

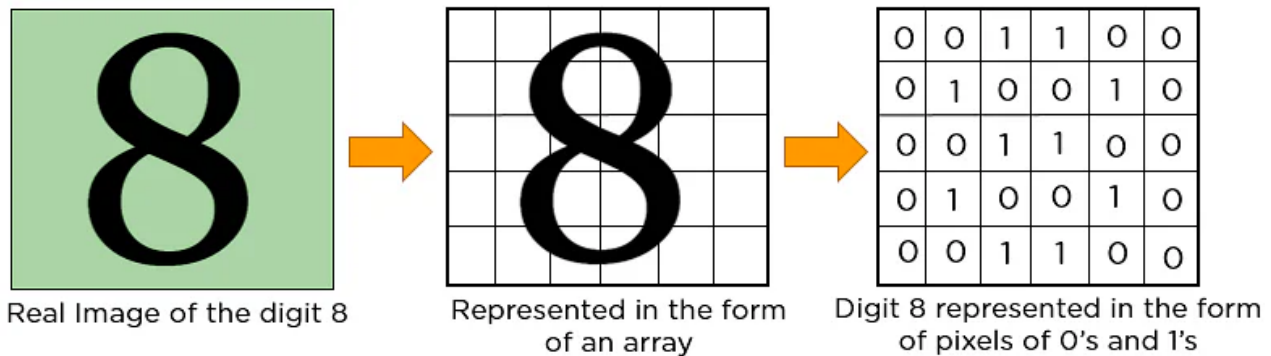
Convolutional Neural Networks

ماهي Convolutional Neural Networks (CNN) ؟

هي خوارزميات ذكاء اصطناعي تعتمد على الشبكات العصبية (artificial neural networks (ANNs متعددة الطبقات و تستخرج الميزات features من الصور و الفيديو , قدرة على أداء العديد من المهام مثل تصنيف (classification) و اكتشاف (detection) الكائنات و تجزئتها (segmentation).

تتميز CNN على خوارزمات التصنيف الأخرى مثل (SVM, KNN) وغيرها , انها تستخرج و تتعلم أفضل الميزات لتمثيل الكائنات في الصور.

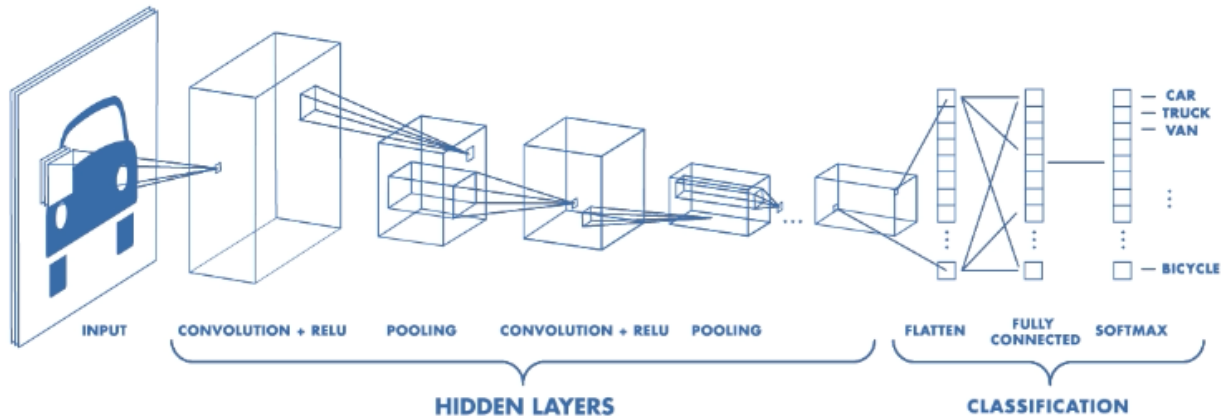
في CNN ، يتم تمثيل كل صورة في شكل مجموعة من قيم البكسل.



بنية Convolutional Neural Networks (CNN):

تحتوي شبكة CNN عادةً من أربع أنواع من الطبقات (Layers) طبقات مخفية متعددة تساعد في استخراج المعلومات من الصورة:

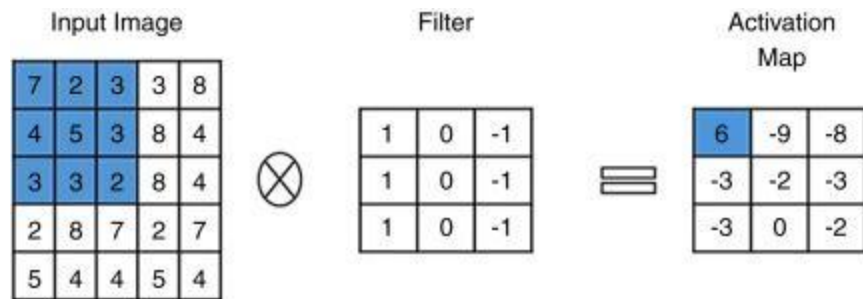
- Convolutional
- Pooling
- Relu
- Fully Connected



