

Невронни Мрежи

проф. В. Младенов, каб. 12518

Използвана литература

- C. Aggarwal, „Neural Networks and Deep Learning”, Springer, 2018, ISBN 978-3-319-94462-3.
- Младенов В., С. Йорданова, „Размито управление и невронни мрежи”, ТУ-София, 2006.
- Йорданова С., В. Младенов, Г. Ценов, Р. Петрова, “Размито управление и невронни мрежи - ръководство за лабораторни упражнения”, ТУ-София, 2008.
- Fausett L., “Fundamentals of Neural Networks”, Prentice-Hall, 1994. ISBN 0130422509.
- Haykin S., Neural Networks and Learning Machines”, 3rd Edition, Prentice Hall, 2016, ISBN-13 : 9789332570313 , ISBN 10: 9332570310.
- Haykin S., “Neural Networks: A comprehensive foundation”, 2nd Edition, Prentice Hall, 1999, ISBN 0132733501.
- Bishop Ch., “Neural Networks for Pattern Recognition”, Clarendon Press, Oxford, 1995.

1. Основни понятия и дефиниции. Биологични невронни мрежи и изкуствени невронни мрежи. Основни компоненти на невронните мрежи: архитектури, правила за обучение и активационни функции. Основни класификации на невронните мрежи.

Подход при работа с НМ

В компютърните науки подходите за решаване на задачи, при които се използват невронни мрежи се различават от класическия алгоритмичен подход.

Програмен алгоритмичен подход

- Прави се *абстрактен модел* на разглежданата задача
- Разработва се *алгоритъм* (прецизна последователност от стъпки) за решаването и
- Алгоритъмът се реализира с *компютърна програма*

Подход базиран на Невронни Мрежи

- Моделиран по *биологично съответствие с човешкия мозък*
- Използването на невронните мрежи се базира на *обучение чрез примери*

Какво е Невронна Мрежа?

Невронната мрежа е система за паралелна обработка на информация, която има свойството за съхранение и използване на експериментални знания. Тя моделира мозъка в следните два аспекта:

- Информацията се натрупва в мрежата в процес на обучение
- Силата на връзките между отделните възли (неврони) се моделира с тегла на съответните връзки, които се използват за съхранение на информацията

Какво е Невронна Мрежа?

Една (изкуствена) невронна мрежа (NN или ANN в литературата) представлява система за обработка на информация, която е математическа абстракция базирана на човешкото знание и невробиологията.

- Най-общо невронните мрежи се състоят от прости елементи за обработка на информация наречени неврони или възли.
- Невроните са свързани и теглата на връзките между тях определят силата на съответните връзки.
- Входната информация за всеки неврон е претеглената сума от сигналите от останалите неврони.
- Тази информация се акумулира в неврона като изходният му сигнал се определя посредством т. нар. активационна или предавателна функция.

Какво е Невронна Мрежа?

- Някои невронни мрежи са модели на биологичните невронни мрежи, а други не са. В исторически аспект идеята за невронните мрежи е свързана с желанието да се разработи интелигентна система за паралелни изчисления, работеща като мозъка и по такъв начин да се подпомогне по-доброто разбиране на дейността на човешкия мозък.
- Мозъкът е много сложна нелинейна система за паралелна обработка на информация. Тази система извършва възприятия, разпознаване на образи, обработка на изображения, управление на движението и др. много по-бързо и от най-бързите съвременни компютри.

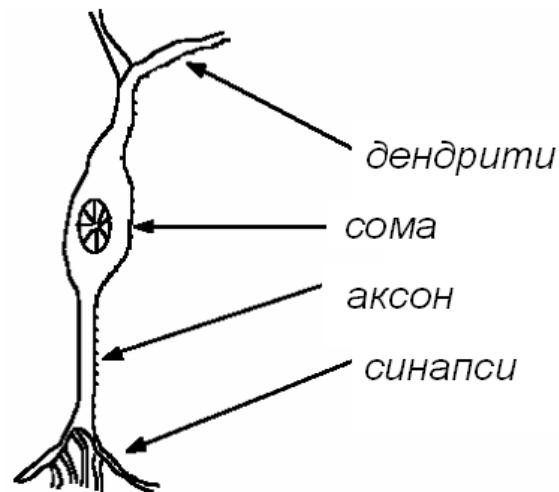
Биологичен Неврон

- Биологичния неврон или нервната клетка се състои от синапси, дендрити, тяло на клетката (сома) и аксон.

Синапсите са места, където се осъществява връзката и се преобразуват сигналите, посредством сложни електрохимични процеси. Сигналите от синапсите се предават по дендритите към тялото на клетката, където се агрегират (сумират се сигналите от синапсите свързани с дендритите на дадена клетка).

