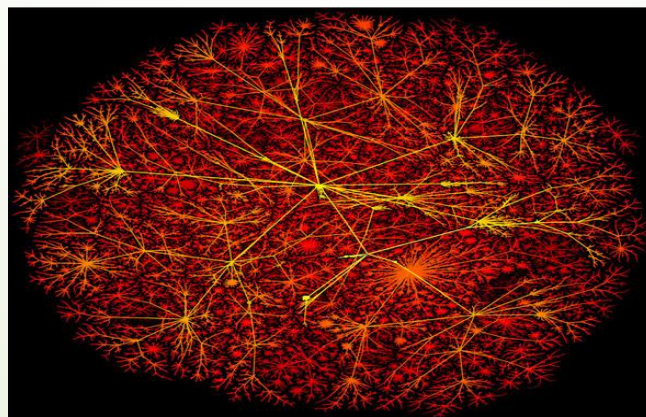


آشنایی با مقدمات شبکه های عصبی مصنوعی

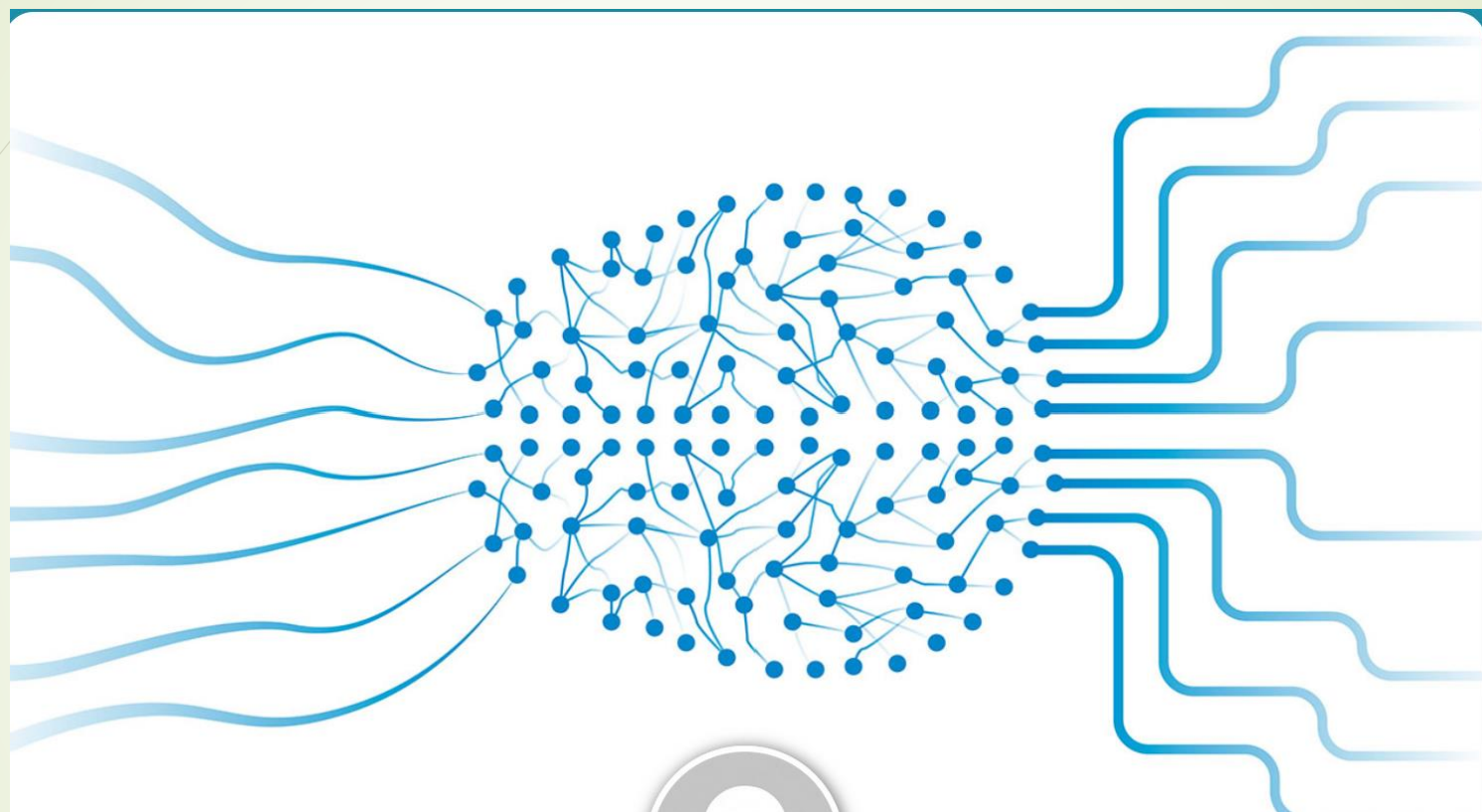
Artificial Neural Network (ANN)

آرزو حبیبی راد
گروه آمار دانشگاه فردوسی مشهد

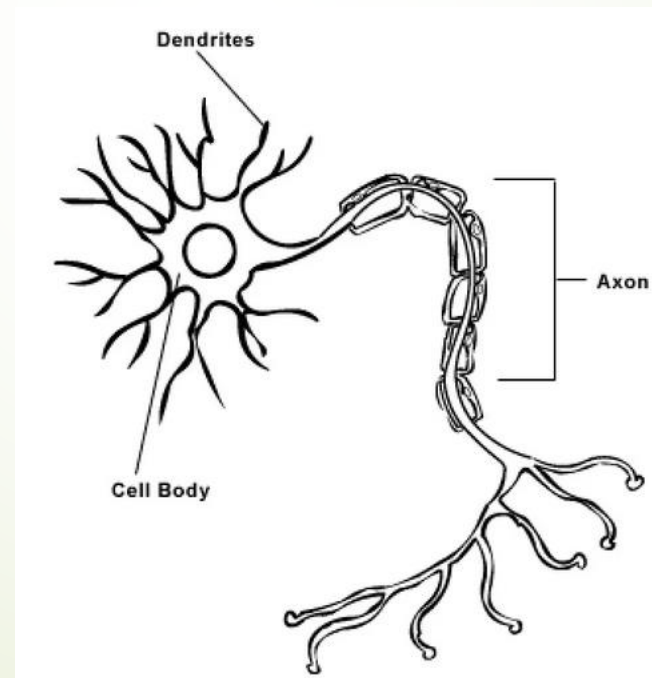
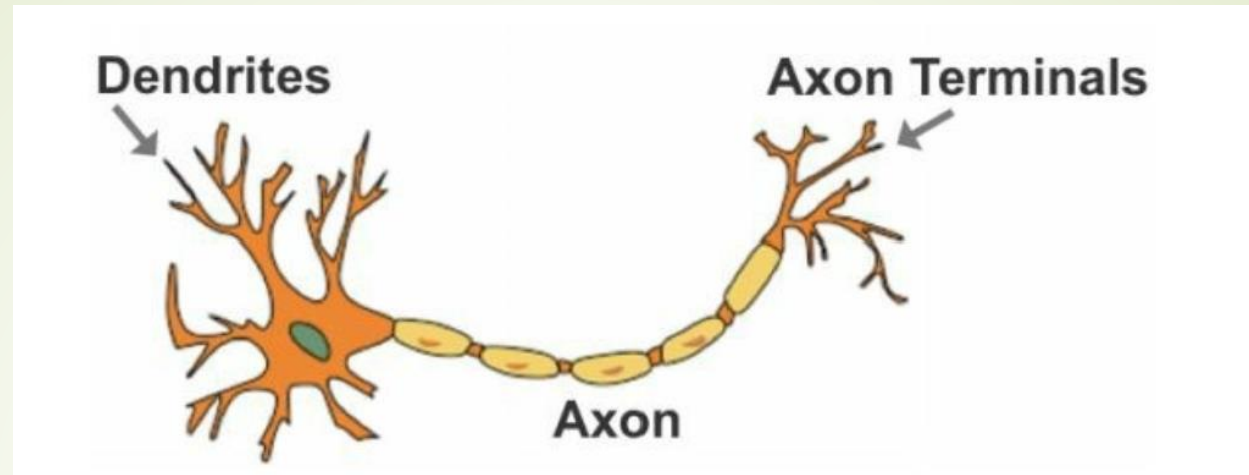


شبکه عصبی (Neural Network) چیست؟





مغز انسان در خود تعداد بسیار زیادی از نورون‌ها را جای داده است تا اطلاعات مختلف را پردازش کرده و جهان اطراف را بشناسد. به صورت ساده، نورون‌ها در مغز انسان اطلاعات را از نورون‌های دیگر به وسیله‌ی دندرویدها می‌گیرند. این نورون‌ها اطلاعات ورودی را با هم جمع کرده و اگر از یک حد آستانه‌ای فراتر رود به اصلاح فعال می‌شود و این سیگنال فعال شده از طریق آکسون‌ها به نرون‌های دیگر متصل می‌شود.



اولین سلول عصبی مصنوعی در سال 1943 به وسیله یک عصب شناس به نام مک کولاک وارن و یک منطق دان به نام والتر پیتس ساخته شد که یک مدل خطی ساده را تفسیر میکرد. اما به دلیل محدودیت های تکنولوژیکی آن زمان مجال کار بیشتر پیدا نکردند.

پس از فرازونشیب هایی در سال 1982، هاپفیلد با معرفی شبکه های چند لایه و الگوریتم یادگیری دارای بازخورد راه حلی برای حل موارد غیر خطی ارائه کرد.

سرانجام در نیمه دوم دهه نود میلادی، با رواج رایانه ها، شبکه های عصبی به طور جدی وارد عرصه کاربردی شدند و تا به امروز با قدرت مورد استفاده قرار میگیرند.

مغز انسان شامل نورون‌های به هم پیوسته با دندریت‌هایی Dendrites است که ورودی‌ها را دریافت می‌کنند. نورون، از این ورودی‌ها، یک سیگنال الکتریکی تولید کرده و از طریق آکسون Axon به خروجی می‌دهد و سپس این سیگنال‌ها را به وسیله ترمینال‌های آکسونی بین دیگر نورون‌ها منتشر می‌کند.

ایده شبکه های عصبی مصنوعی


- کوچکترین واحدهای مغزی یک موجود زنده که وظیفه محاسبات (خاموش یا روشن) را دارند، سلول است.
- این وظیفه را در رایانه ها، ترانزیستور به عنوان کوچکترین واحد محاسباتی بر عهده دارد.
- سرعت پردازش هر واحد سلولی 1 میلی ثانیه و سرعت پردازش یک ترانزیستور 1 نانو ثانیه (یک میلیاردم ثانیه) است (پردازشگر ۱ گیگاهرتز).

نکته‌ی جالب توجه در مورد شبکه‌ی عصبی این است که نیازی نیست آن را برای یادگیری صریح، برنامه‌ریزی کنید. این شبکه در واقع می‌تواند همه‌چیز را مانند مغز انسان، خودش یاد بگیرد.

اما این شبکه‌ی عصبی، مغز نیست. در نظر داشتن این نکته مهم است که شبکه‌های عصبی عموماً شبیه‌سازهای نرم‌افزاری هستند که با برنامه‌نویسی برای کامپیوترهای بسیار ساده و پیش پا افتاده راه می‌افتند و با روش‌های قبلی خود و با استفاده از ترانزیستورها و دروازه‌های منطقی خود کار می‌کنند تا به مانند میلیارد‌ها سلول مغزی متصل و موازی رفتار کنند. هیچ‌کس تا به حال حتی تلاش هم نکرده است تا کامپیوتری بسازد که با ترانزیستورهایی با ساختار موازی مانند مغز انسان کار کند.

