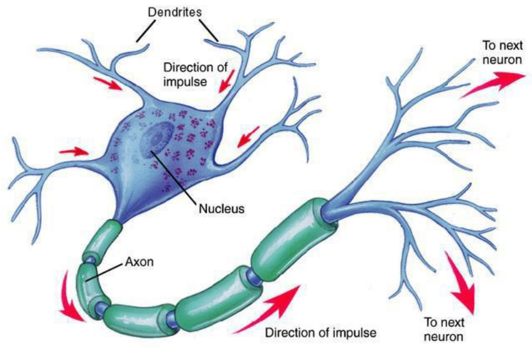
Perceptron

د. ياسر خضرا

مقدمة:
تحدثنا في السابق عن الشبكات العصبونية، ووصفها العام بأنها محاكاة للشبكة العصبية الموجودة في دماغ الإنسان، وفي هذه المحاضرة سنكمل ما بدأناه.

بعض خصائص الخلايا العصبية البيولوجية:



- كل عصبون يستقبل دخل (معلومة) من 50000 إلى 80000 من العصبونات المرافقة.
- التشعبات هي مكان استقبال الدخل.
- المعلومة المنقولة عبر المشبك لها وزن.
- العصبونات هي خلايا دماغية، ويقدّر عددها ب 10^{12} مشبك، ويوجد 10^{14} مشبك اتصال بالدماغ البشري.
- العصبونات عادةً تكون من الشكل MIMO (multi input, multi output)،

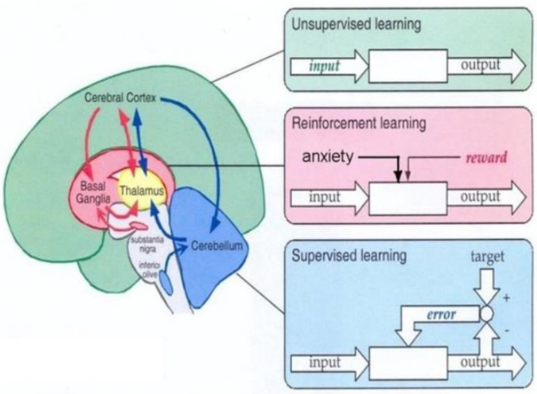
وهي الحالة العامة الأكثر انتشاراً، ولكن في بداية دراستنا سنعمل على الحالة البدائية وهي مداخل مختلفة ومخرج وحيد، وتكون وظيفة هذا العصبون غالباً هي التصنيف، مثال عن الخرج: إما ذكر أو أُنثى، إما ناجح أو راسب.

مشابك الانتقال تمثل مبدلات (transducer)

- وظيفة المبدلات: نقل المعلومات من شكل لآخر.

أمثلة على المبدلات: الميكروفون: يحول تخلخل طبقات الهواء إلى إشارة كهربائية - مكبر الصوت: يحول الإشارة الكهربائية إلى تخلخل في طبقات الهواء (أي صوت).

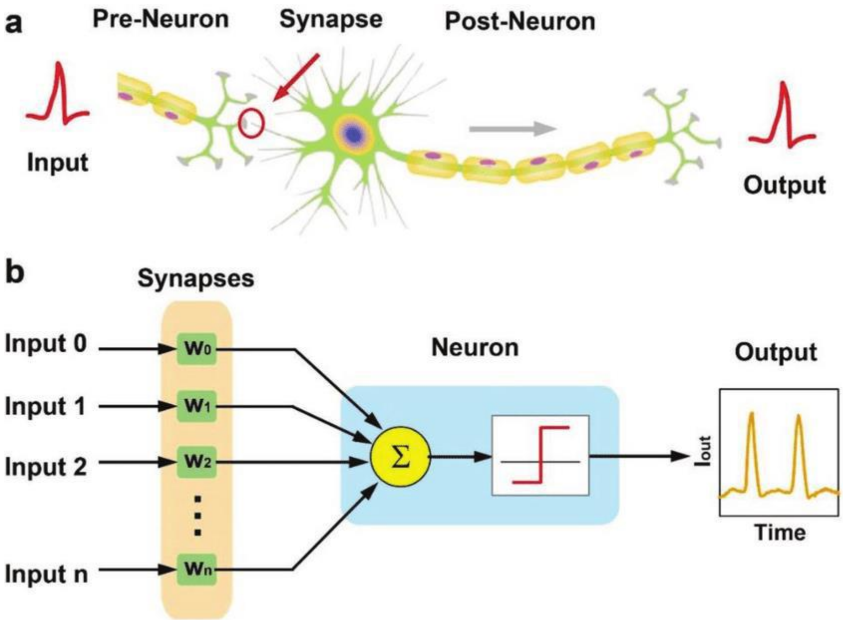
التعلم في الدماغ وربطه بتعلم الآلة:



- ✓ إنّ فهم قواعد عمل الدماغ هو الأساس من أجل فهم كيفية عمل تدريب لشبكة عصبونية، وفهم تطبيقها.