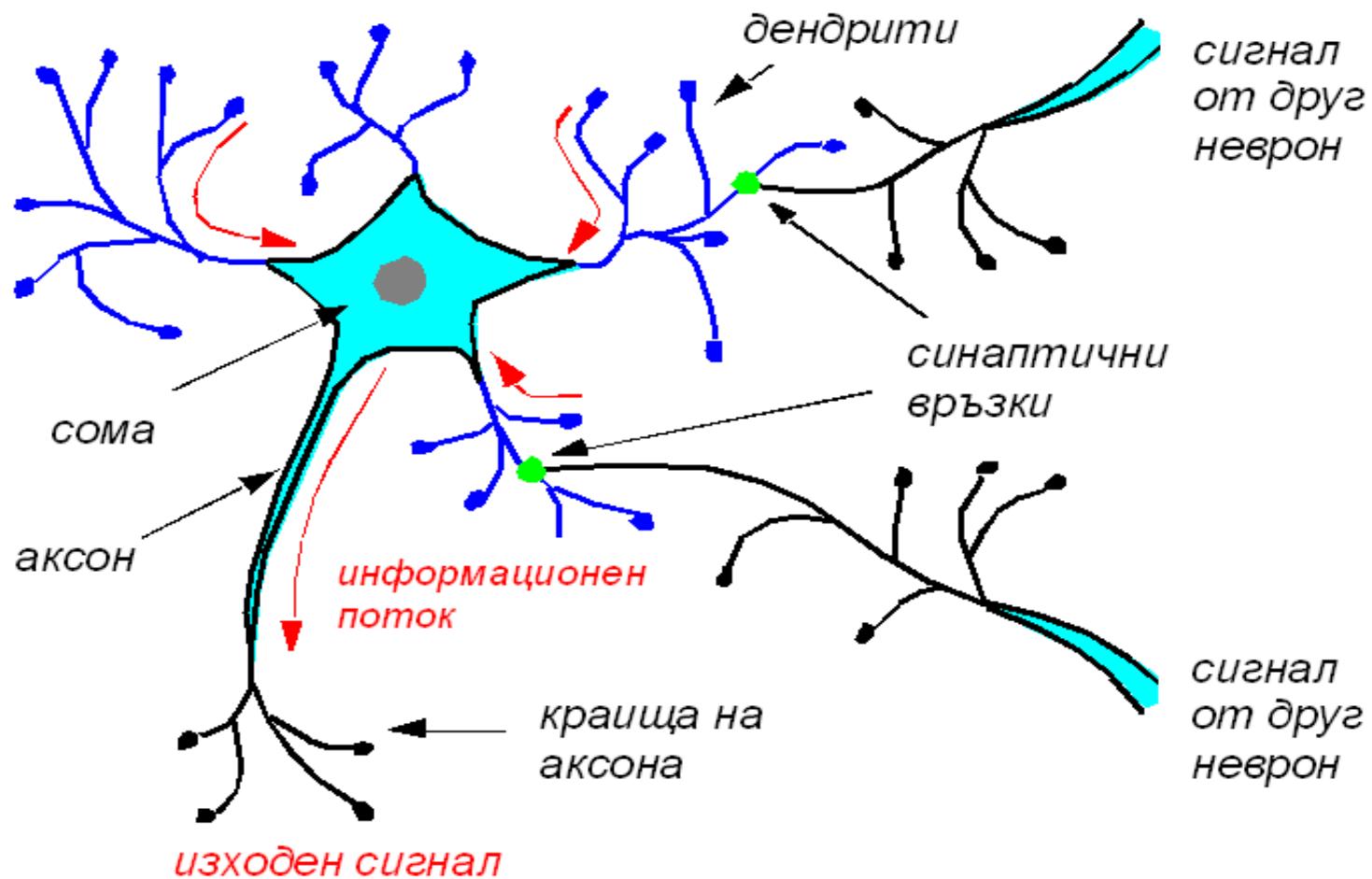
Когато нивото на агрегирания сигнал надхвърли определена граница се генерира сигнал по аксона на клетката, посредством който тя въздейства на останалите нервни клетки:

Дендритите и аксона са израстъци на клетката, чрез които тя е свързана с останалите клетки. Посредством дендритите клетката получава сигнали от аксоните на другите клетки в мозъка.