در علوم کامپیوتر و هوش مصنوعی، از نوروں‌های مغز برای ساخت الگوریتمی به نام شبکه‌ی عصبی مصنوعی بهره می‌گیریم.

با استفاده از الگوریتم شبکه‌های عصبی، می‌توان مدل‌های مختلف و پیچیده‌ای را شناخت. برای مثال می‌توان طبقه‌بندی‌هایی با دقت خوب انجام داده یا خوشه‌بندی‌هایی بر روی داده‌های بزرگ انجام دهیم. در واقع شبکه‌های عصبی نوعی از الگوریتم‌های یادگیری ماشین هستند.



■ یادگیری از دیدگاه زیستی یعنی ایجاد ارتباطات جدید بین سلول های مغزی.

■ این ساختار الهام بخش بستری برای ایجاد شاخه ای جدید از علم تحت عنوان شبکه های عصبی مصنوعی شد.

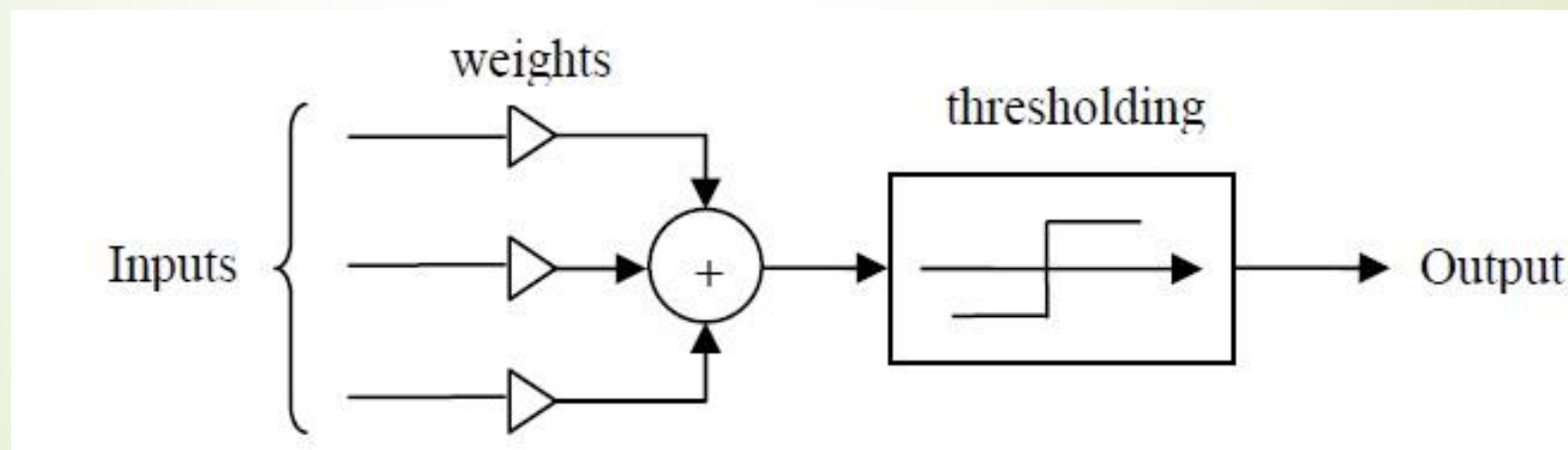
یک شبکه عصبی مصنوعی متشکل از چه چیزهایی است؟

یک شبکه عصبی مصنوعی معمولی ده‌ها، صدها، هزاران یا حتی میلیون‌ها نورون مصنوعی به نام واحد دارد که در مجموعه‌ای از لایه‌ها قرار گرفته‌اند که در هر طرف با بقیه لایه‌ها به هم متصل‌اند. برخی از آن‌ها با نام واحدهای ورودی شناخته می‌شوند. این واحدها برای دریافت شکل‌های مختلف اطلاعات از دنیای خارجی که شبکه سعی در یادگیری، شناسایی و پردازش آن‌ها دارد، طراحی شده‌اند. سایر واحدها که واحدهای خروجی نامیده می‌شوند، در طرف مخالف شبکه قرار دارند و چگونگی واکنش شبکه به اطلاعات یادگرفته‌شده را مشخص و بررسی می‌کنند.

درمیان واحدهای ورودی و خروجی، واحدهای مخفی وجود دارند که به همراه این واحدها، اکثریت مغز مصنوعی را تشکیل می‌دهند. اتصال بین واحدها با عددی به نام وزن ارائه می‌شود. وزن می‌تواند مثبت (اگر یک واحد، واحد دیگر را برانگیخته کند) یا منفی (اگر یک واحد، واحد دیگر را سرکوب یا مهار کند) باشد. هرچقدر میزان وزن بالاتر باشد، تأثیر یک واحد بر دیگری بیشتر می‌شود. این شبیه به راهی است که سلول‌های حقیقی مغز در شکاف‌های کوچکی به نام سیناپس باعث برانگیختگی یکدیگر می‌شوند.

- شبکه عصبی مصنوعی روشی عملی برای یادگیری توابع گوناگون نظیر توابع با مقادیر حقیقی، توابع با مقادیر گسسته و توابع با مقادیر برداری میباشد.
- یادگیری شبکه عصبی در برابر خطاهای داده های آموزشی مصون بوده و اینگونه شبکه ها با موفقیت در مسائلی نظیر شناسائی گفتار، شناسائی و تعبیر تصاویر و یادگیری ماشین بکار گرفته می شوند.

- شبکه عصبی، روشی برای محاسبه است که بر پایه اتصال به هم پیوسته چندین واحد پردازشی ساخته میشود.
- شبکه از تعداد دلخواهی سلول یا گره یا واحد یا نرون تشکیل میشود که مجموعه ورودی را به خروجی ربط میدهند.



انواع نرون هاي يك شبکه عصبی:

نرون هاي ورودی :

این نرون ها، پارامترهاي ورودی خود را از خارج شبکه دریافت میکنند و خروجی خود را به سایر نرون هاي داخل شبکه ارسال میکند.

وظیفه اصلی این نرون ها، دریافت و معرفی اطلاعات ورودی به شبکه می باشد.

نرون هاي میانی :

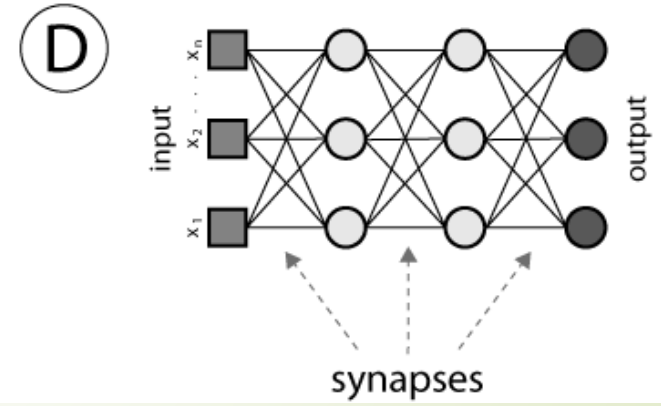
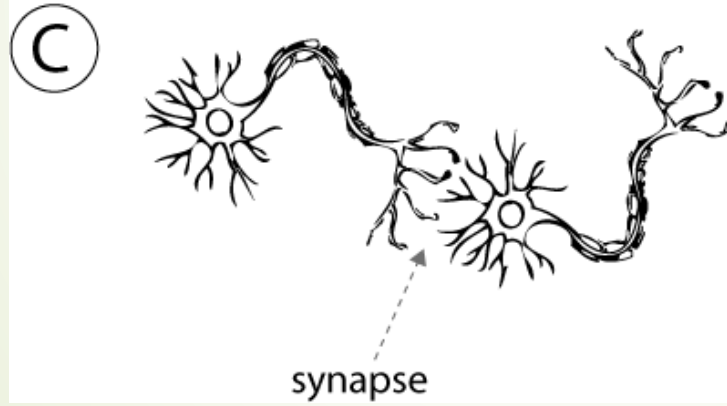
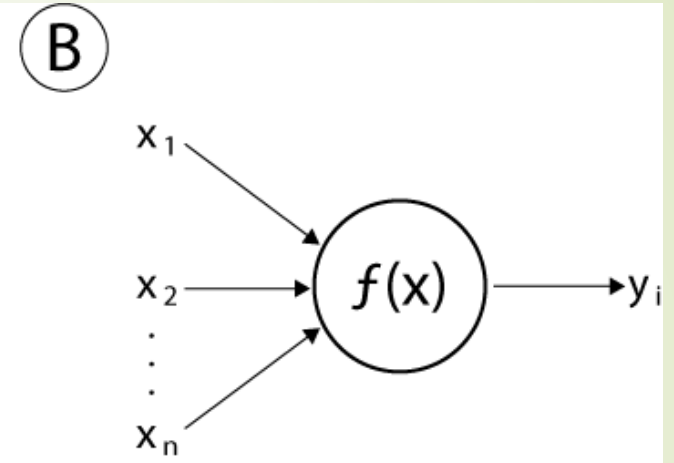
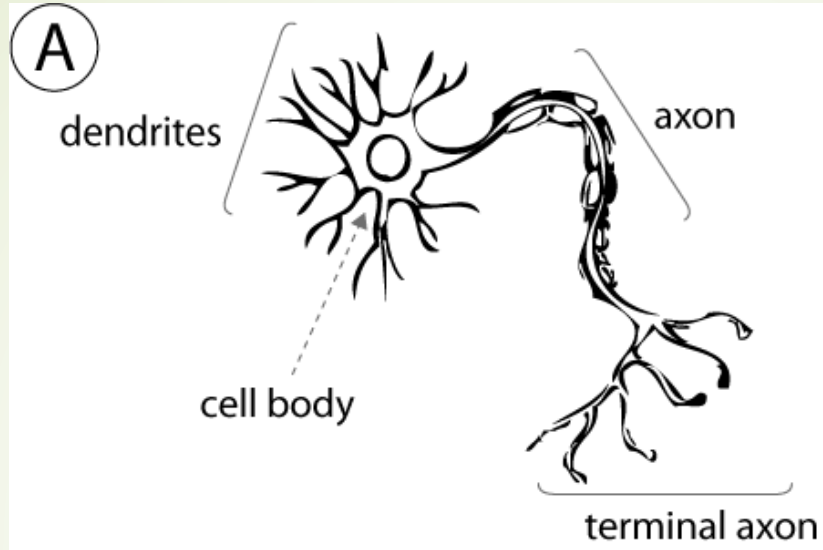
این نرون ها، پارامترهاي ورودی خود را از نرون هاي دیگر در داخل شبکه دریافت میکنند و خروجی خود را نیز به سایر نرون هاي داخل شبکه ارسال می نمایند.

وظیفه اصلی این نرون ها، پردازش اطلاعات می باشد.

