

Data Sample Label Processing Method and Apparatus

ABSTRACT

Disclosed are a data sample label processing method and apparatus. The data sample label processing method comprises: obtaining a set of samples with indeterminate labels and a set of samples with determined labels; performing an iteration process with the following steps until an accuracy rate of learning labels meets a preset requirement: combining the current set of samples with indeterminate labels and set of samples with determined labels for training to obtain a prediction model; inputting a sample with an indeterminate label into the prediction model, and determining a prediction value output by the model as a learning label of the sample with an indeterminate label; obtaining a sampling subset having the learning label, and checking the learning labels of the sampling subset, so as to obtain an accuracy rate of the learning labels, and if the accuracy rate of the learning labels does not meet a preset requirement, labeling the sampling subset with a label checking result, and moving the sampling subset from the set of samples with indeterminate labels to the set of samples with determined labels; and after the iteration process ends, labeling the samples with indeterminate labels with the current learning labels of the samples.

Figure 1