**Páginas consultadas**

* <https://www.youtube.com/watch?v=JDU3AzH3WKg&list=PLqnslRFeH2Upcrywf-u2etjdxxkL8nl7E&index=3>
* Sigmon function

<https://deepai.org/machine-learning-glossary-and-terms/sigmoid-function>

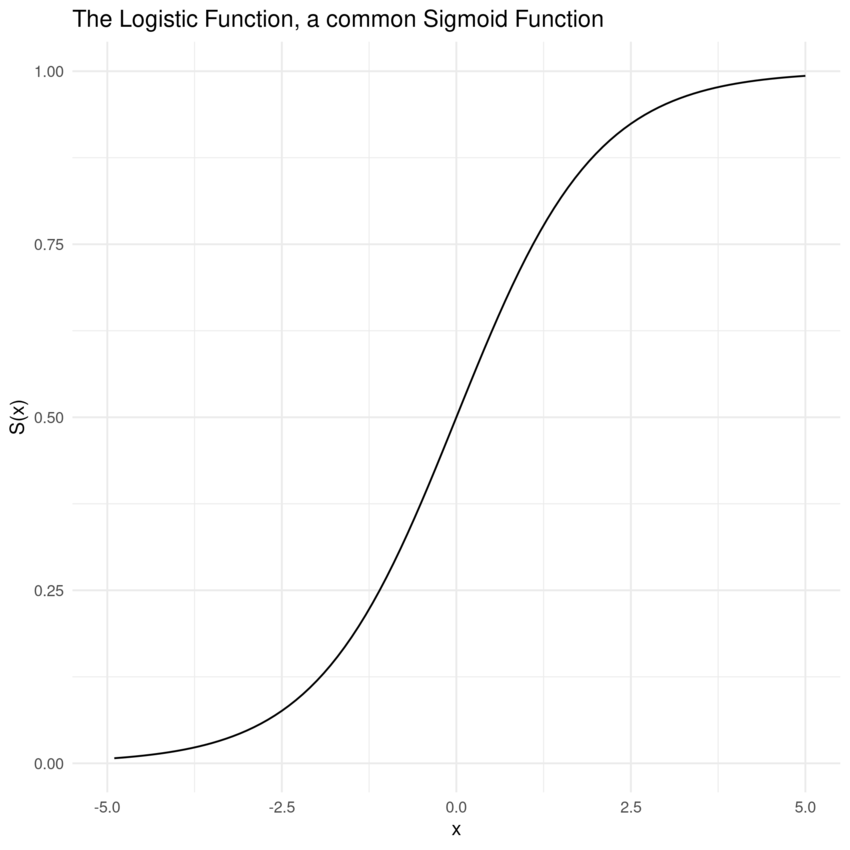
* <https://towardsdatascience.com/logistic-regression-explained-9ee73cede081>

**Apuntes importantes**

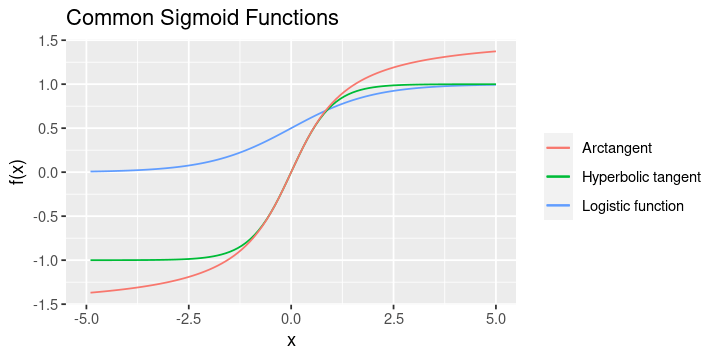
* **Sigmoid function**

A Sigmoid function is a mathematical function which has a characteristic S-shaped curve.

*sigmoid function* is normally used to refer specifically to the logistic function, also called the logistic sigmoid function.

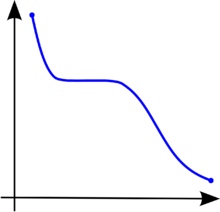
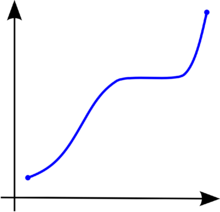


* Sigmoid functions are also useful for many machine learning applications where a real number needs to be converted to a probability.
* All sigmoid functions are monotonic and have a bell-shaped first derivative. There are several sigmoid functions and some of the best-known are presented below.



