



PERCEPTRONŲ TINKLAS



Paskaitos
Turinys

1. Elementarus perceptronų tinklas

- Apibrėžimas
- Struktūra
- Veikimas

2. Perceptronų tinklo struktūros

- Nuosekli struktūra
- Kryžminė struktūra
- Grįžtamųjų ryšių struktūra

3. Perceptronų tinklo mokymas

- Pirmos eilės mokymas
- Antros eilės mokymas

4. Perceptronų tinklo atsakas

5. Perceptronų tinklo savybės

- Konvergencija
- Privalumai
- Trūkumai
- Naudojimas



Pagrindinė
Literatūra

1. Simpson, P.K. (1990). *Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations*. Pergamon Press, pp. 100-105.
2. Rosenblatt, F. (1962). *Principles of Neurodynamics*. Washington: Spartan Books.
3. MATLAB ver. 4.0 (1994). *Neural Networks: Toolbox*. Math Corp., 3 chapter.

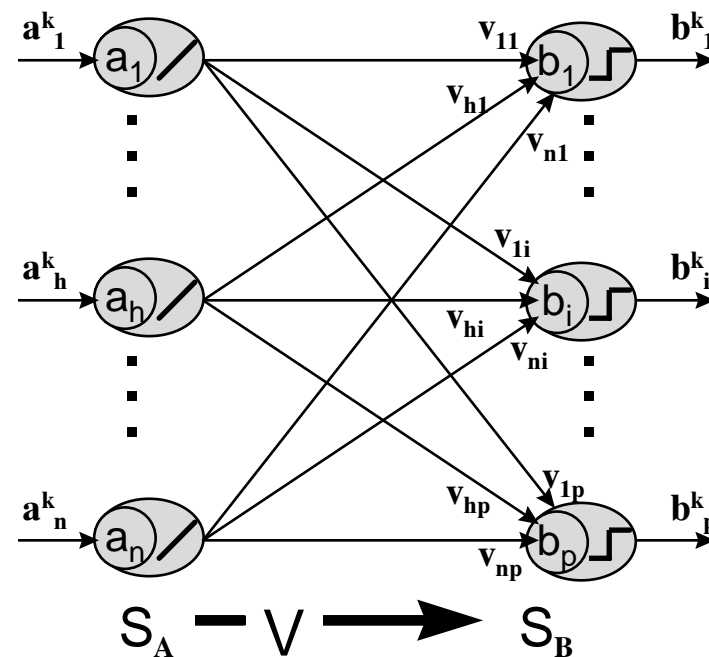


Elementarus perceptronų tinklas

Apibrėžimas

Elementarus perceptronų (angl. k. - perceptron) tinklas yra dviejų sluoksnių heteroasociatyvus artimiausios reikšmės pavyzdžio atitikmens radimo įrenginys, kuris išimena pavyzdžių poras (A^k , B^k), $k=1, 2, \dots, m$ naudodamas perceptrono klaidos korekcijos procedūrą, kur k -toji pavyzdžių pora pateikiama analoginės reikšmės vektoriumi $A^k = (a_1^k, a_2^k, \dots, a_n^k)$ ir dvejetainės $\{-1; +1\}$ reikšmės vektoriumi $B^k = (b_1^k, b_2^k, \dots, b_p^k)$. Elementarus perceptronų tinklas mokomas nerealiame laiko mastelyje, veikia diskrečiame laike ir yra atvaizduojamas dviejų sluoksnių nuoseklių ryšių struktūra.

Struktūra



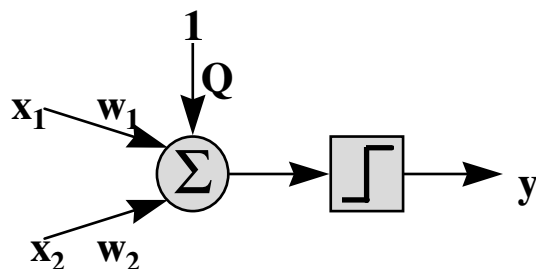
Pav. 4.1



Elementarus perceptronų tinklas (tęsinys)

Veikimas

Vienintelis perceptronas turintis tik du įėjimus jau gali veikti kaip elementarus klasifikatorius:

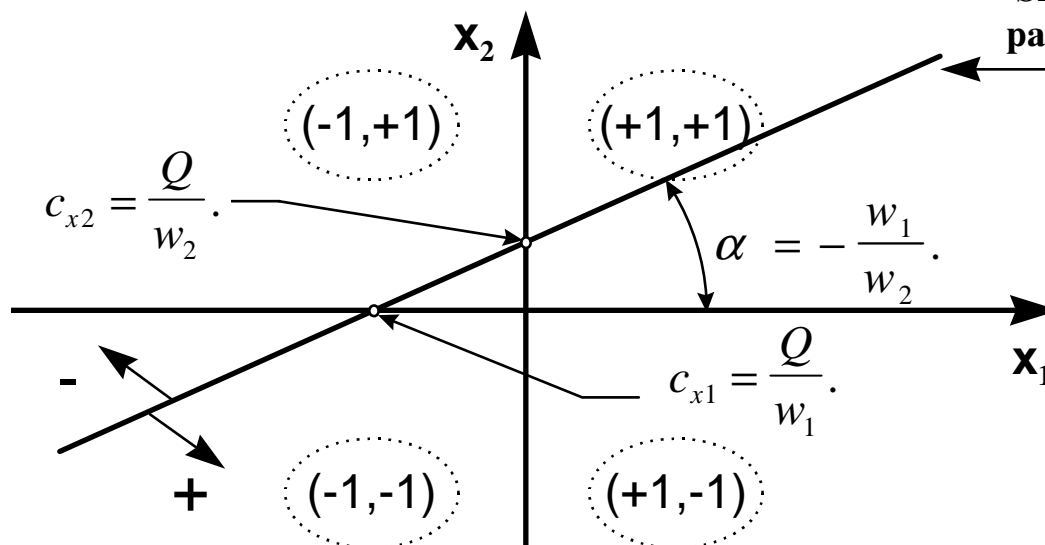


Pav. 4.2

Norint rasti ribą tarp būsimų klasių reikia perceptrono išėjimo signalo lygtį prilyginti nuliui ir ją išspręsti bet kurio įėjimo signalo atžvilgiu:

$$y = x_1 w_1 + x_2 w_2 - Q = 0,$$

$$x_2 = -\frac{w_1}{w_2} x_1 + \frac{Q}{w_2}.$$



Pav. 4.3



Elementarus perceptronų tinklas (tęsinys)

Veikimas (1 pavyzdys)

Turint 4 įėjimo pavyzdžius (X) reikia išmokyti perceptroną juos suklasifikuoti į 2 klases (Y) taip:

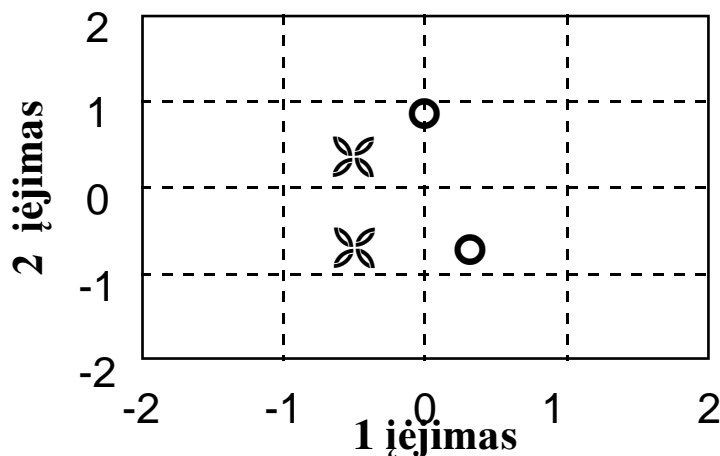
$$X = \begin{bmatrix} -0.5 & -0.5 & 0.3 & 0.0; \\ -0.5 & 0.5 & -0.5 & 1.0 \end{bmatrix}.$$

$$Y = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}.$$

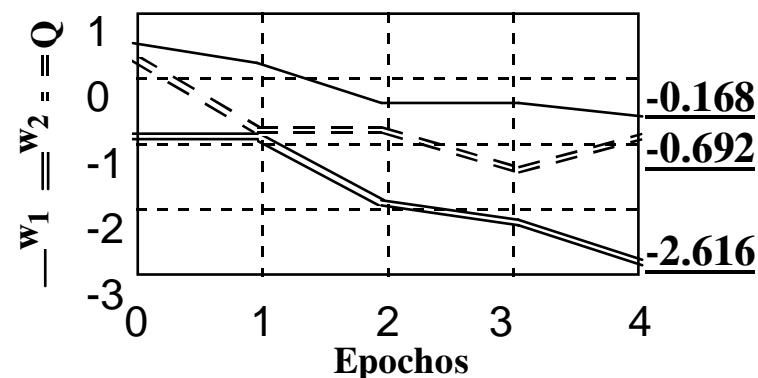
Pažymėjimai:

○ = 0, ✕ = 1.