طبقة الالتفاف (Convolution Layer):

- هي أساس بناء شبكة CNN, هي تتكون من مجموعة من المرشحات (filters) تسمى kernels وعادة ماتكون هذه المرشحات أصغر من الصورة الفعلية.
- كل kernel لها عرض و ارتفاع , العرض \times الارتفاع ينتج لنا الوزن (weight).
- تستخدم الأوزان (weights) لإستخراج الميزات من بيانات الإدخال.
- في خطوة التدريب ، تبدأ الأوزان (weights) في kernel بقيم عشوائية .

فإن هذه الطبقة هي حاصل ضرب نقطي (dot product) بين مصفوفتين هما مصفوفة kernel و مصفوفة البكسل للصورة , يلتف كل مرشح مع الصورة و يقوم بإنشاء خريطة الـ activation, و يتم تكرار هذه العملية لكل عنصر من عناصر الصورة المدخلة.



في المثال التوضيحي نرا أن لدينا صورة بمقاس 5x5 و التي تكون قيم البكسل فيها إما 0 أو 1, و هناك أيضاً مصفوفة المرشح بأبعاد 3x3.

نقوم بتمرير مصفوفة الـ filter فوق الصورة المدخلة ويتم هنا حساب dot product للحصول على مصفوفة الميزات وهي activation map.

الـ activation map هي نتيجة تطبيق المرشحات على الصورة المدخلة و تسمى أيضاً feature map تسمح بفهم الميزات (feature) التي تم اكتشافها.

$$\text{Activation map} = \text{Input} * \text{Filter}$$

$$= \sum_{y=0}^{\text{columns}} \left(\sum_{x=0}^{\text{rows}} \right) \text{Input} (x - p, y - q) \text{Filter} (x, y)$$

Activation map algorithm

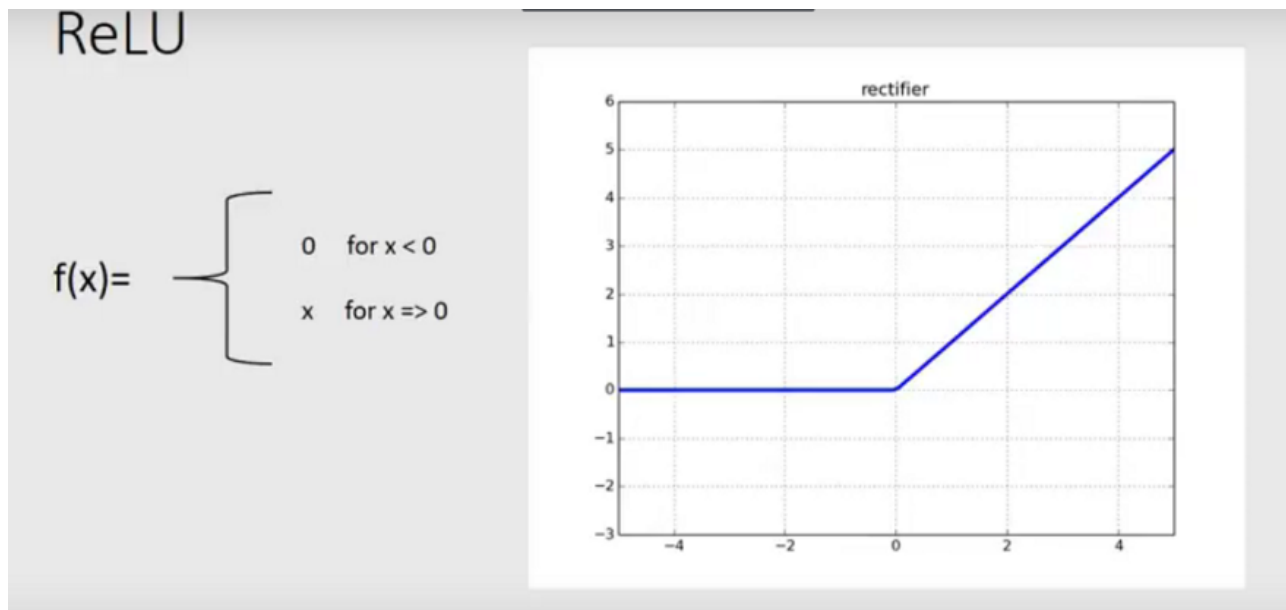
//يمثل كل مرشح في Convolution Layer معلماً (feature)

طبقة (ReLU Layer):

تعني ReLU الوحدة الخطية المصححة (rectified linear unit). بمجرد استخراج الـ activation maps تكون الخطوة التالية هي نقلها إلى طبقة ReLU , أي أنها وظيفة التنشيط (activation function).

