Yanping Chen

Tel: +86-186-0002-1946
Email: ypchen.ruc@qq.com

QUALIFICATIONS SUMMARY

- 10+ years of hands-on experience in planning and delivering data science projects.
- Lead team to provide machine learning, big data, and AI solutions for multi-industries: Telecom, Automotive, Finance Service, Retail etc.
- Lead translator of data science books: "*The Art of R Programming"* and *"Practical Data Analysis"*
- As a technical evangelist, delivered speeches on data science and machine learning topics in technical conferences.
- Microsoft Most Valuable Professional Award in 2017 ~ 2018 and Microsoft FY20 CPE Hero Award in 2020.
- Industry Mentor for Applied Statistics Master Program, School of Statistics, Renmin University.
- Project Management Professional (PMP), Certificate in Quantitative Finance (CQF) and certifications regarding Azure Data and AI domains.

WORK EXPERIENCE

Microsoft China	Cloud Solution Architect – Data & AI	2019/2 – Current	Beijing

- As a Subject Matter Expert in machine leaning and AI, work with Microsoft engagement teams, to plan the technical roadmaps of digital transformation for customers.
- Enable customers to solve complex data science problems by providing consultation and guidance on use case identification, feature engineering, model selection and tuning, model deployment and optimization.
- Architect, design, and deliver Machine Learning and AI solutions. Collaborate with customers and partners leading the Data Architecture and Data Modelling to meet customers' business and IT needs.

Highlight Projects:

- Use text classification algorithms to identify failure engine parts based the equipment maintenance record texts. Naive Bayes, SVM, XGboost, LightGBM are used in this project.
- Machine Tool Predictive Maintenance Platform for Cummins China: Design solution architecture and conduct PoC. This solution collect data from vibration sensors and electrical sensors, transfer the data to the cloud storage via IoT technology, after extracting the features from data, build Deep Learning model (LSTM + CNN Hybrid Algorithm, implemented with TensorFlow) that identify trends to detect when an equipment will require servicing or replacement.
- Intelligent Sales Assistant for Amway China. This solution enables sales agents to search videos, pictures, and other documents to enhance their product knowledge and improve customer interactions. I design the architecture and conduct the PoC, which use chatbot to interact with sales agent, and use Azure Language Understanding service to identify intent from users' queries, improve the searching performance by using Azure Cognitive Search service to extract key phrase and named entities from documents.
- Medicine Literature Searching Platform for AstraZeneca China. Design the architecture, which involves Azure Machine Learning to enable customer to deploy NLP models to Azure Kubernetes.
- Data Science Platform Migration from on-premises to Azure Cloud for Ecolab China.
- Design and develop reusable data science and AI solution assets, cross multi-industry including manufactory, retail, financial service, such as:
- Smart Home Automation Solution

- Customer relationship management analytics
- Predictive maintenance
- Telco customer churn
- QnA Generation & Recommendation for Web Sites -- Excellent Project Awards 1st Prize, Group Project for Microsoft FY21 AI School
- Build MLOps pipeline with Azure DevOps Server, Azure Machine Learning, Jenkins and MLFlow
- Be a technology evangelist, to demonstrate Microsoft AI/ML technologies among customers, partners, internal and external technical communities:
- Build an internal website -- AI/ML Demo Hub, hosting demos and tutorials.
- As a lecturer, delivered AI/ML workshops for internal teams from other departments and won the Microsoft Learning Champ Award of FY20.
- Lead and deliver AI/ML Workshops for customers' data science and IT teams.
- Gave a talk named *Build Computer Vision Applications on Microsoft AI Platforms with Hours* in Microsoft Online Tech Forum 2020.

Beijing Supstat Technology Co., Ltd.	Chief Data Scientist	2014/1 – 2019/1	Beijing

The company is a China subsidiary of Supstat Inc. (https://supstat.com/), which is based in New York. Supstata statistical consulting company specialized in statistical computing and machine learning. Supstat also offer high quality onsite and online training in R/Python, applied statistical methods, machine learning, and data visualization.

Responsibilities:

- **Project Operation**: Lead business as Project Manager, Pre-sales Engineer and Product Manager, responsible for engagement of new potential customers.
- **Project Delivery**: As chief data scientist, lead the whole delivery processes: data preparation, design, develop and implement models according to clients' needs.

Highlight Projects:

Category 1: Developing and Deploying Data Science Algorithms

- Developed Dynamic ROI Model of Multi-Channel Attribution for a market investigation company. The model is evaluated with Bayesian analysis method, which decompose sales time series into long-term impact, seasonal response and short-term response of marketing actions, I developed two R packages: *ssDLM* and *ssDLMServer* to implement the algorithm and developed JAVA middleware to wrap the R packages into RESTful web service, then integrated the algorithm into customer's production systems.
- Developed the vehicle sales analysis platform for Toyota Motor Corporation, to help the company forecast the data of vehicle registration and vehicle sales. Based on the ARIMA-X13 model to decompose the time series data into trend factor, seasonal factor, and irregular fluctuation, adjusted the algorithm considering the seasonal effects of the Chinese New Year problem.
- Developed an R GUI for Nomura Research Institute to computing seasonal factors of vehicle sales data and forecast future sales, also based ARIMA X-13 algorithm.
- "*cigarette quality and style evaluation system*" for Yunnan Tobacco, I took charge of product design and development of dynamic interactive graphics.

