An interpretable ensemble method for deep representation learning

Kai Jiang[1], Zheli Xiong[2], Qichong Yang[2], Jianpeng Chen[3], and Gang Chen[4]

[1]Key Lab of Information System Requirement, Nanjing Research Institute of Electronics Engineering
[2]University of Science and Technology of China School of Data Science
[3]River Delta Information Intelligence Innovation Research Institute
[4]Yangtze River Delta Information Intelligence Innovation Research Institute

February 26, 2023

Abstract

Model ensemble is widely used in deep learning since it can balance the variance and bias of complex models. The mainstream model ensemble methods can be divided into "implicit" and "explicit". The "implicit" method obtains different models by randomly inactivating the internal parameters in the complex structure of the deep learning model, and these models are integrated by sharing parameters. However, these methods lack flexibility because they can only ensemble homogeneous models with the similar structure. While the "explicit" ensemble method can fuse completely different heterogeneous model structures, which significantly enhances the flexibility of model selection and makes it possible to integrate more models with entirely different perspectives. However, the explicit ensemble will face the challenge of averaging the outputs, leading to a chaotic result. To this end, researchers further proposed using knowledge distillation and adversarial learning technologies to perform a nonlinear combination of multiple heterogeneous models to obtain better ensemble performance, however these require significant modifications to the training or testing procedure and are computationally expensive compared to simply averaging. In this paper, based on the linear combination assumption, we propose an interpretable ensemble method for averaging model results which is simple to implement, and conducting experiments on the representation learning tasks of Computer Vision(CV) and Natural Language Processing(NLP). The results show that our method is superior to direct averaging results while retaining the practicality of direct averaging.

Hosted file

An interpretable ensemble method for deep representation learning.docx available at https://authorea.com/users/589939/articles/626562-an-interpretable-ensemble-method-for-deep-representation-learning