Activation Functions

itemize

I f we use identity activation function, i.e.$z^{(i)}(x) = x$, then no matter how many layers are there, the fully connected neural network falls back to a linear regression.

Proof: The output of the first layer, with identity activation function, will be $z^{(1)}(w^{(1)}x+b^{(1)}) = w^{(1)}x+b^{(1)}.Then the output will be the input of the next layer, and hence the output of the second layer will be z^{(2)}(w^{(2)}(w^{(1)}x+b^{(1)})+b^{(2)}) = w^{(2)}w^{(1)}x+w^{(2)}b^{(1)}+b^{(2)}.Similar induction can be obtained for multiple layers.Hence if we use identity activati$

W ith ReLU (Rectified Linear Unit) as activation function i.e. $z^{(i)}(x) = max(0, x), then the fully connected neural networ$

Proof: The output with ReLU activation function, will be $z^{(1)}(w^{(1)}x+b^{(1)}) = max(w^{(1)}x+b^{(1)}, 0).Then the output will b b^{(1)})+b^{(2)}) = max(w^{(2)}w^{(1)}x+w^{(2)}b^{(1)}+b^{(2)}, 0).Similar induction can be obtained for multiple layers.Hence if we use identit$