# LANGUAGE MODEL TRAINING METHODS AND APPARATUSES BASED ON CONTINUAL PRE-TRAINING


## ABSTRACT

Embodiments of this specification provide language model training methods and apparatuses based on continual pre-training. In one example method, a soft prompt feature corresponding to each current training sample in a current domain is obtained. A latent feature corresponding to each piece of text data in the current domain is obtained. A cross-domain loss value is determined based on a difference between the obtained latent feature and a latent feature obtained based on an initial current language model in a previous domain. When a training termination condition for the current domain is not satisfied, model parameters are adjusted based on the cross-domain loss value. When a training termination condition for the current domain is satisfied, the model training process is repeated by continuing using a training sample set in a next domain, until a training termination condition for continual pre-training is satisfied.


[Figure 3]