

Изкуствен интелект

Тема 12: Методи за машинно обучение. Класификация с невронни мрежи без обратни връзки (MLP NN, PNN).

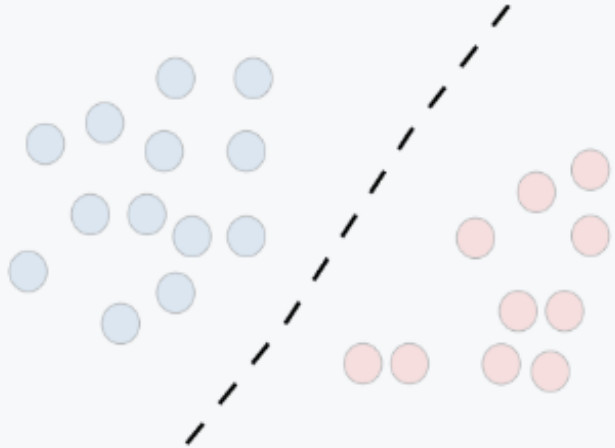
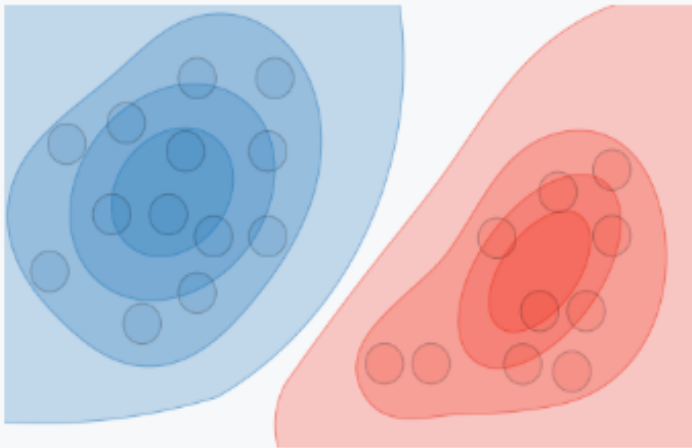


Съдържание

- Таксономия на методите за класификация
 - Дискриминативни класификатори
 - Генеративни класификатори
 - Хибридни методи

Дискриминативни и генеративни методи за класификация



	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 ...

Обща класификация на методите за класификация



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

Невронни мрежи без обратни връзки



- ☐ Класификация с невронни мрежи. Неврони и невронни мрежи без обратни връзки.
- ☐ Многослойна невронна мрежа от перцептрони (MLPNN)
- ☐ Вероятностна невронна мрежа (PNN).

