A Statistical Framework for Machine Learning

Nirdosh Bhatnagar

A statistical framework for machine learning is presented. The following topics are described.

(a) Description of the Phenomenon

(b) Development of the Model

Model 1.

Model 2.

A. Description of the Phenomenon.

Let the input space of a system be specified by X.

Let the output space of a system be specified by Y.

Let $X \times Y \triangleq Z$.

There is an unknown distribution $\mathcal{D}$ over Z.

The learner collects $m \in \mathbb{P}$ samples which are drawn form the distribution $\mathcal{D}$, where $\mathbb{P}$ is the set of positive integers. The sample set is:

$$\{(x_i, y_i) \mid 1 \leq i \leq m\}$$

B. Development of the Model:

Different models for the evaluation of the learning model are developed.

Model 1: Let $\mathcal{F}$ be a collection of models, where for each $f \in \mathcal{F}$, we have $f : X \to Y$ predicts y for a given x.

Using $m \in \mathbb{P}$ observations, the goal is to select a model $f_m \in \mathcal{F}$ which predicts *well*. One way to define the word 'well' is as follows. Let

$$err\left(f\right) = P_{Z \sim \mathcal{D}}\left(f\left(x\right) \neq y\right) = \text{ error probability of } f$$

The model f is said to have learned the phenomenon if

$$f^* = \arg\min_{f \in \mathcal{F}} err\left(f\right)$$

$$\left|err\left(f\right) - err\left(f^*\right)\right| \leq \varepsilon$$

for a specified tolerance level $\varepsilon > 0$.

Model 2: A loss function $\ell\left(f\left(x\right), y\right)$ is used to measure the discrepancy between the predicted output $h\left(x\right)$ and the true output y.

1. The *empirical risk* is:

$$R_{emp}\left(f\right) = \frac{1}{m}\sum_{i=1}^{m}\ell\left(f\left(x_i\right), y_i\right)$$

2. The *theoretical risk* is:

$$R_{theor}\left(f\right) = \mathcal{E}\left(\ell\left(f\left(X\right), Y\right)\right)$$

   where $\mathcal{E}\left(\cdot\right)$ is the expectation operator.

Examples of loss function $\ell\left(\cdot, \cdot\right)$:

1. Classification, 0-1 loss: $\ell\left(f\left(x\right), y\right) = 1\left[f\left(x\right) \neq y\right]$.

2. Regression, square loss: $\ell\left(f\left(x\right), y\right) = \left(y - f\left(x\right)\right)^2$.