وهي وظيفة (function) تسمح بتعلم العلاقات المعقدة في البيانات, تساعد ReLU الشبكة على تعلم وظائف اتخاذ القرار الأصعب وتقليل overfitting , و هي عملية حسابية بسيطة :

$$y = \max(x, 0)$$



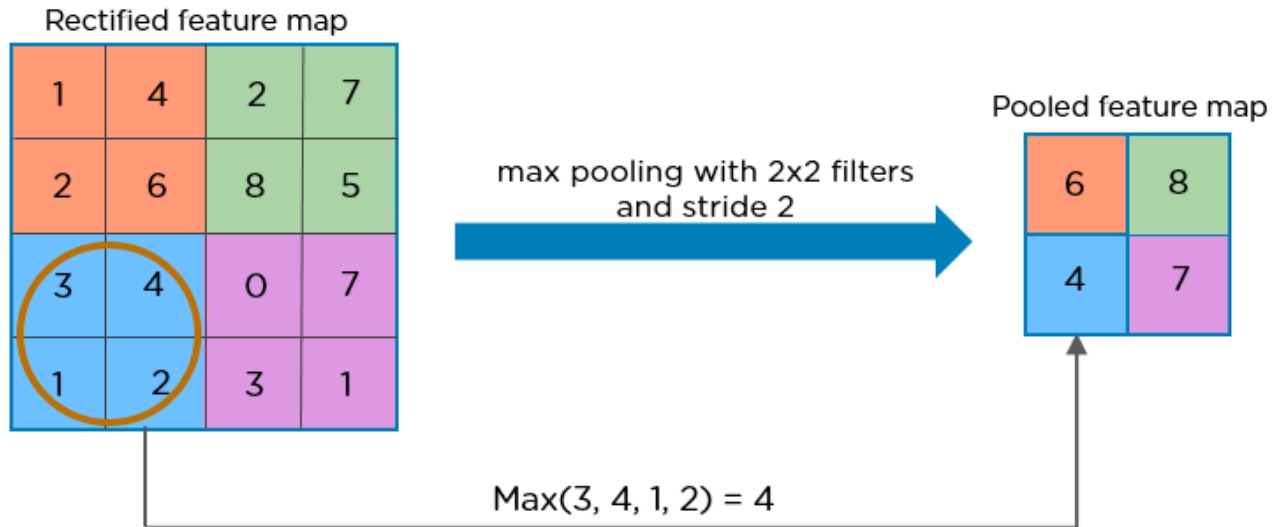
هي مرتبطة بـ convolution function فإذا كان ناتجها 0 أو قيمة سالبة فهذا يعني أن الميزة (feature) غير موجودة فـا نشير إليها بـ 0, إما غير ذاك فـا نحتفظ بالقيمة نفسها.

إذا الوظيفة خطية للقيم الأكبر من الصفر ، مما يعني أن لديها الكثير من الخصائص المرغوبة لوظيفة التنشيط الخطي عند تدريب شبكة عصبية.

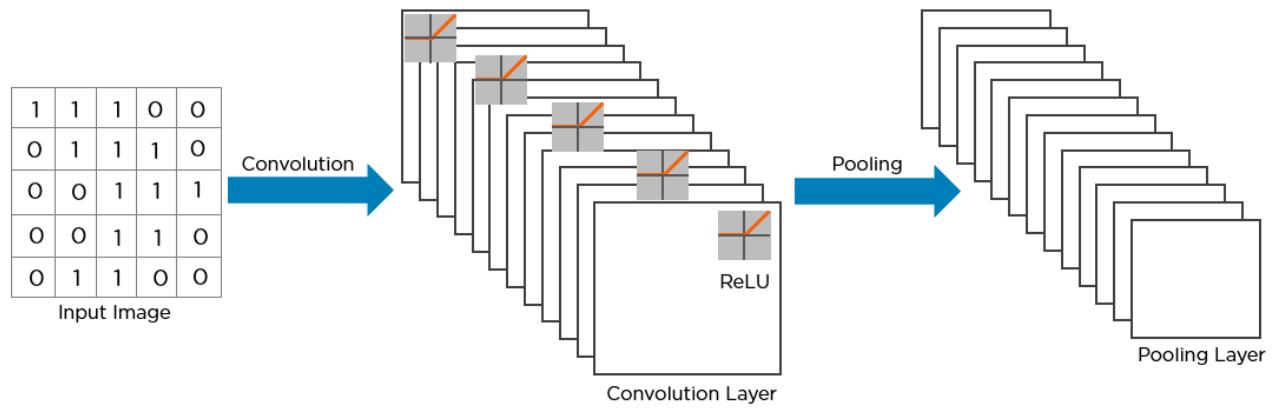
طبقة التجميع (Pooling Layer):

تسمى أيضاً طبقة العينات السفلية (down-sampling layer), تستخدم لتقليل أبعاد activation maps. في طبقة التجميع ، يطبق المرشح فوق بيانات الإدخال ثم بعدها ويطبق عملية التجميع (max, min, avg) لإنشاء pooled feature map, وعادة ما تكون القيمة القصوى (max) أو متوسط القيمة لجميع القيم (avg).