Въведение - терминология

■ Перцептрон

■ Многослойни невронни мрежи

■ Подобро обучение

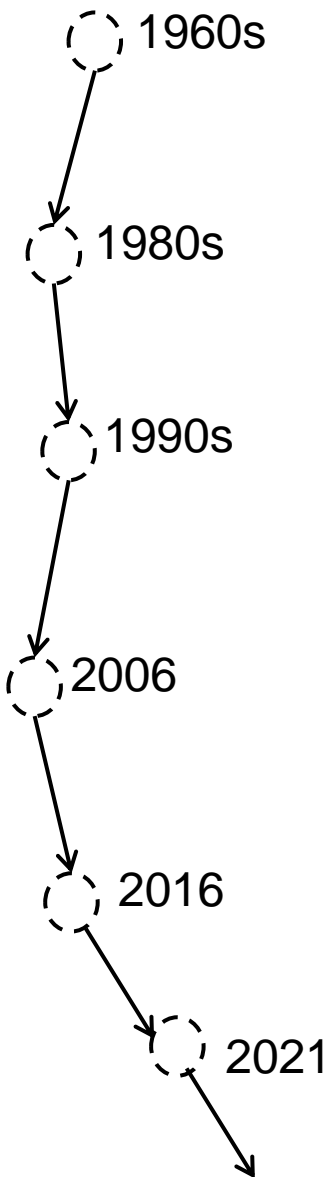
Дълбоки невронни мрежи

■ Deep Learning

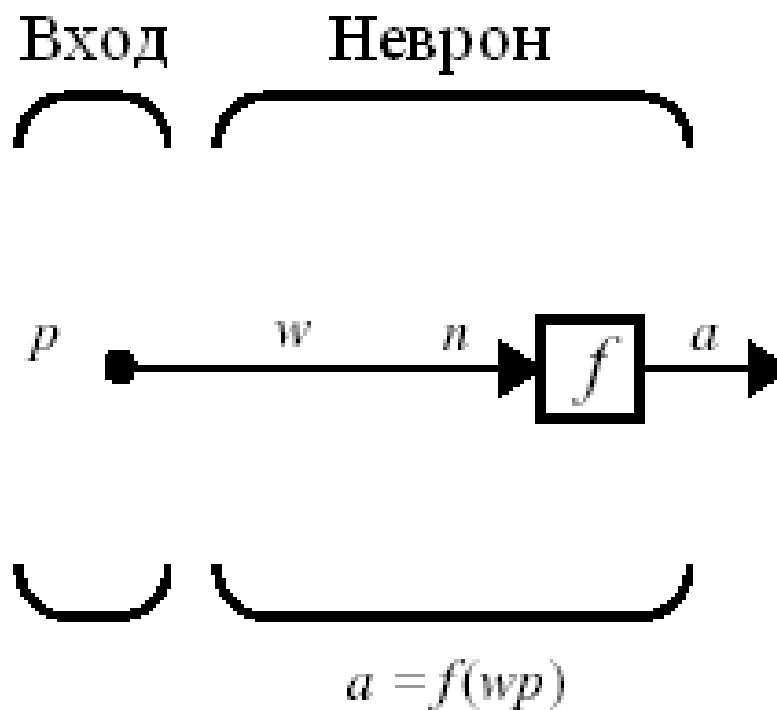
■ Deep Neural Networks, CNNs, GANs

■ Deep Structural Learning

■ Deep Belief Networks

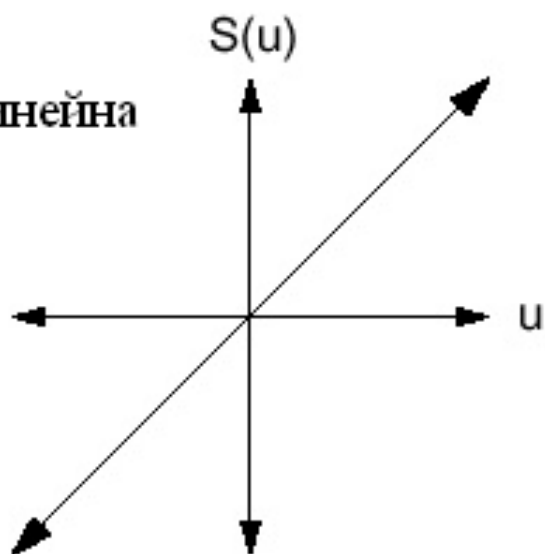


Неврон – изчислителен модел

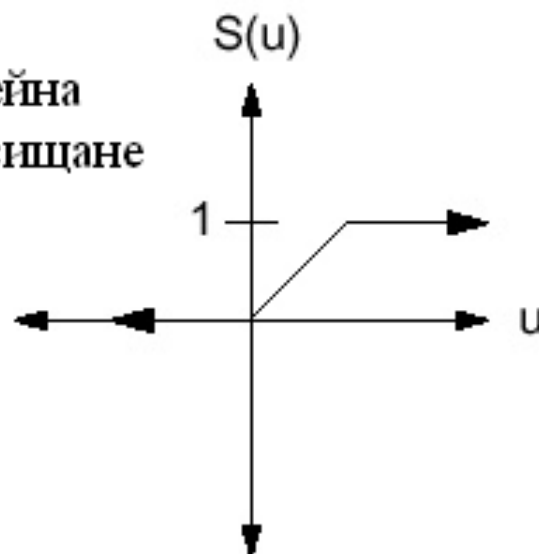


Активиращи функции

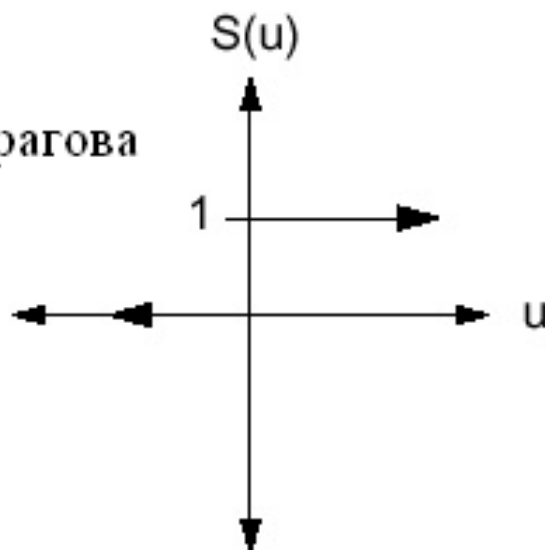
Линейна



