

MULTI-PARTY JOINT TRAINING METHOD AND APPARATUS FOR DATA PRIVACY PROTECTION

ABSTRACT

Embodiments of this specification provide a multi-party joint training method for data privacy
5 protection. Multiple parties hold different feature parts of a common sample. One of the multiple
parties further holds a sample label. The method includes: The multiple parties each determine
first training losses of two local embedding models through contrastive learning based on local
samples. The local samples include a private sample and a common sample. The multiple parties
each calculate stability indicators of the two embedding models for each local sample, use, as a
10 label, an embedding vector that is output by an embedding model with higher stability for the local
sample, and determine a second training loss of the other embedding model for the local sample.
A label party determines a third training loss based on a first embedding vector that is output by a
specified model in the two embedding models for each common sample, a first embedding vector
provided by each of other parties in the multiple parties, a locally deployed fusion model, and a
15 sample label. The above-mentioned training losses are used for training a model locally deployed
on each of the multiple parties.

FIG. 2