1. يمكننا القول بأن الSupervised Learning مشابه لآلية عمل المخيخ، وهو التعلم الذي يحتاج إلى مشرف (معلم)، مثال: تعلم قيادة السيارة للمرة الأولى.
2. ويمكننا القول أيضاً بأن ال Unsupervised Learning مشابه لآلية عمل قشرة الدماغ، وهو يمثل التعلم بالتجربة، مثال: تصنيف الأشياء حسب شكلها.
3. Reinforcement Learning: ويمثل برمجة الإجراءات والمراحل وهو موجود بالعقل الباطن.

ربط شكل العصبونات بشكل ال Perceptron



- إضاءة:
- ال perceptron تم اختراعه من قبل Frank Rosenblatt في 1957 أثناء محاولته لفهم الذاكرة البشرية والتعلم والعمليات المعرفية.
- نشر ورقة بحثية عام 1958 بعنوان: The perceptron نموذج احتمالي لتخزين المعلومات والتنظيم في الدماغ.

الأجزاء الرئيسية لـ Perceptron:

وهو مكون من أربع طبقات:

1) **طبقة الدخل:** تقوم باستقبال الدخل من أجل أن يتم معالجته ويكون الدخل على شكل قيم، صور، الوان...
2) **طبقة الأوزان:** وتتألف من كتلتين:

- Weight: وزن (عامل توليف) كل دخل على ال perceptron، (أكبر من حد معين ← تضخيم، أقل من حد معين ← تخميد) ويضاف له الانزياح (Bias).
- Bias: $\theta = w_0$ يُضاف عامل الانزياح على الجامع في الوسط، يُسمى أيضاً بعامل التوازن وتكون قيمته ثابتة وتضاف من أجل التوازن وتحقيق قيمة (عتبة دنيا مثلاً من أجل تفعيل عنصر التنشيط) ، (ممكن أن تكون قيمة Bias موجبة أو سالبة).

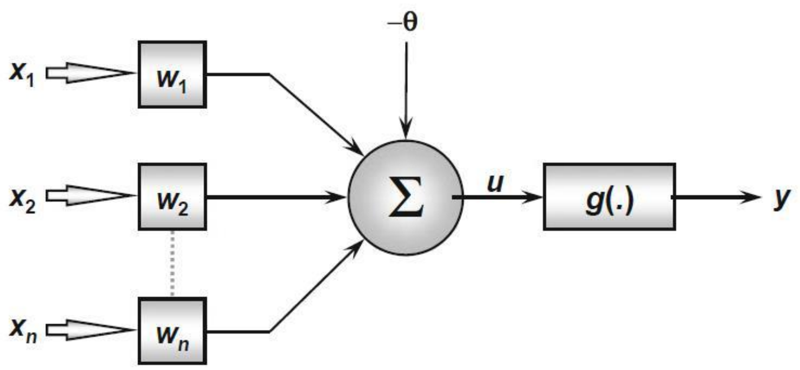
3) الجامع Net Sum: هو نتيجة عملية ضرب كل معامل دخل بالوزن مضافاً إليه الانزياح.

4) تابع التنشيط Activation Function: يحدد فيما إذا كان سيتم تفعيل (تنشيط) العصبون من خلال شرط أو حد معين، وله عدة أشكال: أبسط شكل هو تابع العتبة إذ يتنشط فوق حد معين ويخمد ما دونه.

ولدينا بالإضافة إلى الأجزاء الأربعة الرئيسية الخرج (y):

يجب أن يكون الخرج مكمم (يعطي خرج معروف) ويكون محدد ومرتبط بال Activation function.

- ملاحظة:
- تكون القيم في طبقة الدخل قيم حقيقية أو Binary.
- تكون القيم في طبقة الأوزان (Weight, Bias) قيم حقيقية فقط.
- تكون قيم الخرج (y) قيم Binary (مثال على ذلك: (في نهاية المحاضرة) عند تمثيل seven line segment يتم تمثيل الخرج على ثلاث بتات)



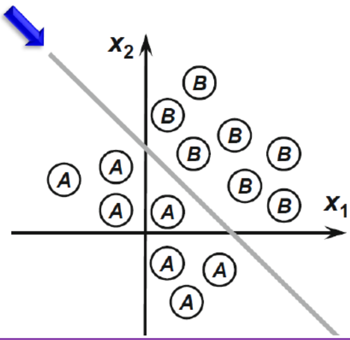
- بعد دراسة المخطط السابق وأقسامه الرئيسية، نجد التالي:
  - ✓ تابع الجمع ثابت.
  - ✓ الدخل والانزياح لا نقوم بالتعديل عليهم.
- ومنه نجد أن عملنا حالياً سوف يقتصر على الأوزان وتابع التنشيط.
  - تابع التنشيط: إذا تغير التابع سوف تتغير النتائج، وسيتم مناقشته لاحقاً.
  - الأوزان: مهمة وآلية التعديل على الأوزان ستكون محور الاهتمام القادم.
- مثال على كيفية استخدام ال perceptron في التصنيف:

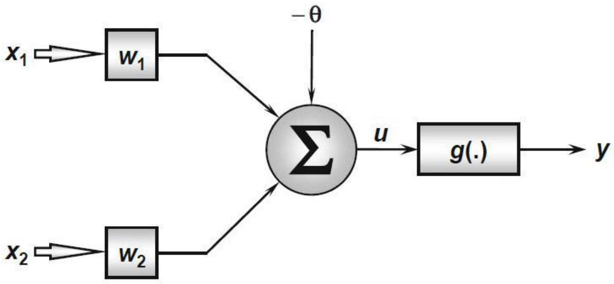
هدفنا هنا إيجاد الحد الفاصل بين المستويين، أي إيجاد حل معادلة هذا المستقيم.

$$y = \begin{cases} 1, & \text{if } \sum w_i \cdot x_i - \theta \geq 0 \Leftrightarrow w_1 \cdot x_1 + w_2 \cdot x_2 - \theta \geq 0 \quad \text{(Class B)} \\ -1, & \text{if } \sum w_i \cdot x_i - \theta < 0 \Leftrightarrow w_1 \cdot x_1 + w_2 \cdot x_2 - \theta < 0 \quad \text{(Class A)} \end{cases}$$

$$w_1 \cdot x_1 + w_2 \cdot x_2 - \theta = 0$$

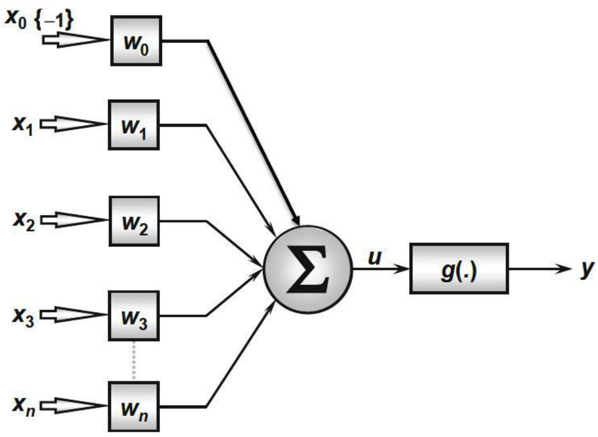
- معادلة المستقيم $= 0$
- إذاً خوارزمية التعلم ستدور في النهاية حول إيجاد معادلة المستقيم والأوزان التي تحققه.
- أي أن كل التعلم لاحقاً سوف يكون إيجاد هذه الأوزان، للحصول بالنهاية على أوزان قادرة على هذا الفرز بدون خطأ (في بعض الحالات تقبل نسبة خطأ صغيرة، وتقدّر حسب الحالة).





إعادة رسم الشبكة العصبونية:

- ✓ أراد مصمموّ الشبكات العصبونية التعديل على بنيتها، وذلك من خلال جعل الانزياح (bias) كدخل للشبكة.
- ✓ في الشبكة التالية وضعنا قيمته كدخل $y_0 = -1$ (تختلف باختلاف المرجع المدروس والحاجة).
- ✓ قيمة الوزن $w_0 = \theta$، ونلاحظ التغير في شكله:



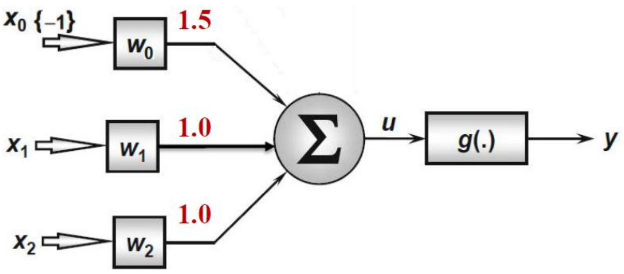
مع ملاحظة التغير في العلاقة الرياضية:

$$u = \sum_{i=1}^{n} w_i . x_i - \theta \Rightarrow u = \sum_{i=0}^{n} w_i . x_i$$

تشكيل البوابات المنطقية

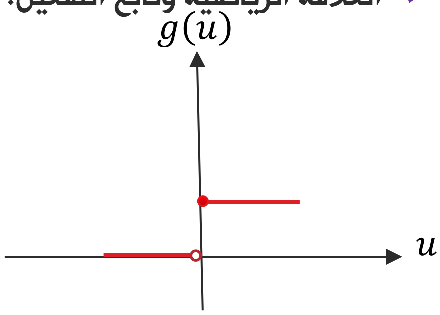
1. AND:

- ✓ نقوم في البداية بكتابة جدول الحقيقة مع إضافة x_0 (تمثل الانزياح) وهو مقدار ثابت.
- ✓ وظيفتنا هي إيجاد الأوزان التي تحقق لنا جدول الحقيقة وتدعى هذه العملية بالتعلم، ولكن من حيث الخبرة المتوافرة لدينا حالياً لا نستطيع فعل ذلك، لذلك نقوم بالبحث عن شبكة مدربة ولها أوزان، ونقوم بتجربتها.



x_0	x_1	x_2	output
−1	0	0	0
−1	0	1	0
−1	1	0	0
−1	1	1	1

✓ العلاقة الرياضية وتابع التفعيل:



$$u = \sum_{i=0}^{n} w_i . x_i$$
$$y = g(u)$$

$$g(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{if } u < 0 \end{cases}$$

نلاحظ أن تابع التفعيل تكون قيمته 1 عندما يكون خرج ال u أكبر أو يساوي الصفر، وتكون قيمته 0 عندما يكون خرج ال u أصغر من الصفر.