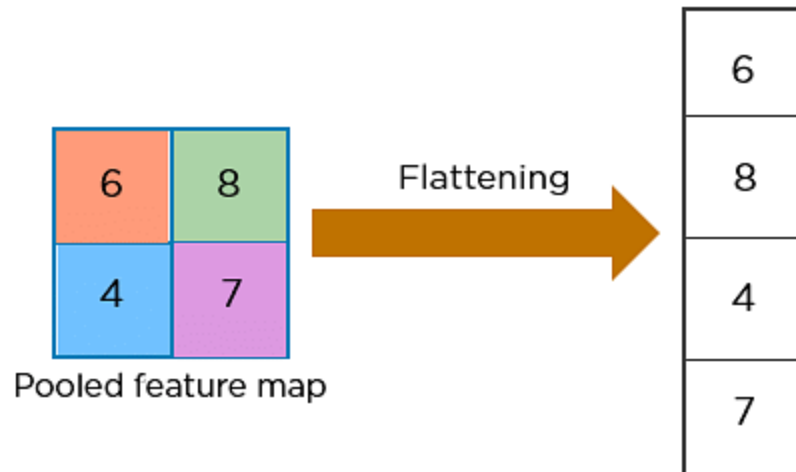
تستخدم طبقة التجميع مرشحات مختلفة لتحديد أجزاء مختلفة من الصورة, أي بمعنى أننا نأخذ في الواقع القيم التي تدل على وجود ميزة (feature) في هذا القسم من الصورة, بهذه الطريقة نتخلص من المعلومات غير المرغوب فيها. فيما يتعلق بوجود ميزة في جزء معين من الصورة ونأخذ في الاعتبار فقط ما هو مطلوب معرفته.



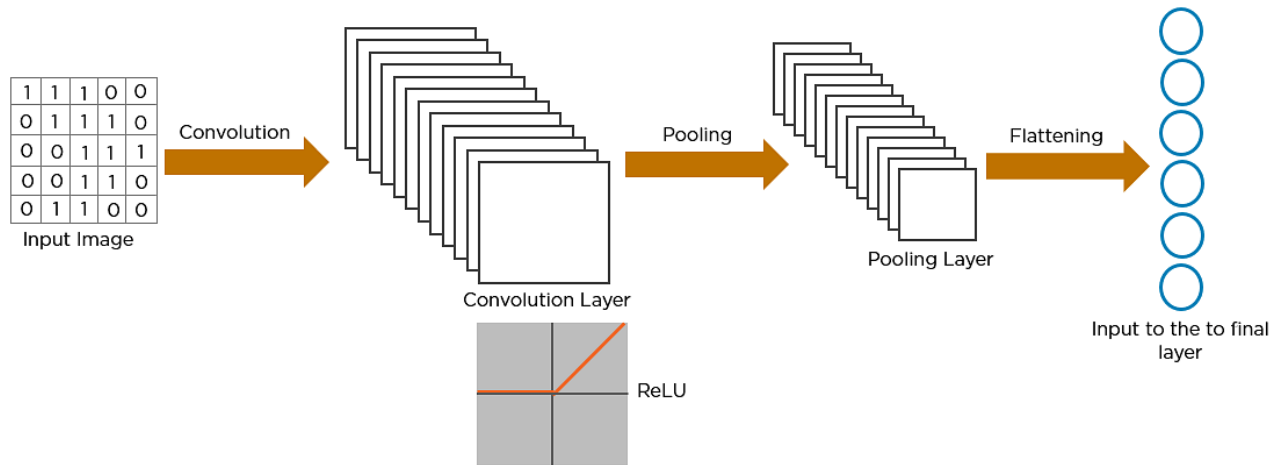
نرا كيف تبدو بنية convolution neural network حتى الآن:



قبل الانتقال الى fully connected layer نمر بمرحلة التسطیح **Flattening**:
يتم استخدامها لتحويل كل مصفوفة ثنائية الأبعاد الناتجة من pooled feature maps الى متجه خطي.



ويتم تعيين المصفوفة المسحطة كمدخل للطبقة المتصلة fully connected layer لتصنيف الصورة .