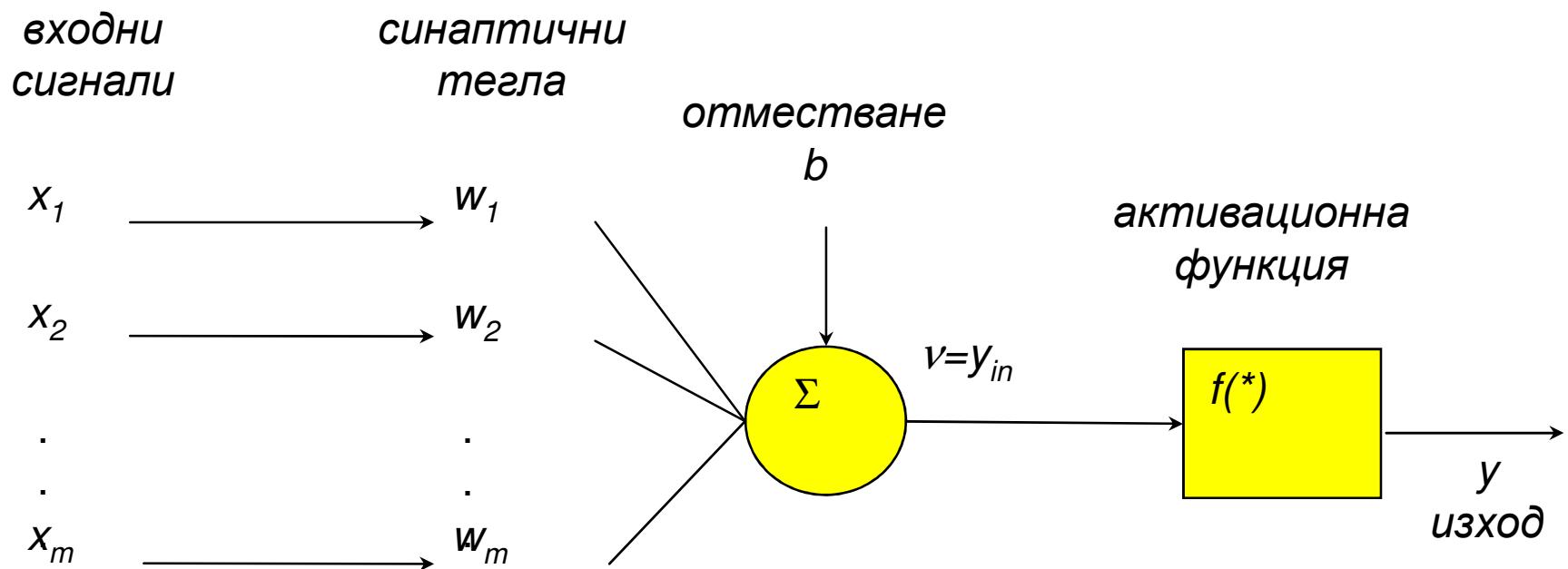
Фиг. 1.1. Биологичен неврон

Биологичен Неврон



Фиг. 1.2. Връзки на нервните клетки в мозъка

Абстрактен неврон



Абстрактен неврон

Всеки неврон има много входове, които са модел на дендритите и един изход, който е модел на аксона. На входовете x_1, x_2, \dots, x_m постъпват сигналите към неврона (външни сигнали или сигнали от изходите на други неврони). С всеки вход е свързано тегло $w_j, j=1, 2, \dots, m$, което понякога се нарича синаптично тегло и моделира силата на връзката при предаване на сигнала посредством синапсиса свързан със съответния дендрит. Агрегирането на входните сигнали в тялото на неврона се моделира със суматора, чийто изходен сигнал се изчислява с формулата:

$$v = y_{in} = \sum_{j=1}^m w_j x_j + b$$

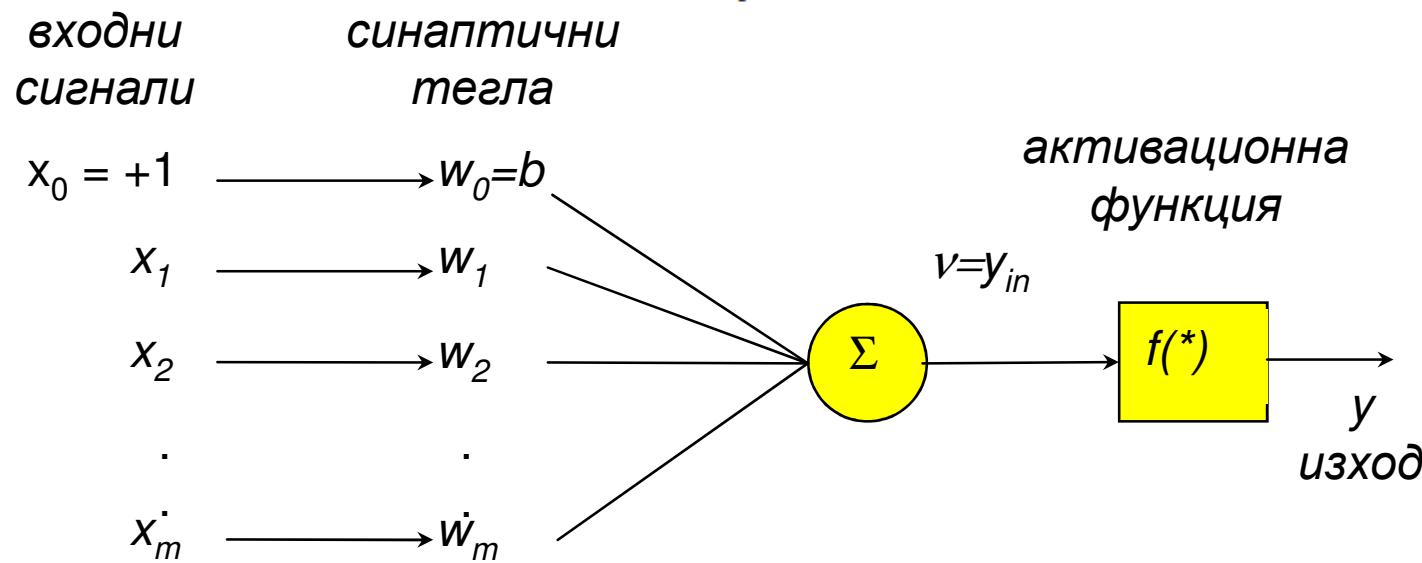
За удобство и възможност за повече степени на свобода, при модела на неврона се въвежда и сигнал с постоянна стойност b , наречен отместване (bias). Тялото на нервната клетка се моделира със суматора и блока на активационната функция, посредством която се получава изходният сигнал y на неврона. В общия случай активационната функция е нелинейна и изходният сигнал се изчислява като

$$y = f(y_{in}) = f\left(\sum_{j=1}^m w_j x_j + b\right)$$

Алтернативна форма

В много случаи за удобство се използва алтернативната форма на модела на неврона, като показаната на фигурата. В този случай отместването b се представя като тегло w_0 , свързано със сигнал с постоянна стойност 1. На база на въведените означения изходът на суматора се изчислява с формулата

$$v = y_{in} = \sum_{j=0}^m w_j x_j$$



Класификация на невронните мрежи

Основните компоненти на невронните мрежи, с които са свързани и различните им класификации са:

- **Архитектура на мрежата** – задава се от начина на свързване между различните неврони;
- **Алгоритъм за обучение** – това е начинът, по който се избират (настройват) теглата на връзките между невроните, така че невронната мрежа да изпълнява желаното преобразуване на сигналите;
- **Активационна функция** (предавателна функция) – правилото посредством което се определя стойността на изходния сигнал.