شرح عملية التأكد من الأوزان:

✓ x^i تمثل مصفوفة الدخل، ومن أجل كل دفعة دخل تمثل سطراً في جدول الحقيقة.

✓ نُطبق الدفعة الأولى من الدخل (السطر الأول في الجدول) مع الأوزان المقترحة ونقوم بإيجاد تابع المجموع، ومن ثم نطبق تابع التنشيط على القيمة الناتجة ونقوم بمقارنتها مع القيمة المرغوبة، فإذا تساوت القيمة الناتجة عن تابع التنشيط مع القيمة المرغوبة فننتقل إلى الدفعة الثانية من الدخل، وهكذا إلى أن نصل إلى نهاية الدخل، وفي حال عدم تساوي القيمة الناتجة عن تابع التفعيل مع القيمة المرغوبة، فسوف نحتاج إلى عملية تعديل على الأوزان، وسنتحدث عنها لاحقاً.

$x^{(1)}$
−1
0
0

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = -1.5$$
$$u < 0 \Rightarrow g(u) = 0 \Rightarrow y = 0$$

✓ نجد أن في القيمة السابقة للدخل تساوت القيم الناتجة، ومنه نكمل العملية بنفس الشكل.

$x^{(2)}$
−1
0
1

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = -0.5$$
$$u < 0 \Rightarrow g(u) = 0 \Longrightarrow y = 0$$

$x^{(3)}$
−1
1
0

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = -0.5$$
$$u < 0 \Rightarrow g(u) = 0 \Longrightarrow y = 0$$

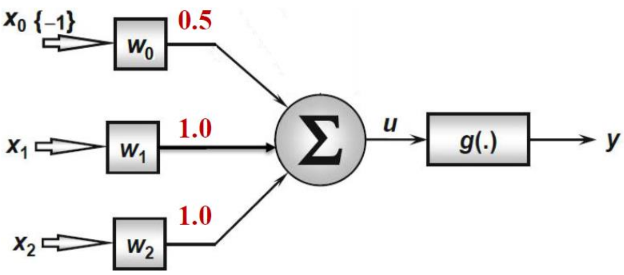
$$x^{(4)} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = +0.5$$
$$u > 0 \Rightarrow g(u) = 1 \Longrightarrow y = 1$$

نلاحظ أن جميع القيم الناتجة السابقة تساوت مع القيم المرغوبة، ومنه نجد أن الأوزان صحيحة.

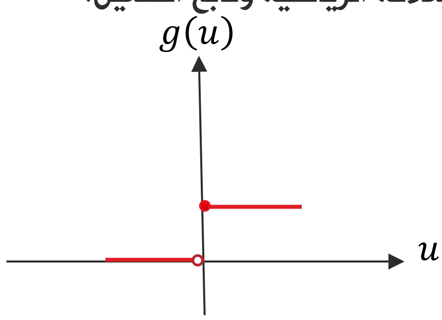
2. OR:

بنفس الطريقة السابقة نقوم بالتحقق من الأوزان الجديدة، مع الانتباه إلى التغيير الحاصل في القيم المرغوبة في جدول الحقيقة.



x_0	x_1	x_2	output
−1	0	0	0
−1	0	1	1
−1	1	0	1
−1	1	1	1

العلاقة الرياضية وتابع التفعيل:



$$u = \sum_{i=0}^{n} w_i . x_i$$
$$y = g(u)$$

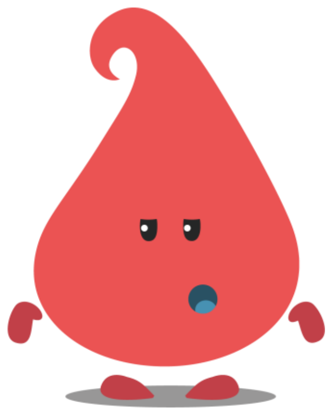
$$g(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{if } u < 0 \end{cases}$$

$$x^{(1)} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = -0.5$$
$$u < 0 \Rightarrow g(u) = 0 \Rightarrow y = 0$$

$$x^{(2)} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = 0.5$$
$$u > 0 \Rightarrow g(u) = 1 \Longrightarrow y = 1$$



$$x^{(3)} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = 0.5$$
$$u > 0 \Rightarrow g(u) = 1 \Longrightarrow y = 1$$

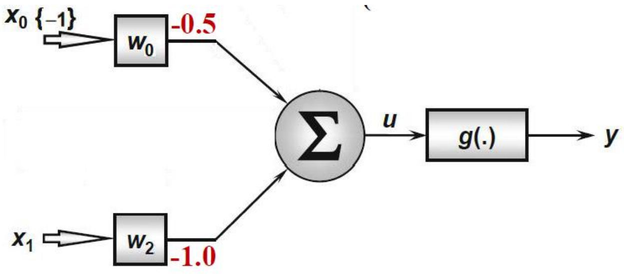
$x^{(4)}$
-1
1
1

$$u = w_0\, x_0 + w_1\, x_1 + w_2\, x_2 = 1.5$$
$$u > 0 \Rightarrow g(u) = 1 \Rightarrow y = 1$$

نلاحظ أيضاً مما سبق أن جميع القيم الناتجة السابقة تساوت مع القيم المرغوبة، ومنه نجد أن الأوزان صحيحة.

