Activation Function: The activation function decides whether a neuron should be activated or not by calculating the weighted sum and further adding bias to it. The purpose of the activation function is to introduce non-linearity into the output of a neuron.

1) **Step Function:** Step function turns on from 0 to 1 at the point where the input crosses zero. The step function is commonly used in digital signal processing and binary classification tasks in machine learning.

   Positive side:
   a) It is easy to calculate and uses less computing power.

   Negative side:
   a) It has a very limited range.

2) **Sigmoid Function:** It is a nonlinear function that maps any input value to a value between 0 and 1. The sigmoid function is defined by the following equation:

   sigmoid(x) = 1 / (1 + e^-x)

   Positive side:
   a) It is bounded and provides a clear output range of 0 to 1.
   b) It is easy to understand and compute.

   Negative side:
   a) Saturated neurons "kill" the gradients.
   b) Sigmoid outputs are not zero centered.

3) **Tanh Function:** It is a mathematical function that maps any real number to a value between -1 and 1. The tanh function is defined by the following equation:

   tanh(x) = (e^x - e^(-x)) / (e^x + e^(-x))

   Positive side:
   a) It is bounded and provides a clear output range of -1 to 1.
   b) It can provide a stronger gradient than the sigmoid function.

   Negative side:
   a) It still kills gradients when saturated.
   b) It is computationally expensive.

4) **Relu Function:** It is a mathematical function used in machine learning and artificial neural networks. The relu function is defined by the following equation:

$$\mathbf{f(x) = max\ (0,\ x)}$$

Positive side:

a) Does not saturate (in +region).
b) Very computationally efficient.
c) Converges much faster than.

Negative side:

a) Not zero-centered output.
b) An annoyance.

5) **Elu Function:** It is similar to the Relu function, but with a more smooth and more continuous output. It has been shown to improve the performance of neural networks, especially for deeper architectures. The relu function is defined by the following equation:

$$f(x) = \begin{cases} 0 & if\ x > 0 \\ \alpha(\exp(x) - 1) & if\ x \leq 0 \end{cases}$$

Positive side:

a) Closer to zero mean outputs.
b) Negative saturation regime compared with Leaky ReLU adds some robustness to noise.

Negative side:

a) It is computationally more expensive.
b) Sometimes negative outputs may not be desired.

6) **Prelu Function:** It is an activation function that generalizes the traditional rectified unit with a slope for negative values.

Positive side:

a) It is flexible.
b) It shows improved performance.

Negative side:

a) It has computational complexity.
b) It is also overfitting.