نرون هاي خروجی :

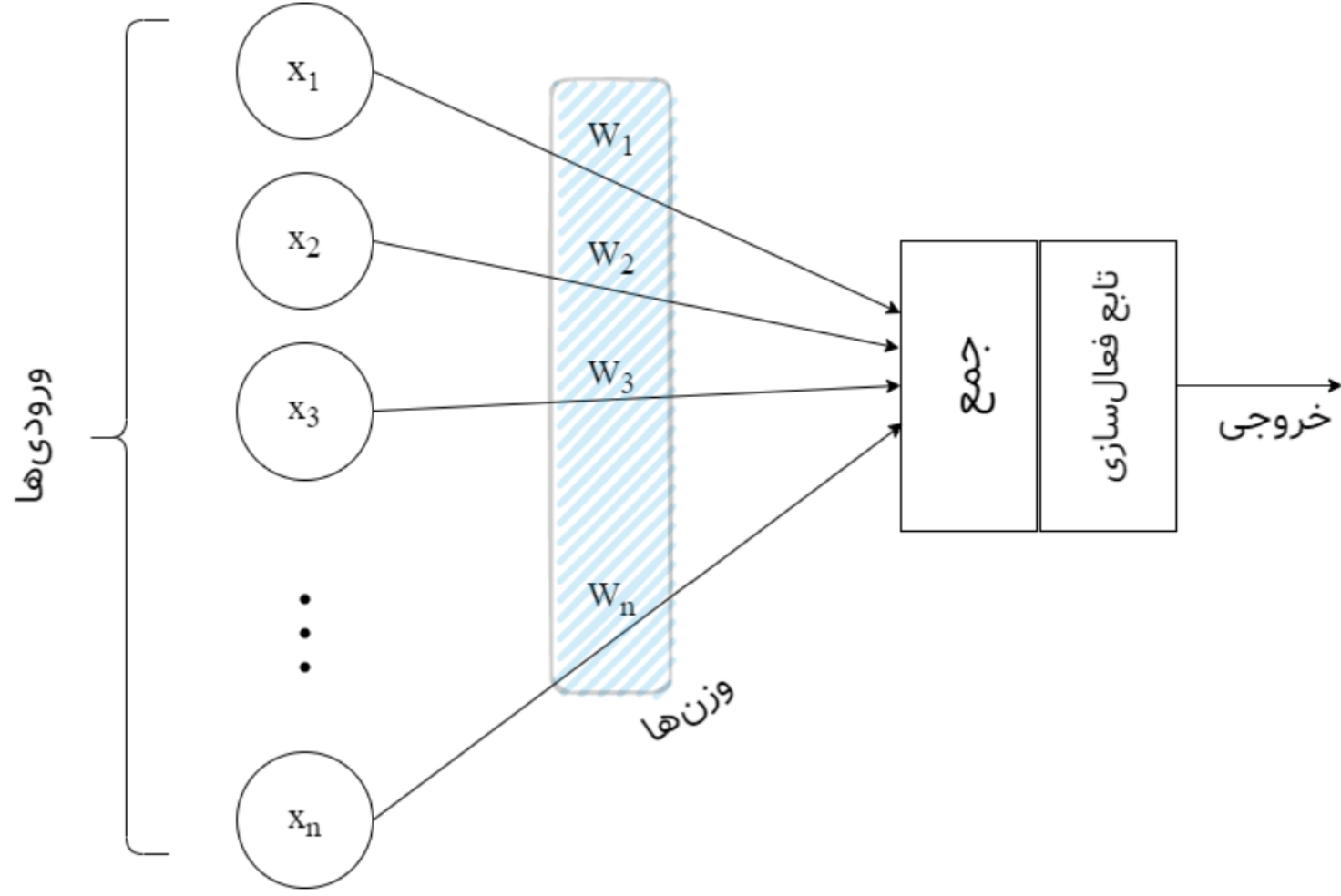
این نرونها، پارامترهاي ورودی خود را از نرون هاي داخل شبکه دریافت می کنند و خروجی خود بعنوان پاسخ نهایی شبکه، به خارج شبکه ارسال می دارند.

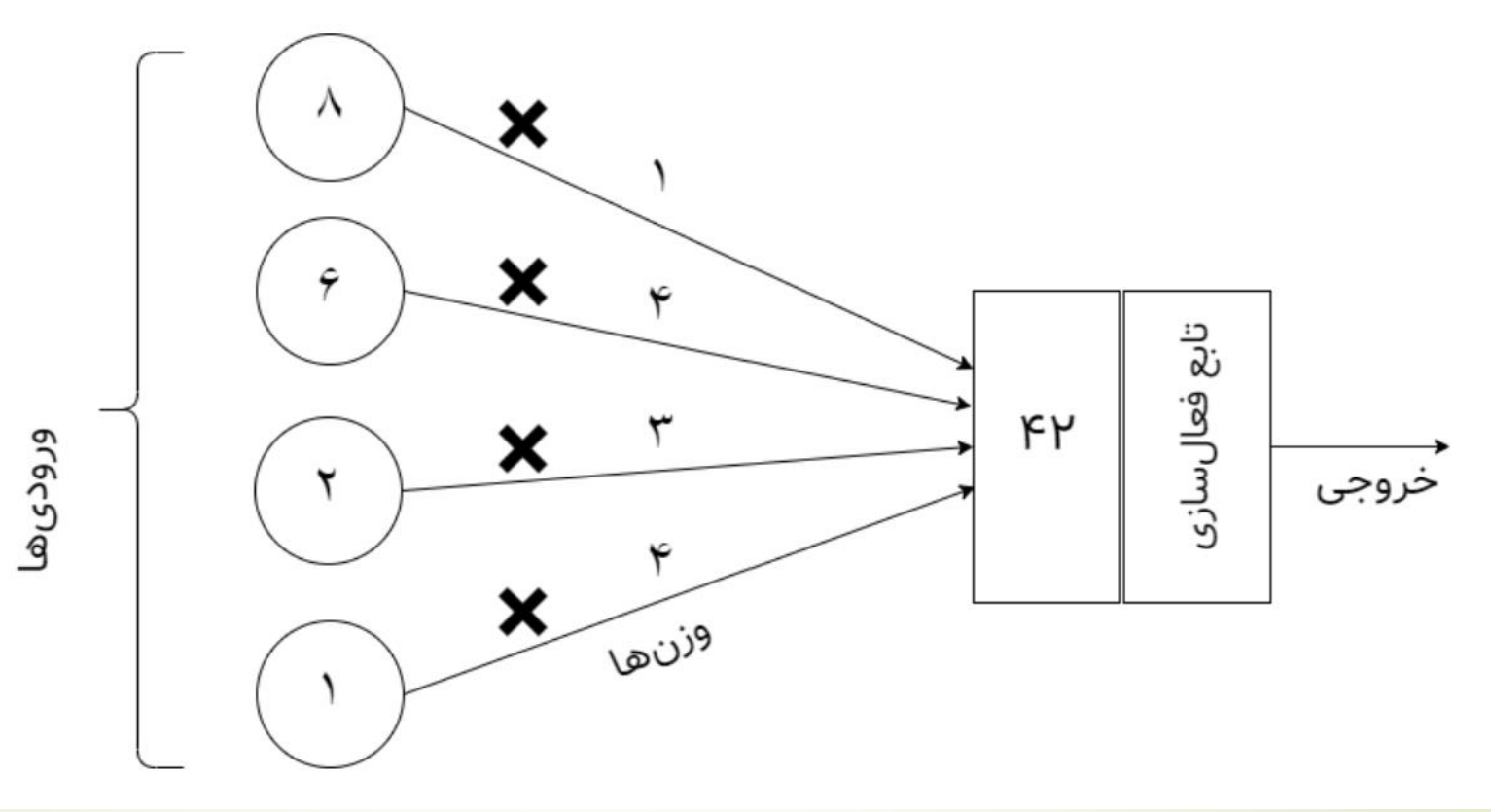
وظیفه اصلی این نرون ها، معرفی پاسخ نهایی شبکه برای بردار ورودی پردازش شده میباشد

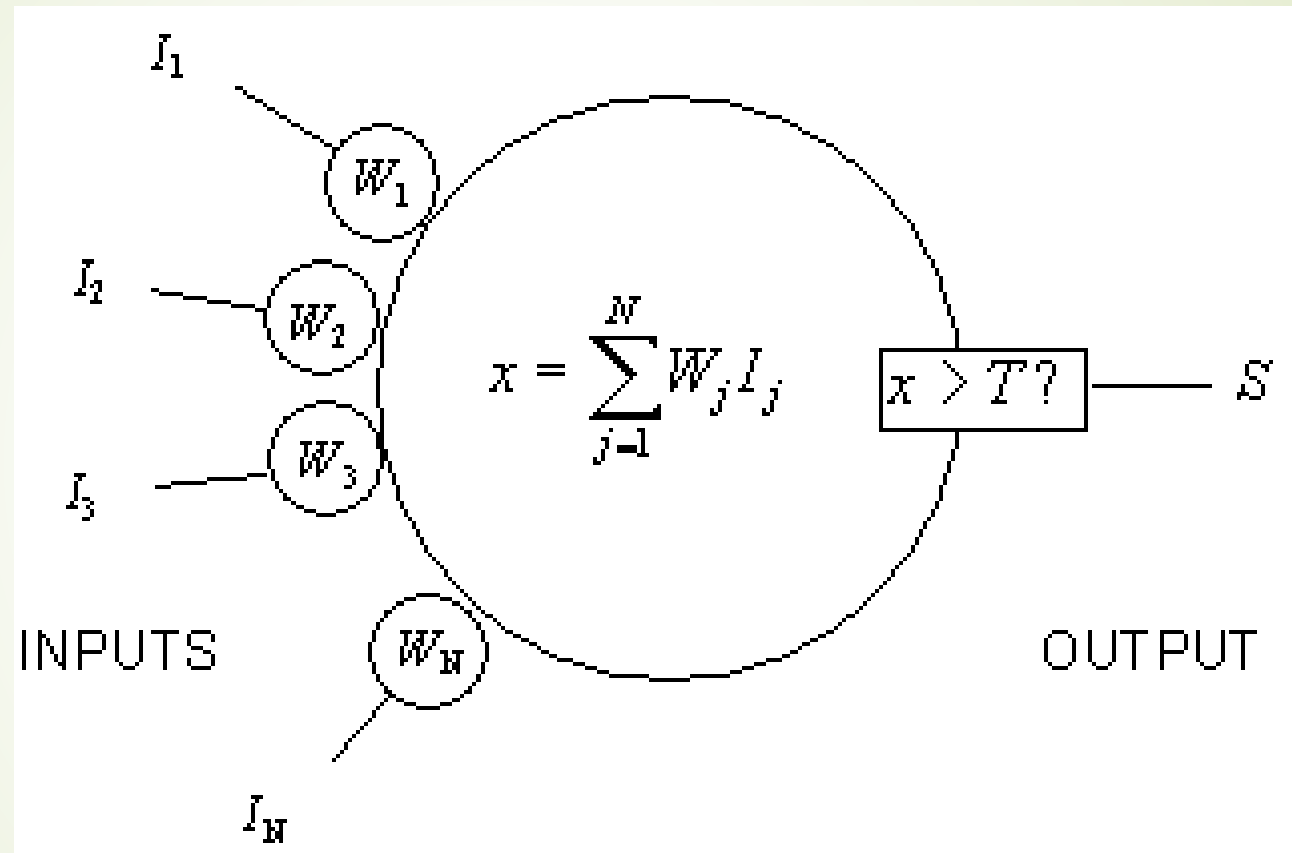


معماری شبکه:

