Some Distance Metrics and Cost Functions

Ethan Graham

21st of June 2023

Distance Metrics

L_1 *Manhattan Distance*

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$

$$d_1(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^{m} |x_i - y_i|$$

L_2 *Euclidean Distance*

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt[2]{\sum_{i=0}^{m} (x_i - y_i)^2}$$

L_p *Minowski Distance*

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$

$$d_p(\mathbf{x}, \mathbf{y}) = \sqrt[p]{\sum_{i=0}^{m} |x_i - y_i|^p}$$

This is just a generalization of the previous two distance metrics. These distances are extended from the concept of p-norms.

P-norm

For $\mathbf{x} \in \mathbb{R}^m$

$$\|x\|_p = \sqrt[p]{\sum_{i=0}^{m} |x_i|^p}$$

Loss Functions

Binary Cross Entropy

If we have sample points $\mathbf{x_i}$, $i = 1, ..., n$ that we want to classify as *red* or *green* with some decision boundary. We can ask the following question: ***What is the probability that a point on the green side of the boundary is, in fact, green?***

i.e., if I have classified a point as *green*, what are the odds of it actually being *green*? Ideally, if a point is classified as green we want its probability of being green to be 1.0, and conversely if a point is classified as red its probability of being green should be 0.0.

With this we can formulate the cross entropy cost function

$$H_p(q) = \sum_{i=1}^{N} \mathbf{y}_i \, log \, p(\mathbf{y}_i) + (1 - \mathbf{y}_i) \, log \, p(\mathbf{y}_i)$$

Where here y_i is the correct label of the data point. In the example, $y_i = 1$ is the point is *green* and $y_i = 0$ if the point is *red*.

Evaluation Metrics

Accuracy

$$accuracy = \frac{\text{True Positive} + \text{True Negative}}{\text{Total Sample}}$$

Precision

$$precision = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

We divide by the number of times we correctly classified an element of this class, plus the number of times we misclassified some *other* class as this one.

Recall

$$recall = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

We divide by the number of times we correctly classified an element from this class, plus the number of times we misclassified an element of this class as something else.

F_1 score

$$F_1 = \frac{2}{recall^{-1} + precision^{-1}}$$

Harmonic mean of precision and recall. Its value lies in the range $[0, 1]$. It tells how precise the classifier is (*how many instances it classifies correctly*) as well as the robustness of it (*did it miss a significant amount of instances?*).

Mean Absolute Error

$$\mathrm{M}AE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|$$

Mean Square Error

$$\mathrm{M}SE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$