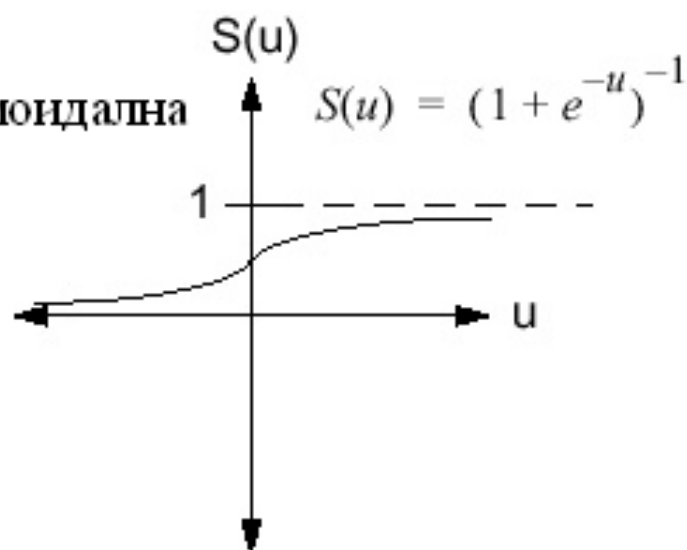
Линейна
с насищане



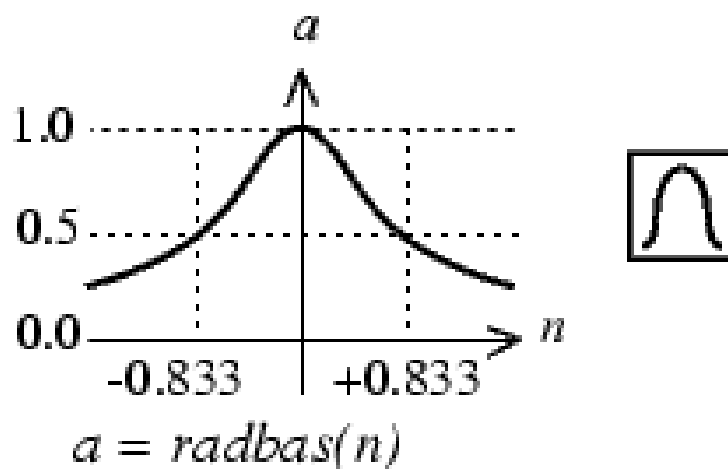
Прагова



Сигмоидална

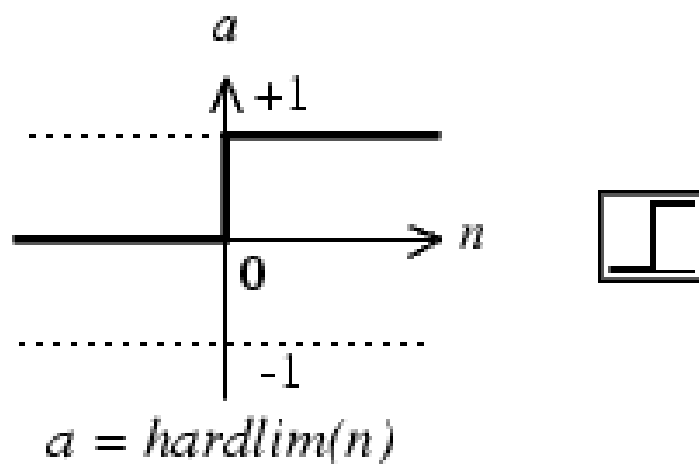


Активиращи функции



Radial Basis Function

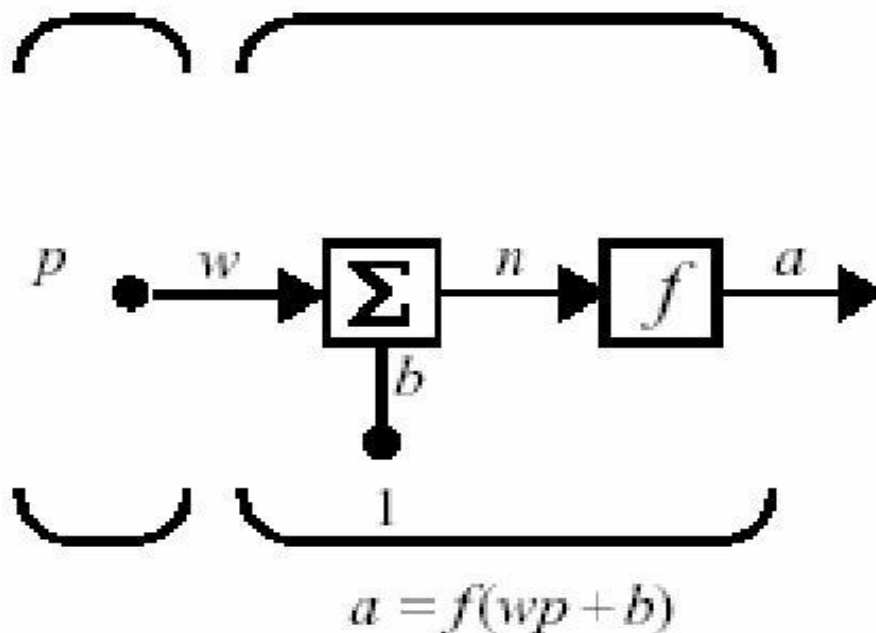
Активиращи функции



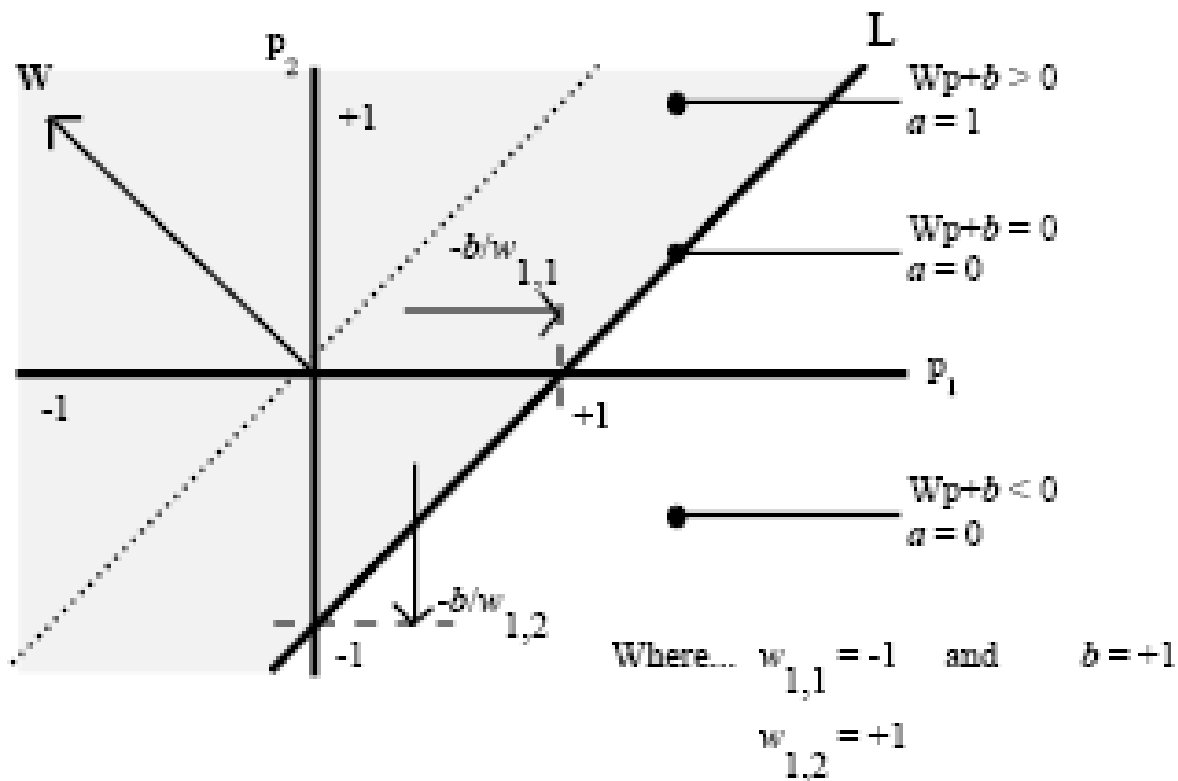
Hard-Limit Transfer Function

Неврон с отместване

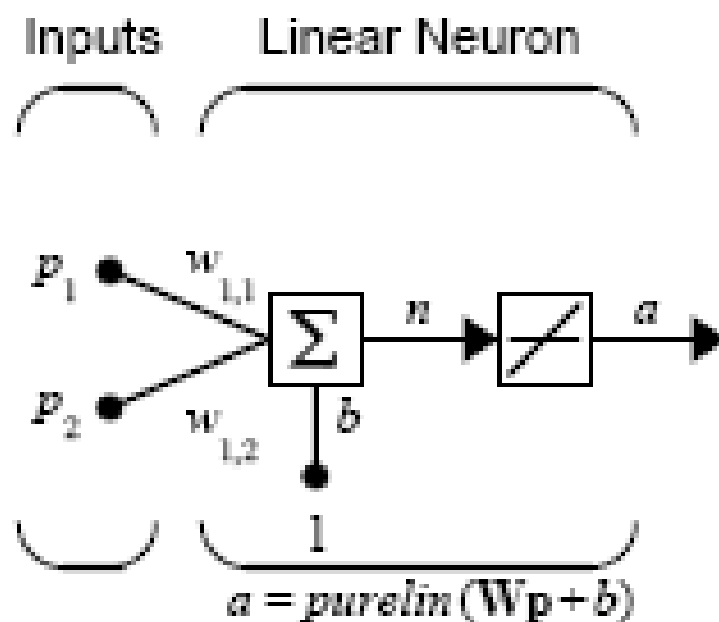
Вход Неврон с отместване



Невронни мрежи



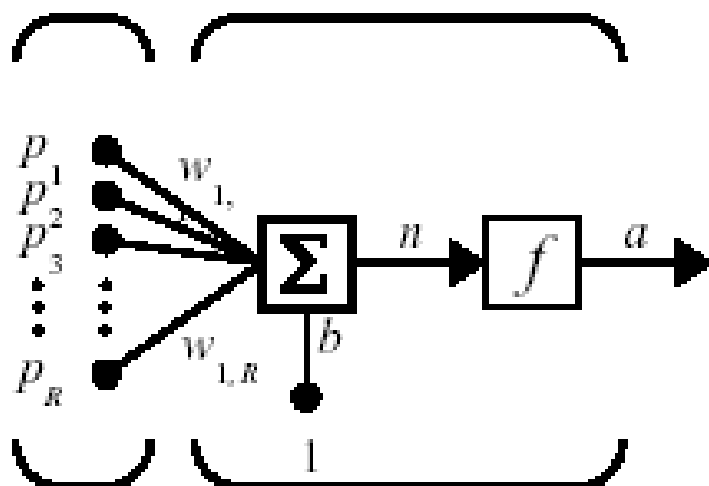
