EE-622

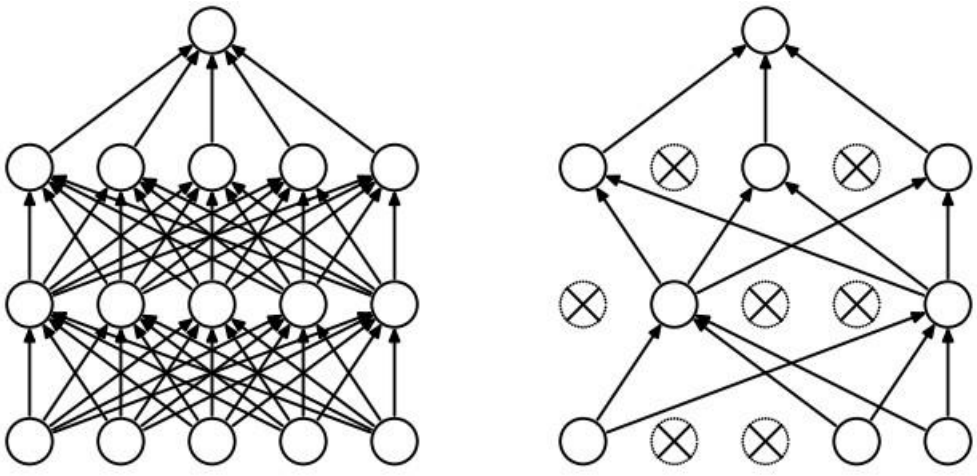
ML Assignment -2

K. Samanth Reddy

130102031

We are given datasets with images of cats and dogs. Each image size is 64*64*3 arranged in rows of trainX matrix .Labels for training is given in trainY matrix .I have used the validation spit command and splitted the training data into training and validation sets.

Before coming to the final model, I have tried various models by changing the parameters such as using the dropout value as 0.5, using different pooling and convolution layers and different activation functions. Intially I got less accuracies and later I learnt that dropout is very useful in deep neural networks. I used dropout later.It helps the network to generalise and give better accuracies since the influence of single dominant node is decreased by dropout. Drop out helps in reducing overfitting in the neural networks.



(a) Standard Neural Net (b) After applying dropout.

For optimization, I have tried Stochastic Gradient descent with momentum.I have tried changing various metrics such as sparse categorical accuracy, but I have learnt that it is only useful when the predictions are for sparse targets.Later I tried mean square error which calculates the mean square error between true and predicted target values. I got different accuracies while changing the parameters and finally I came up with my final model which is mentioned below.

While doing the assignment , I learned about convolution neural network well.For example, in image classification problem ,the first layer can be used to detect some information and in the second layer uses the information from first layer to detect simple shapes and so on. I understand that pooling is used to reduce the output dimensionality but keeps the most salient information.

I have designed a sequential model with two convolution layers and used dropout ,maxpooling and Relu and softmax activations functions in the model.Relu function f(x)=max(0,x).It is widely used in deep neural networks.Relu solves the problem of vanishing gradient in back propagation because its derivative is 1 for x>0.Softmax function is a smooth approximation of Relu function.Its value f(x)=ln(1+e^x).Basically softmax is used to map number to look like probabilities.So, I have used ReLu in the hidden layers and Softmax in the output layer.

While designing the model, I changed the training data set from int type to categorical data type to find categorical cross entropy loss.Brief view of the model is convolution->activation(relu)->maxpooling->dropout(0.25)->convolution->activation->maxpooling->dropout->flatten->dense->activation->dropout->dense->activation(softmax)->compilation

During the compilation ,I have used the loss as categorical cross entropy and used the optimizer as ada delta .I have trained the model for 10 epoches .I have used the batch size of 100 while testing.