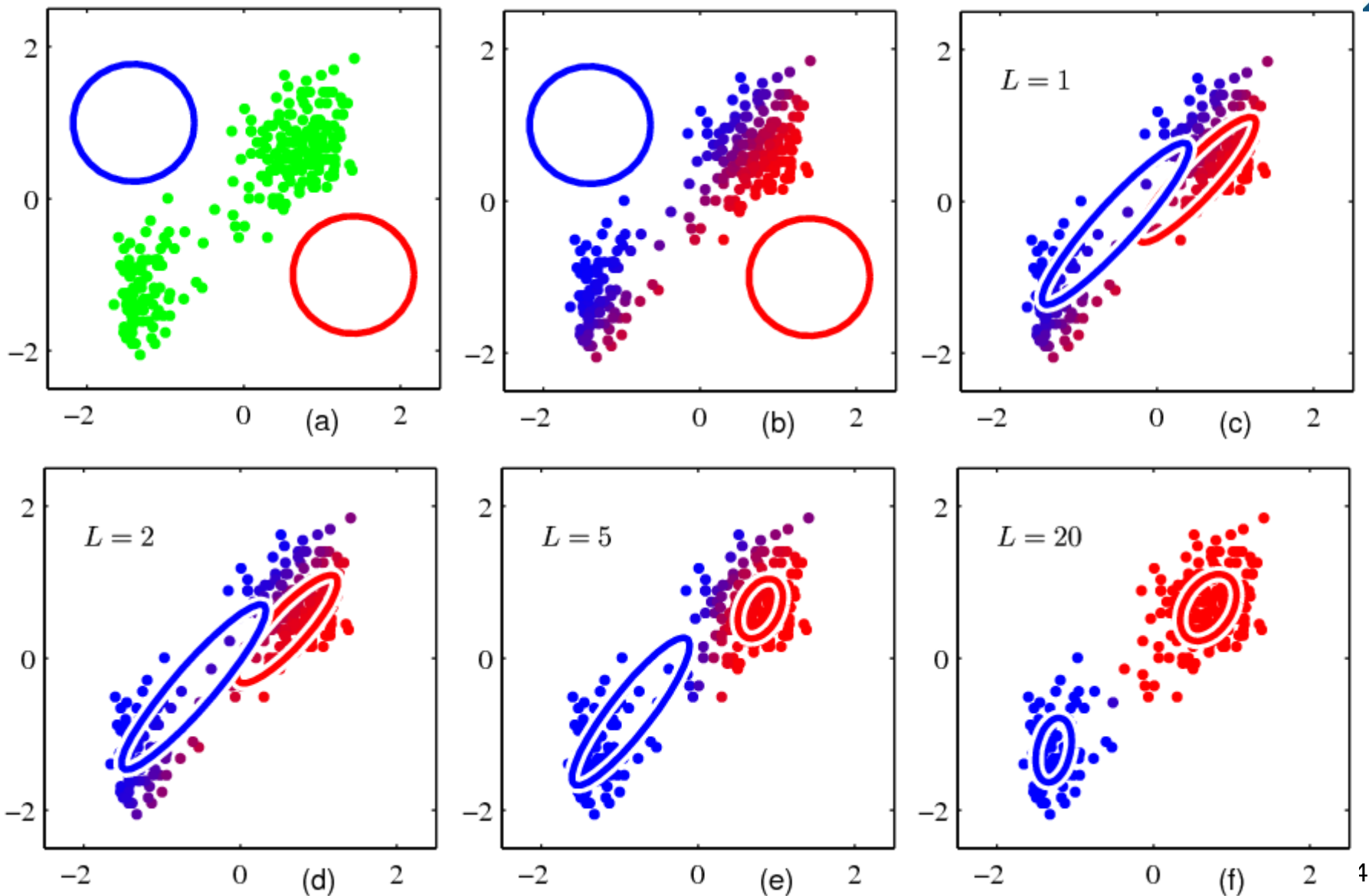
$$= \frac{p(\mathbf{x}) p(\mathbf{x} | k)}{\sum_l p(l) p(\mathbf{x} | l)}$$

$$= \frac{\pi_k \text{Normal}(\mathbf{x} | \mu_k, \Sigma_k)}{\sum_l \pi_l \text{Normal}(\mathbf{x} | \mu_l, \Sigma_l)}$$

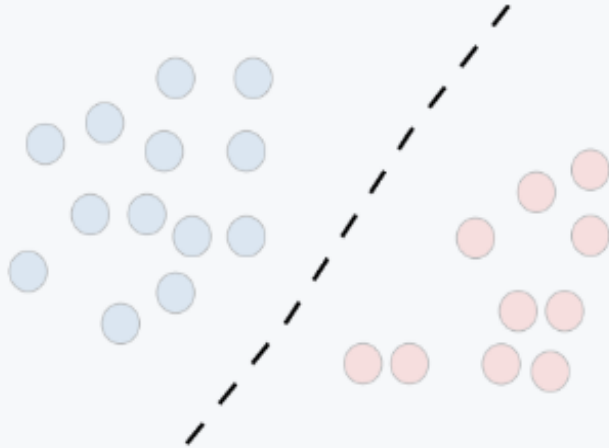
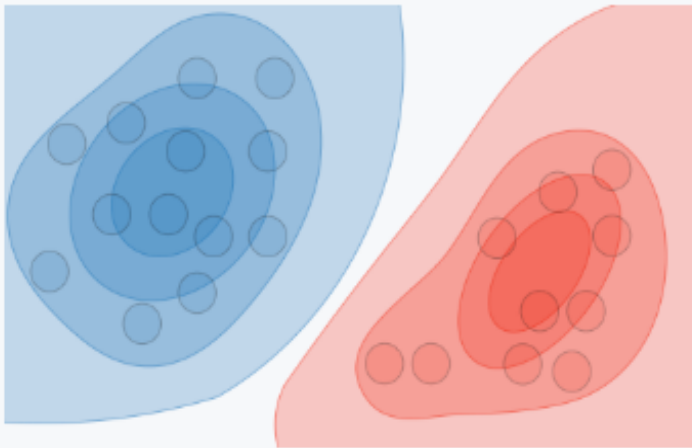
Expectation Maximization (EM)

- Инициализация на параметрите
- Повтаряй до достигане на сходимост
 - **Е стъпка:** изчисляване на класовата принадлежност
 - **М стъпка:** преизчисляване на параметрите
 - Изчисляване на log likelihood за новите параметри

Expectation Maximization (EM)



Дискриминативни и генеративни методи

| | Discriminative model | Generative model |
|----------------|---|--|
| Goal | Directly estimate $P(y x)$ | Estimate $P(x y)$ to then deduce $P(y x)$ |
| What's learned | Decision boundary | Probability distributions of the data |
| Illustration |  |  |
| Examples | MLP NN, SVM, Decision trees, LDA... | PNN, GMM, Bayesian Networks ... |

Съдържание

- ☐ Обща класификация на методите
- ☐ Класификация с невронни мрежи. Архитектури.
MLPNN и PNN.
- ☐ Рекурентни невронни мрежи (RNN, LRPNN,
GLRPNN)
- ☐ Статистически методи (GMM, HMM)
- ☐ Класификация с опорни вектори (SVM)
- ☐ Еволюционни и генетични методи за оптимизация
(DE, PSO)

Скрити модели на Марков