Įėjimo pavyzdžių išsidėstymas atrodo taip:



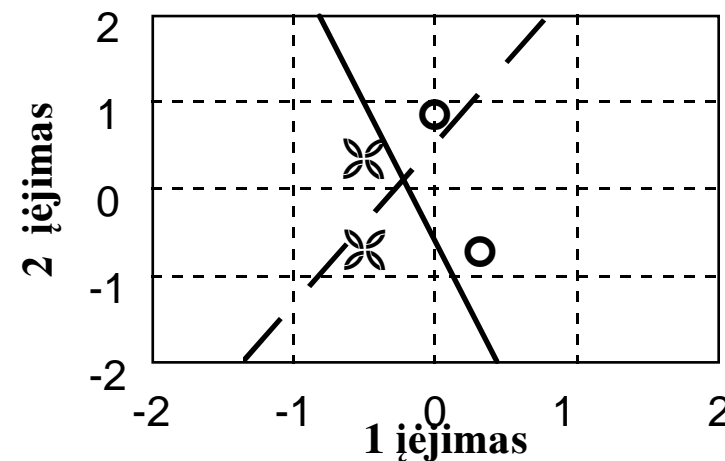
Svorių bei slenksčio reikšmių kitimas perceptronų tinklo mokymo metu



Skaidrės Nr.

4

Perceptrono išmoktas duotų pavyzdžių klasifikavimas

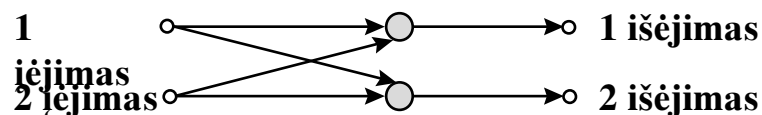




Elementarus perceptronų tinklas (tęsinys)

Veikimas (2 pavyzdys)

Turint dviejų perceptronų tinklą:



reikia išmokyti jį teisingai klasifikuoti pavyzdžius į 4 skirtingas klases, mokymui turint 10 įėjimo pavyzdžių (X) ir juos atitinkančių atsakų (Y):

$X = [0.1 \ 0.7 \ 0.8 \ 0.8 \ 1.0 \ 0.3 \ 0.0 \ -0.3 \ -0.5 \ -1.5;$
 $1.2 \ 1.8 \ 1.6 \ 0.6 \ 0.8 \ 0.5 \ 0.2 \ 0.8 \ -1.5 \ -1.3];$
 $Y = [1 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0;$
 $0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1].$

Žymėjimas:

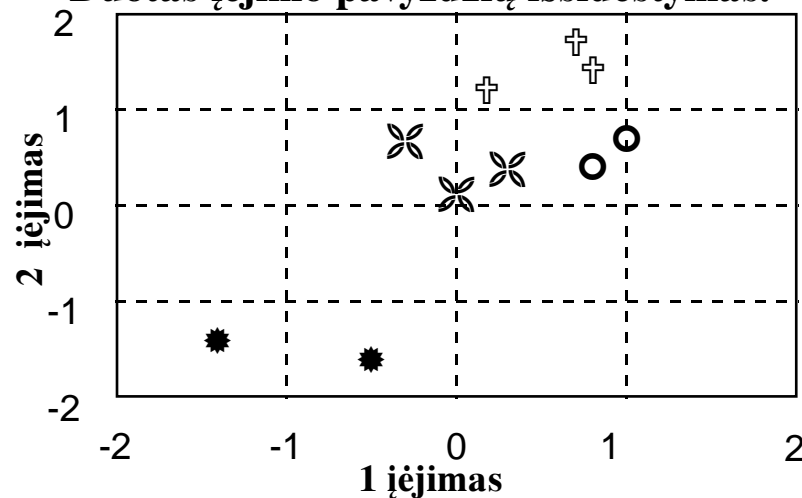
○ = 00;