Компоненти на невронните мрежи

Архитектура на мрежата

Активационна функция

Алгоритъм за обучение

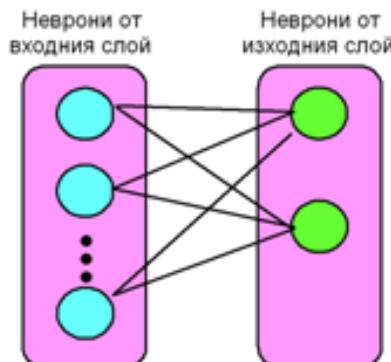
Архитектура на мрежата

Обикновено невроните (възлите) в невронните мрежи се групират в слоеве. Входният слой от възли се свързва с външните сигнали, които се подават към мрежата и той не извършва никаква обработка. Изходните сигнали на невронната мрежа се формират от изходите на невроните от последния слой, наречен изходен слой. Междинните слоеве от неврони се наричат скрити слоеве, т.к. няма пряк достъп до техните изходни сигнали. Невроните от скритите и изходния слой извършват (изчисляват) преобразуване на входния сигнал. Архитектурата на дадена невронна мрежа се определя от **броя на слоевете, броя на невроните** в различните слоеве и **начина на свързване** между слоевете.

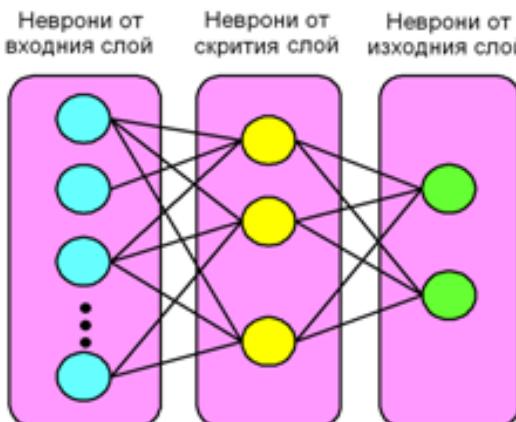
Най-често срещаните архитектури на невронните мрежи са:

- *Еднослойни* мрежи с еднопосочко предаване на сигнала
- *Многослойни* мрежи с еднопосочко предаване на сигнала
- *Рекурентни* мрежи – архитектури с обратни връзки

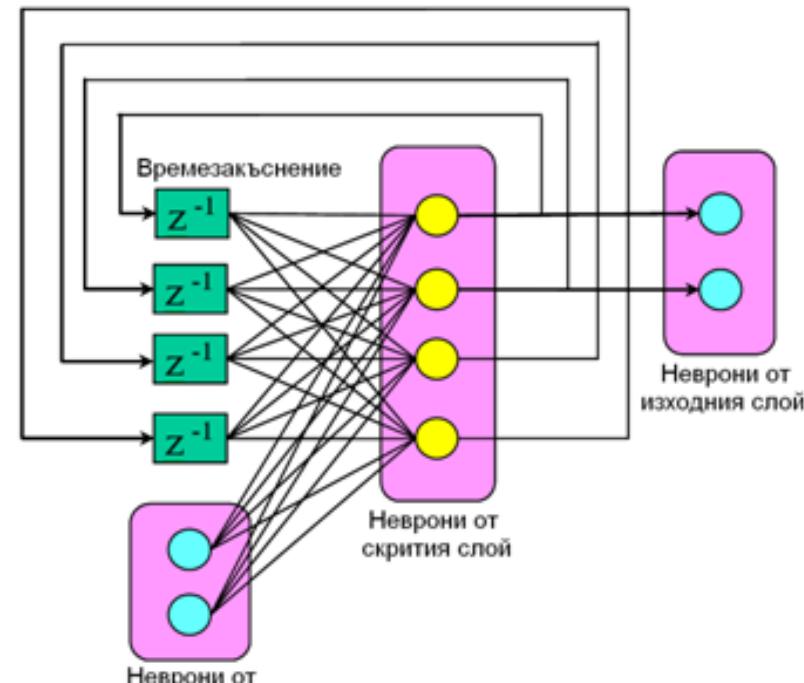
Архитектура на мрежата



*Еднослойна мрежа с
единопосочко предаване на сигнала*



*Многослойна мрежа с
единопосочко предаване на сигнала*



Рекурентна мрежа

Активационна (преходна) функция

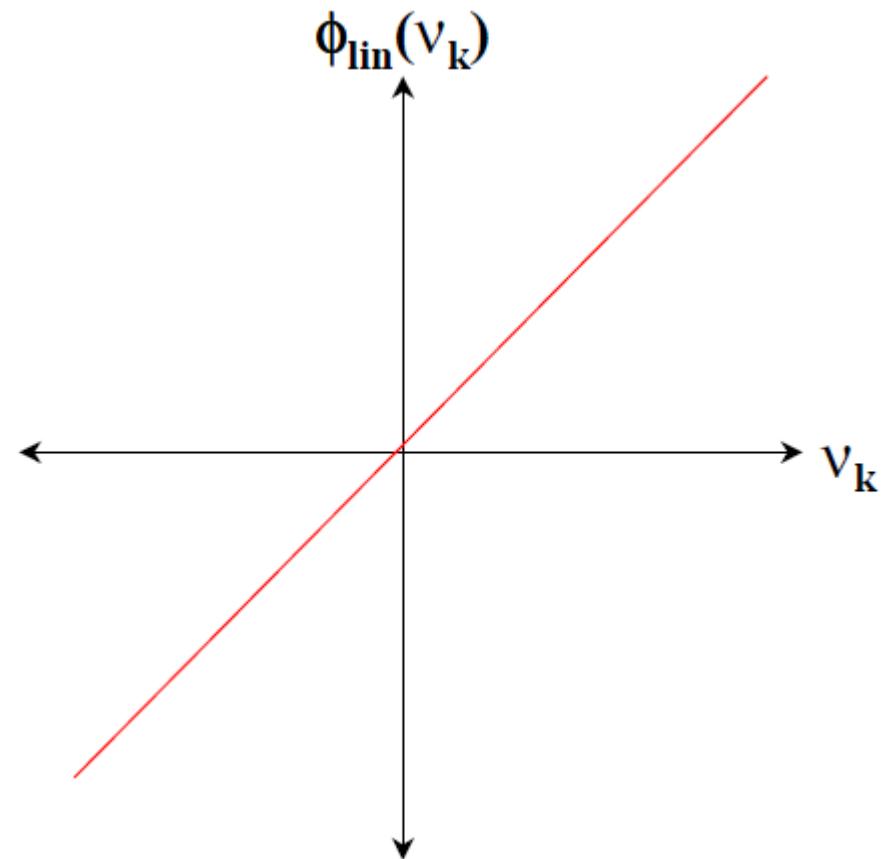
Съществуват много различни типове активационни функции! С използването на повечето от тях се осъществяват структури, които не съществуват биологически в природата (от тук идва и термина изкуствени НМ) .

