GUOCUI GAO

330 Elan Village Ln, San Jose, CA 95134 ⋄ TEL: (617)-470-8789 ⋄ EMAIL: ggcrosemond@gmail.com

SUMMARY

- Software engineer offering a strong foundation in programming principles; Solid Knowledge in Data Structure, Algorithms, Machine Learning, Design Patterns, OOD/DDP
- Experienced in object-oriented programming; developing, testing and debugging code
- Passion for working with big data and professional experience in data mining, statistical analysis, predictive modeling and data manipulation
- Quickly learn and master new technologies; have interest in working in fast-paced, deadline-oriented environment

SKILLS

Language	JAVA (proficient), C++, HTML5/CSS3, JavaScript
Big Data Ecosystem	Hadoop, HDFS, MapReduce, Yarn, Spark
Platforms	Linux, Unix, Ubuntu, CentOS, Shell scripting
Database	PostgreSQL, SQL, NoSQL
Tools	Git/GitHub, Matlab

EDUCATION

International Technological University — *San Jose, CA*
Master of Science in Software Engineer GPA: 3.68 — *Sep. 2015 - Present*

Tufts University — *Boston, MA*
Post-Baccalaureate in Computer Science GPA: 3.4 — *Sep. 2014 - May. 2015*

Boston University — *Boston, MA*
Master of Art and Science in Energy and Environment — *Sep. 2011 - Jun. 2013*

PROJECTS

Target Ads | *Hadoop, MapReduce, Yarn, Java* — *Jan. 2016*
- Analyze massive amount of consumers' weibo contents using MR, and rank the customers by their interests for the product
- Pick out potential target consumers for products, help digital advertisers offer more targeted ads for the potential customers

Seam Carver | *OOD/OOP, Dynamic programming, Java* — *Oct. 2015*
- A Java application that can resize images without distortion, implementing Shai Avidan and Ariel Shamir' algorithm
- Use dynamic programming for finding shortest path in the image, thus speed up the resizing process by 10 times

Percolation Problem Estimation | *Quick-Union, Monte Carlo simulation, Java* — *Sep. 2015*
- Application written in Java; Algorithm: weighted quick-union algorithm
- Estimate the threshold value for this famous scientific problem, percolation, via Monte Carlo simulation. After repeating computation T times (T is sufficiently large), program could give an accurate mean value with a 95% confidence interval

Song Lyrics Search Engine | *File I/O, Quick Search, C++* — *Jan. 2015*
- Built a database system in C++ that can import megabytes of song lyrics
- A search engine with a chained Hash Table to allow quick search through songs, rank songs by the most occurrences of the target word provided by user, and report to user

PROFESSIONAL EXPERIENCE

NextGen Realty | *Database Administrator* — *Boston, MA*
- Generating various reports by querying from database as per need — *Sep. 2013 - Aug. 2014*
- Update data, maintaining archived data, backing up and restoring database

Madison Park Development Corporation | *Statistical Models for Residential Utility Bills* — *Boston, MA*
- Estimated linear regression models for MPDC buildings' natural gas and electricity consumption — *Jan. - Aug. 2013*
- Indicators in models include monthly heating and cooling degree days, according to the public hourly weather documents
- Energy usage efficiency for each building was compared between units where the tenant pays for utilities, and buildings paid by Madison Park Development Corporation, illustrated which payment method helps save energy