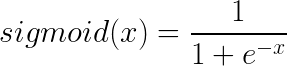
* Monotonic

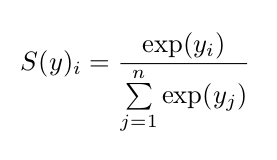
In [calculus](https://en.wikipedia.org/wiki/Calculus), a function f defined on a [subset](https://en.wikipedia.org/wiki/Subset) of the [real numbers](https://en.wikipedia.org/wiki/Real_numbers) with real values is called **monotonic** if and only if it is either entirely non-increasing, or entirely non-decreasing

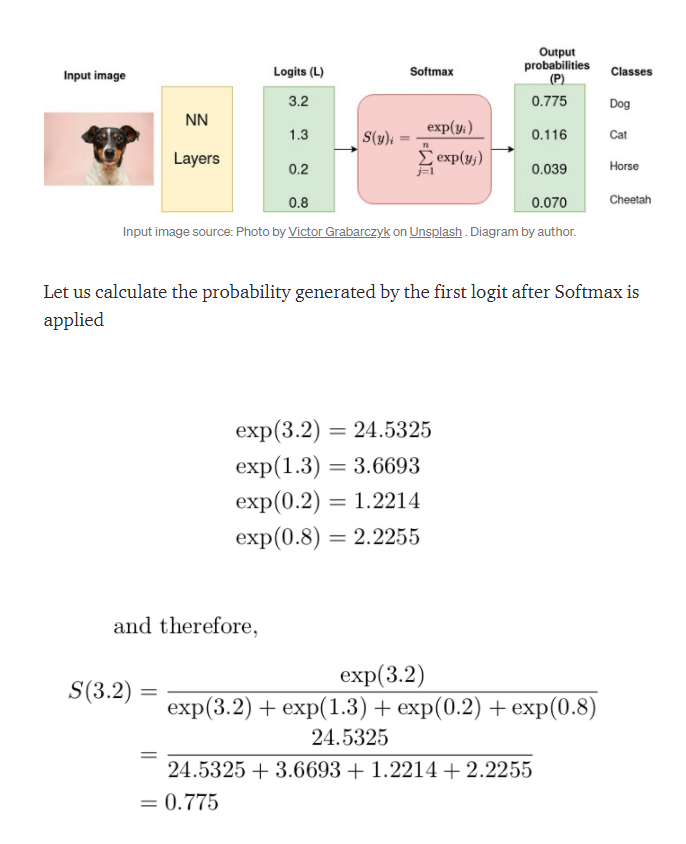


* **Parametric classification model**

In machine learning, a parametric model is any model that captures all the information about its predictions within a finite set of parameters. Sometimes the model must be trained to select its parameters, as in the case of neural networks. Sometimes the parameters are selected by hand or through a simple calculation process. Given the multiple definitions of the word “model,” a parametric model can output either a probability or a value (in some cases a classification)

* **The formula for the *sigmoid* function is the following**:
* **The cross entropy**
* Cross-Entropy loss is a most important cost function. It is used to optimize classification models.
* The lower the loss the better the model.
* **Softmax Activation function**
  + It is often used as the last activation function of a neural network to normalize the output of a network to a probability distribution over predicted output classes.
  + Softmax is an activation function that scales numbers/logits into probabilities. The output of a Softmax is a vector (say v) with probabilities of each possible outcome. The probabilities in vector v sums to one for all possible outcomes or classes.
  + The denominator of Softmax function is a normalization term. It ensures that the output of the function is a value between 0 and 1.





* Softmax converts logits into probabilities. The purpose of the Cross-Entropy is to take the output probabilities (P) and measure the distance from the truth values (as shown in Figure below).
* Integer Encoding (Also called Label Encoding)

In this kind of encoding, labels are assigned unique integer values. For example in our case, we will have,

0 for “dog”, 1 for “cat”, 2 for “horse” and 3 for “cheetah”.

When to use integer encoding: It is used when the labels are ordinal in nature: 0 for “poor”, 1 for “neutral” and 2 for “good”.

The model may take a natural ordering of the labels (2>1>0).

* **One-hot encoding**

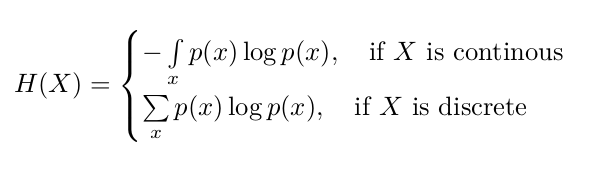
For categorical variables where no such ordinal relationship exists, the integer encoding is not enough. One-hot encoding is preferred:

[1,0,0,0] for “dog”, [0,1,0,0] for “cat”, [0,0,1,0] for “horse” and [0,0,0,1] for “cheetah”.

* Entropy

Entropy of a random variable X is the level of uncertainty inherent in the variables possible outcome.

For p(x) — probability distribution and a random variable X, entropy is defined as follows:



Reason for negative sign: log(p(x))<0 for all p(x) in (0,1) .

**A perfect model has a cross-entropy loss of 0.**

* **Cost function in logistic regression**
* A **cost function** is a mechanism utilized in [supervised machine learning](https://radiopaedia.org/articles/missing?article%5Btitle%5D=supervised+machine+learning&lang=us), the cost function returns the error between predicted outcomes compared with the actual outcomes.
* a cost function is a measure of how wrong the model is in terms of its ability to estimate the relationship between X and y