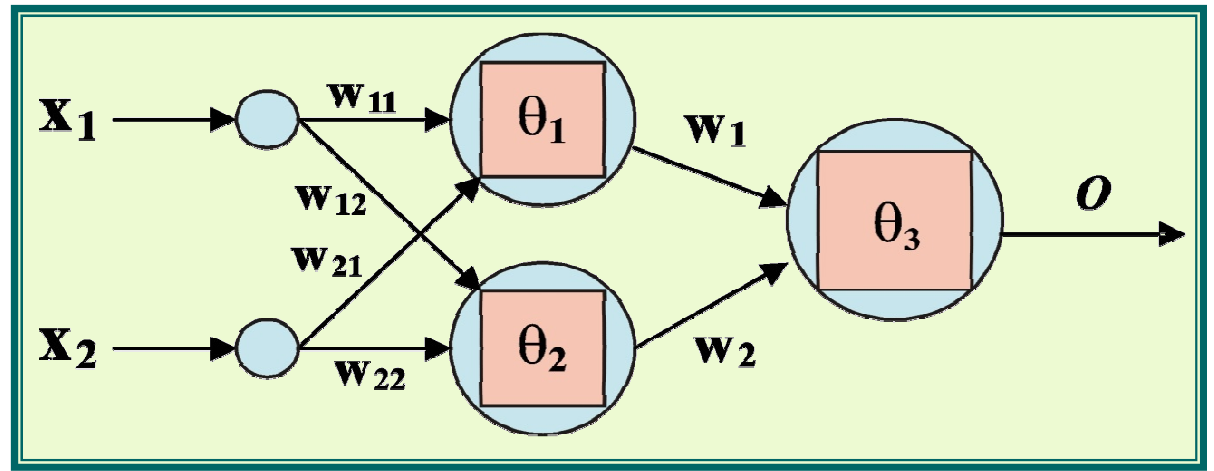
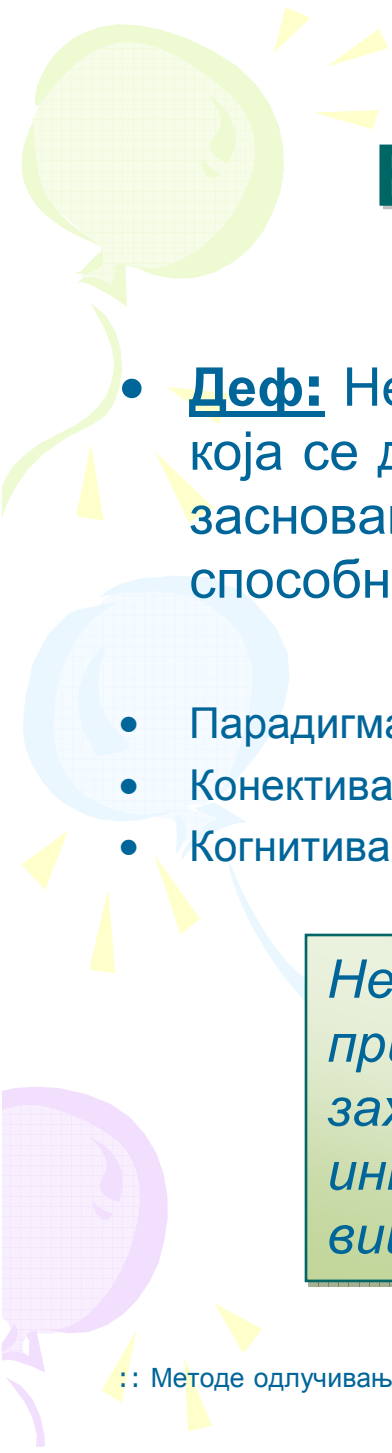