3. NOT:

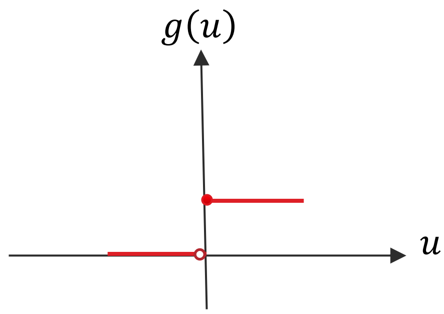
وبنفس الطريقة السابقة للتحقق من الأوزان.



x_0	x_1	output
-1	0	1
-1	1	0

$$u = \sum_{i=0}^{n} w_i . x_i$$
$$y = g(u)$$

$$g(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{if } u < 0 \end{cases}$$



$x^{(1)}$
-1
0

$$u = w_0\, x_0 + w_1\, x_1 = +0.5$$
$$u > 0 \Rightarrow g(u) = 1 \Rightarrow y = 1$$

$x^{(2)}$
-1
1

$$u = w_0\, x_0 + w_1\, x_1 = -0.5$$
$$u < 0 \Rightarrow g(u) = 0 \Rightarrow y = 0$$

ونلاحظ أيضاً صحة الأوزان هنا.

ملاحظة: إن عملية حساب تابع المجموع تتم عن طريق ضرب شعاع الدخل بشعاع الوزن (مع الانتباه بأن الانزياح يمثل العنصر الأول من الدخل x_0).
وتكتب العلاقة خطياً:
Dot product $w.x = w_0x_0 + w_1x_1 + \cdots + w_nx_n = \sum_{i=0}^{n} w_i . x_i$

خوارزمية تعلم ال perceptron

إذا اعطيت n حالة (زوج) دخل وخرج، ويكون هدفنا إيجاد الأوزان المناسبة،

نسمي خوارزمية التدريب بخوارزمية التعلم، وسرعة التدريب ب معدّل التعلم.

الخطوات بشكل عام:

1. أولاً نقوم بتهيئة مصفوفة الأوزان، أي نقوم بفرض قيم للبدء.
2. التكرار والدورات من خلال اللحظات التي نقوم في كل مرة بإدخال x والحصول على قيمة y ونقوم بمقارنتها مع القيمة المرغوبة d، إذا كانت صحيحة نوافق على الأوزان وإذا كانت خاطئة نقوم بعملية تعديل على الأوزان.

$$w = w + y_i\, x$$

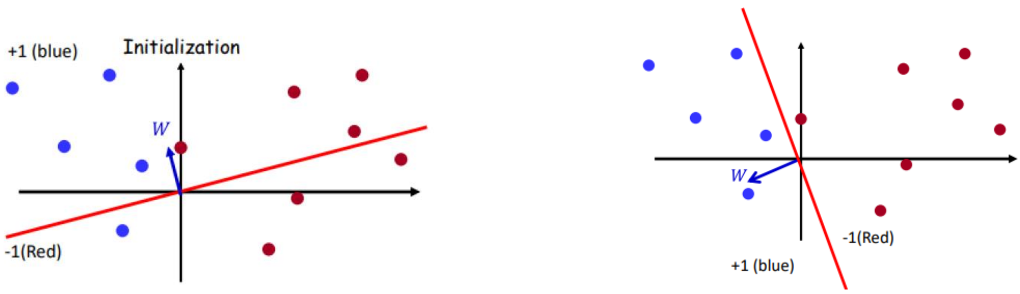
(عملية تعديل الأوزان)

جديدة

قديمة

يختلف مقدار الإضافة من خوارزمية لأخرى

الصورة المجاورة تبين عملية تعديل الأوزان، وذلك من خلال التكرار حتى الوصول لحل مثالي بدون خطأ



عملية تدريب لـ Perceptron:

Hebb's learning rule

$$w \leftarrow w + \eta(d^k - y)x^k$$

شعاع يحتوي على الأوزان

$$w = [\theta\ w_1 w_2 \dots \dots w_n]^T$$

شعاع الدخل

$$x^k = [-1\ x_1^k\ x_2^k \dots \dots x_n^k]$$

d^k **القيمة المرغوبة لشعاع الدخل**

y **الخرج الناتج**

η **معدل التعلم، وتكون قيمته بين ال** $0, 1$ $0 < \eta < 1$

الخوارزمية التفصيلية للتدريب:

