# Liu Bin

DATA MINING · BIG DATA EXPLORER

*200240, Shanghai Jiao Tong University, Dongchuan Rd 800, Shanghai, China*

☐ (+86) 13127810651 | ✉ liupoem@126.com | ⌂ 139.199.14.191/bing | ⌨ liubinpoem | in liubinpoem

*"Stay curious, stay passionate, life is a huge game, never cease to code and explore"*

## Education

**Shanghai Jiao Tong University**                                                                                                     *Shanghai,China*

INSTITUTE ELECTRONIC INFORMATION AND ELECTRICAL ENGINEERING                                              *Sep. 2015-Present*

- Master Degree in Integrated Circuits

**Shanghai Jiao Tong University**                                                                                                     *Shanghai,China*

INSTITUTE OF ELECTRONIC INFORMATION AND ELECTRICAL ENGINEERING                                         *Sep. 2011-July. 2015*

- Bachelor Degree in Micro-Electronics

## Writing

- Proficient:Java, Python, Data Structure, sklearn, Numpy, Pandas, HBase, Hadoop,Linux
- Familiar: C/C++, Operating System, Spark, Scala,Shell,SQL
- Basic:GitHub/SVN, Visual Studio, Vim, Android Studio, Intellij IDEA, Eclipse, Latex
- Language: CET 530

## Experience

**PingAn Technology | Data Mining Group of Big Data**                                                                      *Shanghai,China*

DATA MINING ENGINEER                                                                                                          *Nov. 2016 - Jan. 2017*

- Use Hive to obtain data of insurance salor of PingAn, train SVM, GBDT, Random Forest .etc to mine potential excellent salor
- Use Hive to obtain customers' data of PingAn Bank, then developed a collaborative filter algorithm to recommend financial products for customers. Studied how to solve the cold start problem of collaborative filtering algorithm
- Relevant Skills: Python, Hive, sklearn, MapReduce

## Projects

**Research on Short Text Classification with Naive Bayes Algorithm**                                                  *Shanghai,China*

*Sep. 2015 - Jun.2016*

- Deploy LAMP environment on Ali-cloud elastic server. Deploy Joomla Framework on webserver
- Crawl pages of secondhand board of SJTU BBS website with crawler written in Python. Store pages in local HBase database.
- Develop Naive Bayes algorithm with C++. Use part of text titles to train the algorithm. Analyze accuracy of Naive Bayes algorithm. Compare classification accuracy Naive Bayes algorithm with SVM.
- Deploy crawler and model on webserver. Use Android App to visit server. Transfer classified data in JSON format to Android App.
- Relevant Skills:Python, Java,C++, Crawler, HBase, Naive Bayes algorithm, SVM, LAMP, App Develop

**Take out Food Ranking Website Develop**                                                                                    *Shanghai, China*

*Sep. 2015 - Jun. 2016*

- Crawl raw pages of Eleme, Meituan and Baidu Take Out around SJTU and store all raw pages into HBase
- Write MapReduce program to extract information like food name, sales, price and so on. Then use <Key,Value> structure of MapReduce to rank.
- Upload results of ranking to Ali-cloud server. Develop webpages with PHP to display ranking results.
- Use statistical method and Linear regression method and so on to analyze factors' influence like discount, freight fee, price and positions in website and so on on sales.
- Relevant Skills:Java, MapReduce, HBase, sklearn, Numpy, Pandas, website develop

**Research of Applications of Big Data Analyzation Technology on Status Estimation of Power Transmission Equipment**                                                               *Shanghai, China*

*Sep. 2015 - Present*

- Use local PC and VM to deploy Hadoop cluster. Deploy HBase on top of Hadoop.
- Construct data model to store complex data of State Grid. Model supports JOIN operation, which HBase doesn't support. Use secondary index method to support transactions of State Grid.
- Publicate an academic paper A Versatile Event-Driven Data Model in HBase for Multi-Source Data of Power Grid
- Relevant Skills:Java, Hadoop, HBase, Data Model, Secondary Index, Transaction Oriented Program

## Honors & Awards

SJTU

| | | | |
|---|---|---|---|
| 2015 | **3rd Place in Person, 1st Place in Team**, | New Cup Swim Competition | *Shanghai, China* |
| 2015 | **Outstanding Graduates**, | SJTU | *Shanghai, China* |
| 2014 | **National Inspirational Scholarship**, | SJTU | *Shanghai, China* |
| 2012 | **Grand Prize of Wenzhi Cup Essay Contest**, | SJTU | *Shanghai, China* |