АТ-5 Вештачке неуронске мреже

Проф. др Зоран Миљковић

Методе одлучивања

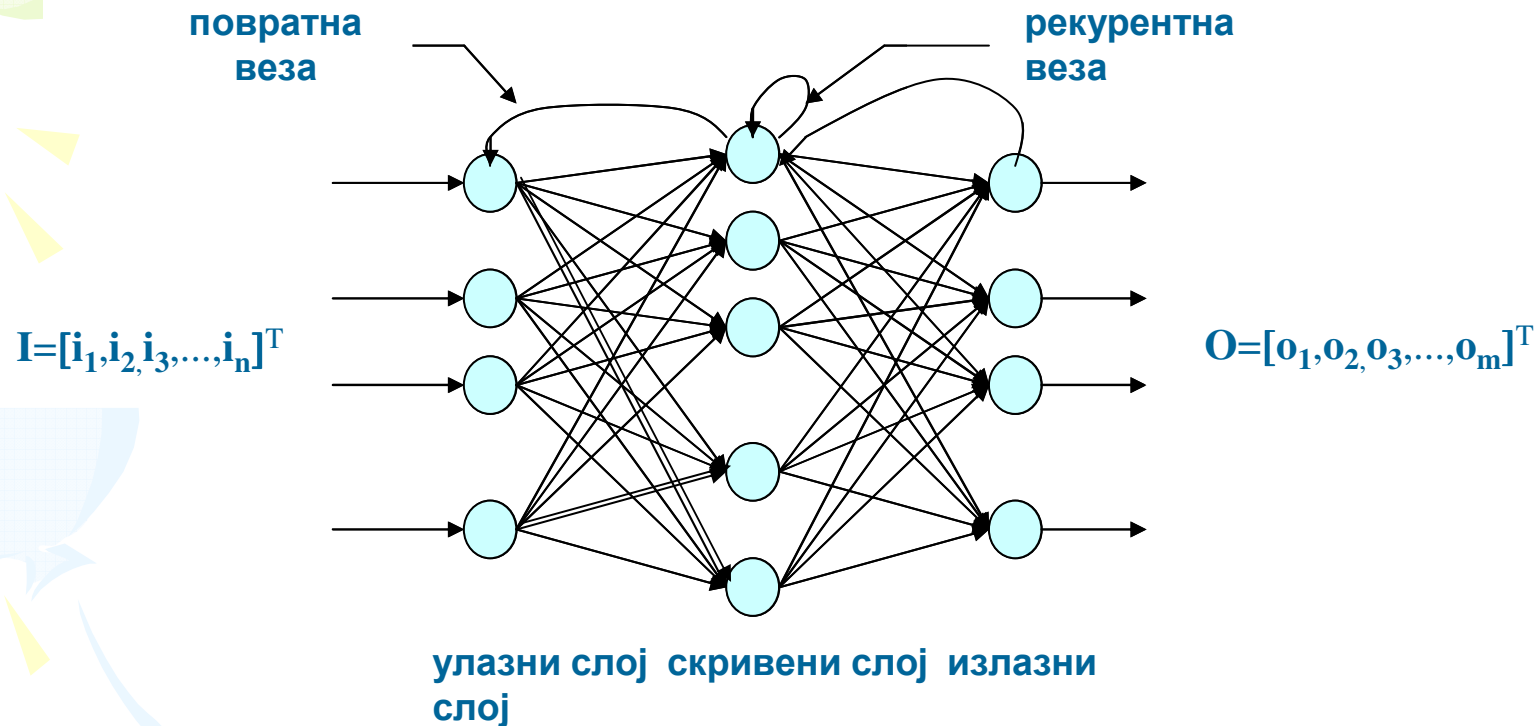


Вештачке неуронске мреже

- **Деф:** Неуронска мрежа је *парадигма* вештачке интелигенције која се дефинише као *конективни* модел за резоновање заснован на аналогiji са мозгом, уз наглашену *когнитивну* способност да учи и врши генерализацију стеченог знања.
- Парадигма – образац, узор, пример за углед;
- Конективан – повезан;
- Когнитиван – који се тиче сазнања, сазнајни, спознајни;

Неуронске мреже обезбеђују значајне предности при решавању проблема процесирања који захтевају рад у реалном времену и интерпретацију односа између променљивих у вишедимензионалним просторима.

Општа структура вештачке неуронске мреже



- У општем смислу, неуронске мреже представљају скуп једноставних процесирајућих елемената-неурона, међусобно повезаних везама са одговарајућим тежинским односима;
- Неуронске мреже имају способност адаптивног понашања према променама, што значи да могу да уче пресликавања између улазног и излазног простора и да синтетизују асоцијативну меморију.