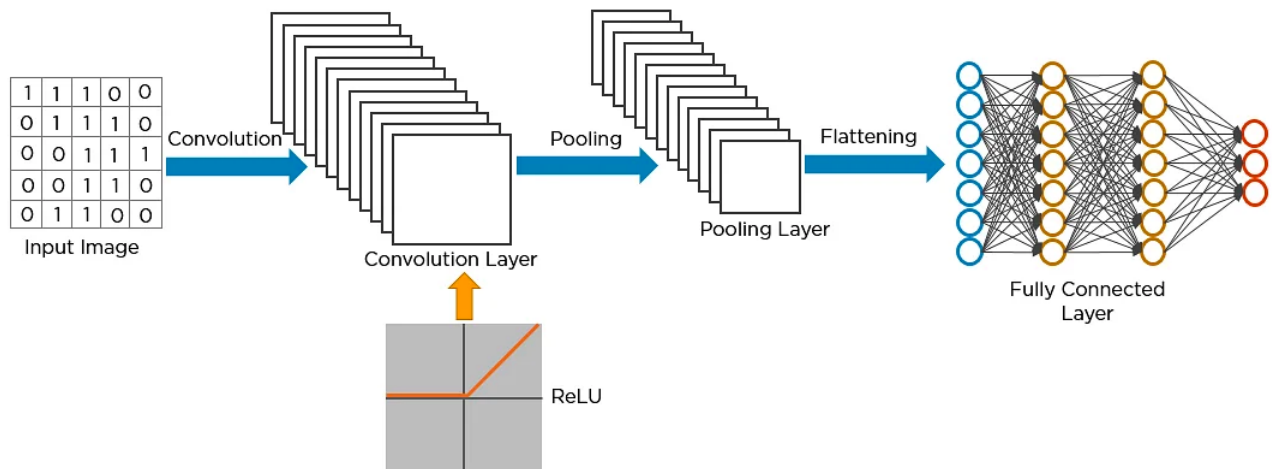


الطبقة المتصلة بالكامل (Fully Connected):

تنقسم شبكة CNN إلى قسمين:

- convolutional
- dense steps

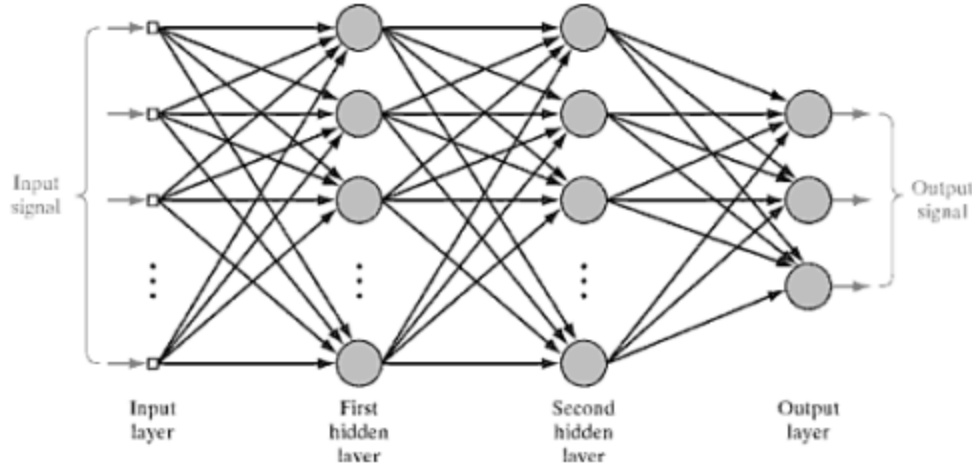
يتعلم الأول أفضل الميزات (features) لاستخراجها من الصور , و يتعلم الثاني كيفية تصنيف الميزات (features) في فئات مختلفة.



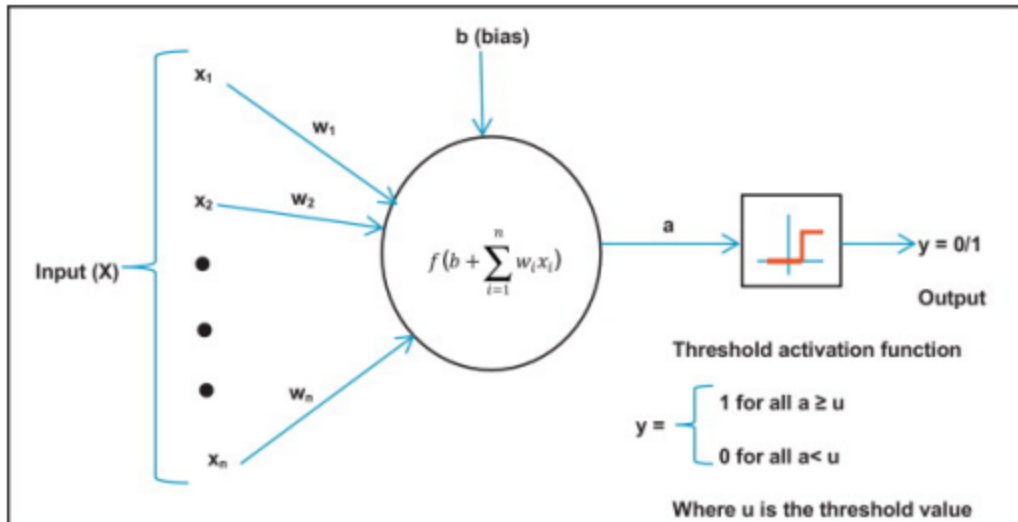
الطبقة المتصلة بالكامل هي مدرك متعدد الطبقات (MultiLayer Perceptron MLP) ، يتكون من ثلاثة أنواع من الطبقات:

