Yichen PAN

Education

2013 – 2017 **The University of Nottingham Ningbo, China**.
BSc Hons Computer Science

Average mark 86 (British education system, rank 1st, GPA: 4.0)

Honors and Awards

08/2016 **Best project in seals competiton**, *Alibaba Group*.
For the best project out of 12 projects

05/2016 **Best group project and group leader**, *The University of Nottingham Ningbo China*.
For the best year 3 group project in Computer Science department and the best group leader

03/2016 **UI design competition**, *The University of Nottingham Ningbo China*.
First award, 1%, Demo link

03/2016 **Outstanding student**, *The University of Nottingham Ningbo China*.
For innovative and excellent academic contribution, 1%

12/2015 **Provost's Scholarship**, *The University of Nottingham*.
For excellent academic performance, 15000RMB, 1.5%

12/2015 **2015 Ningbo Collegiate Programming Contest**, *The Education Committee of Ningbo*.
First award

05/2015 **Outstanding Director of Organization**, *The University of Nottingham Ningbo China*.
For outstanding leadership in social activities, 1%

12/2014 **Dean's Scholarship**, *The University of Nottingham*.
For good academic performance, 6000RMB, 10%

Papers

09/2016 **Learning to detect redundancies with word mover's distance based on distributed representations**, *Association for the Advancement of Artificial Intelligence (AAAI)*.
Applying

11/2016 **Reflections from Student-staff Collaboration to Develop an Open Educational Resource**, *Teaching, Assessment and Learning for Engineering (TALE)*.
Applying

Internship

05/2016 – 09/2016 **Taobao Team**, Alibaba Group, Algorithm Engineer.

- Automatic mobile-based speaker verification system development
  - Realize and compare several state-of-art approaches and models, such as GMM-UBM, I-vector and JFA to realize speaker verification tasks in practice.
  - Optimize and come up with new algorithms to improve the performance.
  - Set up a deep learning based speaker verification system.
- Security Rebot design and development under the funding from GNomeMagic Lab
  - Set up a face detection and verification system.
  - Applying for the national patent.

Technique *Java/C++/Node.js/MPI environment/Hadoop/Mongo DB*

05/2015 – 09/2015 **Big Data and Visual Analytics Lab Internship**, The University of Nottingham Ningbo China, Reserach assistant.

- Configuring and practicing a big data architecture (Hadoop and MPI) for social media data mining and novelty event detection.
- Collecting health-related data from social media websites (Twitter, Weibo).
- Conducting data pre-processing and cleaning work.
- Conducting data analysis (including keywords, sentiments, numerical analysis) using machine learning technology.
- Computer vision of big data.

Technique *Java/C++/Node.js/MPI environment/Hadoop/Mongo DB*

Skills

Basic Image Processing, Computer Vision, HCI, Data Visualization(Gephi), Python, Matlab

Intermediate Web technique, MEAN stack, Machine Learning, Natural Language Processing, Data Mining, Data analysis, R, Weka, Database Management, C/C++, Java, Design Pattern, Software Engineering, Agile Development

Research

09/2015 – Present **First Story Detection (FSD) with multiple approaches**, *coresearcher*.

- Based on the statistics language model, applied LSI and incremental TF-IDF to realize real-time FSD.
- Optimized FSD by applying SVD to solve polysemy and synonymy problems and reduce noise.
- Based on the neural network language model, applied the state-of-the-art word embeddings (word2vec) and earth mover's distance (EMD) in FSD.
  - Optimized the process of calculating EMD.
  - Applied EMD in clustering (including K-means, hierarchical clustering and etc.) to realize highly accurate FSD.
- Applying for a paper in top international conferences.

Technique *Gensim(python)/Word2Vec/neural network/machine learning/statistics modeling*

05/2015 – 09/2015 **Sentiment analysis based on lexicon, supervised and ensemble learning**, Big Data and Visual Analytics Lab.