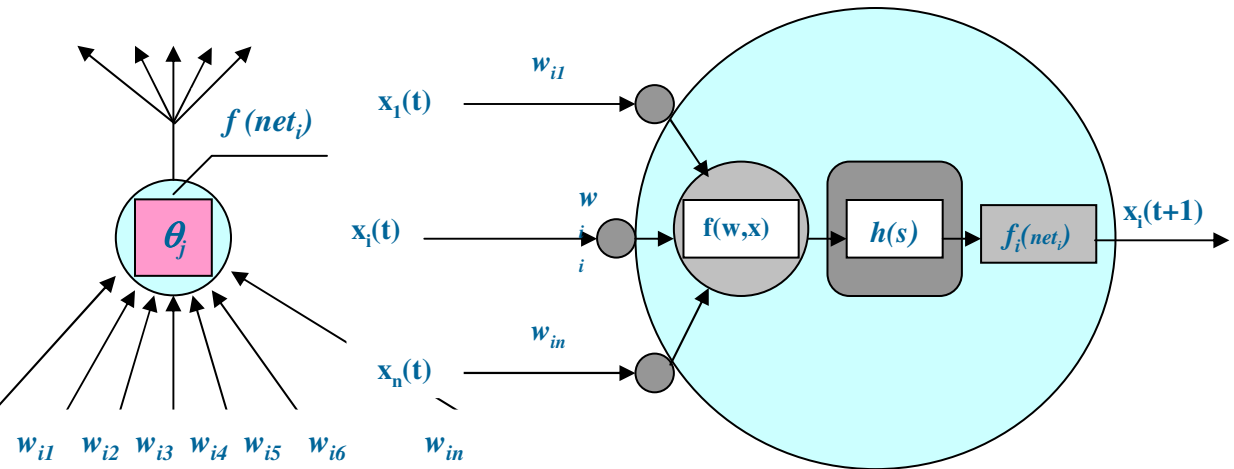
Основне компоненте

Три основне компоненте:

- Неурон;
- Топологија мреже;
- Алгоритам учења;

Додатне компоненте:

- Величина мреже;
- Функционалност неурона;
- Обучавање;
- Имплементација.



- Неурони (*процесирајући елементи*) примају улазне сигнале/информације од окружења или од других неурона преко веза које могу бити *екситационе* (побуђивачке) или *инхибиторне*.
- Праг активације неурона θ_j је његов „окидач”.



Неурон – основни процесирајући елемент

Три саставна елемента:

- **улазни оператор** $f(w, x)$, који обједињује улазе и тежинске односе међусобних веза w и формира јединствену вредност $s = f(w, x) = w^T x$;
- **функција преноса** $h(s)$, која обрађује излаз из неуронског улазног оператора (врши интеграцију), формирајући потребну вредност за активациону функцију;
- **активациона функција** $f_i(net_i)$ која обрађујући излаз функције преноса управља излазном вредношћу неурона.

Математичка формулација улазног оператора неурона

Тип улазног оператора	Математички израз	Напомена
Линеаран	$net_j = \sum w_{ij} x_j$	Линеарна функција
Тежинска сума	$net_j = \sum w_{ij} x_j + \theta_j$	Најчешће коришћен; у неким случајевима праг $\theta_j = 0$
Повратна веза	$net_j = \alpha net_{\text{старо}} + \beta \sum w_{ij} x_j$	α, β су тежинске константе

- Активација неурона (општа једначина):

$$a_i = F_i(a_i(t-1), net_i(t))$$

- Неурон генерише **излазни сигнал** који је у релацији са његовом активацијом преко **активационе функције**, што се математички може изразити на следећи начин:

$$o_i = x_i(t) = f_i(a_i(t)) = f_i(net_i(t))$$

Активационе функције неурона

Основне АФ:

- линеарне,
- бинарне,
- сигмоидне,
- компетитивне и
- Гаусове.