- الإدخال (input layer):
 - تتلقى طبقة الإدخال الميزات (features) التي تم إنشاؤها بواسطة CNN.
- المخفية (hidden layer):

- الطبقة المخفية عبارة عن سلسلة من الخلايا العصبية ذات الأوزان التي سوف تتعلمها في خطوة التدريب.
- ويمكن أن يتكون MLP من طبقة مخفية واحدة أو أكثر.
- المخرجة (output layer):
- أيضاً طبقة الإخراج هي عبارة عن سلسلة من الخلايا العصبية, لكن تحتوي على وظيفة تنشيطية (activation function) مختلفة.
- عادة ما تستخدم وظيفة softmax لتوليد احتمالات كل فئة.



كيفية عمل MLP:



تتكون كل خلية عصبية من :

- المدخلات هي x_0, x_1, \dots, x_n , والتي تمثل لدينا الميزات (features).
- الاوزان هي w_0, w_1, \dots, w_n , و التي سيتم تعلمها في التدريب.

- تتم عملية التحيز (The bias).
- وظيفة التنشيط.
- الإخراج.

يقوم Perceptron بتنفيذ العملية التالية:

- يتم ضرب كل مدخل في وزن للإشارة إلى الأهمية النسبية للمدخلات.
- ببساطة ضرب مصفوفة للوصول إلى مجموع الأوزان للمدخلات , يتم تمثيلها:

$$y = f(x) = \sum x_i w_i$$

- يضاف التحيز (Bias) إلى مجموع المدخلات المرجحة وهي مجموع ضرب المدخلات في الأوزان.

$$output = sum(weights * inputs) + bias$$

- تقوم وظيفة التنشيط داخل الخلايا العصبية بإجراء عملية حسابية على الإجمالي.
- يمكن أن يكون التحيز موجباً أو سالباً ، مما يؤدي إلى زيادة أو تقليل ناتج الخلايا العصبية.
- والنتيجة هي ناتج خلية عصبية ، والذي يتم تمريره إلى الخلايا العصبية الأخرى أو تسليمه كمخرج .
- يمر الإخراج من خلال وظيفة الخسارة (loss function) أو وظيفة التكلفة (cost function) التي تقيم دقة تنبؤ الشبكة العصبية ، ويتم إعادة النتائج مرة أخرى عبر الشبكة ، مما يشير إلى التعديلات التي يجب إجراؤها على الأوزان والتحيزات (biases).
- يتم تكرار العملية لتحسين دقة الإخراج (output) ؛ وفي كل مرة يتم تعديل الأوزان والتحيزات مع كل تكرار.

دالة cost و دالة loss:

تشير cost function أو loss function إلى دقة النموذج.

الوظيفة التنشيطية (Activation Function):

تحتوي كل عقدة في الشبكة العصبية على وظيفة تنشيط تؤدي عملية حسابية على مجموع أوزان مدخلاتها (sum of its weighted inputs) وتحيزها (bias) وتنتج مخرجات. (الوظيفة هي علاقة خاصة يكون فيها لكل إدخال (input) ناتج

واحد(single output), تنتج الوظائف المختلفة مخرجات مختلفة.

الوظائف التنشيط الأكثر شيوعاً :

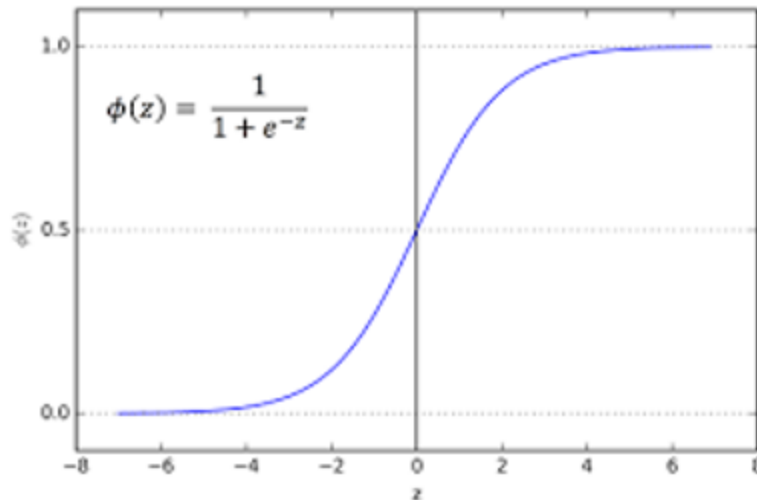
- ReLu
- Sigmoid
- Softmax
- Tanh
- Hardlim