Неврон с паралелен вход



Неврон в статична мрежа с паралелен вход

Неврон с векторен вход

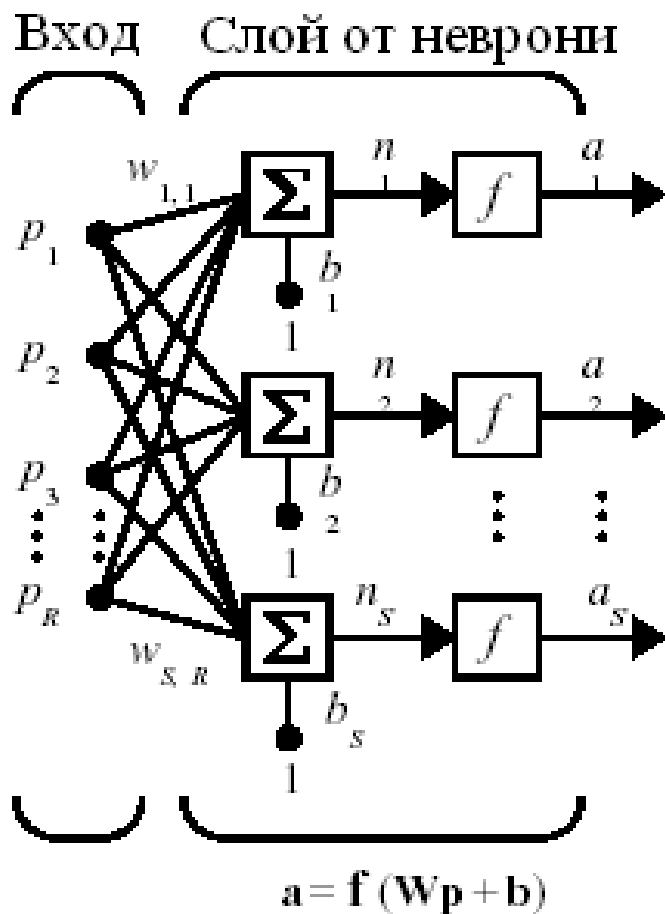
Вход Неврон с векторен вход



R – брой на
елементите
на вх. вектор

$$a = f(\mathbf{Wp} + b)$$

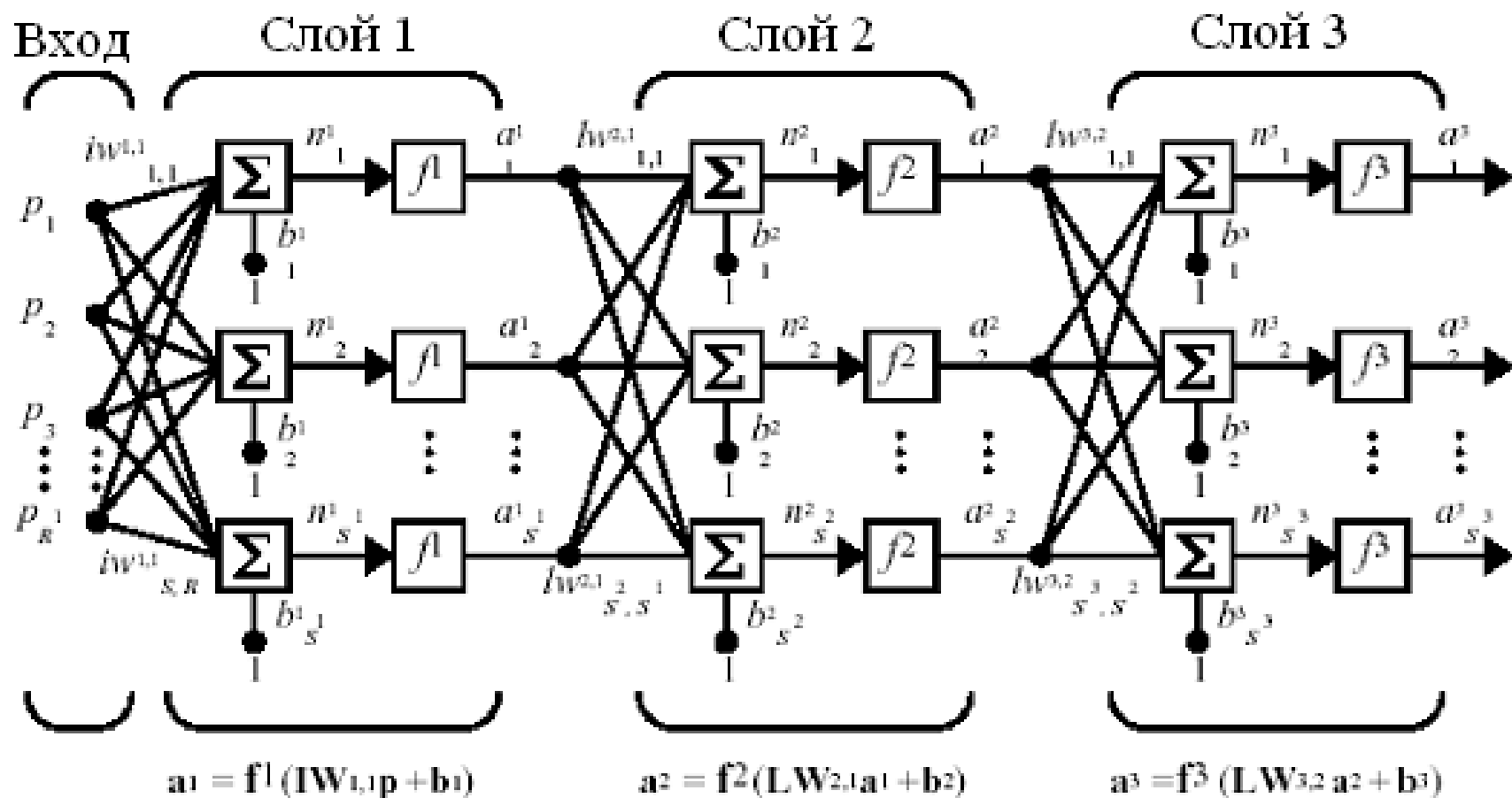
Невронни мрежи – 1-слойна арх.



Еднослойна невронна
мрежа без обратни връзки
→

Feedforward Neural
Network

Невронни мрежи – 3 слоя арх.

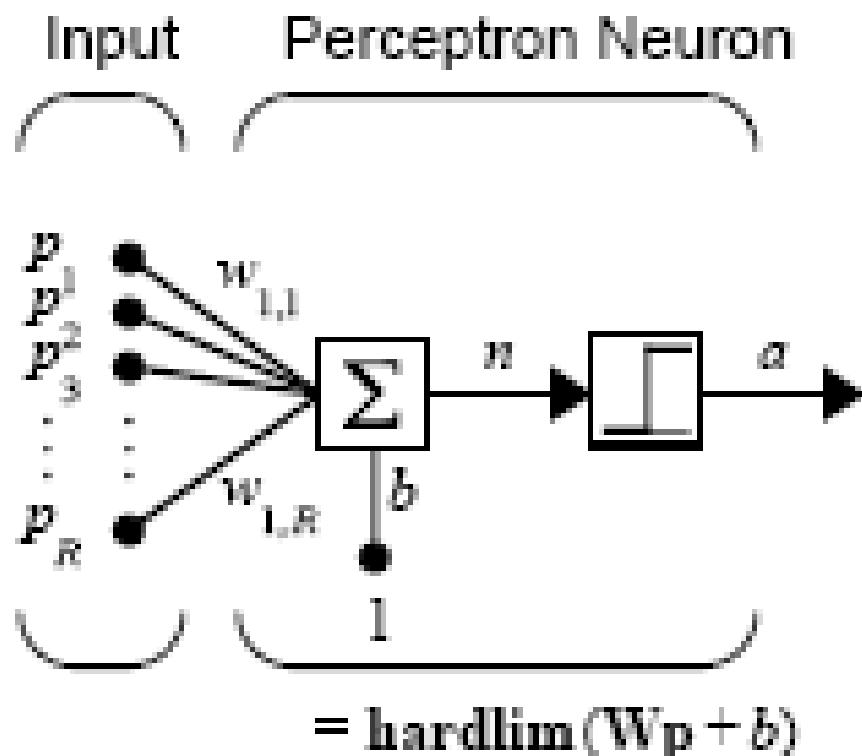


Входен слой

Скрити слоеве (1 и 2)

