PRIVACY-PRESERVING CLUSTERING METHODS AND APPARATUSES

ABSTRACT

Some embodiments of this specification provide privacy-preserving multi-party joint data clustering methods and apparatuses. A first party and a second party hold a first feature matrix and a second feature matrix for forming an overall feature matrix of all sample objects to be clustered, respectively. The method is performed by the first party and includes a plurality of rounds of iterations, and each round of iterations includes: performing a first multi-party secure calculation with the second party based on the first feature matrix and a first centroid partition to obtain a first distance partition of a distance matrix, where a second distance partition of the distance matrix is held by the second party; performing a secure comparison calculation based on the first distance partition and the second distance partition held by the second party to obtain a first index partition of a cluster index matrix, where a second index partition of the cluster index matrix is held by the second party; and performing a second multi-party secure calculation with the second party based on the first feature matrix and the first index partition to obtain an updated first centroid partition in the current round of iterations, where an updated second centroid partition is held by the second party.

[FIG. 3]