**Distribution of Errors:**

In supervised learning the error can be roughly split into two categories, irreducible and reducible errors. The aim of error analysis and development of various techniques to improve model is to decrease the reducible error. The reducible error can be further split into two types namely bias and variance.

In the world of neural networks bias refers to the error between the expected results which mostly is the human error for the task and the predicted results obtained during training that is the training error. On the other hand variance is the difference in the predicted results for different samples from the same distribution.

Roughly speaking there are 3 different entities involved in the training and validation of a neural network. They are:

1. Training Set
2. Dev/Validation Set
3. Test Set

The training set is used to teach the model to recognise and learn various features and parameters to produce predictions. The dev set is used to optimize the trained model so as to maximise the accuracy of the model. The test set is used to run the optimised model and see its performance in the real world.

**Note:** No optimizations are made to the model when it runs on the test set.

It is important to note that the dev and the test set must come from the same distribution otherwise all the optimizations done on the dev set will yield poor results when the model is run on the test set.

The following examples taken from Andrew Ng’s course**“Structuring Machine Learning Projects”**demonstrates the meaning and fix of errors for the case when the training, dev and test set share the same distribution.

Following are the equations needed for the analysis:

1. Bias = (Training Error — Human Error)
2. Variance = (Dev Error — Training Error)

**Assumption:**training error < human error otherwise we have to consider bayesian error instead of human error

Say, we consider a task and following are the errors obtained at each step:

**Example 1:**

a) Human Error = 1%

b) Training Error = 8%

c) Dev Error = 10%

In this example clearly the value of bias is significantly higher than variance. This means that the model isn’t able to learn the features in the dataset properly. In deep learning bias mostly refers to improper gradient descent.

**Example 2:**

a) Human Error = 1%

b) Training Error = 2%

c) Dev Error = 10%

In this example the variance value is higher than bias. This means that the model isn’t able to generalise well that is it is overfitting on the training data.

**Fixes**

**Bias:**

a) Train a bigger model that is model with more layers

b) Train for a longer time

c) Use better gradient descent optimizations such as [RMSprop](https://insidebigdata.com/2017/09/24/rmsprop-optimization-algorithm-gradient-descent-neural-networks/" \t "_blank), [Momentum](https://distill.pub/2017/momentum/), [Adam](https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/) etc

d) If the above don’t work, change the network architecture

**Variance:**

a) Get more data

b) Use regularization techniques like Lasso , [Ridge](https://brilliant.org/wiki/ridge-regression/), [Dropout](https://en.wikipedia.org/wiki/Dropout_%28neural_networks%29) etc to reduce the overfitting

c) If none of the above work, change the neural network architecture

Next let us see how to interpret the errors when the training and dev + test sets come from different distributions. In this case another entity named **train-dev set**is defined on the same distribution as that of the training set. This will serve the purpose of detecting variance as the dev set comes from a different distribution as that of the training set.

The idea is to use the train-dev set to tune the model to fix the variance introduced in the model. Tuning on the dev set which comes from another distribution ensures that the model learns to make accurate predictions on both distributions.

Analysis equations:

1. Bias = (Training Error — Human Error)
2. Variance = (Train-Dev Error — Training Error)
3. Mismatched Data = (Dev Error — Train-Dev Error)
4. Overfitting on Dev Set = (Test Error — Dev Error)

Bias and Variance can be fixed using the techniques mentioned above. Mismatched Data can be fixed by adding more samples from the dev set distribution to the training set distribution. Overfitting of any kind can be solved by regularization.