

Large-scale Distributed Machine Learning at Tencent

SPEAKER Mr Honglin Zhang

Senior Engineer
Platform and Content Group
Tencent
Beijing, China

DATE 11 March 2019 (Monday)

TIME 2:00 pm - 3:00 pm

VENUE CS Seminar Room, Y6405
6th Floor, Yellow Zone
Yeung Kin Man Academic Building
City University of Hong Kong
83 Tat Chee Avenue
Kowloon Tong

ABSTRACT

When applying deep learning to predict the CTR (click-through rate) in a production recommendation system, one challenge is that the number of features is very large – 100 billion features for example in the feeds system at Platform and Content Group of Tencent. Embedding all the features and using a deep model such as WDL to predict the CTR would result in a large model with TBs of parameters. How to train such a big model and serve it online is then an interesting and important problem. In this talk I will share some theory and practical experience on this topic.

BIOGRAPHY

Mr Honglin Zhang is a Senior Engineer at Platform and Content Group, Tencent. He obtained his master's degree in CS at HUST in 2011 and joined Tencent in the same year. In 2014 he started working on machine learning, in particular recommendation tasks using CF, GBDT+LR, and so on. From the middle of 2017, he has been involved in the development of Numerous – a large scale distributed machine learning system that trains and serves models at TB scale for Tencent. Numerous has improved the CTR by 50% in just one year's time. He is currently working on the Instant Search business, responsible for the knowledge graph creation, question-answering systems, and document mining.

All are welcome!



In case of questions, please contact Dr Henry Xu at Tel: 3442 4840, E-mail: henry.xu@cityu.edu.hk, or visit the CS Departmental Seminar Web at <http://www.cs.cityu.edu.hk/>.