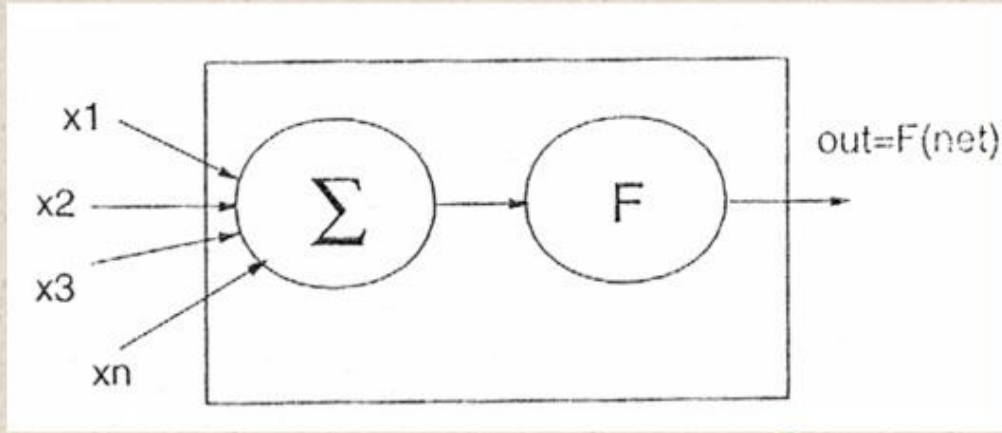
نحوه اتصال نوروں ها (اجزای تشکیل دهنده شبکه) را ساختار یا معماری شبکه گویند که این ساختار ارتباط تنگاتنگی با شیوه آموزش شبکه دارد. معماری شبکه های مصنوعی به صورت گراف های جهت دار موزونی است که در آن نوروں های مصنوعی، گره ها بوده و پیکانهای جهت دار به همراه وزن ها ارتباط بین ورودی و خروجی های نوروں را نشان میدهند.





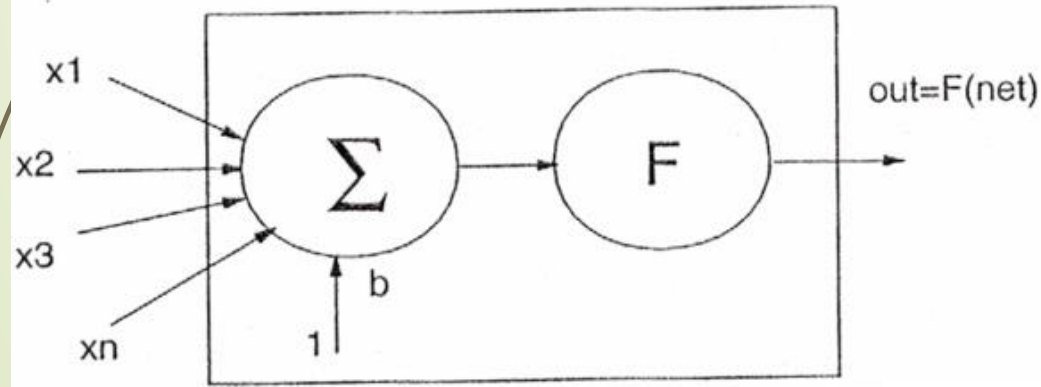


گراف یک شبکه عصبی:

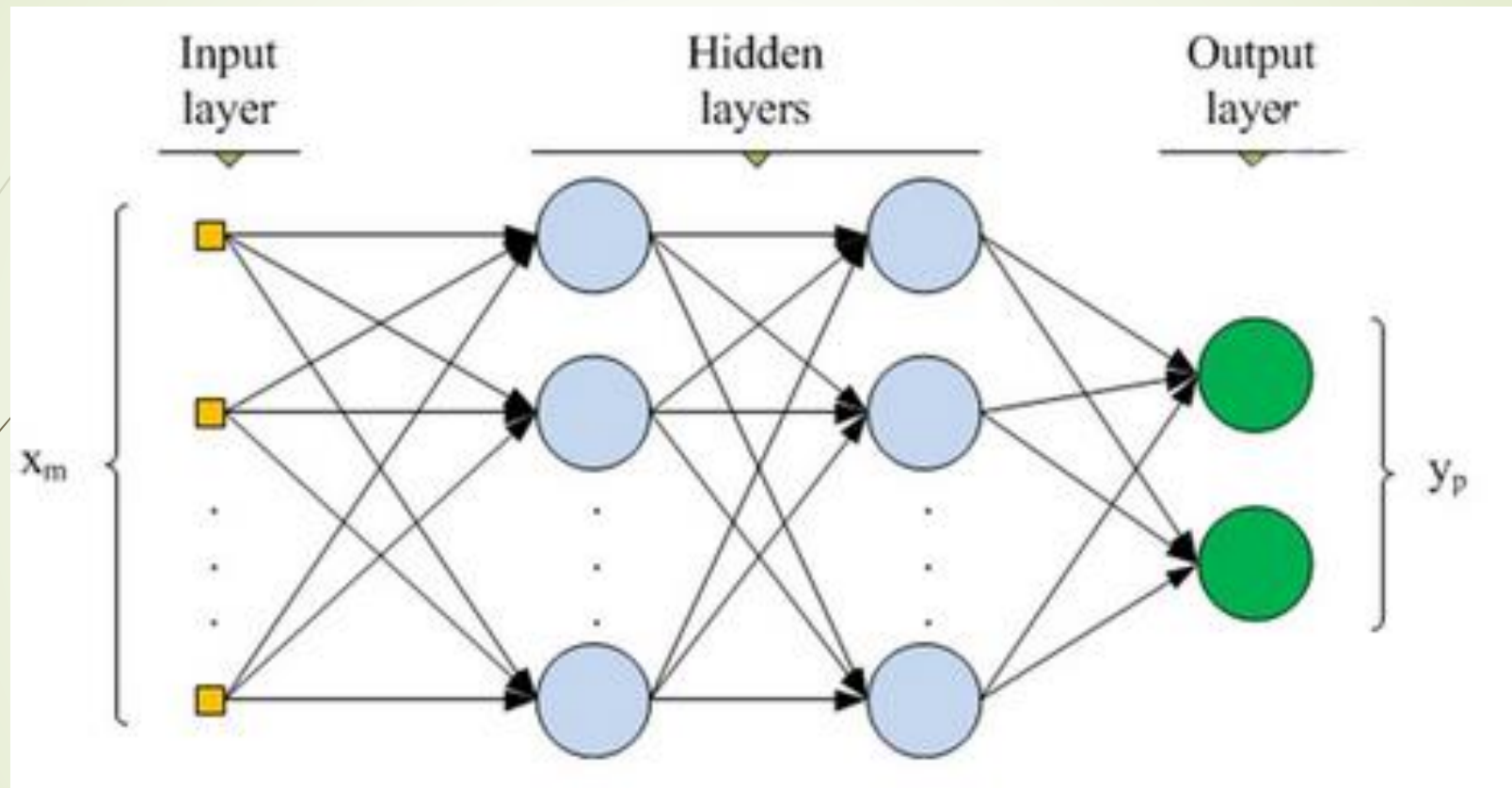


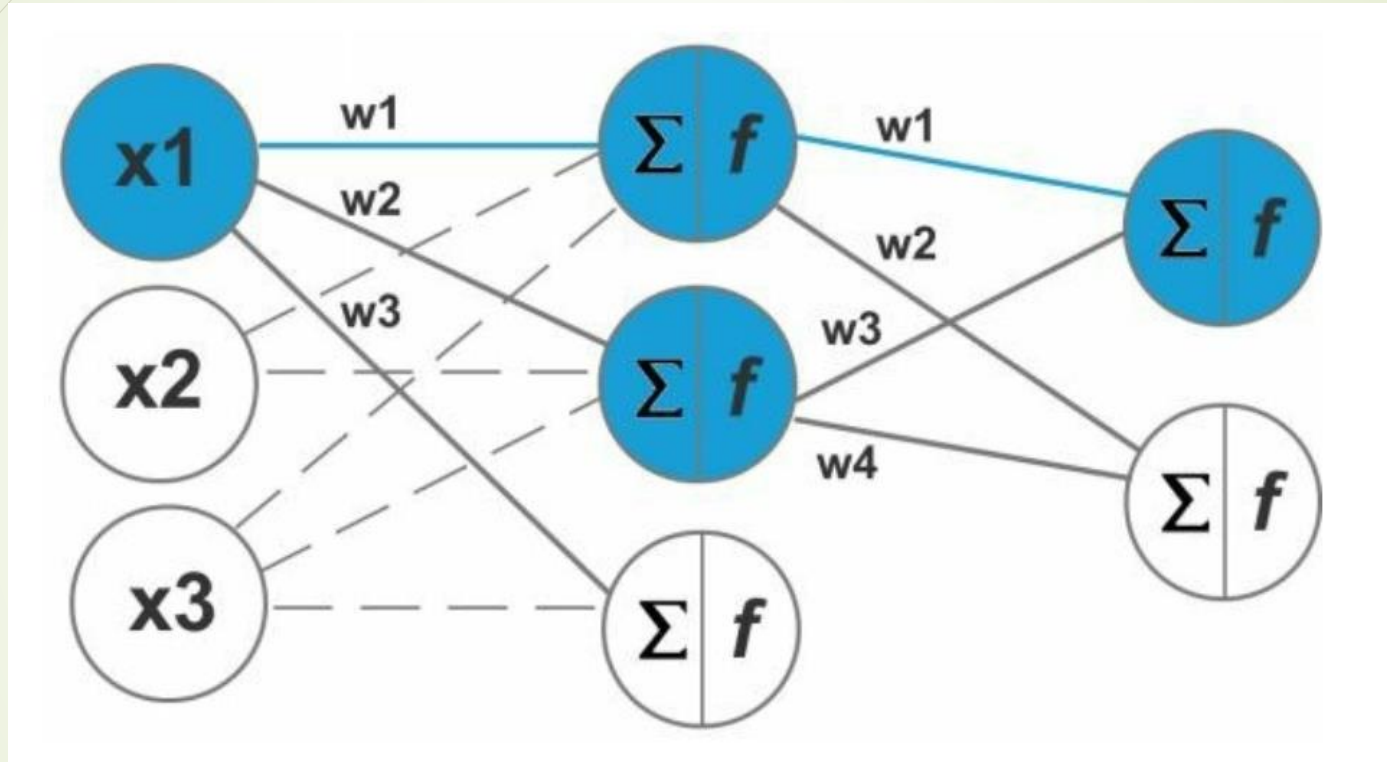
$$net = \sum_{i=1}^n x_i w_i$$

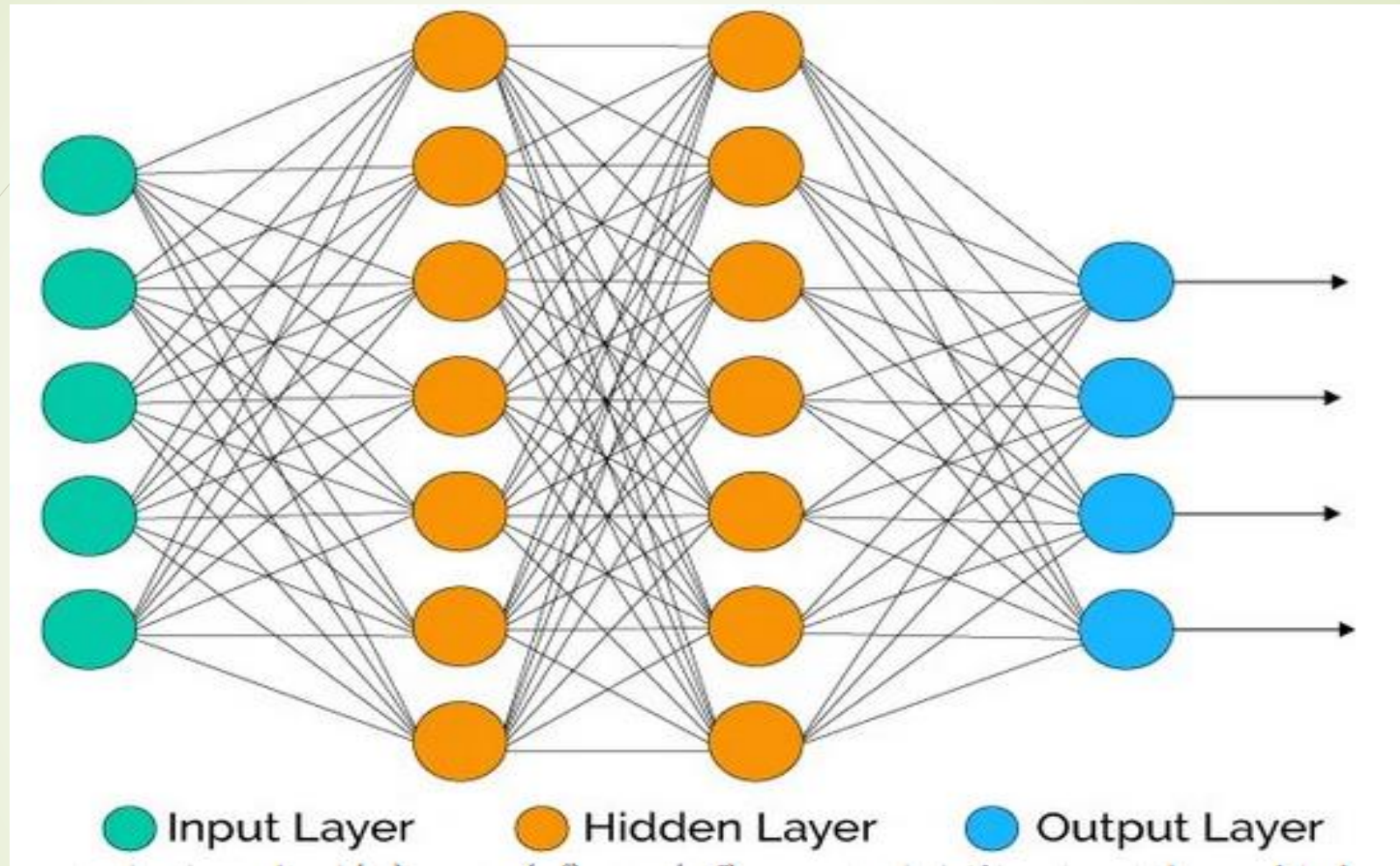
$$out = F(net)$$



$$net = \sum_{i=1}^n x_i w_i + b$$







انواع توابع فعال ساز:

ردیف	نام	تعریف تابع	علائم قراردادی
۱	آستانه‌ای دو مقداره	$a = 0, n < 0$ $a = 1, n \geq 0$	Sign
۲	آستانه‌ای دو مقداره متقارن	$a = -1, n < 0$ $a = 1, n \geq 0$	Sign
۳	خطی	$a = n$	Lin
۴	آستانه‌ای خطی متقارن	$a = -1, n < 0$ $a = n, -1 \leq n \leq 1$ $a = 1, n > 0$	SSA+1
۵	آستانه‌ای خطی	$a = 0, n < 0$ $a = n, 0 \leq n \leq 1$ $a = 1, n > 0$	
۶	زیگموئیدی	$a = \frac{1}{1 + e^{-n}}$	sig
۷	تانژانت هیپربولیکی	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	Ssig
۸	خطی مثبت	$a = 0, n < 0$ $a = n, n \geq 0$	Post 1

شبکه عصبی چگونه یاد می‌گیرد؟

اطلاعات به دو طریق در شبکه‌ی عصبی جریان دارند: زمانی که در حال یادگیری است؛ یا بعد از اینکه عمل یادگیری انجام شد. در این زمان‌ها الگوهای یادگیری به‌وسیله‌ی واحدهای ورودی وارد شبکه می‌شوند و لایه‌های واحدهای مخفی را برانگیخته می‌کنند و این لایه‌ها به واحدهای خروجی می‌رسند. به این طراحی رایج، شبکه عصبی پیش‌خور می‌گویند. هر واحدی اطلاعات ورودی را از واحدهای سمت چپ خود دریافت می‌کند و ورودی‌ها در وزن اتصالات مربوطه خود ضرب می‌شوند. هر واحدی تمامی ورودی‌هایی را که دریافت می‌کند به این طریق جمع می‌زند و (در ساده‌ترین نوع شبکه) اگر جمع بیش‌از یک مقدار آستانه مشخص شد، این واحد به هدف میرسد و واحدهای متصل به‌خود را (که در سمت راست هستند) راه می‌اندازد.

ویژگی های شبکه های مصنوعی

1- قابلیت یادگیری

تنظیم وزنهای الگو بر اساس داده ها و نزدیکتر کردن الگو به واقعیت

2- پردازش اطلاعات به صورت موازی

3- قابلیت تعمیم (پیش بینی)

4- پردازش اطلاعات به صورت متنی

چون در یک سیستم عصبی عمل پردازش توسط مجموعه ای از سلول ها و وزن های مربوطه انجام میشود، هر سلول متاثر از فعالیت سایر سلول هاست (پردازش به صورت متنی انجام میشود). با از بین رفتن بخشی از شبکه، کل شبکه مختل نمیشود.

5- مقاوم بودن

در شبکه های عصبی هر سلول به طور کاملاً مجزا عمل میکند و عملکرد شبکه حاصل از عملکرد کلی سلول هاست. یعنی اگر یک سلول اشتباه کند سلول های دیگر آن اشتباه را جبران میکنند.



کاربردهای های شبکه های عصبی:

1- پردازش تصویر


2- شناسایی الگوها

شناسایی چهره، اثر انگشت، تشخیص نوع صدا و دستخط و... .

3- پزشکی

تجزیه تحلیل علائم دستگاه ضربان نگار قلب و... .

4- سیستم های نظامی



5- سیستم های تجاری

6- برنامه ریزی، کنترل و جستجو

مثل کنترل رباط ها، موتورهای جستجوگر و ...

7- هوش مصنوعی

8- سیستم های قدرت

9- محاسبه یک تابع معلوم

10- تقریب یک تابع ناشناخته

و ...



معایب شبکه های عصبی:

قواعد مشخصی برای طراحی شبکه جهت یک کاربرد وجود ندارد.

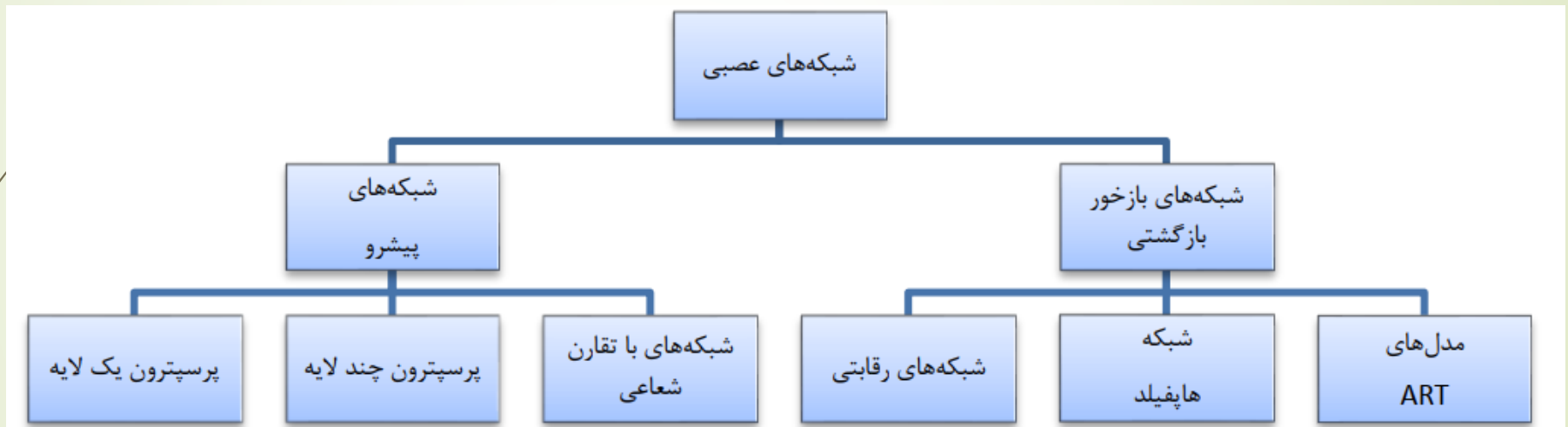
دقت نتایج بستگی به اندازه مجموعه آموزش دارد.

آموزش شبکه ممکن است مشکل یا حتی غیر ممکن باشد.

پیش بینی عملکرد آینده شبکه به سادگی امکانپذیر نیست.

در مورد مسایل مدل سازی، نمیتوان صرفاً با استفاده از شبکه عصبی به فیزیک مسئله پی برد.

انواع شبکه های مصنوعی بر اساس معماری



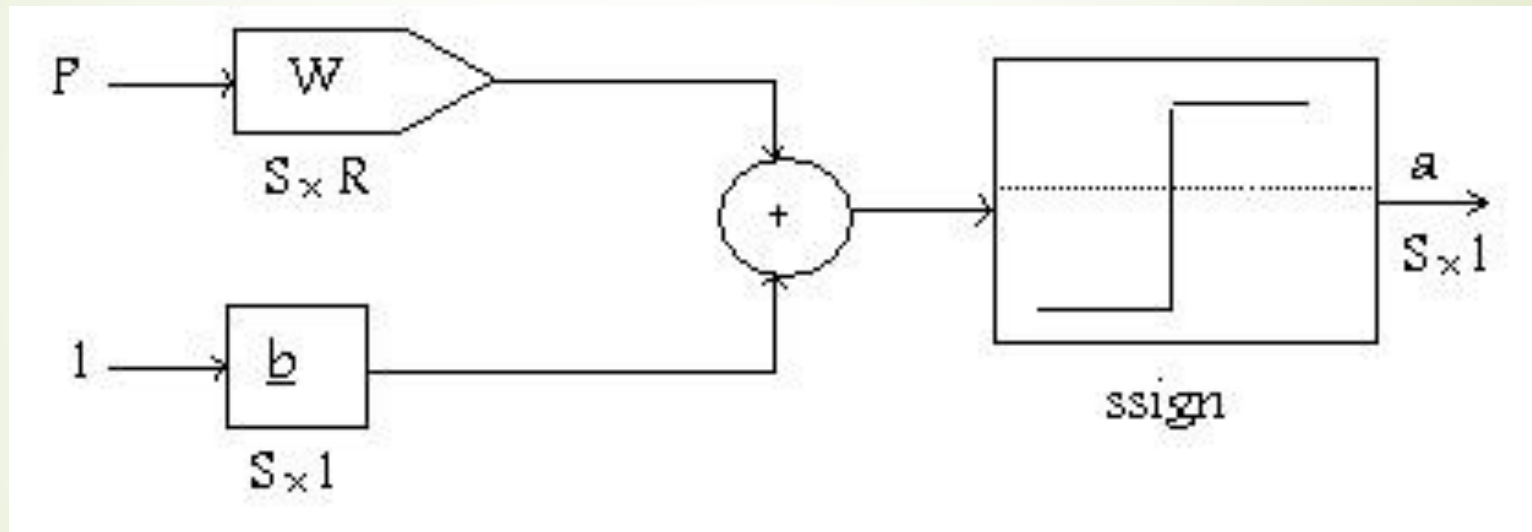
انواع شبکه عصبی:

- شبکه عصبی پرسپترون
- شبکه عصبی هاپفیلد
- شبکه عصبی همینگ
- شبکه عصبی کوهنن
- شبکه عصبی انتشار رو به عقب
- شبکه عصبی تاخیر زمانی
- و ...

