

EDUCATION

- **Shanghai University of International Business and Economics(SUIBE)** Shanghai, China
B.S. in Applied Statistics; Overall GPA: 3.35/4.00 ; Major GPA: 3.86/4.00 Sept. 2013 – Jul. 2017

WORK EXPERIENCE

- **Qutoutiao Inc.** Shanghai, China
Machine Learning Engineer Nov.2018 - Present
 - Built Deep Learning models (Item2vector, DSSM, Graph Embedding) using Python. It learned multidimensional vector representation of users and items.
 - Built Deep Learning models (DNN, DeepFM) using Python to predict click-through rates(CTR), and the CTR improved by 5% over the traditional machine learning models (FM, Xgboost).
 - Researched how to extract text features and classify large-scale texts. Completed a Python program with Deep Learning models (FastText, Text-CNN, BERT) to compute text multi-label classification.
 - Developed an online model prediction service based our Recommendation Engine using Golang and C++. Deployed and maintained the service on a Linux server.
- **Ctrip.com International, Ltd.** Shanghai, China
Machine Learning Engineer Nov.2017 - Nov.2018
 - Developed the LambdaMART model in Python based on Xgboost. It predicted the score of products, and the Top-3 recommendation accuracy increased by 10%.
 - Built a Matrix Factorization model of Spark Milib package in Scala. It calculated the implicit vector of users and items. Managed data of implicit vectors on Redis.
 - Derived user behavior data from report log with PySpark/Hive, and completed an online recommendation service in Java.
- **QianLong InfoTech Ltd.** Shanghai, China
Data Analyst Mar.2017 - Nov.2017
 - Provided data support and analysis for business needs. Built market models (RFM clustering model) using R, it classified users into different levels.
 - Collected data with Python crawler, analyzed the data and used tableau to make BI reports.

PUBLICATIONS

- [1] Cheng Gang. (2017). **Analysis of Influence Factors of Teaching Quality Evaluation Based on Bayesian Theory** (In Chinese). *Modern Economic Information*(05), 384-386.

RESEARCH EXPERIENCE

- **Challenge Cup Competition** Aug.2016 - Nov.2016
 - Developed a statistical software (VISAP) in C#. Implemented some basic statistical functions, including parameter estimation, hypothesis testing, linear regression, etc.
- **Kaggle: Home Credit Default Risk (Top 2% | 137/7200)** Jun.2018 - Aug.2018
 - Designed feature engineering, feature selection (Lasso, FeatureTools), unbalanced data processing (ENN, Tomek), model optimization (GridSearch, BayesOpt, LightGBM, XGBoost, DNN), and model blend(Stacking).
- **The 8th Mathematical Modeling Competition of SUIBE (First Prize)** May.2016
 - Proposed a Bayesian Generalized Stochastic Petri Net to solve optimization problem.

TECHNICAL SKILLS

- **Coding:** Python, Golang, Java, Scala, Sql, R, C#, etc.
- **Machine Learning:** Proficient in Statistical/Machine Learning algorithms and Deep Learning models for Recommendation System and Natural Language Processing.
- **Frameworks & Tools:** Scikit-learn, Tensorflow, Pytorch, Hadoop, Spark, Mysql, Redis, Maven, Git