Which Activation Function Should I Use?

Why there is a need of Non-Linear transformation??

① As linear function are not strong enough to model many kinds of data. But by using a 'Non-Linear Activation' the mapping of the input with output are non-Linear.

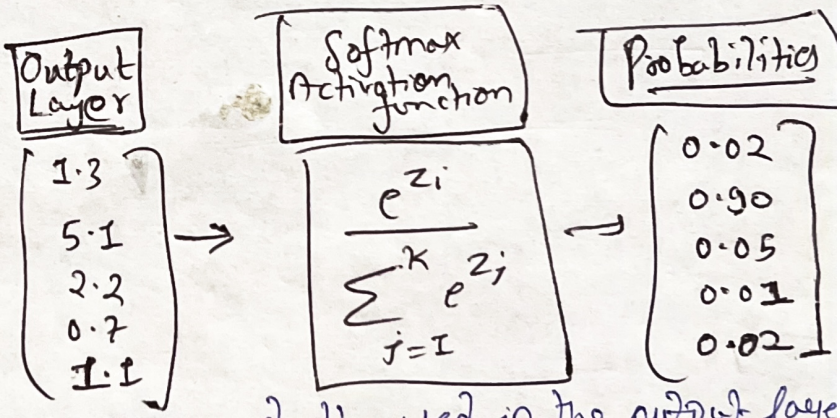
② Also, activation functions should be differentiable. So, that we can perform back-propagation optimization strategy.

: Activation Function

④ **Softmax Function**: It is also a type of Sigmoid function but is handy when we are trying to classify/handle

Multiple Classes

Range: 0-1



Output: The softmax function is ideally used in the output layer of the classifier.

---

⑤ **Tanh Function**: It almost always work better than sigmoid function. also known as Tangent Hyperbolic Function.

$\tanh(x) = 2 * \text{sigmoid}(2x) - 1$ $\quad \text{Sigmod} = \frac{1}{1+e^{-x}}$

Range → -1 to 1 ~~Hence~~ Usually used in hidden layers & value comes out to be -1 to 1 hence the mean of for the hidden layer comes out to be 0 or very close. Hence, centering the data by bringing mean close to 0