Theory of learning

June 10, 2018

1 Learning Feasibility

Generalization: It follows from hoeffdings inequality (related to law of large numbers), Out of sample error is equivalent to In-sample error usually means generalization is good.

$$E_{in} \approx E_{out} \tag{1}$$

Fitting sample data: Getting sample error down is an example of fitting data well and corresponds to

$$E_{in} \approx 0 \tag{2}$$

What we want Ideally we want our learning system to perform best on out-of-sample data i.e real-world or test data, which is signified by:

$$E_{out} \approx 0 \tag{3}$$

1.1 Input distribution and Target distribution

Input Distribution THis is the distribution that generates the input examples i.e. $x_1, x_2, \cdots x_N$. This is present, but this is not what we are trying to learn.

Target Distribution This can be considered as target function combined with noice also given as $P(y|x)$. This is what we are aiming to learn.

Together using the product rule, input distribution and Target distribution get together to produce joint distribution, i.e. the example pairs that we have i.e $(x_1, y_1), (x_2, y_2) \cdots (x_N, y_N)$.