الوظيفة التنشيطية السينية Sigmoid:

- الوظيفة السينية هي شكل خاص من وظائف logistic وعادة ما يتم الإشارة إليها بواسطة $\sigma(x)$ أو $\text{sig}(x)$.

$$\sigma(x) = 1/(1+\exp(-x))$$

- ودائماً ما يكون مدى الـ output هو (0,1).
- تعتبر دالة غير خطية , فأكيد سوف يكون output غير خطي ايضاً.
- تسمى هذه الخلايا العصبية التي تستخدم وظيفة السينية كوظيفة تنشيط كوحدة سيني (sigmoid unit).
- لا يفضل استخدام وظيفة التنشيط السيني في مشاكل التصنيف متعدد الفئات.
- تشكل حرف S عند الرسم :

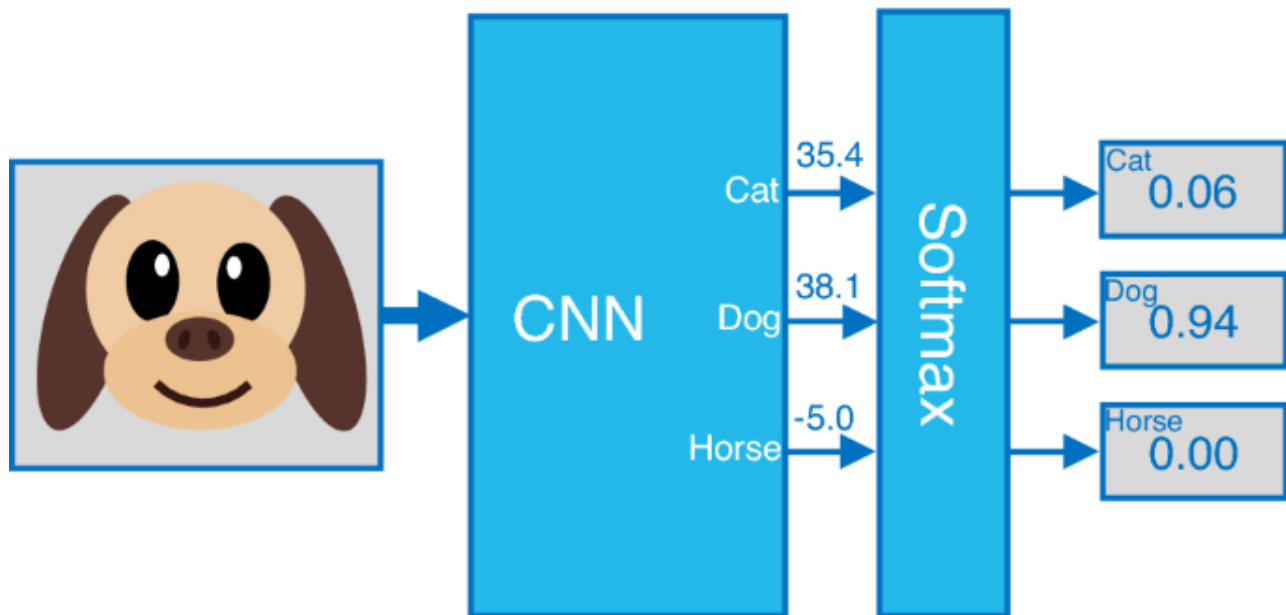


الوظيفة التنشيطية softmax:

- تحسب دالة تنشيط Softmax الاحتمالات النسبية.

$$s(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

- يمثل X هنا عدد الاحتمالات لكل نود.
- تعتبر دالة غير خطية ايضاً.
- ينصح استخدامها في مشاكل التصنيف متعدد الفئات.



