Step function:

$$step(z) = \begin{cases} 0 & \text{if } z < 0 \\ 1 & \text{otherwise} \end{cases}$$

Rectified linear unit (ReLU):

Rectified linear unit (ReLU):
$$0 \quad \text{if } z < 0$$

$ReLU(z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{otherwise} \end{cases} = max(0, z)$

Sigmoid function: Also known as a *logistic* function, can be interpreted as probability, because for any value of
$$z$$
 the output is in $(0,1)$

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

$$tanh(z) =$$

$$o(z) = \frac{1}{1 + e^{-z}}$$
Hyperbolic tangent: Always in the range $(-1, 1)$

 $tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$ **Softmax function:** Takes a whole vector $\mathsf{Z} \in \mathbb{R}^{\mathsf{n}}$ and generates as output a vector $\mathsf{A} \in \mathsf{R}$ $(0,1)^n$ with the property that $\sum_{i=1}^n A_i = 1$, which means we can interpret it as a

$$tann(z) = \frac{1}{e^z + e^{-z}}$$
nax function: Takes a whole vector $Z \in \mathbb{R}^n$ and generates $[0,1)^n$ with the property that $\sum_{i=1}^n A_i = 1$, which means $[0,1]^n$ which identification expansions the second $[0,1]^n$.

Example 1.1 Takes a whole vector
$$Z \in \mathbb{R}^n$$
 and generat $(0,1)^n$ with the property that $\sum_{i=1}^n A_i = 1$, which mean probability distribution over n items:

softmax(z) =
$$\begin{bmatrix} \exp(z_1)/\sum_i \exp(z_i) \\ \vdots \\ \exp(z_n)/\sum_i \exp(z_i) \end{bmatrix}$$