Изходен слой (3)

Неврони -- Перцептрон

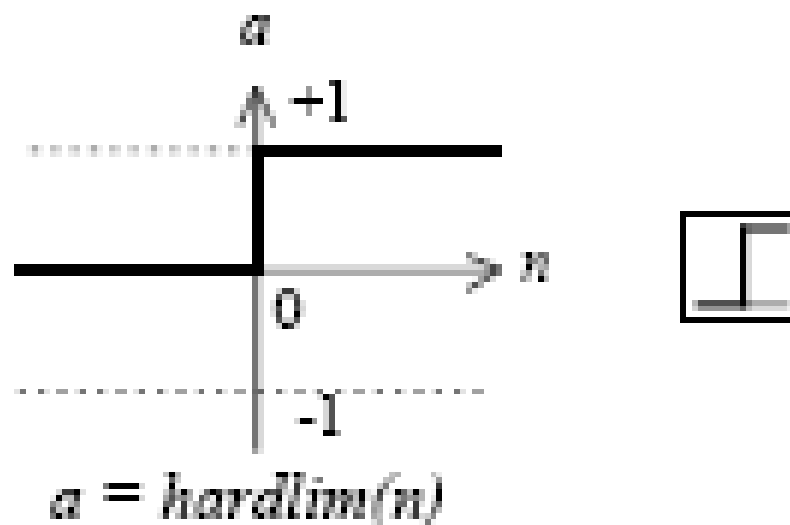


Where...

$R = \#$ Elements
in input vector

Когато многослойната невронна мрежа е изградена от перцептони се използва понятието
Multilayer Perception Neural Network (MPLNN)

Перцептрон – Активираща функция

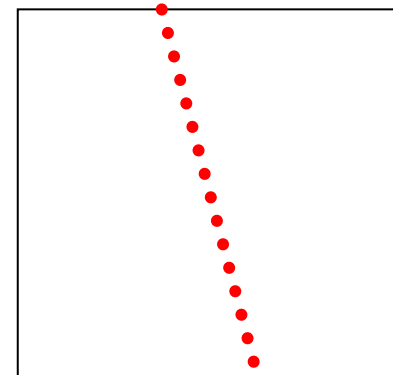
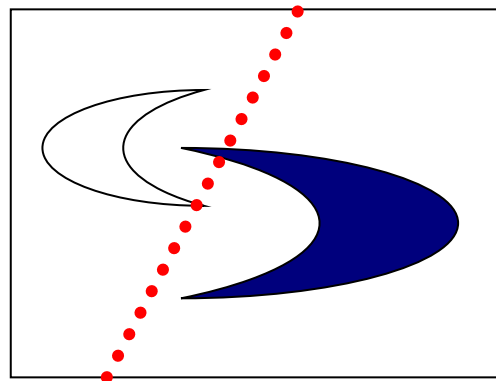
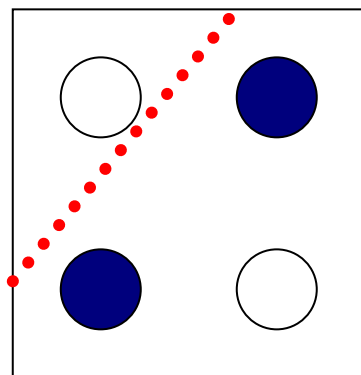
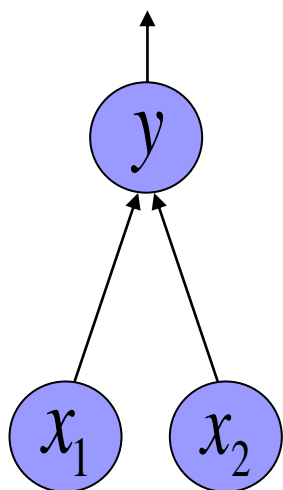


Hard Limit Transfer Function

Какви проблеми решават MLP NN?

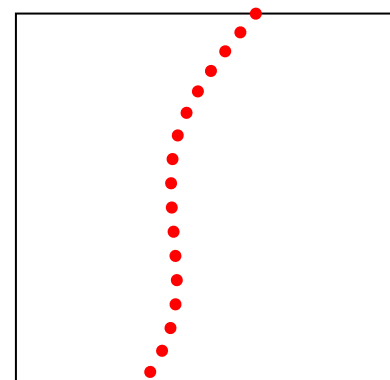
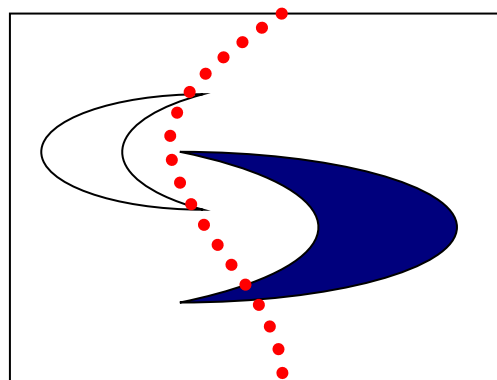
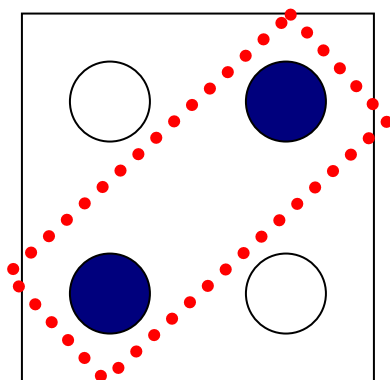
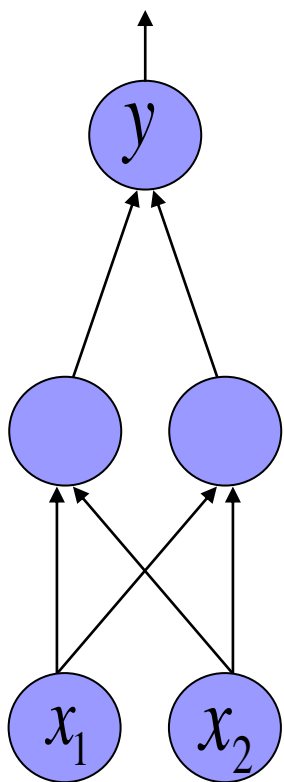
- При 0 скрити слоя: линеен класификатор
 - Хипер-равнини

