

Is ReLU a linear, piece-wise linear, or non-linear activation function?

<http://cs231n.github.io/neural-networks-1/#actfun>

This question previously had details. They are now in a comment.

Answer

Request ▾

Follow 4 Comment 1 Downvote

Promoted by Interana

Did you know analytics tools are still a challenge for many?

Learn how 170 product and data professionals like yourself are using analytics to gain insights.

[read more at interana.com](#)

2 Answers



Chomba Bupe, develops machine learning algorithms

Answered Feb 28

A linear rectified unit (ReLU) is piece-wise linear. Here is a mathematical proof using definition of a linear system^[1].

Given two inputs x_1 and x_2 and an operation $h()$ such that.

$$y_1 = h(x_1)$$

$$y_2 = h(x_2)$$

Then $h()$ is linear if and only if the property of superposition is fully satisfied.

$$\alpha * y_1 + \beta * y_2 = h(\alpha * x_1 + \beta * x_2)$$

So let $h = \text{ReLU}$.

- **Condition 1:** α, β, x_1 and x_2 being all positive. Then the above expression is satisfied. Thus the ReLU is linear given those conditions.
- **Condition 2:** α, β, x_1 and x_2 being all negative, the ReLU does not satisfy the linearity check. Thus it is non-linear in those conditions.
- **Condition 3:** If either α, β, x_1 or x_2 is negative. The ReLU may or may not be linear.

Thus a ReLU is definitely piece-wise linear but it is not a linear function because it does not totally satisfy the property of superposition in some conditions. So

$$\alpha * y_1 + \beta * y_2 \neq h(\alpha * x_1 + \beta * x_2)$$

For a ReLU.

The rectification is what introduces the non-linearity.

Hope this helps.

Footnotes

[1] [Linear system - Wikipedia](#)

1.1k Views · 6 Upvotes · Answer requested by Sahil Juneja

Upvote 6

Downvote Bookmark



Recommended [All](#)

Promoted by Time Doctor

Software for productivity tracking.

Time tracking and productivity improvement software with screenshots and website and applications.

[free trial at timedoctor.com](#)



Conner Davis, PhD student Mathematics, The University of Texas at Dallas

Answered Feb 28

It's both.

A function $f : D \rightarrow S$ is linear if $\forall x, y \in D, f(x) + f(y) = f(x + y)$.

$\max(0, x)$ is a piece-wise linear function because if we restrict its domain to $(-\infty, 0]$ or $[0, \infty)$, it is linear. It is not, however, linear over its entire domain.

$f(-1) + f(1) \neq f(0)$.

Okay, so it's not linear, but it's so close to linear that you may be wondering if it still has all the approximation power of other activation functions like sigmoid.

If so, check out [Conner Davis' answer to Is a single layered ReLu network still a universal approximator?](#) You'll see exactly how we can use ReLU to do anything a sigmoid can.

1.2k Views · 14 Upvotes