⊛ = 01;

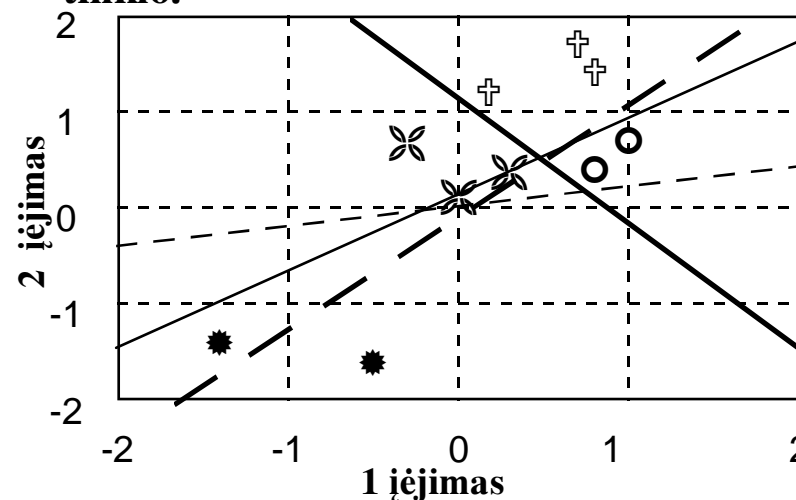
⊕ = 10;

⊗ = 11.

Duotas įėjimo pavyzdžių išsidėstymas:



Klasifikavimas išmoktas perceptronų tinklo:



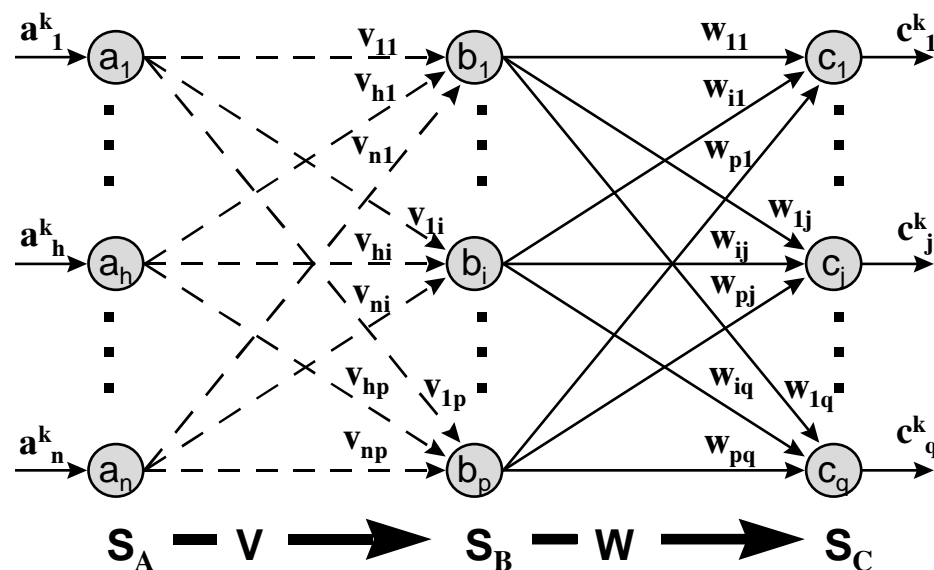
Skaidrės Nr.
5



Perceptronų tinklo struktūros

“Principles of Neurodynamics” (Rosenblatt, 1962) veikale yra aptariamos trys pagrindinės perceptronų tinklo struktūros: *nuosekli, kryžminė ir grįžtamųjų ryšių*.

