

Yuan Long

Current Position: ZTE Big Data Developer

Phone: 18118801655

Site: [yiddishkop's blog](#)

Applied For: ML Algorithm Engineer

Add: NanJing

Email: yiddishkop@163.com

Summary

- 8+ years work experience in the fields of CS includes **Big Data Spark development** and **Machine Learning** technologies
- Good knowledge of Data Mining including **Perceptron, LR, GBDT, SVM, RandomForest**
- Good knowledge of Deep Learning including **CNN, RNN, GAN, LSTM**
- Good knowledge of Reinforcement Learning including **Q-learning, SARSA**
- Hands on experience in Machine Learning packages including **Numpy, Scipy, Matplotlib, Pandas, SparkML**
- Hands on experience in Machine Learning framework including **Scikit-learn, Tensorflow**
- Hands on experience in Spark ecosystem including **Spark, MapReduce, SparkSQL, SparkML**
- Hands on experience in RDD architecture, implementing **Spark** operations on RDD and optimizing transformations
- Hands on experience in **High Concurrency** by **Scala** and **AKKA**
- Experience in Functional Programming Language including **Scala, Lisp**

Technical Skills

Machine Learning Framework

Numpy, Scipy, Matplotlib, Pandas, Scikit-learn, Tensorflow

Spark Eco-system

Spark, SparkSQL, SparkML

Programming Language

Python, Scala, Java, C, lisp

Machine Learning Algorithms

LR, GBDT, SVM, RF, CNN, RNN, GAN, LSTM, Q-learning

Operating Systems

Linux, Mac OS

Develop tools

Emacs, Vim, Git

Work Experience

Data Engineer @ [ZTE, NanJing](#)

Spark, SparkML

03-2017 ~ present

PROJ: IPTV GoldenData, a analytics platform for China Telecom

- Migrate MapReduce jobs into Spark RDD transformations using Scala
- Design and implement statistical models including K-means, LR, GBDT using SparkML
- Implement Spark using SparkSQL for faster processing of data
- Develop Spark jobs to parse the JSON and XML data
- Optimizing of existing algorithms in Hadoop using Spark Context, SparkSQL, Data Frame and Pair RDD.
- Used SBT scripts to fetch, build, deploy application to development enviroment

Big Data Engineer @ ZTE, NanJing

Spark, Hadoop

03-2016 ~ 12-2016

PROJ: PON manage system, a real-time event-driven web server for China Telecom

- Involved in implementing API of real-time streaming ETL module using Scala and utilizing Spark Core, Spark Streaming and Spark SQL API for faster processing of terabytes log data.
- Involved in Completing high concurrency API of RPC for management of PON's connection, request, uploading error log.
- Debugging and identifying issues reported by QA with Hadoop jobs.

J2EE Developer @ DaKun, ShenZhen

Java, J2EE, Hibernate, Spring

07/2009 ~ 12/2014

PROJ: DaKun is a start-up focuses on company management consulting, this project is a HR management system

- Website persistence layer development using Hibernate framework.
- Design and implement ORM layer of web server, using Hibernate.
- Involved in Controller(of MVC) component in Spring framework.
- Developed database schema and SQL queries for querying, inserting and managing database.
- Implemented various design patterns in the project such as Builder, Singleton, Proxy patterns.

Education

study program

Master of Science	Computer Science
University	NanJing University of Science and Technology

ML mooc courses

NTU	<u>ML Foundations & Technical</u>	my lecNotes, HW, Final Project
NTU	<u>DL with deep and structured</u>	my lecNotes, HW, Final Project
Ucberkeley	<u>cs-189: Intro to ML</u>	my lecNotes, HW, Final Project
Ucberkeley	<u>cs-294: Deep Reinforcement Learning</u>	my lecNotes, HW, Final Project
Stanford	<u>cs-20SI: DL with Tensorflow</u>	my lecNotes, HW, Final Project