حيث أن:

$$\Omega^{(\boldsymbol{x})} = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{cccc} \boldsymbol{x}^{(1)} & \boldsymbol{x}^{(2)} & \boldsymbol{x}^{(3)} & \boldsymbol{x}^{(4)} \\ \left[\begin{array}{c:c:c:c} -1 & -1 & -1 & -1 \\ 0.1 & 0.3 & 0.6 & 0.5 \\ 0.4 & 0.7 & 0.9 & 0.7 \\ 0.7 & 0.2 & 0.8 & 0.1 \end{array}\right] \end{array};$$

تمثل مجموعة تدريب وكل عامود فيها يمثل عينة تدريب

$$\Omega^{(d)} = \begin{array}{cccc} d^{(1)} & d^{(2)} & d^{(3)} & d^{(4)} \\ \end{array}$$
$$\Omega^{(d)} = [\boxed{1} \; \boxed{-1} \; \boxed{-1} \; \boxed{1}]$$

مجموعة الخرج المرغوب وكل قيمة تمثل فيه عمود تدريب.

شرح الخوارزمية: نقوم بالبداية بالحصول على عينات التديب (الدخل)، ومن ثم نقوم بربط قيم الدخل هذه مع القيم المرغوبة ، بعد ذلك نقوم بتهيئة شعاع الأوزان بقيم عشوائية صغيرة، ويتم بعدها تحديد قيمة معدّل التعلم. بعد ذلك ندخل بحلقة تكرار ومن أجل كل قيمة دخل نقارنها مع الخرج المرغوب، وفي حال التساوي ننتقل إلى الدخل التالي، وفي حال عدم التساوي نقوم بعملية تعديل الأوزان، ونعمل على الأوزان الجديدة حتى يختفي الخطأ والفرق بين القيم.

Begin {PERCEPTRON Algorithm – Training Phase}

- <1> Obtain the set of training samples $\{\boldsymbol{x}^{(k)}\}$;
- <2> Associate each desired output $\{d^{(k)}\}$ to each sample;
- <3> Initialize vector $\boldsymbol{w}$ with small random values;
- <4> Specify the learning rate $\{\eta\}$;
- <5> Initialize the epoch counter $\{epoch \leftarrow 0\}$;
- <6> Repeat the following instructions:
  - <6.1> *error* ← "none";
  - <6.2> For all training samples $\{\boldsymbol{x}^{(k)}, d^{(k)}\}$, do:
    - <6.2.1> $u \leftarrow \boldsymbol{w}^T \cdot \boldsymbol{x}^{(k)}$;
    - <6.2.2> $y \leftarrow \text{signal}(u)$;
    - <6.2.3> If $y \neq d^{(k)}$
      - <6.2.3.1> then $\begin{cases} \boldsymbol{w} \leftarrow \boldsymbol{w} + \eta \cdot (d^{(k)} - y) \cdot \boldsymbol{x}^{(k)} \\ error \leftarrow \text{"existent"} \end{cases}$
  - <6.3> *epoch* ← *epoch* + 1;
- Until: *error* ← "none"

End {PERCEPTRON Algorithm – Training Phase}

خوارزمية التصنيف:

وهي خوارزمية بسيطة يمكن تلخيصها كالآتي: نقوم بالحصول على عينات الدخل ومن ثم شعاع الأوزان، ونقوم بعملية التصنيف بناءً على القيم الناتجة من تابع التنشيط، ففي حال كانت القيمة الخارجة $-1 \Longleftarrow \text{class A}$،

وفي حال كانت القيمة الخارجة $1 \Longleftarrow \text{class B}$

وهكذا حتى انتهاء الدخل.

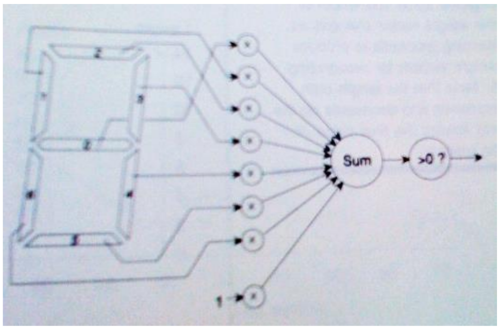
Begin {PERCEPTRON Algorithm – Operation Phase}

- <1> Obtain one sample to be classified $\{\boldsymbol{x}\}$;
- <2> Use the vector $\boldsymbol{w}$ adjusted during training;
- <3> Execute the following instructions:
  - <3.1> $u \leftarrow \boldsymbol{w}^T \cdot \boldsymbol{x}$;
  - <3.2> $y \leftarrow \text{signal}(u)$;
  - <3.3> If $y = -1$
    - <3.3.1> then: sample $\boldsymbol{x} \in$ *{Class A}*
  - <3.4> If $y = 1$
    - <3.4.1> then: sample $\boldsymbol{x} \in$ *{Class B}*

End {PERCEPTRON Algorithm – Operation Phase}

استخدام ال perceptron للتعرف على الأرقام

في الصورة المجاورة نلاحظ أن الدخل عبارة عن ال 7-segment مضافاً لها الانزياح، حيث يقوم نظام رؤية بالإبلاغ عن الأجزاء المضاءة وبناءً عليه يتم تحديد الرقم.



