

PING LI

✉ lipingict@gmail.com · ☎ (+86) 130-200-48987 · 🌐 <http://www.lipingict.com>

🎓 EDUCATION

University of Chinese Academy of Science(UCAS), Institute of Computer Technology(ICT), Beijing, China
2014 – 2017

Master student in Computer Science (CS)

Jilin University (JLU), Changchun, China

2010 – 2014

B.S. in Computer Science (CS)

⚙️ SKILLS

- Programming Languages: C++ == Python > scala > Shell > Java
- Platforms: Spark, Hadoop
- Languages: Chinese - Native Speaker; English - Fluent; Japanese - JLPT N2.

👥 EXPERIENCE

✓ **Baidu FCR, Beijing, China**

2015.05 – 2015.08

Role Data Mining Intern Engineer

Project Overview Mainly Responsible for Model Evaluation, Feature Selection and Feature Investigation.

- *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's 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 thorough report on the strategy of the CTR Prediction competition winners on Kaggle. Concerning data include the procession of contiguous features, model utilization and combination.

✓ **Graphical Big Data Machine Learning Platform BDA**

2015.08 – present

Role Algorithm Engineer

Project Overview Consists of **BDA Studio**(dragable machine learning platform) and BDA Lib(machine learning algorithm library).

- *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.

✓ **Tianyi Big Data Application Competition**

2015.12 – 2016.03

- *Overview*: 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

Session Segmentation Method Based on COBWEB, Indexed in EI

2012.06

National Scholarships for three successive years, JiLin University

2011 – 2013

School Outstanding Student for four successive years, JiLin University

2011 – 2013