Multiclass Classification

Multilabel and Multiclass classification

Multiclass classification deals with more than two classes where an instance is classified into a single class. For example, given a dataset with a set of features that describe the weather such that the classes are sunny, rainy and windy, a multiclass classification task will only give a single class as the result. In contrast, multilabel classification classifies an instance into a set of target labels. Articles and movies are examples where this can apply. An article can discuss a single topic but can also be about politics, religion, education and many more while movies are commonly tagged to multiple genres such as comedy, adventure, action.

The Sigmoid and the Softmax function

The softmax function is quite similar to the sigmoid explained earlier. It is used for multiclass classification because it can obtain the probabilities for various classes such that the probabilities of each class sum to 1. This means that an increase in the probability of a class causes a decrease in the probability of at least one of the other classes. It can also be referred to as a generalization of logistic regression or the sigmoid function and can be used for multi-class classification while the sigmoid function is used in multi-label classification. The softmax function is popularly used in the output layers of neural networks. Although the sum of the outputs of the softmax must be 1, this is not the same for the sigmoid function.

$$softmax(z_j) = \frac{e^{z_j}}{\sum_k e^{z_k}} \quad for\ j = 1, \ldots, k$$