- Някои функции са линейни; други са нелинейни
- Някои функции са непрекъснати и диференцируеми
- Изборът на активационната функция се определя в зависимост от решаваната задача, типа използван неврон или от типовата структура на мрежата

Линейна Активационна Функция

Линейна Активационна Функция

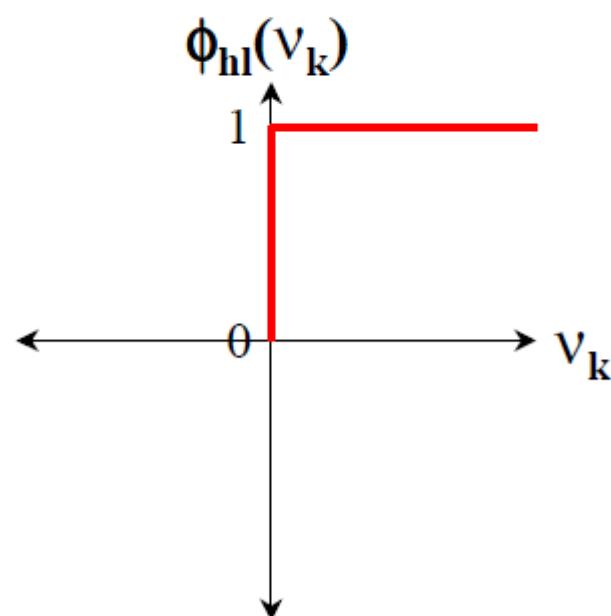
$$y_k = f(v_k) = \phi_{lin}(v_k)$$



Релейна Активационна Функция

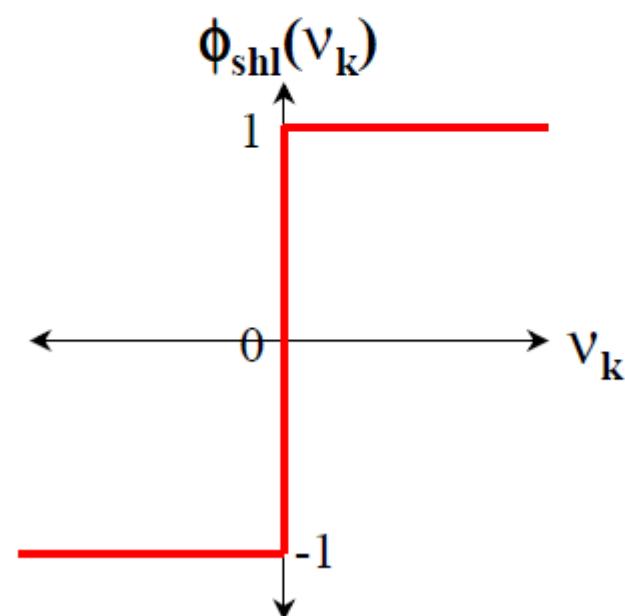
Релейна

$$y_k = f(v_k) = \phi_{hl}(v_k) = \begin{cases} 0 & \text{при } v_k < 0 \\ 1 & \text{при } v_k \geq 0 \end{cases}$$



Симетрична Релейна

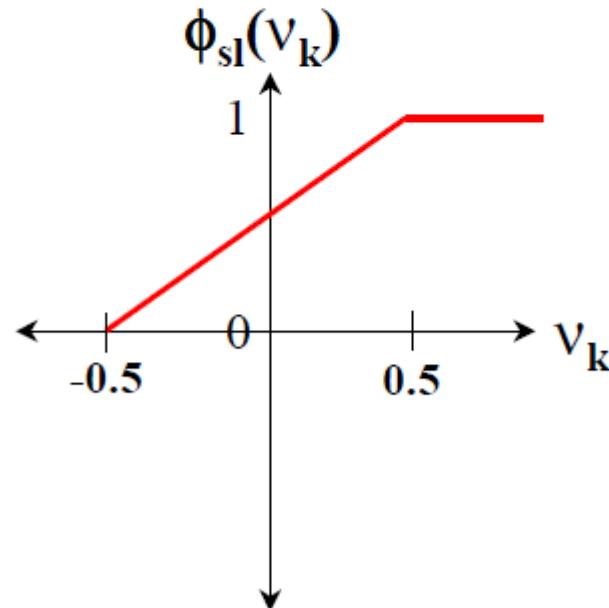
$$y_k = f(v_k) = \phi_{shl}(v_k) = \begin{cases} -1 & \text{при } v_k < 0 \\ 0 & \text{при } v_k = 0 \\ 1 & \text{при } v_k > 0 \end{cases}$$



