Li Ping

Education

2014–2017 **M.S. in Computer Science**, *University of Chinese Academy of Science(UCAS)*, Institute of Computer Technology(ICT), China, Beijing.

2010–2014 **B.S. in Computer Science**, *Jilin University (JLU)*, College of Computer Science, *top1%*.
JiLin,ChangChun

Miscellaneous

Programming C++ == Python > scala > Shell > Java

Languages Chinese,Mothertongue; English, fluent; Japanese,JLPT N2

Experience

Internship

15.05–15.09 **Data Mining Intern Engineer**, BAIDU, Beijing.
Mainly Responsible for Model Evaluation, Feature Selection and Feature Investigation.
Detailed achievements:

- *Feature Analysis Tool*:Developed a Feature evaluation tool based on the feature extraction system Adfea and model training system Platform to assist feature selection. With no uniform feature analysis tools beforehand, this work brings an end to independent assortment and thus remarkably raised the overall efficiency in FCR-Model.
- *Sug Model Evaluation*: Experimented with half-year Baidu suggestion query dataset to verify the model performance when the strategy is applied with different shrinkage factor and different time-window size. Time shrinkage strategy is widely utilized in the model to cut down the time consumption of the online CPM prediction model for suggestion query (i.e., sug).
- *Feature Survey*: Developed a through report on the strategy of the CTR Prediction competition winners on Kaggle. Concerning data include the procession of contiguous features, model utilization and combination.

Project

15.08–Now **Algorithm Developer**, **Graphical Big Data Machine Learning Platform BDA**, Beijing.
Consists of BDA Studio(dragable machine learning platform) and BDA Lib(machine learning algorithm library).
Detailed achievements:

- *Graph Algorithm*: Developed three graph algorithm (Pagerank, ICmodel, KShell), compared with intrinsic algorithms in spark graphx, they have several distinguishing features including fast convergence, scalability, and quickness.
- *Recommendation Algorithm*: Implemented Factorization Machine and NMF with local version, spark shared version, spark graphx version. Completed the Movie Rating task on MovieLens dataset from Netflix with the algorithm and got a decent RMSE.
- *ETL*: Implemented ETL module on BDA Studio, supported data loading and exporting from different data sources like Mysql, Hive and Json etc.

Competition

15.11–16.03 **Tianyi Big Data Application Competition**, *Beijing*.
Predict how many times each user visit 10 specific video websites in the 8th week given the last 7 weeks' visiting log.

- *Duties*: Feature survey, Feature extraction, Feature Evaluation.
- *Achievement*: Survived the first season as rank 6 and finally got the **best prediction result** among the **1111** participated teams.

Honors and Awards

2012 **Sesssion Segmentation Method Based on COBWEB** *Indexed by EI*

2011–2013 **National Scholarships for three successive years** *JiLin University*

2016 **Implementation of Distributed Machine Learning Algorithms: Data Distribution and Model Distribution** *CCIR*

2016 **Ease the Process of Machine Learning with Dataflow** *CIKM*

☎ (+86) 13020048987 • ✉ lipingict@gmail.com • www.lipingict.com