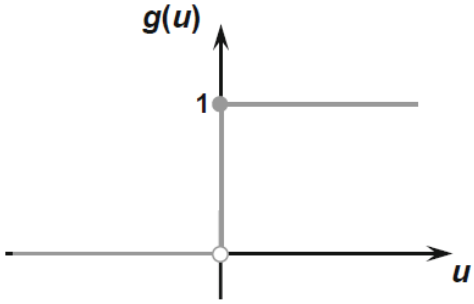
Activation Function توابع التنشيط

حالياً سنقوم بذكر أنواع التوابع، وسنذكر وظيفة كل منها لاحقاً.

1. Step function:

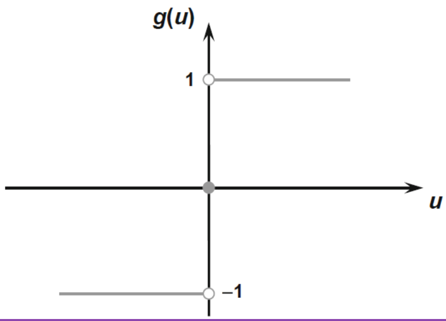
هو أبسط تابع، حيث أن كل القيم التي تكون أصغر من الصفر تصبح صفر، والقيم الأكبر أو تساوي الصفر تصبح واحد.

$$g(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{if } u < 0 \end{cases}$$



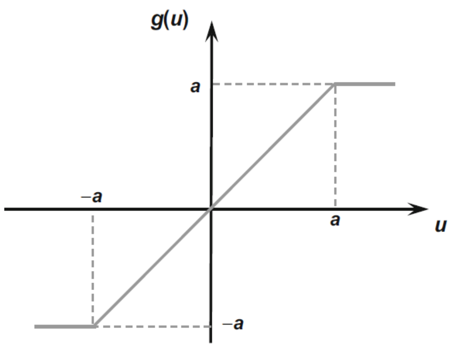
2. Bipolar function

$$g(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ -1, & \text{if } u < 0 \end{cases}$$



3. Symmetric ramp function

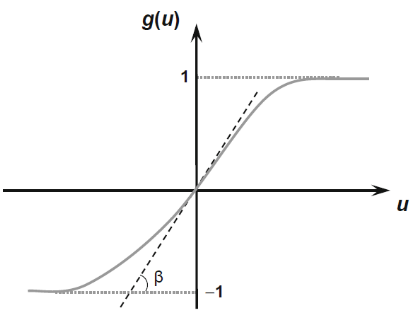
$$g(u) = \begin{cases} a, & \text{if } u > a \\ u, & \text{if } -a \leq u \leq a \\ -a, & \text{if } u < -a \end{cases}$$



التوابع الثلاث الماضية كانت تعطي قيم حديّة، والتوابع الثلاث التالية تعطي قيم مستمرة.

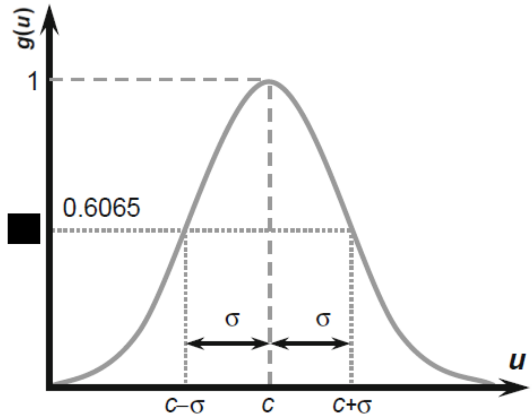
1. Logistic function

$$g(u) = \frac{1 - e^{-\beta \cdot u}}{1 + e^{-\beta \cdot u}}$$



2. Gaussian functionتابع التوزيع الطبيعي :

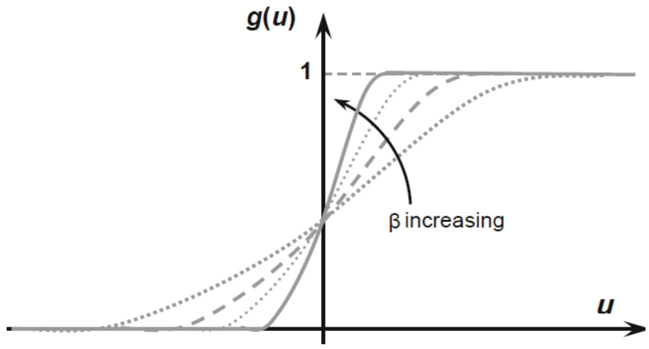
$$g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}}$$



ملاحظة: عند دراسة أي عينة عشوائية بدون تدخل للإنسان أو عوامل خارجية مؤثرة نجد أنها تخضع للتوزيع الطبيعي.