2D пример:



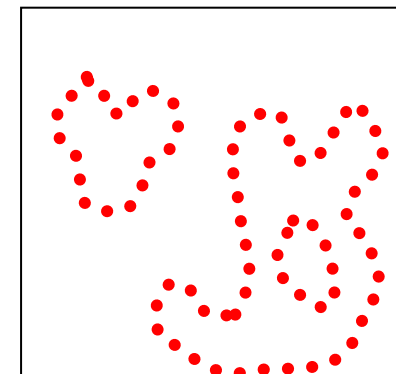
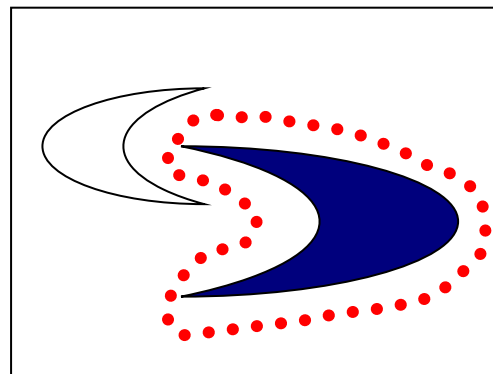
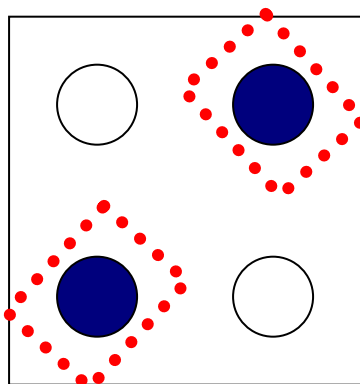
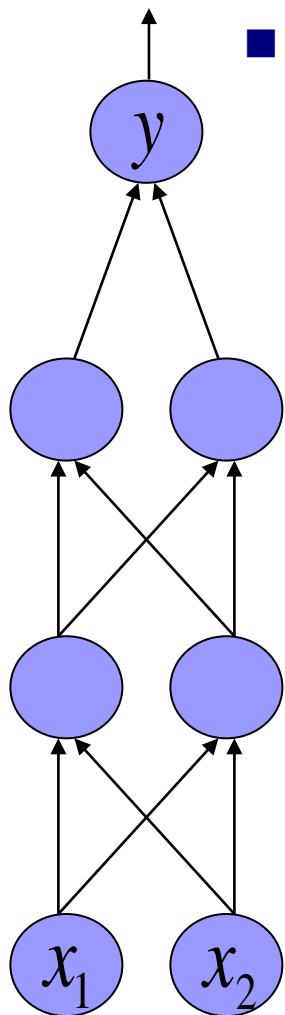
Какви проблеми решават MLP NN?

- При наличие на 1 скрит слой
 - Граница на конвексна област (отворен или затворен контур)



Какви проблеми решават MLP NN?

- При наличие на 2 скрити слоя
 - Комбинации от конвексни области



Съдържание

- ❑ Обща класификация на методите
 - Дискриминативни класификатори
 - Генеративни класификатори
 - Хибридни методи
- ❑ Класификация с невронни мрежи. Архитектури.
Невронни мрежи без обратни връзки.
- ❑ Многослойна невронна мрежа от перцептрони (MLPNN)
- ❑ Вероятностна невронна мрежа (PNN).

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

Вероятностна невронна мрежа

Вероятностна невронна мрежа (Probabilistic Neural Network)

- изчислява функция на разпределение на данните във всеки клас и
- взема решение въз основа на правилото на Бейс за класификация с минимален риск

Правило на Бейс

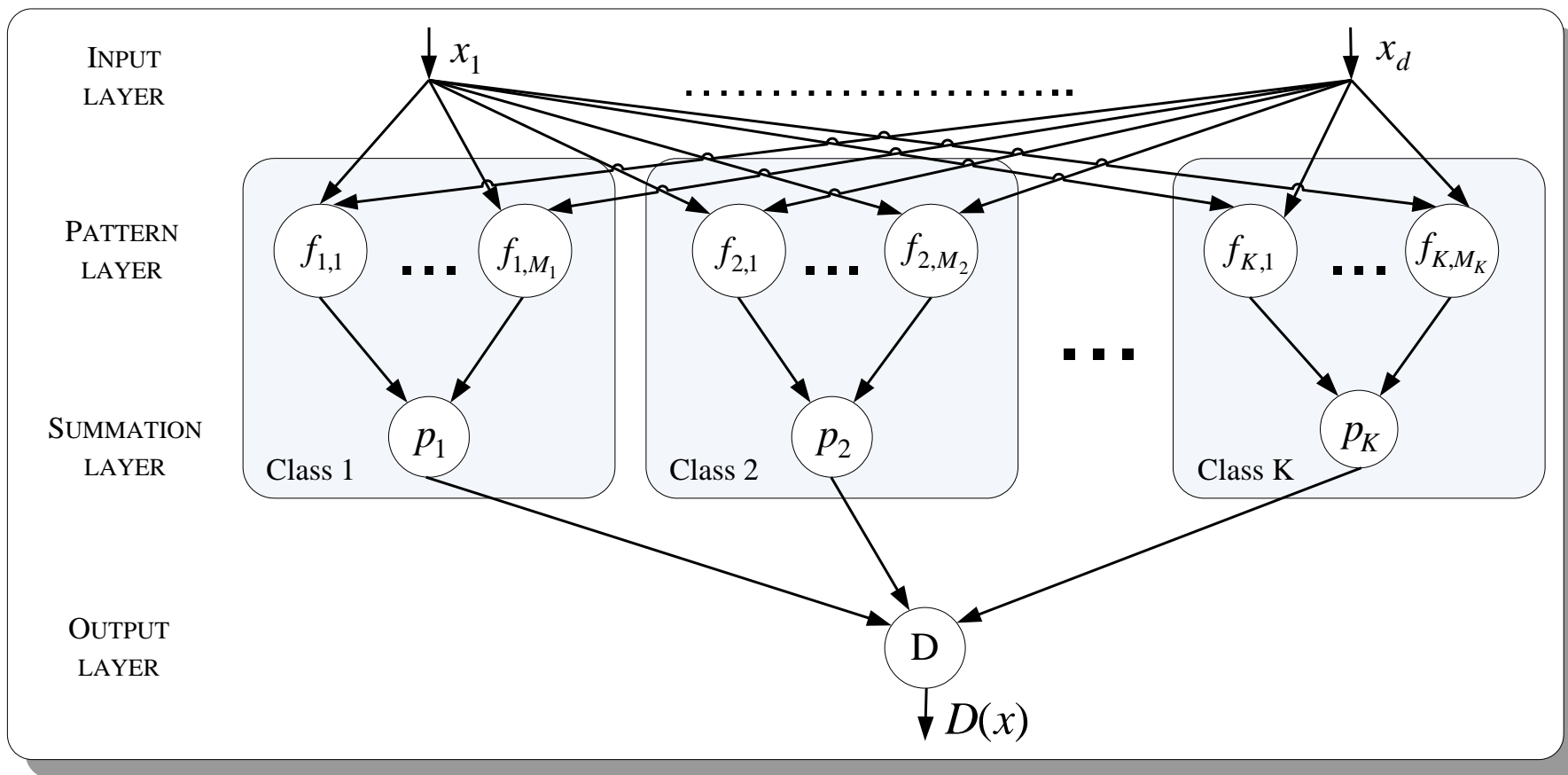
$$P(\text{Hypothesis} \mid \text{Evidence}) = P(\text{Hypothesis}) \times \frac{P(\text{Evidence} \mid \text{Hypothesis})}{P(\text{Evidence})}$$