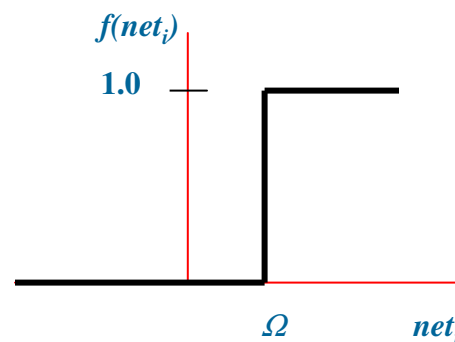
Линеарна активациона функција

- најједноставнија;
- сумира све улазе у неурон;

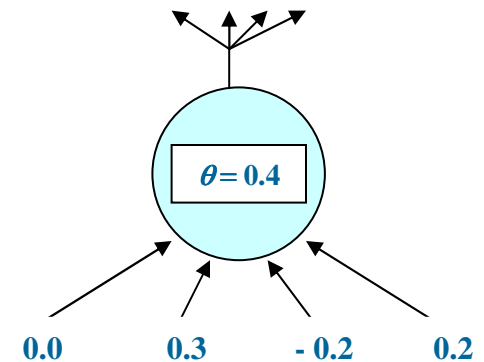
$$f_i^l(net_i(t)) = net_i(t)$$

Бинарна активациона функција

$$f_i^b(net_i(t)) = \begin{cases} 1 & \text{ако је } net_i(t) > \Omega \\ 0 & \text{остало} \end{cases}$$



a)



b)

Активационе функције неурона

Сигмоидна активациона функција

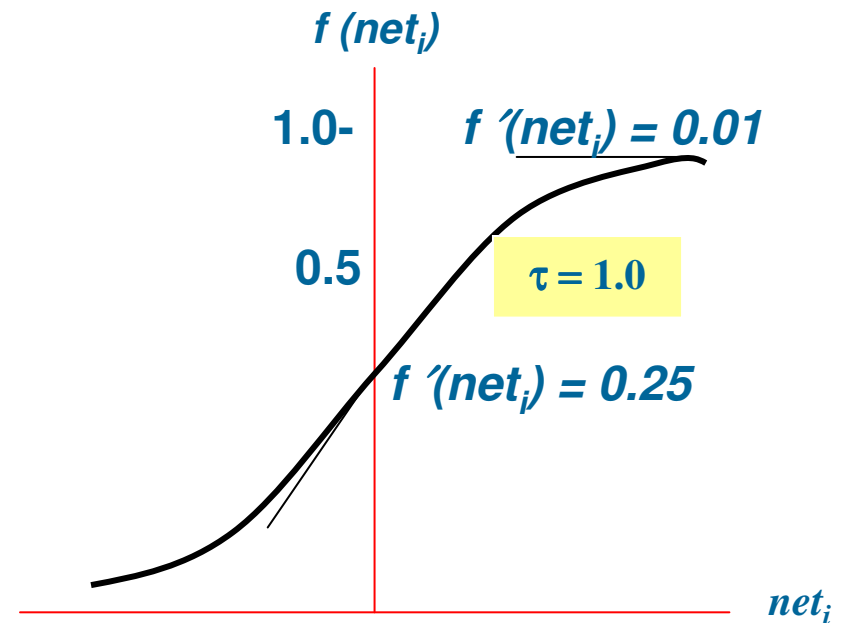
- Генерише излазни сигнал који има такође два стабилна стања;
- Континуална и диференцијабилна;
- Нелинеарна и неопадајућа.

Математичка формулација:

$$f_i^s(net_i(t)) = \frac{1}{1 + e^{-(net_i(t) - \Theta) / \tau}}$$

или...

$$f_i^{s*}(net_i(t)) = \frac{1}{2} (1 + \tanh(\lambda net_i(t)))$$



Алгоритми учења

- Модели неурона су динамички => општа диференцијална једначина:

$$\dot{x} = g_i(x_i, net_i)$$

- излаз i -тог неурона, при чему је очигледно да се, због **динамичких неуронских модела**, ради о диференцирању по времену;
- улаз i -тог неурона net_i , зависи од излаза многих неурона са којима је у вези, практично се развија *систем спрегнутих нелинеарних диференцијалних једначина*.

$$\dot{w}_{ij} = A_i(w_{ij}, x_i, x_j, \dots)$$

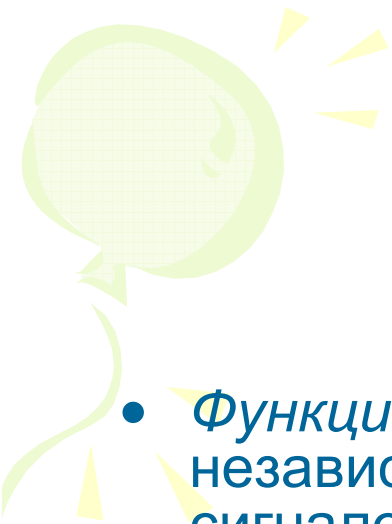
- И модификацију тежинских односа потребно посматрати као динамички систем => систем диференцијалних једначина за тежинске односе
- A_i представља **алгоритам учења**;
- Процес учења: одређивање тежинских односа кроз итеративни поступак њихове модификације;
- Избор** одговарајућег **алгоритма учења**, сходно апликацији за коју се мрежа користи, представља значајан проблем при пројектовању мреже.