Линейна с Насищане

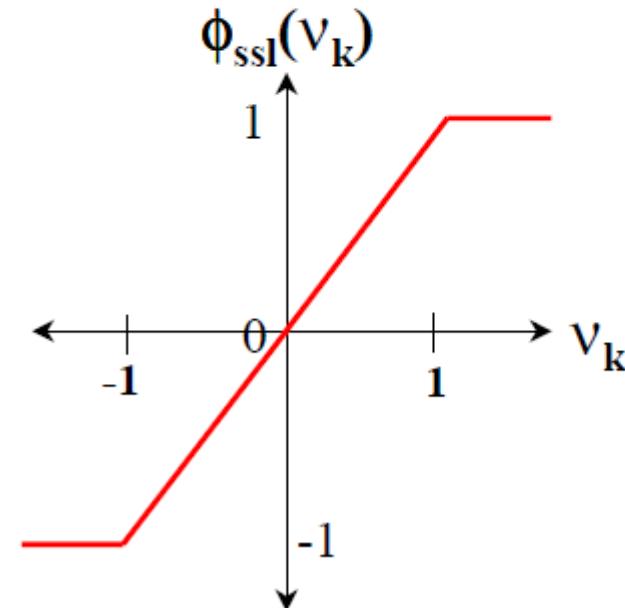
Линейна Активационна Функция с Насищане

$$y_k = \phi_{sl}(v_k) = \begin{cases} 0 & \text{при } v_k < -\frac{1}{2} \\ v_k + \frac{1}{2} & \text{при } -\frac{1}{2} \leq v_k \leq \frac{1}{2} \\ 1 & \text{при } v_k > 1 \end{cases}$$



Линейна Симетрична с насищане

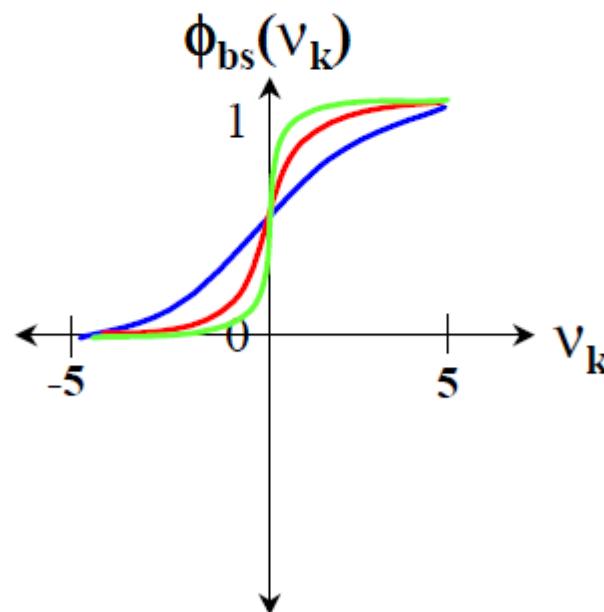
$$y_k = \phi_{ssl}(v_k) = \begin{cases} -1 & \text{при } v_k < -1 \\ 0 & \text{при } -1 \leq v_k \leq 1 \\ 1 & \text{при } v_k > 1 \end{cases}$$