شبکه عصبی پرسپترون

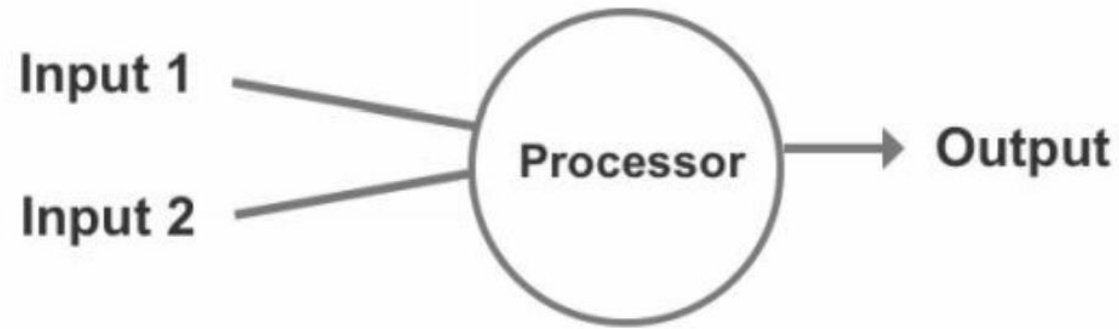
این شبکه عصبی بر مبنای یک واحد محاسباتی به نام پرسپترون ساخته میشود. یک پرسپترون برداری از ورودیها با مقادیر حقیقی را گرفته و یک ترکیب خطی از ورودیها را محاسبه میکند. اگر حاصل از یک مقدار آستانه بیشتر بود خروجی پرسپترون برابر با 1 و در غیر اینصورت معادل 1- خواهد بود.

پرسپترون تک لایه



پرسپترون (Perceptron)

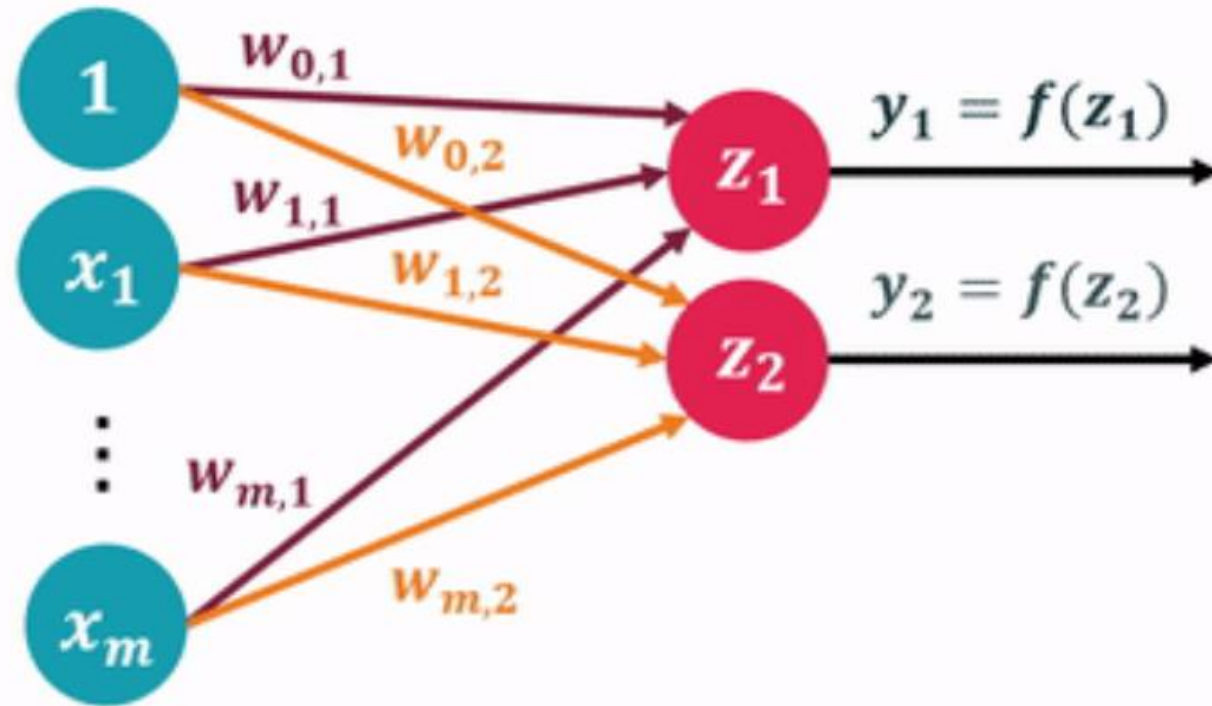
ساده‌ترین شکل یک شبکه عصبی روبه‌جلو، پرسپترون (Perceptron) است که در شکل زیر مشاهده می‌کنید.



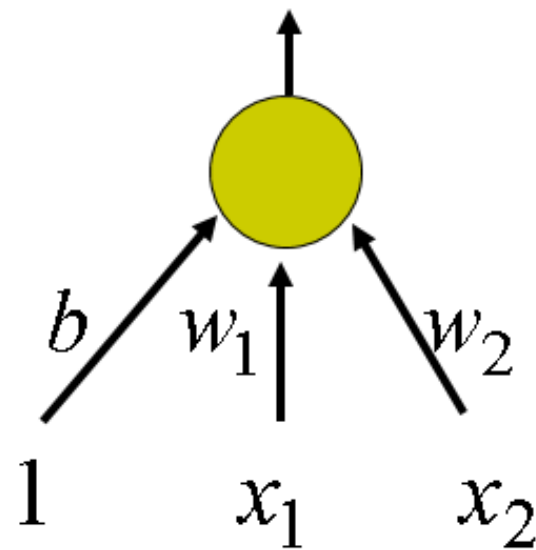
یک پرسپترون از یک یا چند ورودی، یک پردازنده (Processor) و یک خروجی تشکیل شده است. در یک

آموزش داده ها:

- در یادگیری بانظارت، پرسپترون ها برای آموزش داده و توسعه مدل پیش بینی کننده استفاده می شوند. مراحل آموزش داده به شرح زیر است:
- ورودی ها به پردازنده وارد می شوند (نورون ها/گره ها).
- پرسپترون مقدار ورودی ها را تخمین می زند.
- پرسپترون خطای بین تخمین و مقدار واقعی را محاسبه می کند.
- پرسپترون وزن هایش را باتوجه به خطا تنظیم می کند.
- تا وقتی که دقت مدل رضایت بخش نباشد، چهار مرحله قبلی را تکرار می کند.
- پس از آن می توان مدل آموزش داده شده را بر روی داده آزمایشی اعمال کرد.



$$\hat{y} = b + \sum_i x_i w_i$$



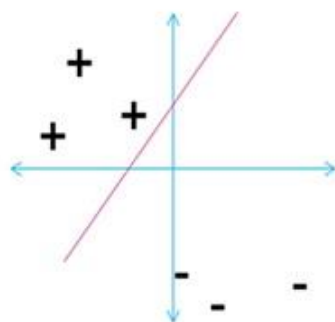
یادگیری پرسپترون:

یادگیری پرسپترون عبارت است از پیدا کردن مقادیر درستی برای وزن‌ها. بنابراین فضای یادگیری پرسپترون عبارت است از مجموعه تمام مقادیر حقیقی ممکن برای بردارهای وزن. خروجی پرسپترون توسط رابطه زیر مشخص میشود:

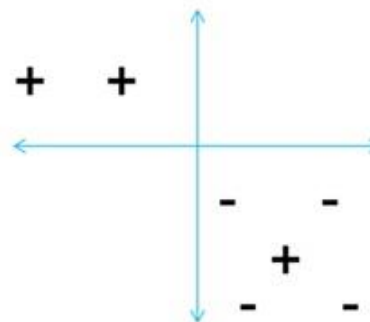
$$(x_1, x_2, \dots, x_n) = \begin{cases} 1 & w_1x_1 + w_2x_2 + \dots + w_nx_n > 0 \\ -1 & \text{o.w} \end{cases}$$

توابعی که پرسپترون تک لایه قادر به یادگیری می باشد:

یک پرسپترون فقط قادر است مثالهایی را یاد بگیرد که بصورت خطی جداپذیر باشند.



خطی جدا شدنی



خطی جدا نشدنی

مثال 1: یک مثال ساده:

فرض کنید می‌خواهیم شبکه ای را آموزش دهیم که تفاوت بین یک اتومبیل کوچک و یک اتوبوس را تشخیص دهد. برای این منظور هم مشخصات 7 اتومبیل را به شرح زیر داریم:

شماره	طول	ارتفاع	نوع وسیله	η
1	7	4	اتوبوس	15
2	6.5	4.5	اتوبوس	15.5
3	7.5	4.5	اتوبوس	16.5
4	99	4.5	اتوبوس	18
5	3	1.5	وسیله شخصی	6
6	2.5	1.7	وسیله شخصی	5.5
7	2	1.6	وسیله شخصی	5.8

مثال 1:

- در این مثال وزن ها را به صورت زیر در نظر گرفتیم:

- $W_1 = 1$
- $W_2 = 2$
- $\eta = W_1 * X_1 + W_2 * X_2$

- تابع فعال ساز:

- $$f(\Sigma) = \begin{cases} -1, & \eta < 9 \\ 1, & \eta \geq 9 \end{cases}$$

ارتفاع

#۱

#۲ #۴

#۳


#۶

#۷ #۵

معادله خط جدا کننده

طول



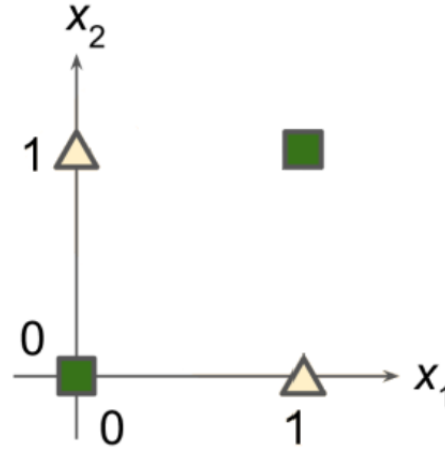


نقطه‌ضعف پرسپترون، خروجی دودویی است و همین نقطه‌ضعف باعث قطبی شدن نتایج Polarizing Result در دو حالت زیر می‌شود:

- تغییرات کوچک در وزن‌ها.
- اریبی در هر کدام از پرسپترون‌ها در یک شبکه عصبی بزرگ تر، این ضعف ممکن است باعث تغییرات ناگهانی در داخل شبکه و چرخش کامل خروجی نهایی بشود.