- Compared the difference in feature selection methods (Information gain, CHI and Document frequency).
- Optimized the lexicon based sentiment analysis:
  - Optimized the algorithm by considering intensifying words, negation words, position and punctuations.
- Compared and optimized the supervised approach using machine learning classification algorithms:
  - Compared the difference in feature weight algorithms (Presence, TF and TF-IDF).
  - Compared the difference in classification by using different supervised algorithms (including Multinomial Naive Bayes, KNN, C4.5, SVM and etc.) under different feature selection methods and feature weight algorithms.
- Classified sentiment using lexicon-based ensemble learning methods:
  - Generated features based on various lexicons.
  - Enhanced classification by using ensemble classifiers (including Boosting, Bagging and etc.).

02/2016 – Present **Data mining and visual analytics**.
- Data visualization using Gephi to explore small-world network.

03/2016 – Present **Deep learning**, Big Data and Visual Analytics Lab and Nvidia, team leader.
- Cooperating with Nvidia Tech. to expand DIGITS (deep learning open source package) for practical application.

Projects

01/2016 – Present **Scientific note-taking app development**, *team leader, scrum master*.
- Product description:
  - Base on MEAN stack and node-webkit technique
  - Monitor and record users' activity data for data analysis and mining to find a scientific way of taking notes
  - Base on machine learning and NLP algorithms to classify notes to build a university ranged but major or topic specific knowledge database
  - Build a user-friendly notes and knowledge sharing environment to facilitate and manage innovation and organisational learning
- Embraced and promoted the spirit of open-source.
- Efficiently led the development team to complete a full cycle of the software engineering process.
- Acquired the experience of being a MEAN full-stack developer.
- Highly understood the conflict and change management.
- Highly developed the understanding in the agile development.

Collecton of works

Interactive games using Python, *https://github.com/PAN001/Interactive-Game-using-Python*.
- Helped myself learn Python better.

Super Delivery (UI design), *http://pan001.github.io/SuperDelivery*.
- Won the first award in the university UI competition and highest mark in the HCI module.
- Developed deeper understanding in HCI, UI design and software engineering process.

Scores

GRE 326 (Verbal: 159, Quantitative: 167)

TOEFL 100 (Reading: 29, Listenning: 26, Speaking: 20, Writing: 25)

Selected Courses at university

AE1DBS **Database Systems**, first class.
- In-depth knowledge of the relational models of database systems.
- Be able to considerable and effective use MySQL and MongoDB for practical usage.

AE1WPS **Web Programming and Scripting**, first class.
- In-depth knowledge of the network protocols, layers and models.
- Adept at designing static and dynamic websites and complete web applications using MEAN stack.
- Skilled in using scripts to conduct web crawler and information retrieval.

AE2ADS **Algorithms and Data Structures**, first class.
- In-depth knowledge of common data structures and algorithms.
- Adept at applying knowledge into practical usage.

AE2SAD **Software Application Development**, first class.
- In-depth knowledge of modern interface paradigms and design patterns.
- Adept at designing and writing object-oriented programs and GUIs.

AE2SEM **Software Engineering Methodologies**, first class.
- In-depth knowledge of modern software development strategy, change and conflict management.
- Skilled in agile development and productive teamwork.
- Skilled in software testing and quality assurance.

Selected Completed online Courses

Johns Hopkins **Exploratory Data Analysis** *Verified certification URL*

Duke **Data Analysis and Statistical Inference** *Verified certification URL*

Stanford **Mining Massive Datasets** *Certification URL*

Stanford **Cryptography** *Certification URL*

Organization

05/2014 – 05/2015 **Student Society Center (SSC)**, *The director of the secretariat department*, The University of Nottingham Ningbo China.

- Organize and managed the secretariat department.
- Handled the public relations among the organization.
- Regulated and supervised more than 80 societies and organizations at campus.
- Organized and directed several university level activities, including: UNNC Open Day, Sports Festival, Art Month and etc.
- Taking charge of editing and designing UNNC society books;
- Highly developed the capability of the operation and coordination.

09/2014 – Present **Learning Community Fora (LCF)**, *LCF for Computer Science Department*, The University of Nottingham.

- Liaised with all students and represented as students to raise their issues and requests to the academic faculty.
- Supported the improvement of academic teaching and learning.
- Obtained the ability of communication and negotiation.
- Highly developed the sense of responsibility and integrity.