Math 789 - Rethinking Generalization

Hitesh Ulhas Vaidya
hv8322@rit.edu

30th March 2020

1 Rademacher complexity

The Rademacher complexity is used to measure complexity by how accurately a hypothesis space can match a random labeling of the data. In other words, we assign random labels to datapoints in a hypothesis space and check if there exists a classifier that matches these datapoints to the randomly assigned labels. And for this purpose, we use the measure as Rademacher complexity which will tell us how generalizable a classifier is.

For this purpose we assign we assign labels, $\sigma_i \in \{+1, -1\}$ with equal probability,
$P(\sigma_i = +1) = \frac{1}{2} = P(\sigma_i = -1)$. And the formula for calculating Rademacher complexity is given by,

$$\hat{R}_n(H) = \mathbb{E}_\sigma \left[\sup_{h \in H} \left(\frac{1}{n} \sum_{i=1}^{n} \sigma_i h(x_i) \right) \right]$$

where, $h(x_i)$ = hypothesis output
$\frac{1}{n} \sum_{i=1}^{n} \sigma_i h(x_i)$ gives us the correlation between hypothesis output and the random labeling of datapoints.
example, consider, $X = \{(+1, -1)\}$,

$$\hat{R}_n(H) = \frac{1}{2} \left(\frac{1}{4}(+1+1) + \frac{1}{4}(+1-1) + \frac{1}{4}(-1+1) + \frac{1}{4}(-1-1) \right) = 0$$

2 Uniform Stability

It gives us the measure of how stable a machine learning algorithm is to a single change of datapoint in the dataset. In simple words, if you train a classifier to detect handwritten characters where you give 1000 examples for along with their label ('A' to 'Z') as training set. One way to modify this training set is to leave one example and just provide 999 training examples. A stable algorithm would still predict similar output for both 1000 and 999 training examples. The formula for this measure can be explained as follows,

- Domain Z (eg. $X \times Y$)
- Dataset $S = (z_1, ..., z_n) \in Z^n$
- Learning algorithm $A : Z^n \to W$
- loss function $\ell : W \times Z \to \mathbb{R}^+$

A has uniform stability γ w.r.t. ℓ if
For all neighbouring $S, S', z \in Z$

$$|\ell(A(S), z) - \ell(A(S'), z)| \leq \gamma$$