Chen Yanjun
betochenyanjun@163.com|86-181-100-244-60
GitHub: https://github.com/Yanjun-Chen

Education Background

University of International Business and Economics (UIBE) ***Sept.2016-Jul.2020***
Major: E-Commerce; GPA: 3.63/4.0, Major GPA: 3.78/4.0, **Rank: 4th/60** *Beijing, China*
·Honors: H Award of 2018 Mathematical Modeling Competition (*Apr.2018*); Comprehensive Scholarship at UIBE & Outstanding College Student (Top 10%) (*2019/2018/2017*)

Internships

Xiaoice AI Team, Microsoft ***Nov.2019-present***
Software Engineer Intern *(NLP Algorithm)* *Beijing, China*

- Participate in R&D tasks of Xiaoice AI products, including designing **NLP algorithms** with **C#** and **Python**
- Facilitate data mining and data analytics with Python and **Spark**

Trading Engine Technology Team, Didi Chuxing ***Jul.2019-Nov.2019***
Algorithm Engineering Intern *Beijing, China*

- Conducted the optimization and feature engineering of 3 large **Machine Learning** models covering all users in China with daily processing capacity of more than 200+ million rows
- Collaborated with data scientists to develop **recommendation system** in DiDi App and improved its accuracy by 2.5%
- Developed the cross-business order allocation model based on **Bipartite Graph** and **Reinforcement Learning**, wrote Python codes of more than 9,000+ lines and improved the GMV of models by 3%

Global Department, Didi Chuxing ***Dec.2018-Jan.2019***
Data Analysis Intern *Beijing, China*

- Facilitated data analysis, including data processing with **Hadoop** and data visualization with **Python**
- Wrote **Hive SQL** codes of 3 large-scale search projects with 1,500+ lines of codes each

Research

"Claim Frequency Predicting Based on Lightgbm" ***Feb.2019-Aug.2019***
First Author, Co-authored with Prof. Yuantao Xie *UIBE, Beijing*

- Submitted to SCI journal *Decision Support Systems*, presented a new model containing higher veracity and stability by combining LightGBM algorithm and Youden index and thereby improved the AUC by 13%
- Responsible for data pre-processing, feature engineering, algorithm design & development, visualization, and mathematical modeling with Python and Matlab

Text Data Mining Based on Social Media ***Dec. 2016- Apr. 2017***
Research Assistant *UIBE, Beijing*

- Coordinated with 5 research scientists, graduate students and undergraduates to build Hidden Markov models for **Natural Language Sentiment Analysis** (NLP)

Awards & Competitions

DNS & DGA Recognition System, Chinese Collegiate Computing Competition ***Mar.2019-Jul.2019***
Group Leader, 3rd Prize Winner (top 0.2%) *Beijing, China*

- Independently designed the whole recognition system frame, algorithm and data pre-processing modules
- Built **Deep Learning** models **(LSTM/CNN1v/XGBoost)** with **Attention** Mechanism and pre-processing module and realized the automation of the system with Python on **Linux**

The Third Tsinghua-Hualiang Cup Big Data Quantification Competition ***Dec.2018-Feb.2019***
Team Leader, Ranked 5th/150 *Beijing, China*

- Established the stock forecasting model with the processing capacity of more than 4 million+ rows
- Utilized Python and Hadoop for data pre-processing, feature engineering, and parameter adjustment
- Built 10 **Machine Learning** models for prediction of 2,831 stocks in 5 years, with the model's R-square being 0.641

Additional

IT Skills: C, C++, C#, Python, Java, R, Hadoop, SQL, Hive, Linux
Extracurricular Activities: Deputy Head and Lead Clarinet Player, UIBE Philharmonic Orchestra (*Sept.2016-Jun.2018*); First Baseman, UIBE Baseball Team (*Dec.2017-present*)