

AML SEE LAB REPORT

Image Classification using Auto Encoders.

What is AutoEncoder ?

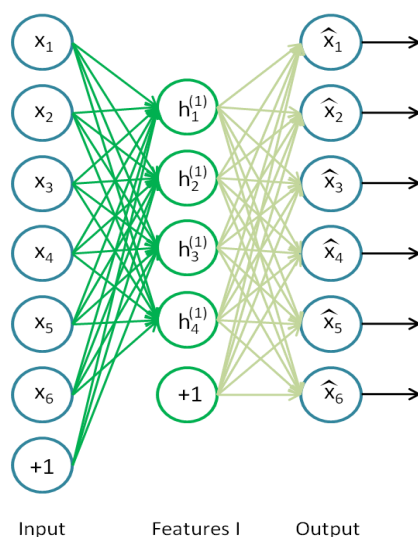
An autoencoder neural network is an unsupervised learning algorithm that applies backpropagation, setting the target values to be equal to the inputs. The aim of an autoencoder is to learn representation of a data used mostly for the purpose of dimensionality reduction.

What is Stacked AutoEncoder ?

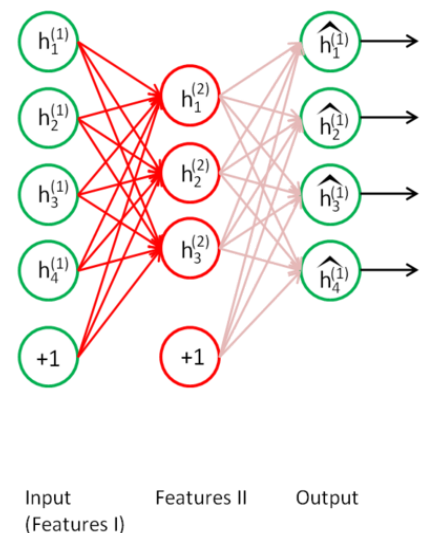
A stacked autoencoder is a neural network consisting of multiple layers of sparse autoencoders in which the outputs of each layer is wired to the inputs of the successive layer. Formally, consider a stacked autoencoder with n layers.

A stacked autoencoder network is a deep neural network which needs to be trained layer-wise such that the input and targets to the network are the same.

Image Classification

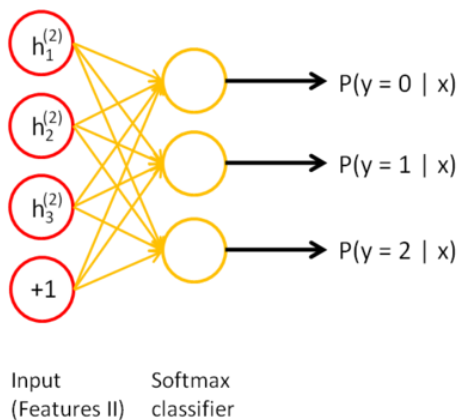


Training: First train the first layer on raw input to obtain parameters



$W(1,1), W(1,2), b(1,1), b(1,2)$. Use the first layer to transform the raw input into a vector consisting of activation of the hidden units, A . Train the second layer on this vector to obtain parameters $W(2,1), W(2,2), b(2,1), b(2,2)$. Repeat for subsequent layers, using the output of each

layer as input for the subsequent layer.

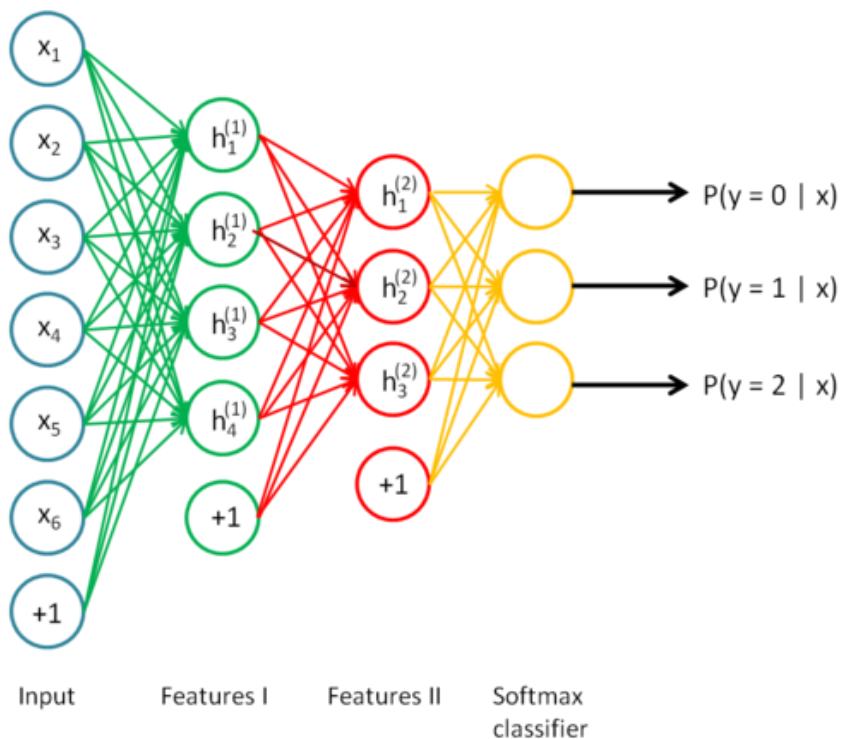


Following this, you would feed the primary features into the second sparse autoencoder to obtain the secondary feature activations $h(2)$ for each of the primary features $h(1)$ (which correspond to the primary features of the corresponding inputs $x(k)$). You would then treat these secondary features as "raw input" to a softmax classifier, training it to map secondary features to Images.

$P(y=0 | x)$ signifies that the image belongs to a class

0 when x is given

So finally our model will look like this with 5 output classes rather than 3.



Observations

h1	h2	h3	Accuracy
128	64	-	84%
128	64	32	82%
32	16	8	80%
16	8	-	78%

Conclusions

1. Using larger number of hidden units was giving a better result, might be possible because 768 input features was not properly learned by smaller size units
2. Adding more hidden layers to the stack does not guarantee better results