Алгоритми учења

Алгоритам учења	Математичка формулација	Напомене
Hebb-ово правило (1949)	$\partial \Phi / \partial t = O(t)[I_i(t) - O(t)\Phi_i(t)]$	$O(t)$ – излаз неурона $I_i(t)$ – улази у неурон $\Phi_i(t)$ – јачина синаптичких веза
Widrow-Hoff-ово правило (1960) (LMS-„least mean square“ алгоритам, тј. алгоритам најмање квадратне разлике)	$\mathbf{w}(t+1) = \mathbf{w}(t) + 2\mu \varepsilon_k \mathbf{x}_k$	$\mathbf{x}_k$ – k -ти улазни вектор $\mathbf{w}$ – тежински вектор μ – позитивна константа ε_k – тренутна разлика (грешка) између захтеване вредности излаза (d_k) и актуелне вредности (y_k)
Генералисано делта правило (1986) (простирање грешке уназад)	$\mathbf{w}_{ji}(t+1) = \mathbf{w}_{ji}(t) + \eta \delta_{pj} \mathbf{x}_i$	$\mathbf{x}_i$ – i -ти улазни вектор $\mathbf{w}_{ji}$ – вектор тежинских односа η – параметар учења δ_{pj} – грешка везана за излаз неурона $\delta_{pj} = (y_{pj} - o_{pj}) - y_{pj}$ је захтевана вредност излаза, а o_{pj} је актуелна вредност $\mathbf{p}_{\text{тог}}$ обучавајућег вектора за j -ти неурон

Алгоритми учења – основна питања

- Које перформансе треба да има систем неуронске мреже у погледу обучавања?
- Да ли је потребно развити чисто динамички систем неуронске мреже?
- Хоће ли фаза учења бити реализована у „on-line” или „off-line” режиму?
- Да ли је брзина конвергенције важна за учење неуронске мреже?



Примена вештачких неуронских мрежа

- *Функционална апроксимација* је представљена скупом независних променљивих $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$ и излазним сигналом $y = f(x_1, x_2, \dots, x_m)$, који су груписани у парове $\{y, f(\mathbf{x})\}$.
- *Предикција* је представљена претходним вредностима стања система, $\mathbf{X} = [x(k), x(k-1), \dots, x(k-n)]^T$, као и тренутним и претходним вредностима улаза, $\mathbf{U} = [u(k), u(k-1), \dots, u(k-m)]^T$, које са будућим стањима система, $x(k+1) = f(\mathbf{X}, \mathbf{U})$ формирају парове $\{x(k+1), (\mathbf{X}, \mathbf{U})\}$.
- *Класификација* је представљена скупом улаза $\{^c x_i, i=1, \dots, N_j\}$, који са специфицираном класом C_j формира парове $\{C_j, ^c \mathbf{x}_j\}$.



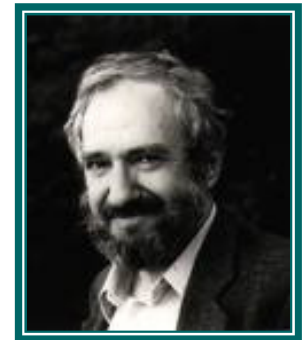
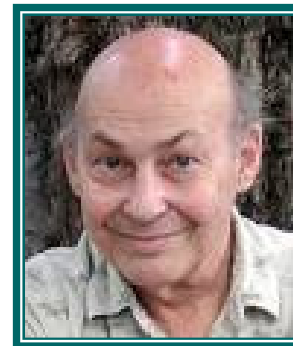
„Најраспрострањенији“ модели вештачких неуронских мрежа

- Перцептрон;
- „**Backpropagation**“ (BP) неуронска мрежа;
- Асоцијативне неуронске мреже;
- *Hopfield*-ове неуронске мреже;
- ART неуронске мреже (ART-1, ART-2, ART-3);
- *Fuzzy* асоцијативне неуронске мреже;
- Самоорганизујуће неуронске мреже.