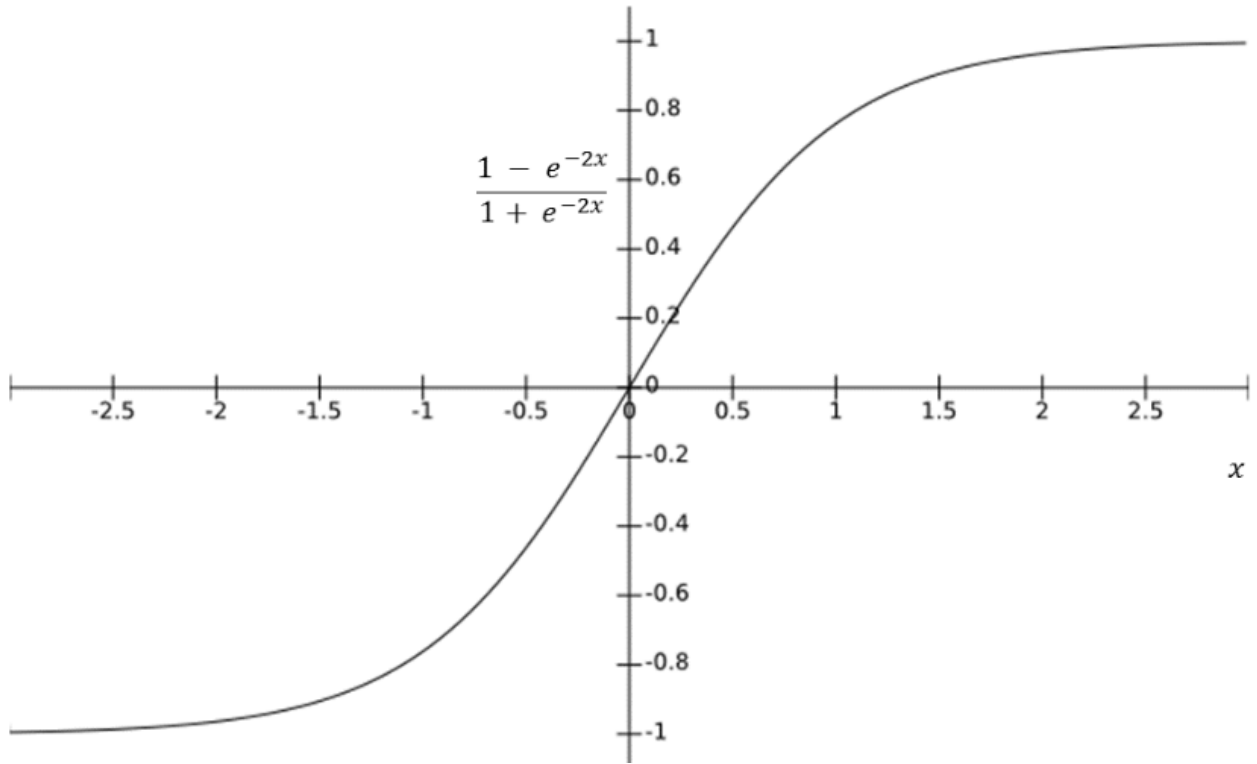
الوظيفة التنشيطية Tanh or hyperbolic tangent:

- هو مثل sigmoid ولكن أفضل.
- يعتبر دالة غير خطية ايضاً.

$$y = f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- نطاق الدالة tanh من (-1 إلى 1).

- يستخدم بشكل أساسي التصنيف بين فئتين.
- تشكل حرف S عند الرسم :

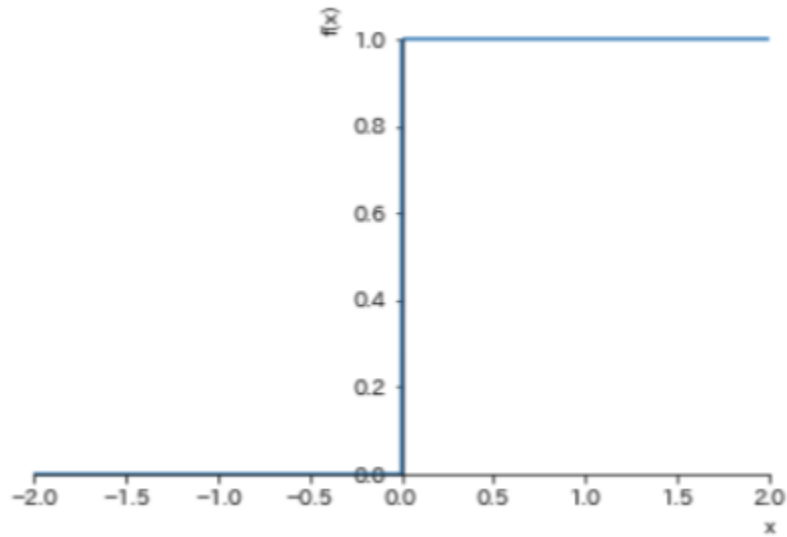


الوظيفة التنشيطية Heaviside:

- هي دالة غير خطية.
- تستخدم لمشاكل التصنيف ثنائي الفئات.

$$f(x) = \begin{cases} 0, & \text{for } x < 0 \\ 1, & \text{for } x \geq 0 \end{cases}$$

- تنتج هذه الوظيفة 1 إذا كان مدخلها المرجح (weighted input) + bias موجباً أو صفراً ، وتخرج 0 إذا كان مدخلها المرجح (weighted input) + bias سالباً.
- يكون output واحدًا أو صفراً ، ولا شيء بينهما.
- عندما نقوم برسم هذه الوظيفة ، تحصل على شيء يشبه درجة السلم:



: Different types of CNN Architectures - CNN البنى الشائعة في

تنقسم إلى نوعين

• الشبكات الكلاسيكية (Classic network):

- LeNet-5
- AlexNet
- VGG 16

• الشبكات الحديثة (Modern network):

- Inception
- ResNet
- ResNeXt
- DenseNet