Nuosekli struktūra



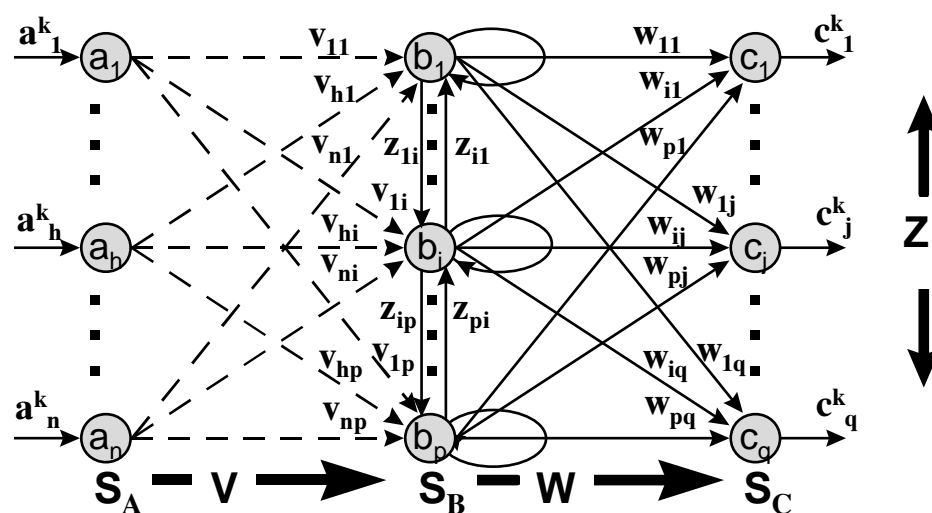
Pav. 4.4

✎ Kokia bebūtų perceptronų tinklo struktūra ryšiai tarp įėjimo sluoksnio S_A ir paslėptojo sluoksnio S_B yra *sudaromi atsitiktiniu būdu*, o jų *svoriai yra nekintantys*. Nuoseklių ar grįžtamųjų ryšių tarp paslėptojo sluoksnio S_B ir išėjimo sluoksnio S_C svoriai yra *adaptyvūs*, o tuo tarpu vidinių ryšių paslėtajame sluoksnyje S_B svoriai yra *nekintantys*.

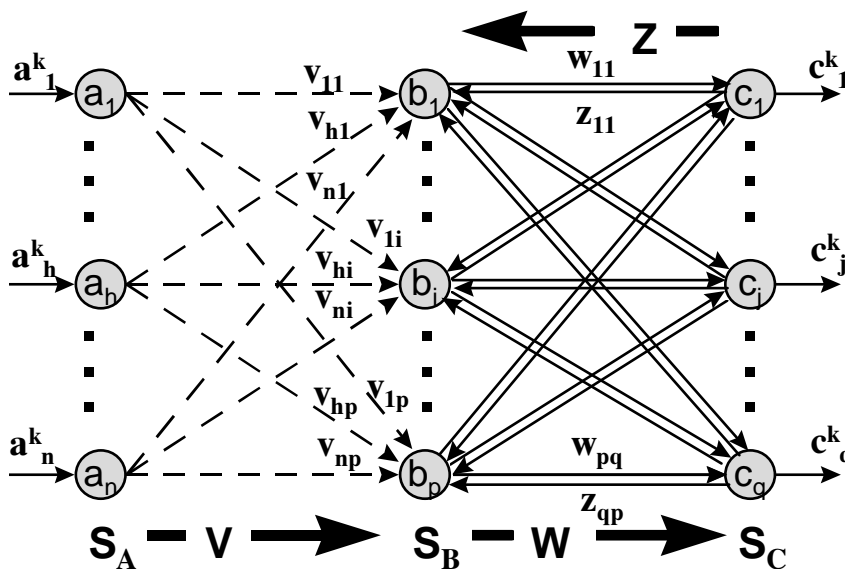


Perceptronų tinklo struktūros (tęsinys)

Kryžminė struktūra



Grižtamųjų ryšių struktūra



Pav. 4.5, 4.6



Perceptronų tinklo mokymas

Pirmos eilės mokymas (perceptrono klaidos korekcijos procedūra)

A. Ryšių bei slenksčių inicializavimas

Visiems adaptuojamiems perceptronų ryšių w_{ij} bei slenksčių Q_j svoriams priskiriamos atsitiktinės reikšmės (+1,-1) ribose.

B. Įėjimo pavyzdžių pateikimas

Įėjimo pavyzdžiai A^k , $k=1,...,n$ pateikiami perceptronų tinklui ir gaunami DNT atsakai B^k :

$$b_j^k = F\left(\sum_{i=1}^n w_{ij}^k \cdot a_i^k - Q_j^k\right), \text{ kai } F(x) = \begin{cases} +1, & x > 0; \\ -1, & x \leq 0. \end{cases}$$

C. Tikrinimas

Skaičiuojama klaida D^k tarp esamų B^k ir norimų gauti B'^k atsakų:

$$d_j^k = b_j'^k - b_j^k,$$

jei klaida norimai maža arba lygi nuliui, tai peršokama į algoritmo E žingsnį.