Вероятността хипотезата да е вярна, при наблюдаваните данни, е равна на предварителната вероятност хипотезата да е вярна, умножена по вероятността доказателствата да са налице, като се има предвид, че хипотезата е вярна, разделена на пределната вероятност доказателствата да са налице при всякакви обстоятелства.

Терминология: Posterior = Prior x (Likelihood over Marginal probability)

$$\underbrace{P(A|B)}_{\text{posterior}} = \underbrace{P(A)}_{\text{prior}} \times \frac{\underbrace{P(B|A)}_{\text{likelihood}}}{\underbrace{P(B)}_{\text{marginal}}}$$

Probabilistic Neural Network



Probabilistic Neural Networks

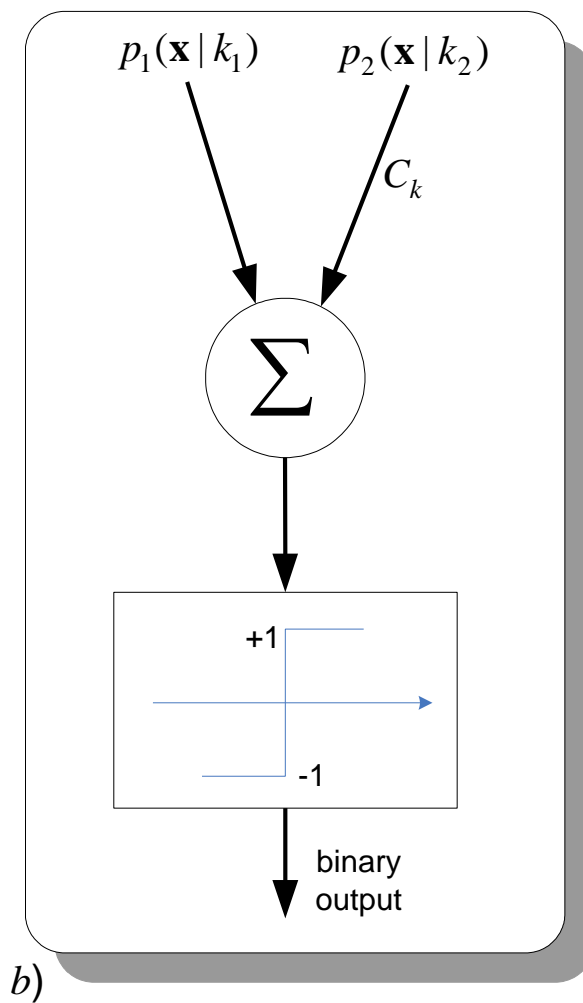
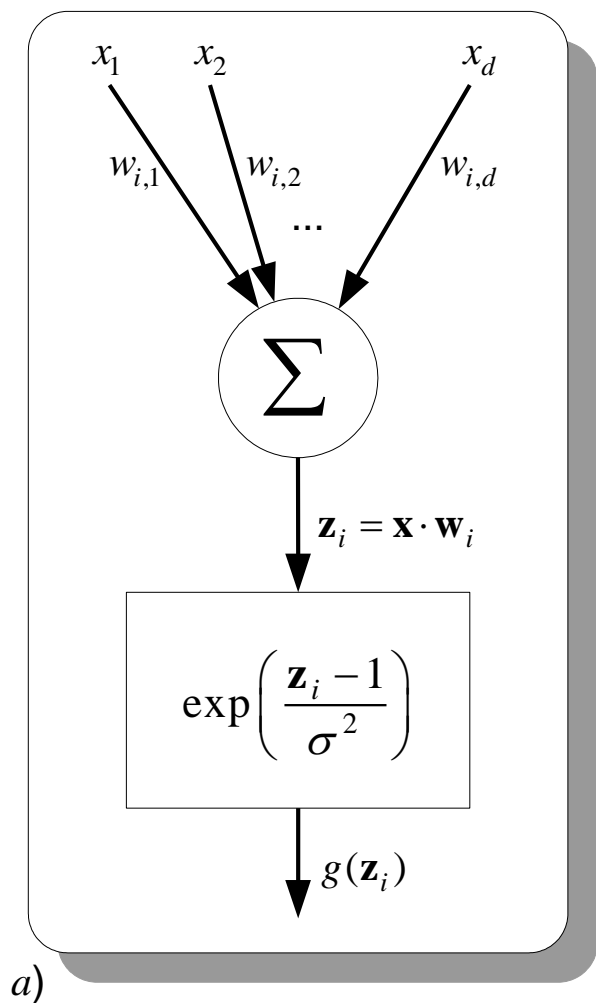
$$p_i(\mathbf{x} | k_i) = \frac{1}{(2\pi)^{d/2} \sigma^d} \cdot \frac{1}{M_i} \sum_{j=1}^{M_i} \exp\left(-\frac{1}{2\sigma^2} (\mathbf{x} - \mathbf{x}_{ij})^T (\mathbf{x} - \mathbf{x}_{ij})\right) \quad i = 1, \dots, K,$$

$$D(\mathbf{x}) = \underset{i}{\operatorname{argmax}} \{P(k_i) p_i(\mathbf{x} | k_i)\}, \quad i = 1, \dots, K,$$

$$P(k_i | \mathbf{x}) = \frac{P(k_i) p_i(\mathbf{x} | k_i)}{\sum_{j=1}^K P(k_j) p_j(\mathbf{x} | k_j)}, \quad i = 1, \dots, K.$$