■ محدود بودن آن ها به برآورد توابع خطی ساده است.

در نتیجه، آموزش یک مدل بادقت بالا که بتوان آن را با موفقیت روی داده آزمایشی و داده‌های ورودی آینده، اعمال کرد، بسیار سخت می‌شود.

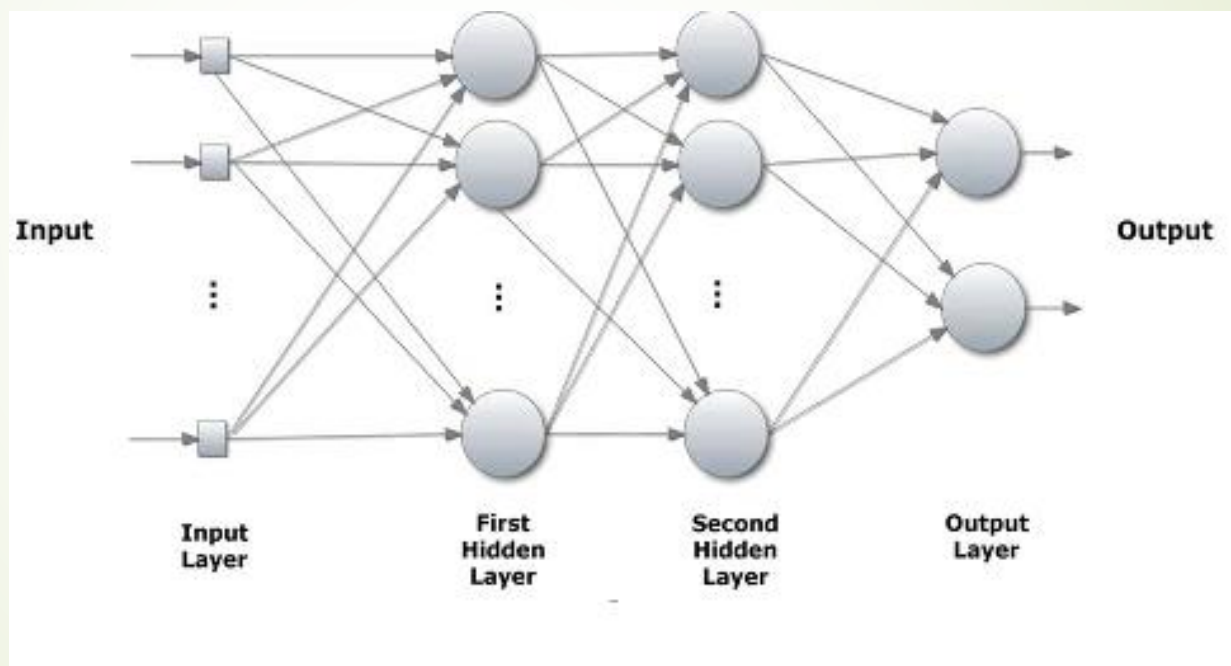


شکل ۳: مسئله XOR

نمی‌توان با یک خط مربع‌ها را از مثلث‌ها جدا کرد، درست است؟ بنابراین پرسپترون نمی‌تواند این مسئله را حل کند.

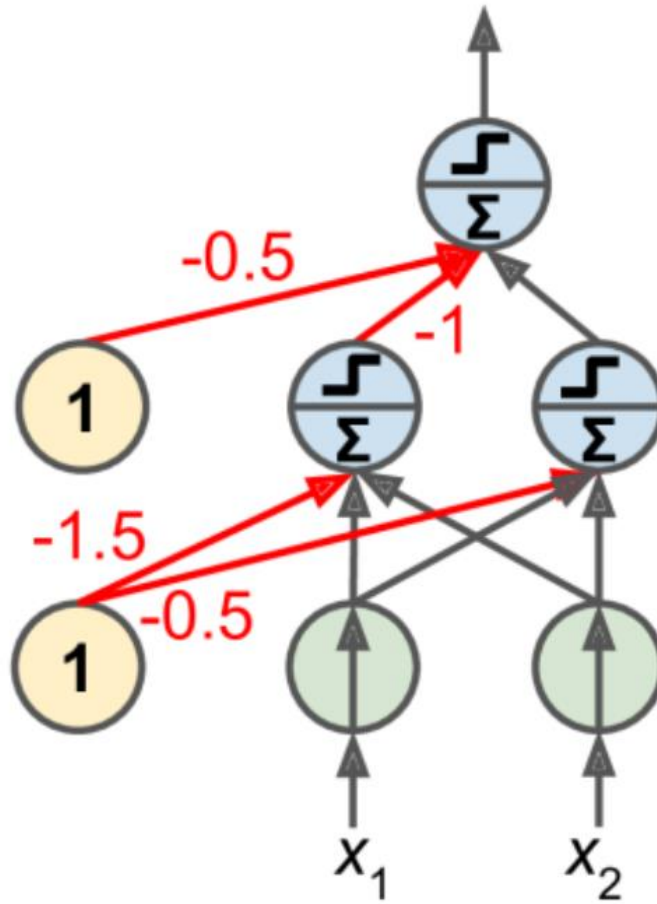
پرسپترون های چند لایه:

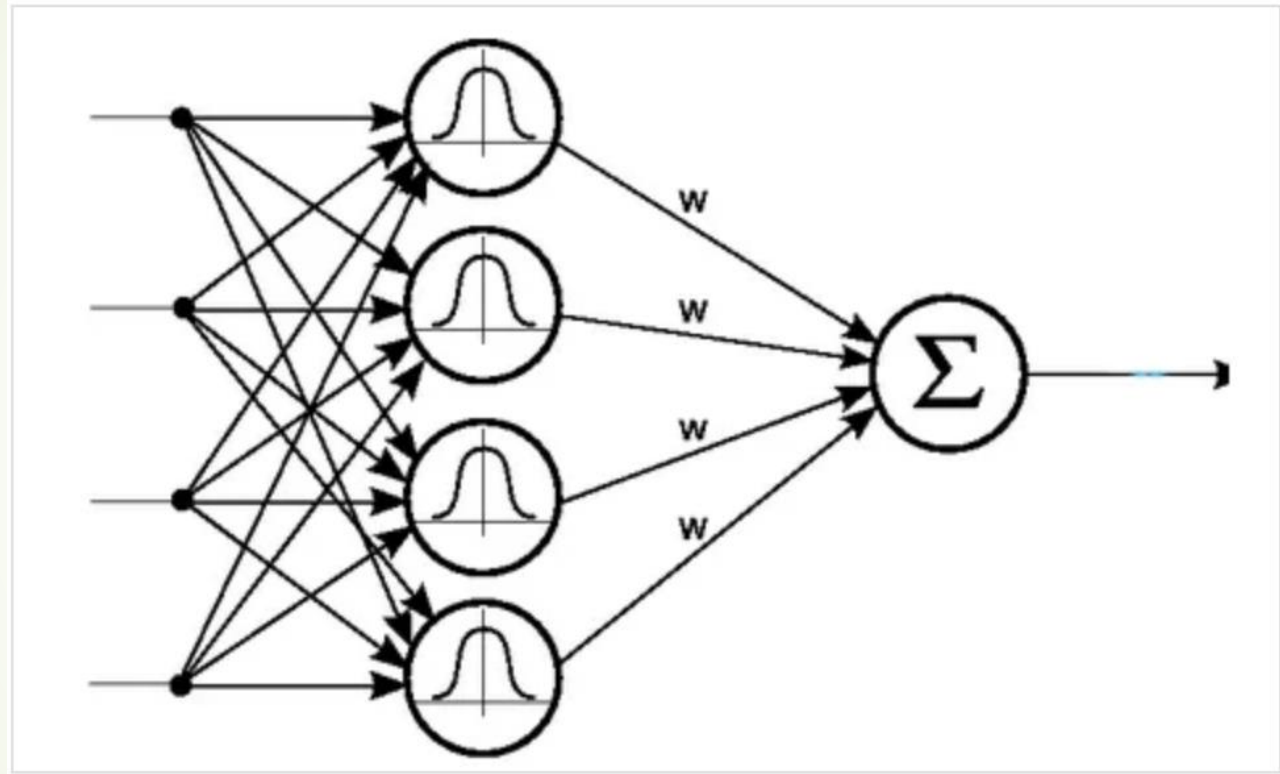
برای حل این مشکل از شبکه های عصبی چند لایه استفاده می کنند.

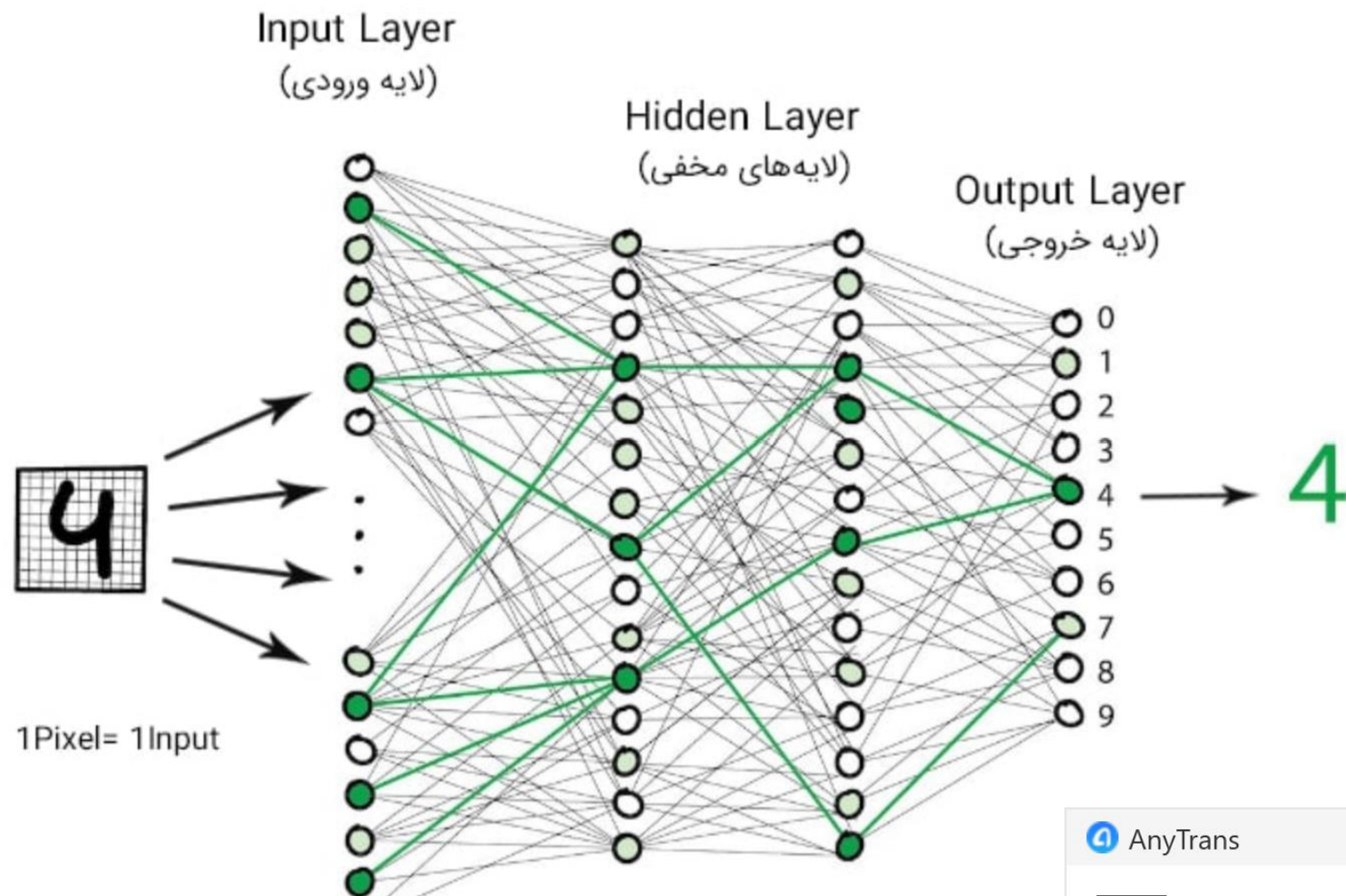


پرسپترون چندلایه یا MLP:

یکی از پایه‌ای‌ترین مدل‌های عصبی موجود، مدل پرسپترون چند لایه یا Multi-Layer Perceptron به اختصار (MLP) است که عملکرد انتقالی مغز انسان را شبیه‌سازی می‌کند. در این نوع شبکه عصبی، بیشتر رفتار شبکه‌ای مغز انسان و انتشار سیگنال در آن مد نظر بوده است و از این رو، گهگاه با نام شبکه‌های پیش‌خورد Feedforward Networks نیز خوانده می‌شوند. هر یک از سلول‌های عصبی مغز انسان، موسوم به نورون Neuron پس از دریافت ورودی (از یک سلول عصبی یا غیر عصبی دیگر)، پردازشی روی آن انجام می‌دهند و نتیجه را به یک سلول دیگر (عصبی یا غیر عصبی) انتقال می‌دهند. این رفتار تا حصول نتیجه‌ای مشخص ادامه دارد، که احتمالاً در نهایت منجر به یک تصمیم، پردازش، تفکر و یا حرکت خواهد شد.

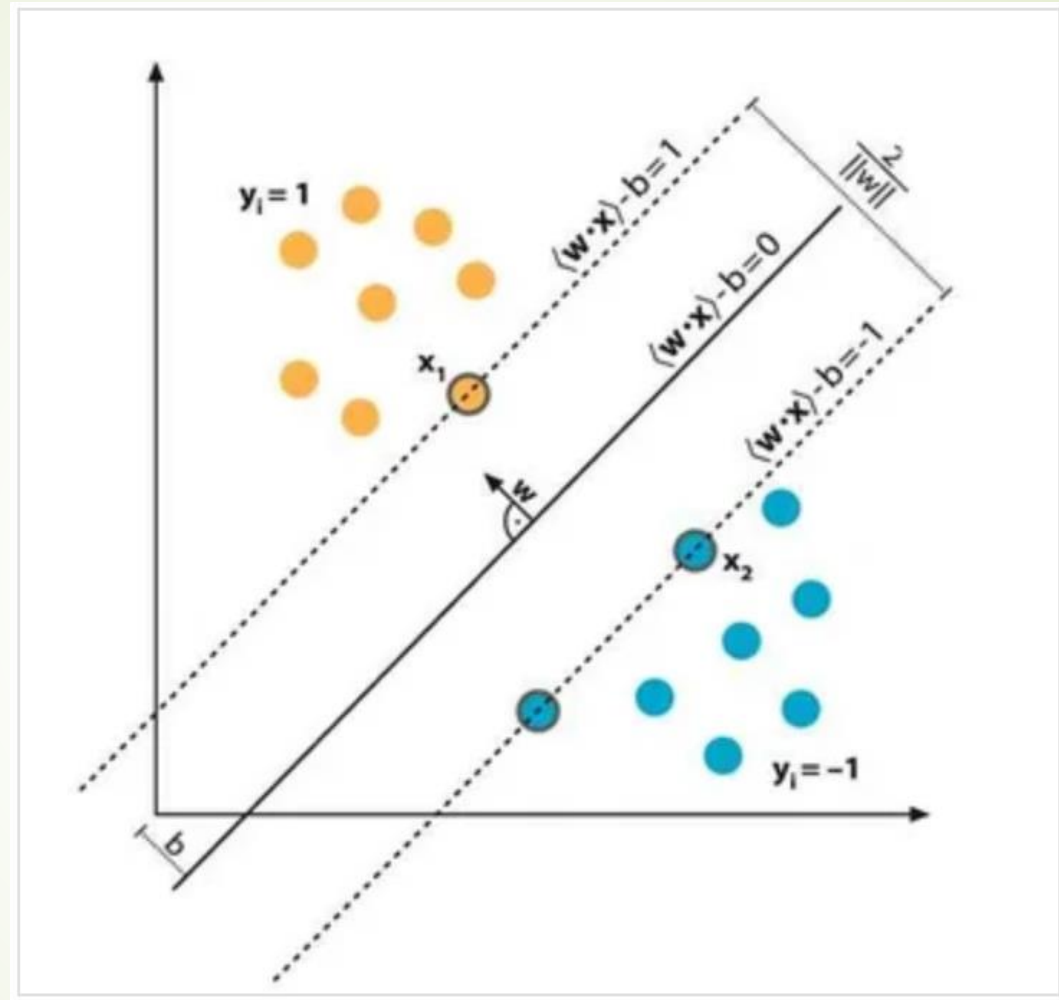






شبکه‌های عصبی شعاعی یا RBF:

مشابه الگوی شبکه‌های عصبی MLP، نوع دیگری از شبکه‌های عصبی وجود دارند که در آن‌ها، واحدهای پردازنده، از نظر پردازشی بر موقعیت خاصی متمرکز هستند. این تمرکز، از طریق توابع شعاعی یا Radial Basis Functions به اختصار RBF مدل‌سازی می‌شود. از نظر ساختار کلی، شبکه‌های عصبی RBF تفاوت چندانی با شبکه‌های MLP ندارند و صرفاً نوع پردازشی که نورون‌ها روی ورودی‌هایشان انجام می‌دهند، متفاوت است. با این حال، شبکه‌های RBF غالباً دارای فرایند یادگیری و آماده‌سازی سریع‌تری هستند. در واقع، به دلیل تمرکز نورون‌ها بر محدوده عملکردی خاص، کار تنظیم آن‌ها، راحت‌تر خواهد بود.

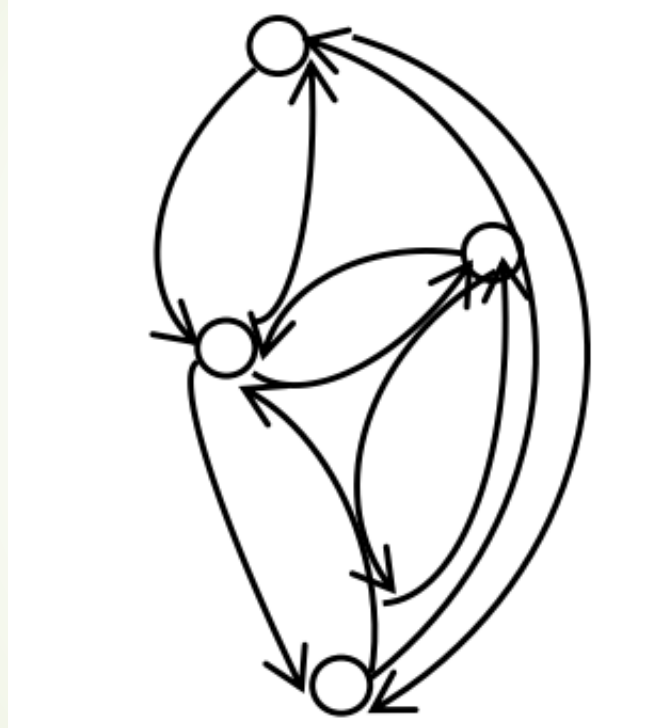


شبکه عصبی هاپفیلد یا Hopfield:

این نوع شبکه عصبی، بیشتر دارای ماهیتی شبیه به یک سیستم دینامیکی است که دو یا چند نقطه تعادل پایدار دارد. این سیستم با شروع از هر شرایط اولیه، نهایتاً به یکی از نقاط تعادلش همگرا می‌شود. همگرایی به هر نقطه تعادل، به عنوان تشخیصی است که شبکه عصبی آن را ایجاد کرده است و در واقع می‌تواند به عنوان یک رویکرد برای حل مسائل طبقه‌بندی استفاده شود. این سیستم، یکی از قدیمی‌ترین انواع شبکه‌های عصبی است که دارای ساختار بازگشتی است و در ساختار آن فیدبک‌های داخلی وجود دارند.

گرافی از شبکه های بازگشتی

• هاپفیلد:



مثال 2 (cereals)

➤ حل مثالی در R (neuralnet):

➤ بررسی ارتباط کالری و پروتئین مصرفی و میزان چربی، سدیم و فیبر 76 نفر با درجه بندی سلامتی آنها.

`NN = neuralnet (rating ~ calories + protein + fat + sodium +
fiber, trainNN, hidden = 3 , linear.output = T)`

➤ `plot(NN)`

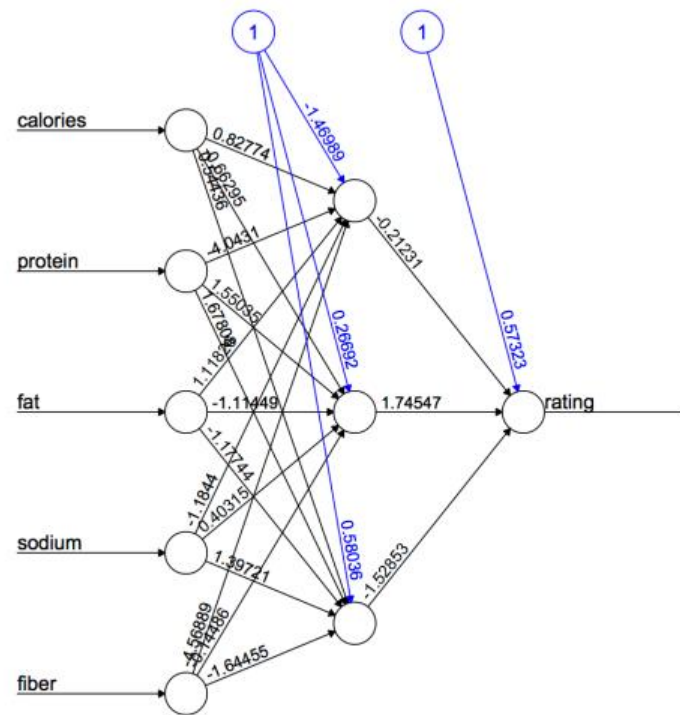
مثال 2 (cereals)

- حل مثالی در R (neuralnet):
- بررسی ارتباط کالری و پروتئین مصرفی و میزان چربی، سدیم و فیبر 76 نفر با درجه بندی سلامتی آنها.

```
NN = neuralnet(rating ~ calories + protein + fat + sodium + fiber,  
trainNN, hidden = 3 , linear.output = T )
```

➡ plot(NN)

شمای کلی شبکه برآورد شده:



Error: 0.088385 Steps: 146

مقادیر واقعی در برابر مقادیر پیش بینی