Сигмоидална функция

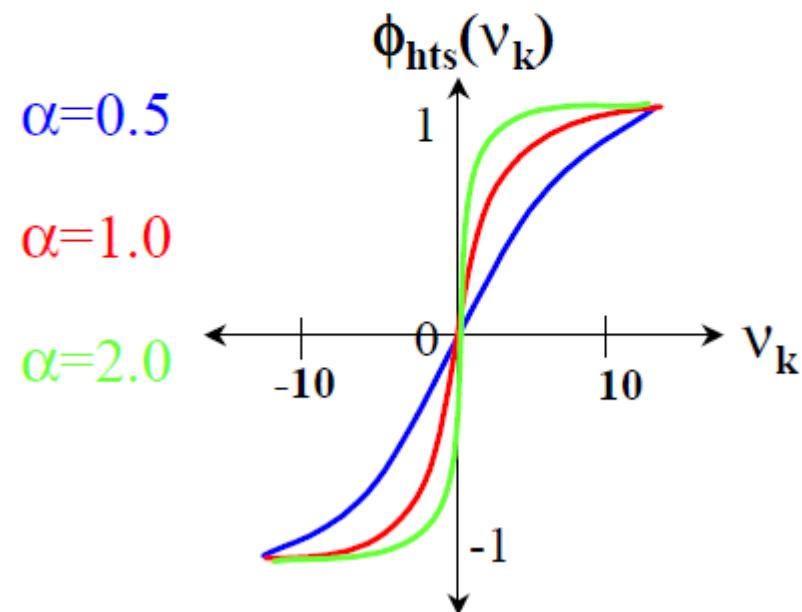
Бинарна Сигмоидална Функция

$$y_k = f(v_k) = \phi_{bs}(v_k) = \frac{1}{1 + e^{-\alpha v_k}}$$



Хиперболична Тангенциално-Сигмоидална

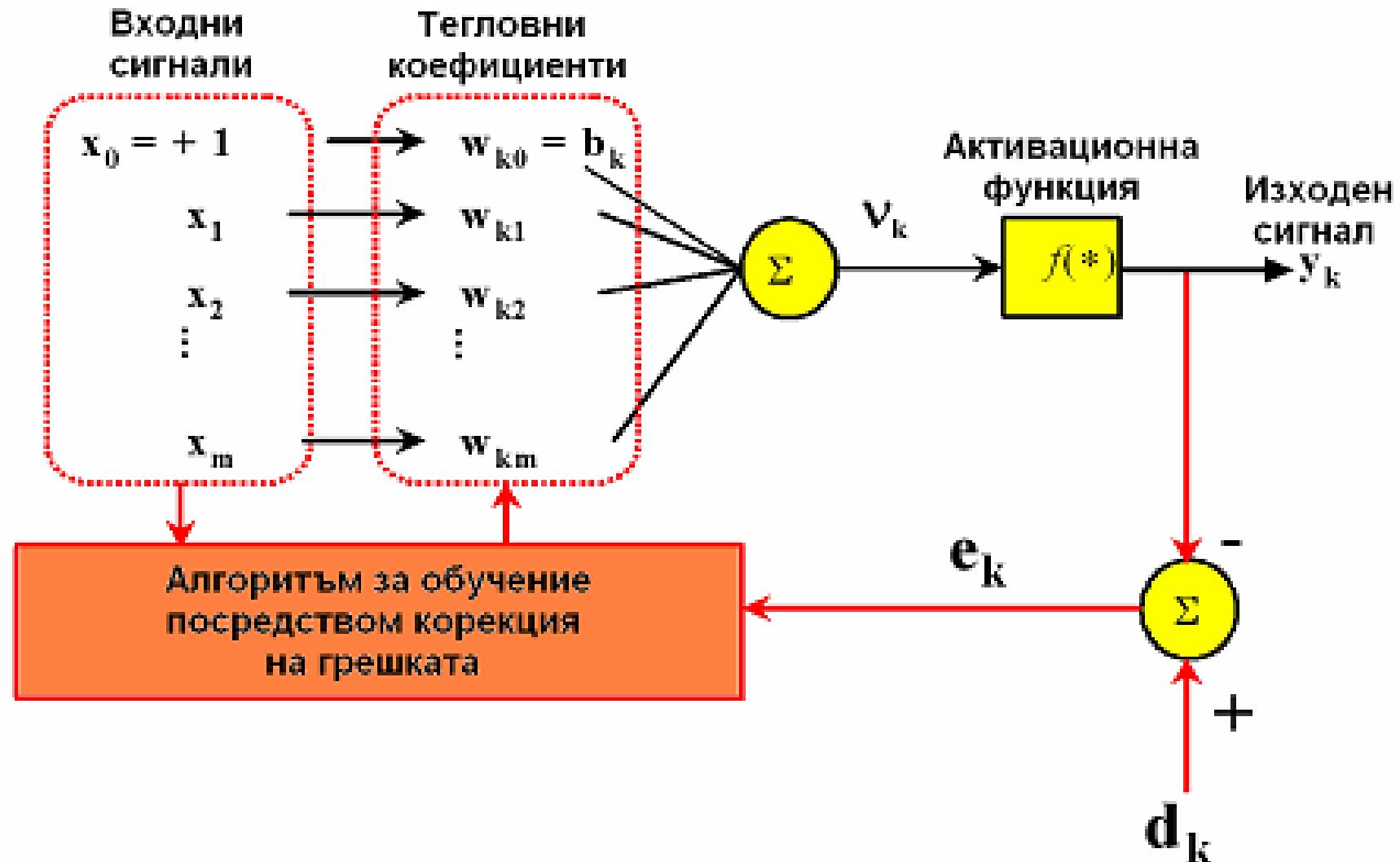
$$y_k = f(v_k) = \phi_{hts}(v_k) = \frac{1 - e^{-2\alpha v_k}}{1 + e^{-2\alpha v_k}}$$



Алгоритми за обучение

- . **Error correcting** (Принцип на корекция на грешката)
- . **Memory based** (Принцип на минималното разстояние)
- . **Hebbian** (Правило на Hebb)
- . **Competitive** (Принцип на съревнованието)
- . **Boltzman** (Правило на Boltzman)

Принцип на корекция на грешката



Memory based (Принцип на минималното разстояние)

- Формират се образци (памети) от правилно класифицирани входни вектори, които могат да се разглеждат като входни образци за невронната мрежа.
- Дефинира се критерий за близост (съседство) до всеки от образците
- Правилото за обучение се прилага само за входни вектори, които са в съседство на съответните образци.
- Примери за алгоритми използващи този принцип са алгоритмите на *най-близките съседи, k-те най-близки съседи, радиални базисни функции*.

Hebbian (Правило на Hebb)

Правилото на Hebb гласи: Ако активирането на неврон k (генериране на изходен сигнал) е свързано с активирането на неврон j , то силата на връзката между тях се увеличава, а ако активирането на неврон k не е свързано с активирането на неврон j , то силата на връзката между тези неврони отслабва.

Корелираните активации усилват връзките, а некорелираните ги отслабват.