Category 2: Data Mining and Predictive Analytics Projects

- Conducted customer retention and precision marketing projects for China Telecom, involving customer churn analysis, customer segmentation, customer life-time value, and data-driven

marketing strategy-making. Involve algorithm: logistics regression, decision tree, random forest, SVM, XGBoost, K-means, etc.

- Analyzed the behavior of mobile phone replacement with data of telecommunication operators, including the prediction model of the impact of changing phone type on flow promotion, the phone changing probability analysis of users, the policy preference analysis on the phone changing of users, and the channel preference analysis on the phone changing of users. Involve algorithm: Discrete Choice Models, decision tree, SVM, LightGBM, etc.
- Performed customer relationship management data analysis project for PetroChina, which involved RFM analysis on gas station membership card data, customer segmentation and natural gas sales forecast. Involve algorithm: K-means, ARIMA, Facebook Prophet, etc.

Category 3: Providing Onsite and Online Training Courses of Machine Learning

- Developed training materials, cooperating with team members in New York, operate the training school in the brand of NYC Data Science Academy (https://nycdatascience.com)
- Delivered onsite training course for enterprise Customers, e.g., HP, Roche, China Eastern Airlines, Citi Bank, China Merchants Bank, China Telecom etc.
- Published online courses, such as *"Introduction to Data Analysis with R"* and *"Large-scale Data Analysis with R"*, on several online learning platforms, with the cumulative number of students exceeding 50,000.

Xinhua News Agency, Index Research Center	Data Analyst	2010/7 – 2013/12	Beijing

Responsibilities:

- Responsible for the research and development of Index products: designing the architectures of indicator systems, establishing the database of indicator data, developing the algorithms of index computation, and writing the analysis reports. Major products include:
  - International Financial Centre Development Index
  - Iron ore Price Index
- Served as an IT leader
  - design and implement the RMB Index Analysis Platform, which compute and visualize RMB index based on foreign currency rates.
- Wrote comparative analysis reports based on carefully study of major providers of economics indices.

EDUCATION

School of Statistics	**Renmin University of China**	2007/9 – 2010/7	**Beijing**
Master of Science in Mathematical Statistics

School of Science	**Beijing Jiaotong University**	2003/9 – 2007/7	**Beijing**
Bachelor of Science in Information and Computing Science

AWARDS & CERTIFICATIONS

- 2021 – Generation & Recommendation for Web Sites: Excellent Project Awards 1st Prize, group project for Microsoft FY21 AI School
- 2021 – Project Management Professional certification
- 2021 – Microsoft FY21 Learning Champ Award
- 2020 – Microsoft FY20 CPE Hero Award
- 2020 – Microsoft Certified: Azure Data Engineer Associate

- 2020 – Microsoft Certified: Azure Data Scientist Associate
- 2019 – Microsoft Certified: Azure AI Engineer Associate
- 2017 ~ 2018Microsoft Most Valuable Professional Award
- 2018 ~ 2021 Microsoft Certified Trainer
- 2018 – Microsoft Certified Solutions Expert: Data Management and Analytics
- 2018 – Microsoft Certified Solutions Associate: Machine Learning
- 2008 – CQF (Certificate in Quantitative Finance)
- 2005 – 1st Prize, CSEE Cup Mathematical Contest in Modeling

PROFESSIONAL SKILLS

- Big data technologies: Hadoop, Hive, Spark, etc.
- R/Python: data analysis, machine learning and data visualization, web scraping
- PHP: Familiar with web applications development, more than 6 years' experience in PHP MVC frameworks such as CodeIgniter and Yii
- More than 8 years of experience in frond-end techniques: HTML, CSS, Ajax, jQuery, Bootstrap
- C/C++: Familiar with numerical computing, Monte Carlo Simulation, derivative pricing, and risk evaluation
- MATLAB: Have experience to develop the MATLAB codes to evaluate the risk parameters of interest rate derivative financial products.
- About 3 years of experience in Azure Cloud: Cognitive Services, Azure Data Services, DevOps
- Languages: Mandarin (Native Speaker) and English (Fluent, IELTS overall score: 6.5)

PUBLICATIONS

- Translated the book *The Art of R Programming* (author: Norman Matloff) and *Practical Data Analysis* into Chinese, which have been published by China Machine Press.

CONFERENCE SPEECH TOPICS

- *Machine Learning Explainability at a glance.* 12th China R Conference (Shanghai), December 22, 2019
- *Build Flexible and Powerful Web Crawler with RSelenium*. 10th China R Conference (Shanghai), December 3, 2017.
- *Building Big Data and Machine Learning Applications with R, HDI and PowerBI*. Microsoft Tech Summit in Beijing, Nov. 2, 2017.
- *Business Applications of Dynamic Linear Models*. 9th China R Conference (Beijing), 29 May 2016.
- *New Direction for Big Data, New data Industry under the Open Sharing Trend*. 7th China R Conference (Beijing), May 25, 2014.
- *Application of Financial Data Mining with R: an example of stock return forecast*. 4th China R Conference (Shanghai), November 13, 2011.
- *Non-parametric Regression with R*. 3rd China R Conference (Beijing), June 15, 2010.
- *Stochastic Differential Equations: Simulation and Inference with R*. 1st China R Conference (Beijing), December 13, 2008.

RESEARCH PROJECTS

- Participated in the sub-project of the key research project on humanities and social sciences of the Ministry of Education: *Multi-level analysis of factors affecting the fertility intention of Chinese women of childbearing age* In 2009, and was responsible for data processing and model algorithm.