Yanjie He

Mobile: +1 (202)733-7796
Portfolio of Projects: https://yanjiehe.github.io/
Email : heyanjie0@outlook.com
Address: 1255 New Hampshire Ave NW, Washington DC 20036, USA
Github: https://github.com/YanjieHe
LinkedIn: https://www.linkedin.com/in/yanjiehe/

Education

- **The George Washington University** — Washington D.C., USA
  Master of Science in Data Analytics (Computer Science track); GPA: 3.80/4.00 — *Anticipated May. 2020*
- **Shanghai University of International Business and Economics** — Shanghai, China
  Bachelor of Arts in Economics; GPA: 3.71/4.00 — *Sept. 2013 – June. 2017*

Skils

- **Coding**: C, C++, C#, Java, Scala, Python, R, SQL, Scheme/Racket, HTML/CSS, JavaScript
- **Web Development**: React.js, D3.js, Spring Boot, MyBatis, RESTful API
- **Modeling Skills**: Machine Learning, Statistics, Social Network Analysis, NLP
- **Databases**: MySQL, SQL Server, Sqlite3
- **Frameworks & Libraries**: Qt5, Spark, JUnit, Eigen, OpenCV
- **Toos & Technologies**: Compiler Design, Recommender System, Linux, Git, Web Scraping, Database Systems

Experience

- **Machine Learning Researcher** — Washington D.C. Metro Area, USA
  George Washington University — *Oct. 2018 - Anticipated Apr. 2019*
  - **Project Title**: Artificial Intelligence and User Behavior for Robust Near Real-Time Recommendations
  - **Position**: Group research collaborator for Dr. Benjamin Harvey, faculty of George Washington University
  - **Recommender System**: Developing a graph-based recommender system, providing real-time query service and utilizing collected user behaviors data.
  - **Back-end Development**: Developing the back-end with Spring Boot. Using Github for collaborative development and version control. Utilizing RESTful API to provide interoperability between front-end and back-end. Operating MySQL database to organize the data.
  - **Information Retrieval System**: Developing a web scraper with Python to collect data. Applying methods in natural language processing to retrieve information from user activities in the browser.
- **Computational Social Scientist** — Washington D.C. Metro Area, USA
  George Washington University — *Sept. 2018 - Anticipated Apr. 2019*
  - **Social Network Analysis and Data Analysis**: Working for Professor Vontrese Pamphile's social science research project at George Washington University. Applying mathematical and statistical techniques to novel data.
  - **Analytical Programming**: Reviewing academic papers. Using Python, R and Scala to clean datasets and run the models. Utilizing packages including NetworkX and igraph.
  - **Reporting**: Writing and submitting data analysis report bi-weekly.
- **Data Engineer Intern** — Shanghai, China
  Kantar Media CIC — *July. 2016 - Feb. 2017*
  - **Data Collection**: Collected online comments with team members utilizing Python, and provided data cleaning solution with C# and SQL Server for Chanel APAC project.
  - **Data Visualization**: Developed Data Visualization Solution in C# for GroupM television show, to plot Venn Graph according to given data automatically.
  - **Text Mining System**: Developed a text mining system in C# with team members. The system was used by more than 50 data analysts in the company to make data analysis solution for L'Oréal, Chanel, Volkswagen, and Dell. Saved more than 10,000 dollars of outsourcing cost. The system was processing millions of text data records every month.
  - **Sentiment Analysis**: Improved the accuracy of sentiment analysis by 17% compared to the previous tool used in the company.

Selected Projects

For more details about my projects, please click the link https://yanjiehe.github.io/.

- **A Compiler and a Virtual Machine**: Developed a compiler for a statically typed language, a bytecode disassembler, and a virtual machine in C++.
- **A Content-based Movie Recommender System**: Developed a content-based movie recommnder system in C++ and Python. Extracted image and audio features from movie trailers, and utilized the network relationship between movies.
- **Text Co-occurrence Network Analysis for The Hunger Games**: Implemented text co-occurrence network to visualize the relationship between the main characters in the novel *The Hunger Games* with Python.