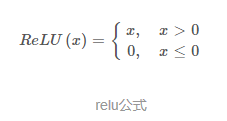
**ReLu**

ReLu来自于对人脑神经细胞工作时的稀疏性的研究，在 Lennie,P.(2003)提出人脑神经元有95%－99%是闲置的，而更少工作的神经元意味着更小的计算复杂度，更不容易过拟合。修正线性单元(Rectified linear unit,ReLU)公式：



ReLU具有线性、非饱和性，而其非饱和性使得网络可以自行引入稀疏性。

ReLU的使用解决了sigmoid梯度下降慢，深层网络的信息丢失的问题。

ReLU在训练时是非常脆弱的，并且可能会“死”。例如，经过ReLU神经元的一个大梯度可能导致权重更新后该神经元接收到任何数据点都不会再激活。如果发生这种情况，之后通过该单位点的梯度将永远是零。也就是说，ReLU可能会在训练过程中不可逆地死亡，并且破坏数据流形。如果学习率太高，大部分网络将会“死亡”（即，在整个训练过程中神经元都没有激活）。而设置一个适当的学习率，可以在一定程度上避免这一问题。