K-means clustering algorithm is a clustering analysis algorithm with iterative solution. It divides the sample set into k subsets, constituting k classes, and divides samples into classes. The distance between the center of each sample and its class is minimum, and each sample belongs to only one class (which is the characteristic of hard clustering algorithm).

Its specific construction process is as follows:

**Step 1:** Initialization. The sample of is randomly selected as the center of the initial clustering.

**Step 2:** Cluster the samples. For the clustering center selected during initialization, the distance between all samples and each center is calculated, the default Euclidean distance, and each sample is aggregated into the class of its nearest center to form the clustering result.

**Step 3:** Calculate the class center after clustering, and calculate the centroid of each class, namely the mean value of samples in each class, as the new class center.

**Step 4:** Then perform steps (2) and (3) again until the clustering results no longer change.

图示

描述已自动生成