

## Deep learning

### A. Dropout Layers

Cross Validation or Hold-out Validation

Use Regularization

Early Stopping stop training model before overfitting happened

decrease the complexity of the model

Increasing variation of data such as data augmentation

B. In the classification , example want to classify number 1-10. Output of softmax layer is a probability of number 1-10 that the greatest probability value is a valid number. Sum of probability output layer is 1. Relu is linear function that the function returns 0 if it receives any negative input, but for any positive value it returns that value [  $F(x) = \max(x,0)$  ]

C. Batch size is one of the most important hyperparameters to tune in modern deep learning systems. Advantages of small batch size is uses less memory than large bath size ,in this case use a [Oxford-IIIT Pet Dataset by O. M. Parkhi et al., 2012](#) compare between 2 bath size,32 bath size, 64, bath size ,512 bath size, 32 and 64 batch have a high accuracy than 512 but 2 batch is too small batch size has a poor accuracy than 32,64,512 batch size, small batch size take a lot of time more than large bath size during training.

Batch Size	Time	Accuracy	Memory
512	08:35	24%	Large
2	34:00	2%	Small
64	07:09	54%	Medium
32	07:23	62%	Medium

D. Use Rectified Linear Unit (ReLU Function) to fill the missing data because ReLU Function return value 0 to positive value and model converge faster. Real Data(no missing value -99) in Part 1 not have a negative value, Using this function doesn't make output is negative value.