

FANGFANG (FIONA) SHENG

617-8493-417 | fangfangsheng46@gmail.com | github.com/fangfangsheng | linkedin.com/in/fangfangsheng

EDUCATION

Northeastern University

Master's Degree in Bioinformatics (3.8/4)

Boston, MA

December 2020

Relevant Coursework: Advanced Statistics, Bioinformatics Methods, Algorithm, Machine Learning, Complex Network

University of California, Davis

Global Study Program (3.6/4)

Davis, CA

September 2016-June 2017

Relevant Coursework: Programming and Problem Solving, Machine Dependent Prog, Applied Statistics for Biological Sciences

Jiangnan University

Bachelor of Engineering in Food Science and Engineering

Wuxi, China

June 2017

TECHNICAL SKILLS/PUBLICATIONS

Languages Python, SQL, R, HTML5, CSS3, JavaScript, MATLAB, C and LATEX

Packages Pandas, NumPy, Scipy, Scikit-Learn, TensorFlow, Matplotlib, Seaborn, Altair, D3.js and NetworkX

Softwares Visual Studio, RStudio, Jupyter Notebook, Tableau and Docker

Databases MySQL, Oracle, MongoDB, Redis

Cloud Amazon Web Services (EC2, S3)

Machine Learning Supervised Learning, Semi-supervised Learning, Neural Networks and Active Learning

Certificates Python for Data Analysis and Visualization, Deep Learning Specialization, Information Visualization at Coursera

Others Native in Mandarin, Fluent in English

Publications [1] Evaluating Alignment Approaches in Superimposed Time-Series and Temporal Event-Sequence Visualizations.

[2] Sequence Braiding: Visual Overviews of Temporal Event Sequences and Attributes. To appear in IEEE VIS 2020.

PROFESSIONAL EXPERIENCE

Novartis Institutes for BioMedical Research (NIBR)

Cambridge, MA

AI and Machine Learning Co-Op

January-August 2020

- Completed 2 highly collaborative projects using machine learning and big data processing technologies to develop a high performance automated Electroencephalography (EEG) analysis system.
- Built and optimized a genotype classification model with feature selection on rodent EEG signals data helping identify novel EEG biomarkers for psychiatric, neurodevelopmental and neurodegenerative diseases.
- Designed and built a multi-parametric signal collecting pipeline that integrated three EEG recording systems to ensure that data were saved in a unified and highly usable format which support future Machine Learning and AI projects.
- Communicated and presented data insights, model key metrics to project teams, senior data scientists, and subject-matter experts daily.

Northeastern University

Boston, MA

Data Visualization Lab Research Assistant

April 2019-Present

- Work with student researchers on 2 research projects aimed to create new visualization of event sequences for decision making. These projects led to 2 accepted publications in IEEE VIS.
- Contributed in the design of experiments and conducted interviews both online and in-person to validate visualizations' effects.
- Performed analysis of outcome survey data, including contributing to the development of data analysis reports and publications.

Inspur Group Co., Ltd.

Beijing, China

ERP Consultant

October 2017-July 2018

- Tested and upgraded company financial data visualization software, identified software problems and implemented solutions.
- Ensured clear communication between clients and software developers for on-time delivery and customer satisfaction.
- Conducted training to clients and provided relevant technical support.
- Streamed Oracle Database maintenance with daily imports and backup.

ACADEMIC PROJECTS

Twitter Emoji Co-mention Network

September-December 2019

- Collected and preprocessed more than one million tweets with streamline Twitter API calls containing keywords "art" to build the Emoji Co-mention network.
- Built a directed emoji network with 1408 nodes and 12977 edges. Conducted the network analysis on the Co-mention network.
- Created a fixed network graph with dynamic labels using d3.js.

Reimagining Pathway Similarity Networks

August 2019

- Collected data from public databases, embedded pathways, implemented analysis, and designed network-based visualizations to connect pathways.
- Built a pipeline for early discovery questions of pathway similarity, pathway labeling, and inferred associations, including drug-repurposing by similarity and completed a case study involving protein pathway clustering.
- Presented the project at Novartis - Academia Hackathon 2019.