Нека $y_k(n)$ е изходният сигнал на неврон k , а $y_j(n)$ е изходният сигнал на неврон j на стъпка n . При най-простата форма на алгоритъма за обучение с правилото на Hebb старата стойност на теглото на връзката между двета неврона w_{kj} се донастройва с $\Delta w_{kj} = \eta y_k(n)y_j(n)$, където η е параметър, който се нарича скорост на обучение. В резултат за новата стойност на това тегло се получава $w_{kj} + \Delta w_{kj}$. Примери за използване на този принцип са невронните мрежи Hebb, дискретните мрежи на Hopfield и др.

$$\Delta W_{kj}(n) = \eta(Y_{kj}(n) - Y_{kj,av})(X_{kj}(n) - X_{kj,av})$$

Competitive (Принцип на съревнованието)

- Този принцип се използва при т.нар. *методи за обучение без учител*.
- Входните вектори (образи) за невронната мрежа трябва да се самоорганизират по подходящ начин.
- Невроните от мрежата съответстват на класовете, в които са се самоорганизирали входните вектори (образи).
- При нов входен вектор, невроните на мрежата се съревновават, кой от тях е “най-сходен” (в определен смисъл, разбира се) до този входен вектор.
- В най-простите случаи на прилагане на принципа се модифицират само теглата на връзките свързани с неврона “победител”.
- Примери за използване на този принцип са невронните мрежи базирани на съревнование, мрежите на Kohonen, невронните мрежи базирани на адаптивния резонанс и др.

Фази при приложение на НМ

- **Фаза на обучение** – в тази фаза се настройват теглата на мрежата;
- **Фаза на използване** – в тази фаза теглата вече са определени и се извършва реалната обработка на сигналите като се задава необходимият брой входни вектори и на тяхна база се изчисляват съответните изходни вектори.

Класификация на НМ

- По отношение на **фазата на използване**, невронните мрежи могат да бъдат класифицирани като **невронни мрежи с еднопосочно предаване на сигнала** (статични) и **невронни мрежи с обратни връзки** (рекурентни, динамични).
- По отношение на **фазата на обучение**, невронните мрежи могат да бъдат класифицирани като **невронни мрежи, които използват т.нар. обучение с учител (*supervised learning*)** и такива, които използват т.нар. **обучение без учител (*unsupervised learning*)**.

При обучението с учител за всеки входен вектор (образ) е известен съответният изходен вектор, докато при обучението без учител няма изходни вектори свързани с входните и в процеса на обучение се реализира самоорганизация на входните вектори (образи) по подходящ начин. Такива мрежи са мрежите на Hebb, самоорганизиращите се мрежи на Kohonen, невронните мрежи базирани на съревнование и др.

Използване на НМ

- Невронните мрежи с еднопосочко предаване на сигналите, използващи обучение с учител основно се използват за **решаване на задачи свързани с апроксимации**.
- Примери за такива невронни мрежи са линейните рекурсивни мрежи базирани на най-малки квадрати, мрежите с обратно разпространение на грешката, мрежите с радиални базисни функции и др.

Използване на НМ

- Основното приложение на невронните мрежи с еднопосочко предаване на сигнала, използващи обучение без учител е при **извлечането на свойства (признания) от входните данни**, както и при преобразуване на входните вектори (данни) от една в друга по-подходяща област.

Примери за такива невронни мрежи са мрежите на Hebb, самоорганизиращите се мрежи на Kohonen, невронните мрежи базирани на съревнование и др.

Използване на НМ

- Рекурентните невронни мрежи основно се използват при **работка с временни редове**. При тях вътрешното им състояние се променя с времето.
- Типични примери за такива мрежи са рекурентните невронни мрежи с обратно разпространение на грешката, асоциативните памети, мрежите базирани на адаптивния резонанс и др.

Какво може да се направи с НМ?

- По принцип с невронни мрежи може **да бъде изчислена всяка изчислима функция**, т.е. с невронни мрежи може да се прави всичко, което се прави със стандартните компютри.
- За разлика от компютрите обаче, невронните мрежи използват принципа на паралелната обработка на информацията и поради това те могат да бъдат използвани успешно при задачите за обработка на данни (сигнали, изображения и др.) в реално време.

Какво може да се направи с НМ?

- Невронните мрежи са полезни при **решаване на задачи за класификация и апроксимации**, при наличие на **много информация** (данни от измервания напр.) за обучение, когато е трудно да бъдат намерени достатъчно прости аналитични зависимости за описание на съответните обекти и процеси.
- При наличие на достатъчно данни и компютърни ресурси при обучението, **почти всяко изображение между две векторни пространства може да бъде апроксимирано** с желана точност посредством многослойна невронна мрежа с еднопосочno предаване на сигнала (на практика най-често използваният невронни мрежи).

Какво може да се направи с НМ?

- Невронните мрежи работят на принципа на черната кутия като се използва информацията от разглеждания процес или явление, с цел обучение при настройване на теглата на мрежата. Поради това те са удобно средство за моделиране на процеси и явления, за които може да бъде събрана много входно-изходна информация (от измерване и др.), но трудно могат да бъдат описани аналитично.
- Основно преимущество на невронните мрежи е, че те са робастни по отношение на неточности в коефициентите на мрежата.
- Основен техен недостатък е, че не съществуват методи за обучение, посредством които да се генерира информация, която не се съдържа в данните, с които е обучена мрежата.

Къде се прилагат НМ?

- В **инженерните и компютърните науки**, невронните мрежи се използват за обработка на сигнали и изображения, за филтрация, при системи с обучение и разпознаване на образи, при моделиране, идентификация и управление на сложни многосвързани обекти и др.
- В **приложната математика** невронните мрежи се използват за решаване на различни видове задачи като апроксимация, оптимизация без и с ограничения, решаване на различни видове уравнения и системи уравнения и др.
- В **статистиката** невронните мрежи намират приложение като модели за нелинейна регресия и класификация, прогнозиране на временни редове и др.
- В **науката за познанието** невронните мрежи се използват за описание и моделиране на мисленето и съзнанието.
- В **икономиката и финансите** невронните мрежи се използват за изграждане на сложни икономически модели, за моделиране и прогнозиране на финансови пазари и много други.

Исторически преглед на развитието на теорията на невронните мрежи

