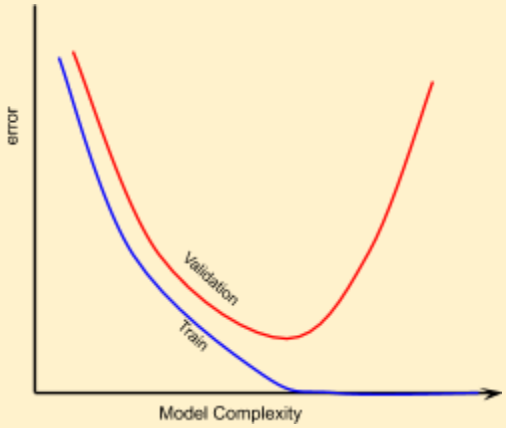
PadhAI: Regularization

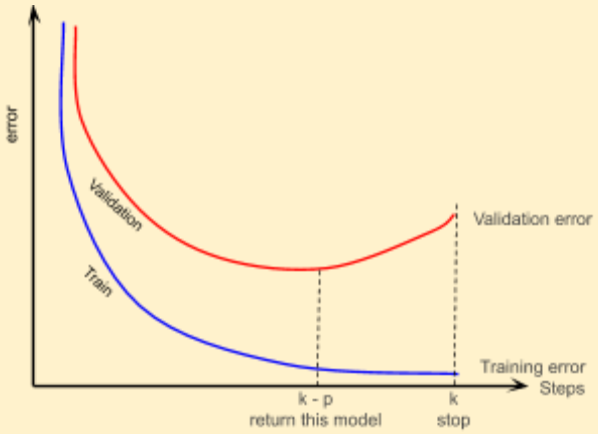
One Fourth Labs

Dataset Augmentation and Early Stopping

1. What is the intuition behind dataset augmentation?
   a. Let's look at the train-validation error curves as drawn in the previous explanations



   b. If our original dataset size is small, then it becomes easy to drive the training error to zero (too many parameters for very little data). This is because the parameters learn from the data too well to the point of overfitting. Here we will see a low train error and a high validation error
   c. Augmenting with more data will make it harder to drive the training error to zero
   d. Data Augmentation could be used to obtain multiple data points from a single input, by performing operations such as blurring, cropping, translating (move horizontally or vertically) etc. The benefit of this is that no extra effort needs to be made in labelling the data, as all the augmented images have have the same label as the original image
   e. By augmenting more data, we might also end up seeing data which is similar to validation/test data (hence, effectively reduce the validation/test data)
2. What is early stopping?
   a. Look at the image of the error curves to see how early stopping works



   b. First we keep training our model for a large number of epochs, and keep monitoring the loss.
   c. With a patience parameter p, say p = 5 epochs, monitor the validation error after a large number of epochs k.
   d. If the training error continues to decrease but the validation error stays constant in the patience period of 5 epochs, then we can avoid any more steps and revert back to k-p epochs.
   e. This can be compared to losing patience while waiting for the loss to decrease.
   f. Thus, we return the weights corresponding to the no. of epochs with lowest error.