

Han Zhu

- University: Tsinghua Univ.
 - Address: 11-419 East Main Building, Tsinghua Univ., Bei Jing, 100084
 - Mobile: +86 15201519155
 - Email: zhuhan10@gmail.com
-

Education

- 2010.8-2014.7 Bachelor student at School of Software, Tsinghua Univ.
 - 2014.8-2017.7 Master student at Institute of Information System and Engineering, School of Software, Tsinghua Univ.
-

Projects

Wanda box office prediction

In this project, our mission is to predict the box office of a movie in the first week of its release. We have all historical movie informations and their booking data in Wanda. I contributed to data cleaning, feature selection and algorithm design in this project. The final prediction result in APE (Absolute Percentage Errors) is 12% in testing data. The results are very instructive when arranging the film row piece volume.

Ocean plankton recognition

It's a National Data Science Bowl competition in Kaggle. In this competition, Participants are given 30,000+ images of 121 different kinds of planktons. The task is to train a classification model. I designed several deep nueron networks and the best one in which achieves an classification accuracy of 88%.

Transaction system optimization of HSQLDB

HSQLDB is a SQL relational database written in Java. In the earlier version of HSQLDB, it only support "SERIALIZABLE" and "READ COMMITTED" transaction levels. I lead our team to add another two transaction levels "REPEATABLE READ" and "SNAPSHOT" to HSQLDB. Besides, we extend its table level lock to row level lock to improve query efficiency.

Research

Research Area

Machine Learning, Data Quality, Deep Learning, Computer Vision

Publications

- Deep Hashing Network for Efficient Similarity Retrieval. **Han Zhu**, Mingsheng Long, Jianmin Wang and Yue Cao (AAAI 16)
 - Deep Quantization Network for Efficient Image Retrieval. Yue Cao, Mingsheng Long, Jianmin Wang, **Han Zhu**, Qingfu Wen (AAAI 16)
 - Constraint-Variance Tolerant Data Repairing. Shaoxu Song, **Han Zhu**, Jianmin Wang (SIGMOD 16)
 - Probabilistic Correlation-based Similarity measure on text records. Shaoxu Song, **Han Zhu**, Lei Chen (Information Science 289, 8-24)
-

Skills

- Data analysis
- Machine learning, especially deep learning
- Computer Vision
- C/C++, Java, Python, JS, Matlab