Qirun Chen

https://github.com/OopsRyan

Email : ryan.qirun.chen@gmail.com
Mobile : 353-087-0937260

Education

- **University College Dublin** — Dublin, Ireland
  Master of Science in Computer Science; GPA: 3.48 — *Sep. 2017 – Sep. 2018*
  - **Main Modules**: Data Mining (Rapid Miner), Machine Learning (Weka), Text Analytics, Information Visualisation (d3.js and Tableau), Advanced Machine Learning (scikit-learn, Pandas, and Keras), Connectionist Computing (ANNs), Big Data Programming (bash, Hadoop MapReduce, and Spark), Advanced Data Structures in Java (Algorithms), Recommender Systems & Collective Intelligence (Collaborative Filtering and Crowdsourcing)
- **Ningbo University** — Ningbo, China
  Bachelor of Engineering in Software Engineering; GPA: 3.84 — *Sep. 2012 – Jun. 2016*

Work Experience

- **Zalando SE** — Dublin, Ireland
  Data Science Intern — *Summer 2018*
  - **Customer Segmentation**: Zalando is the biggest online fashion store in Europe. Built models to classify customers into one of pre-defined preference types according to their browse data on Zalando online store.
  - **Dev Environment**: Analytics environment based on **AWS**, Apache **Spark** and **Scala** dependencies. Deployed Spark applications on clusters in Amazon **EMR**, and interacted with customer data stored in **S3**. Version control on **Github**. Worked in **Agile** environment.
  - **Modelling**: Performed data cleaning and transformation to create features. Also set thresholds to reduce noise further during label generation. Built a multinomial model(based on **Logistic Regression** Model in Spark Machine Learning library). Tuned hyper-parameters using CrossValidation and GridSearch.
  - **Evaluation**: Applied basic evaluation metrics on each target zType. Used **lift analysis** on top N customers and visualised the lifts to compare with the benchmark.
  - **Job Monitoring**: Used **Ganglia** to monitor the conditions of each worker node, like free memory percentile and CPU IO wait. Analysed log in **Scalyr**, a log monitoring system, to solve problems.
- **Uni-tech Zhejiang** — Ningbo, China
  Software Engineer Intern — *Sep 2015 - Mar 2016*
  - **Web application - Java EE**: A task management system for local authority supervisors to assign officers tasks(normally checking local restaurant industry legal requirements like sanitary conditions), and to analyse follow-ups by visualisations.
  - **Android Application**: A mobile application by which officers can perform tasks and record details on site. Made several customised UI by reusing the Fragments in different layouts to enhance user experience, and optimised coordinate display on maps(ArcGIS).
  - **Web Service**: Service based on **REST** for providing the interaction between the mobile app and **Oracle** database. It was the bridge between the mobile app and database. The mobile app rendered the content from database by requesting WebService.

Skills

- **Programming**: Python, Scala, Java, SQL, Spark, JavaScript, Shell, scikit-learn, Numpy, Pandas, Ruby
- **Other**: Git, Hadoop, Cassandra, Rapid Miner, Tableau, Play Framework, REST

Projects **(more at www.qirunchen.com)**

- **Gapminder Bubbles - Data Visualisation**: An interactive bubble chart presents an animation with statistical and historical data about the development of the countries of the world. I replicated this bubble chart using **d3.js**.
- **Super Learner Classifier - Stacked Ensemble Algorithm**: Implemented the stacked ensemble classifier described in Super Learner In Prediction(van der Laan et al, 2007) based on **scikit-learn**. A classification model that uses a set of base classifiers of different types, the output of which are combined in another classifier at the stacked layer.
- **Multi-layer Perceptron**: Implemented a multi-layer perceptron using **Backpropagation** algorithm with stochastic gradient descent. It can learn XOR and Sin functions very well. Also trained the model on public Letter Recognition Data Set to recognise people's handwriting.