Probabilistic Neural Networks

(a) i -th neuron in the pattern layer, (b) a neuron in the output layer

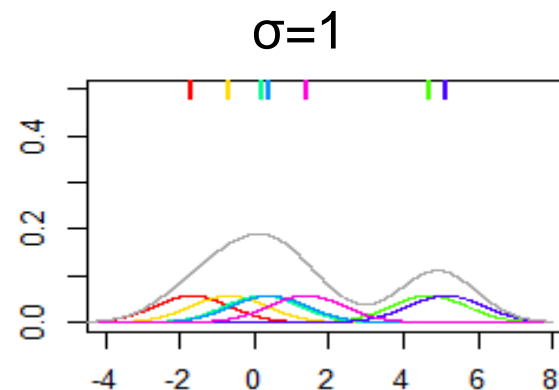
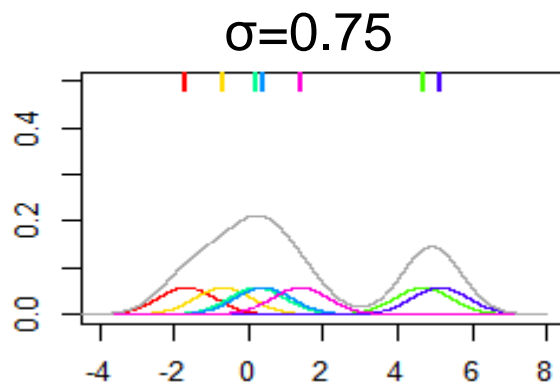
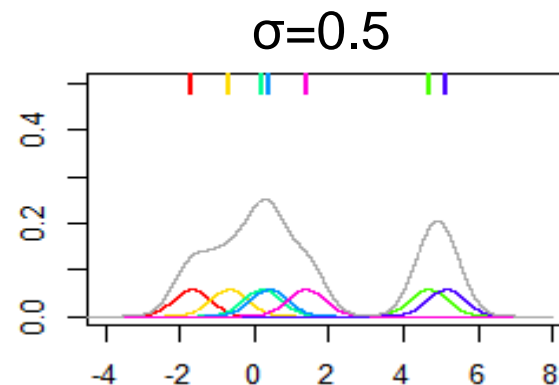
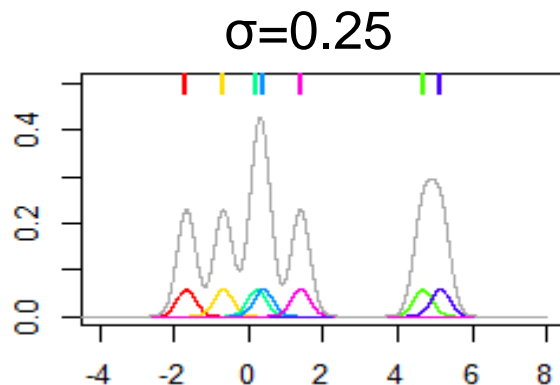


$$g(\mathbf{z}_i) = \exp\left(\frac{\mathbf{z}_i - 1}{\sigma^2}\right)$$

$$C_k = -\frac{P_2(k_2) \cdot c_{2,k}}{P_1(k_1) \cdot c_{1,k}} \cdot \frac{M_1}{M_2}$$

Probabilistic Neural Networks

$$f_i(\mathbf{x}_p) = \frac{1}{(2\pi)^{d/2} \sigma^d} \frac{1}{M_i} \sum_{j=1}^{M_i} \exp \left(-\frac{1}{2\sigma^2} (\mathbf{x}_p - \mathbf{x}_{ij})^T (\mathbf{x}_p - \mathbf{x}_{ij}) \right),$$



Probabilistic Neural Networks

$$f_i(\mathbf{x}_p) = \frac{1}{(2\pi)^{d/2} \sigma^d} \frac{1}{M_i} \sum_{j=1}^{M_i} \exp \left(-\frac{1}{2\sigma^2} (\mathbf{x}_p - \mathbf{x}_{ij})^T (\mathbf{x}_p - \mathbf{x}_{ij}) \right),$$

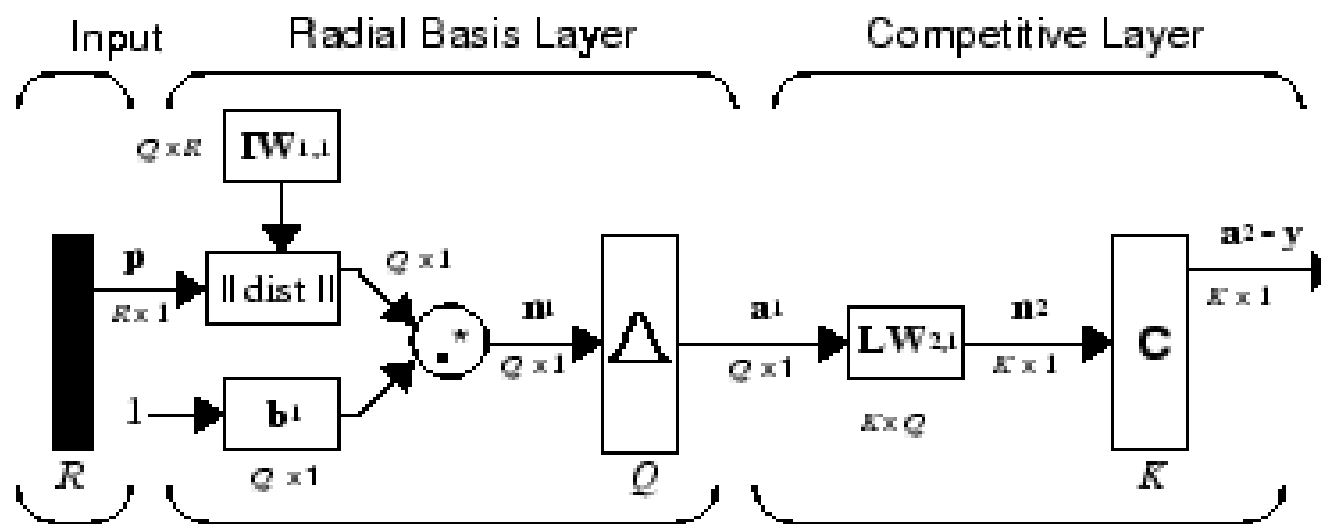
$$D(\mathbf{x}_p) = \underset{i}{\operatorname{argmax}} \left\{ h_i \ c_i \ f_i(\mathbf{x}_p) \right\}, \quad i = 1, 2$$

$$P(k_i | \mathbf{X}) = \frac{N(D(\mathbf{x}_p) = k_i)}{\sum_{j=1}^K N(D(\mathbf{x}_p) = k_j)}, \quad i = 1, 2,$$

$$\chi = \eta \left(P(k_1 | \mathbf{X}) - \beta \right)$$

$$O(\theta) = \begin{cases} 1 & \text{for } \chi \geq \theta \\ 0 & \text{for } \chi < \theta \end{cases}$$

Probabilistic Neural Networks



Where...

R = number of elements in input vector

$$a_i^1 = \text{radbas}(\|\mathbf{IW}_{1,1} - \mathbf{p}\| \mathbf{b}^1)$$

$$\mathbf{a}^2 = \text{compet}(\mathbf{LW}_{2,1} \mathbf{a}^1)$$

a_i^1 is i th element of \mathbf{a}^1 where $\mathbf{IW}_{1,1}$ is a vector made of the i th row of $\mathbf{IW}_{1,1}$

Q = number of input/target pairs = number of neurons in layer 1
 K = number of classes of input data = number of neurons in layer 2

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