Chen **Lu**

PhD in Chemistry · MS in Computer Science

605 Leland Ave, St. Louis, MO, USA, 63130

☐ (+1) 314-402-8727 | ✉ lvchen727@gmail.com | 🏠 lvchen727.github.io | ☐ lvchen727 | ☐ chen-lu

"Live as if you were to die tomorrow. Learn as if you were to live forever. – Mahatma Gandhi"

Summary

- PhD researcher with strong enthusiasm for artificial intelligence, machine learning and deep learning.

- Software engineer with solid expertise in data structures and algorithms and 2+ years experience in frontend/backend/cloud technologies.

- Data scientist with knowledge in statistics, probability, supervised and unsupervised learning, reinforcement learning and deep learning.

- Biochemist with expertise in instrumental analysis, nucleic acid chemistry, molecular biology and computational biology.

- Self-learner and problem-solver with willingness to work in a team and learn new technologies.

Experience

Washington University in St. Louis   *St. Louis, USA*

PhD researcher, Advisor: John-Stephen Taylor   *2012 - PRESENT*

- Thesis: Effect of DNA sequence context, DNA structure, and excitation method on cyclobutane pyrimidine dimer (CPD) formation for illustrating epigenetics mechanism underlying skin cancers.

Washington University in St. Louis   *St. Louis, USA*

Graduate Teaching Assistant   *2012 - 2015*

- Independently taught General Chemistry Lab I and II(5 semesters, 6 lab sections, 110 students).
- Grader for CSE584A, Algorithm for Biosequence Matching

BALSA   *St. Louis, USA*

Consultant   *2017 - Present*

- Performed market sizing and competitive analysis for local biotech and life science companies.
- Worked and communicated with team, and reported and presented results weekly.

Anhui University   *Hefei, China*

Undergraduate Researcher, Advisor: Shikuo Li   *2010 - 2012*

- Thesis: One pot synthesis of Cu2O-Graphene hydrogel for its application in electrochemistry

Education

Washington University in St. Louis   *St. Louis, USA*

Ph.D. in Chemistry and M.S. in Computer Science   *May 2018*

- GRE: Verbal 670/800(96%), Quantitative 800/800(94%), Writing 3.5/6
- GPA: 3.7/4.0, Research and Teaching Scholarship
- Courses: Fundamentals of Computer Science (A), Algorithm and Data Structures (A+), Advanced Algorithm (A), Algorithm for Biosequence Matching (A-), Computational Biology (A-). Introduction to Machine Learning, Introduction to Artificial Intelligence, Data Mining, Algorithms for Nonlinear Optimization, Cloud Computing with Big Data Applications etc

Anhui University   *Hefei, China*

B.S. in Applied Chemistry   *July 2012*

- GPA: 3.7/4.0, Rank 1/58
- National Scholarship, First Prize Scholarship

Independent Study   *USA*

Coursera/Udacity/SOA   *2014 - Present*

- Coursera: Genomic Data Science Specialization, Machine Learning Specialization, Python for Everybody Specialization
- Udacity: Machine Learning Nanodegree
- SOA: Probability, Financial Mathematics

Skills

Programming	proficient in JAVA and Python; experience with C++, Mathematica, C, R, Matlab, SQL, HTML, CSS, PHP, JavaScript
Frameworks/Libraries	Hadoop, Spark, Impala, Hive, Scikit-learn, Pandas, Numpy, Matplotlib, jQuery, Tensorflow, SQLite3
Software/Others	Git/GitHub, AWS, Google Cloud Platform, Atom, Eclipse, LaTeX, Linux, MAC OS, Bioinformatics
ML/DM	supervised learning, unsupervised learning, reinforcement learning, statistics, probability, financial mathematics
Biochemistry	HPLC, Gel Electrophoresis, PCR, Molecular Biology, CD spectroscopy, mass spectroscopy

Projects

Machine Learning Projects
Udacity

HTTPS://GITHUB.COM/LVCHEN727/UDACITY-MACHINE-LEARNING
2016

- Developed a decision tree model to predict the value of a given house in the Boston real estate market and identified the best selling price for the customers using statistical analysis tools.
- Trained and tested supervised machine learning models(SVM, Decision Tree, Naive Bayes) on a given dataset to test the factors that affect a student's performance in high school.
- Successfully employed unsupervised learning algorithms (Gaussian Mixture Model) to build customer segments from unstructured data with the help of PCA to reduce the data dimensionality.
- Applied reinforcement learning(Q-learning) to build a simulated vehicle navigation agent to drop off the passenger to the goal state in the shortest time possible.

K-mer Kernels
WashU CSE584A Final Project

HTTPS://GITHUB.COM/LVCHEN727/ALGORITHMS-FOR-BIOSEQUENCING-COMPARISON
2016

- Wrote python scripts to process raw next generation sequencing data.
- Implemented the linear-time DC3 suffix array construction algorithm to build 2BWT in JAVA.
- Developed 2BWT-based algorithm to construct distance table by counting unique k-mers in genomes.
- Successfully generated the phylogeny tree for 4 real genomes through built program.

Computational Biology
WashU CSE587A Course Projects

HTTPS://GITHUB.COM/LVCHEN727/ALGORITHMS-FOR-COMPUTATIONAL-BIOLOGY
2016

- Implemented MEME algorithm for finding multiple motifs in biopolymers.
- Implemented the Viterbi algorithm for HMM decoding.
- Implemented Baum-Welch algorithm for unsupervised parameter estimation.
- Implemented the Smith-Waterman algorithm for local alignment of two DNA sequences.

Algorithms and Data Structures
WashU CSE502N Course Projects

HTTPS://GITHUB.COM/LVCHEN727/ALGORITHMS-AND-DATA-STRUCTURES
2015

- Implemented the divide-and-conquer algorithm to find the closest pair of points in the inputs.
- Implemented hashing as part of a tool for comparing genomic DNA sequences.
- Implemented skip list as a database indexing strategy to achieve quick queries.

Publication

1. Lu C., Smith J., Taylor J.S. (2017). Evidence for the involvement of reverse Hoogsteen hairpin structures in the photocrosslinking of human telomeric DNA sequences *In preparation*.

2. Lu C. and Taylor J.S. (2017). A dipyrimidine sequence library for determining the sequence dependence of cyclobutane pyrimidine dimer formation. *In preparation*.

3. Smith J., Lu C., Taylor J.S. (2014). Effect of sequence and metal ions on UVB-induced anti cyclobutane pyrimidine dimer formation in human telomeric DNA sequences. *Nucleic Acids Res.* 8, 5007-5019.

4. Li S.K., Huang F.Z., Guo X., Yue X., Lu C., Shen Y. H., Xie A. (2012). Morphology-controlled synthesis of hierarchical ball-flower metallic Co superstructures and their thermal catalytic property. *J. Materials Research Bulletin* 11, 3499-3507.