D. Ryšių bei slenksčių adaptavimas

Prie esamų ryšių bei slenksčių svorių reikšmių pridedama:

$$\Delta w_{ij}^k = \lambda \cdot a_i^k \cdot d_j^k \quad \text{ir} \quad \Delta Q_j^k = \lambda \cdot d_j^k, \text{ kur } \lambda - \text{mokymo greičio konstantė.}$$

Grįžtama prie algoritmo B žingsnio.

E. Pabaiga



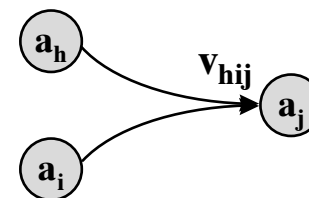
Perceptronų tinklo mokymas (tęsinys)

Antros eilės mokymas

Pirmos eilės mokymo algoritme koreguojami du žingsniai leidžiantys įvertinti antros eilės (naujus) ryšius tarp neuronų trejetų:

Nauja algoritmo B žingsnio atsako skaičiavimo formulė:

$$b_j^k = F \left(\sum_{i=1}^n w_{ij}^k \cdot a_i^k + \sum_{h=1}^n \sum_{i=1}^n v_{hij}^k \cdot a_h^k \cdot a_i^k - Q_j^k \right).$$



Algoritmo C žingsnį papildanti antros eilės ryšių svorių adaptavimo formulė:

$$\Delta v_{hij}^k = \lambda \cdot a_h^k \cdot a_i^k \cdot d_j^k.$$

Atsakas

- Naudojant pirmos eilės mokymą

$$b_j^k = F \left(\sum_{i=1}^n w_{ij}^k \cdot a_i^k - Q_j^k \right).$$

- Naudojant pirmos ir antros eilės mokymą drauge

$$b_j^k = F \left(\sum_{i=1}^n w_{ij}^k \cdot a_i^k + \sum_{h=1}^n \sum_{i=1}^n v_{hij}^k \cdot a_h^k \cdot a_i^k - Q_j^k \right).$$



Perceptronų tinklo savybės

Konvergencija

Rosenblatt (1960) įrodė, kad perceptronų mokymo algoritmas visada suras sprendinį bet kokiam tiesiškai atskiriamam uždaviniui per baigtinį laiko intervalą.

Privalumai

- lengvai suprantamas veikimas
- adekvati saugojimo talpa
- betarpiškas atsakas

Trūkumai









- blogos apibendrinimo savybės
- ilgas mokymas nerealiame laiko mastelyje
- negali įsisavinti netiesiškai išskiriamų klasių

Naudojimas

- žinių apdorojime
- pavyzdžių atpažinime
- kalbos atpažinime
- sekų atpažinime
- vaizdo apdorojime
- valdymo uždaviniuose



Pagrindiniai teiginiai

-  Perceptronų tinkluose naudojamos slenkstinės neuronų aktyvavimo funkcijos, o jie mokinami naudojant perceptronų mokymą.
-  Yra išskiriamos trys pagrindinės daugiasluoksnių perceptronų struktūros: nuosekli, kryžminė bei grįžtamųjų ryšių.
-  Perceptrono klaidos korekcijos procedūros esmė: svorių pokyčiai yra tiesiogiai priklausomi nuo tinklo klaidos dydžio.
-  Mokant perceptronų tinklą vyksta hiperplokštumos pozicionavimas įėjimo pavyzdžių erdvėje siekiant reikiamai atskirti pavyzdžius.
-  Aukštesnės eilės perceptronų tinklo mokymas gali žymiai pagerinti mokymo charakteristikas.
-  Perceptronų tinklų konvergencija yra įrodyta.
-  Pagrindinis perceptronų privalumas - galimybė paprastai ir greitai klasifikuoti, didžiausias trūkumas - nesugebėjimas klasifikuoti tiesiškai neišskiriamų klasių.
-  Nuo 1960 m. kai perceptronų tinklai buvo išrasti, jie naudojami daugelyje sričių.