3. Hyperbolic tangent function

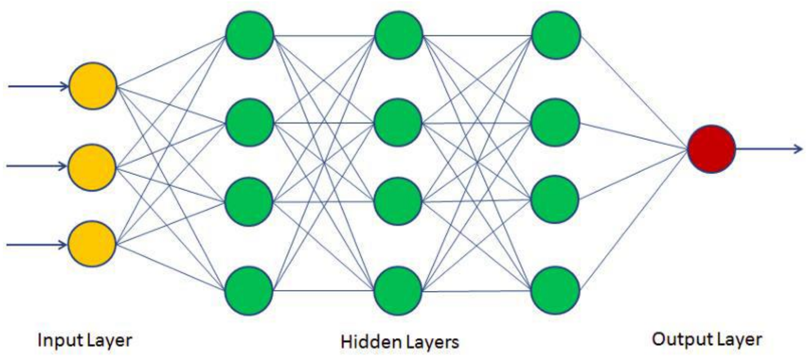
$$g(u) = \frac{1}{1 + e^{-\beta \cdot u}}$$



Combining multiple perceptron

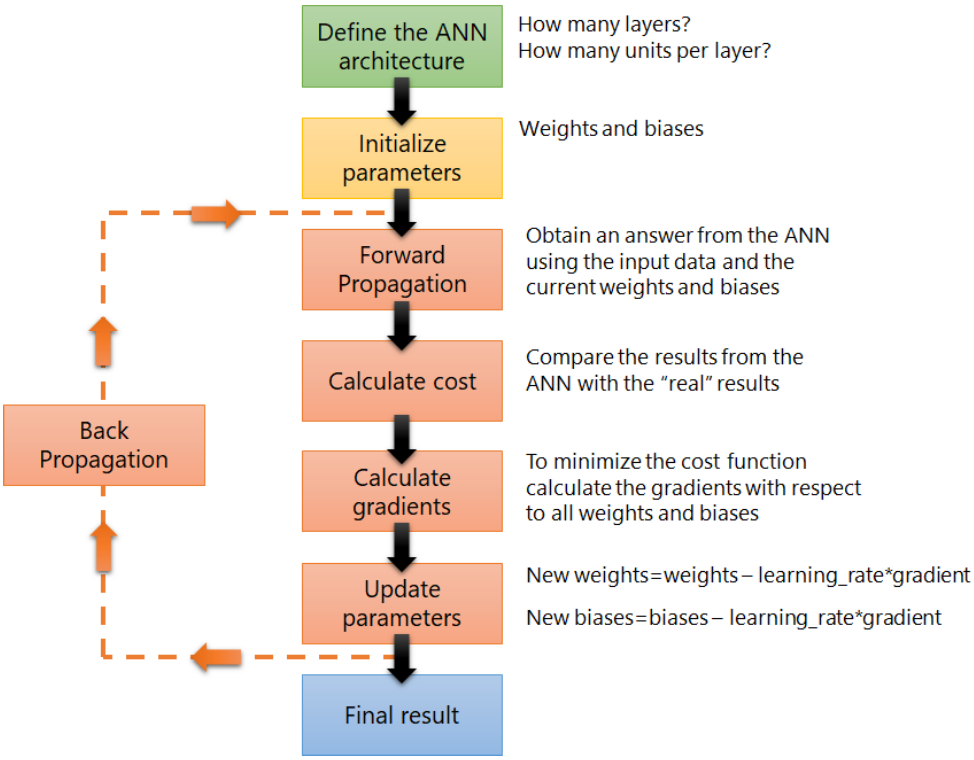
درسنا في هذه المحاضرة حالة ال perceptron بطبقة واحدة، ولكن هذا ليس هدفنا في المادة، حيث أننا سنقوم ببناء شبكات عصبونية بعدّة طبقات، وكل عصبون فيها يستقبل الدخل من جمع باقي العصبونات، وتكون هذه الطبقات مخفية وتدعى بال (hidden layers).

كلما كانت الشبكة العصبونية أعمق (تحوي طبقات أكثر)، كلما كنا قادرين على إنجاز مهام أكثر تعقيداً.



كيف يتم تدريب الشبكة العصبونية فعلياً:

- نبدأ بتحديد بنية الشبكة، مثلاً: تابع التفعيل وكم طبقة عميقة تحوي الشبكة.
- تهيئة القيم: مثل الوزن والانزياح.
- تدفق المعلومات الأمامي: وهي الوصل إلى الخرج بدءاً من الدخل.
- حساب الكلفة: الكلفة هنا لا تقتصر على النقود، وإنما دورات انشغال الحاسب تعتبر تكلفة، والذاكرة تكلفة.
- تصغير تابع الكلفة، وسنتحدث عنه لاحقاً.
- تحديث القيم: أوزان وانزياح.
- التدفق الخلفي: الأوزان الجديدة تصبح هي المعتمدة، ونقوم بعملية التراجع من أجل العودة والعمل على الأوزان الجديدة.
- في حال عدم وجود أخطاء فإننا ننتهي من عملية التدريب، ونحصل على النتائج النهائية.



The end