Перцептрон

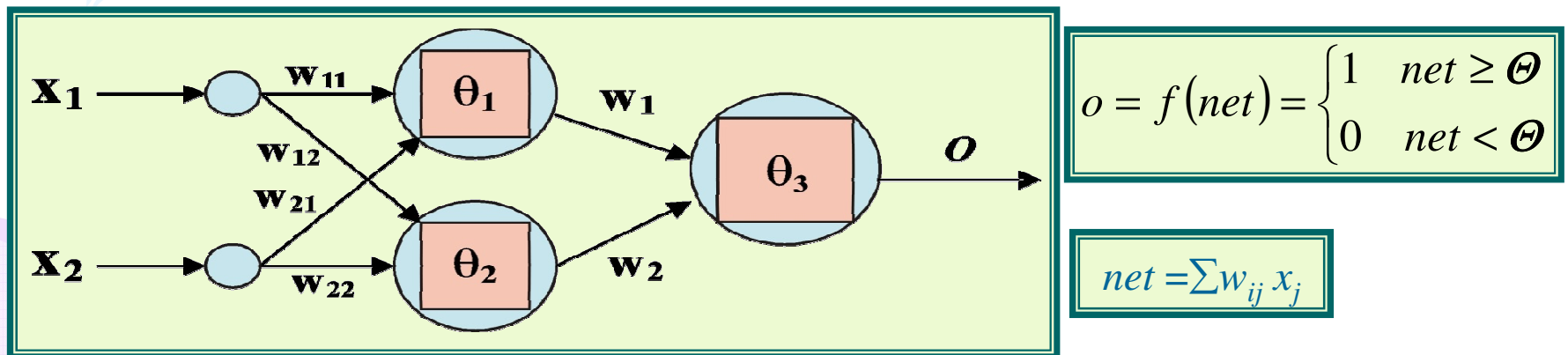
- Представља најранији модел вештачке неуронске мреже;
- Пионирски радови:

McCulloch&Pitts (1943.); Rosenblatt (1958,1962.); Minsky&Papert (1969).



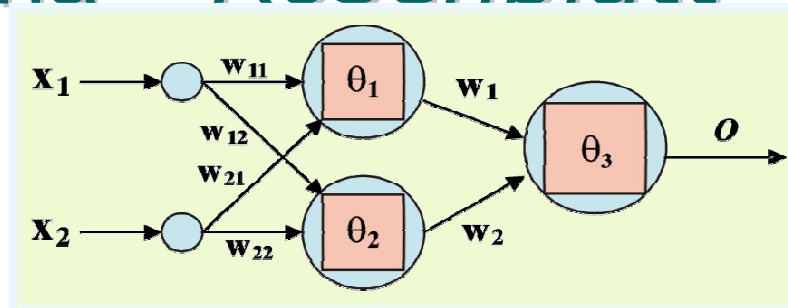
Перцептрон

- Једноставна архитектура;
- Користи простирање сигнала у једном смеру („*feedforward*”);
- Има један или више слојева неурона између улазних и излазних неурона (најчешће цео перцептрон има два слоја);
- Функционише на бази супервизорског учења;
- Линеарни тип неурона са прагом активације θ .



Структура перцептрона - *Rosenblatt*

- Структуру перцептрона чини:
 - скуп улазних неурона ($\{s\}$ -неурони),
 - скуп скривених неурона ($\{a\}$ -неурони) и
 - један излазни неурон ($\{r\}$ -неурон);



- Rosenblatt*-ова основна идеја: јачину веза између улазног слоја неурона и скривеног слоја неурона случајно изабрати по неком закону вероватноће и фиксирати њихове вредности током целог процеса учења;
- Алгоритам учења реализује подешавање тежинских односа између скривеног слоја неурона и јединог излазног неурона;
- Ако има довољно различитих типова неурона у скривеном слоју (репрезентују различите логичке исказе из улазног слоја), онда је процес учења од неурона из скривеног слоја до излазног неурона у стању да реализује комплексно логичко закључивање.