

Fangfang Sheng

◆ 617-849-3417 ◆ github.com/fangfangsheng ◆ fangfangsheng46@gmail.com ◆ [linkedin.com/in/fangfangsheng/](https://www.linkedin.com/in/fangfangsheng/)

MISSION

Seeking to leverage technical and professional expertise to effectively help user understand the data by doing the actual research and reporting it. Contribute to Human-computer interaction (HCI), Data Visualization and Personal Health.

EDUCATION

Northeastern University

Master's Degree in Bioinformatics

Relevant Coursework: Statistics for Bioinformatics, Algorithm, Machine Learning, Complex Network

GPA: 3.8/4.0

Boston, MA

2019-Present

University of California Davis

Global Study Program

GPA: 3.6/4.0

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

Davis, CA

September 2016-June 2017

Jiangnan University

Bachelor of Engineering in Food Science & Engineering

GPA: 3.5/4.0

Wuxi, China

June 2017

PUBLICATION

[1] Yixuan Zhang, Sara Di Bartolomeo, **Fangfang Sheng**, and Cody Dunne. "Evaluating alignment approaches in superimposed time-series and temporal event-sequence visualizations". In: Proc. VIS 2019 Short Papers. Vol. 25. 1. Accepted to IEEE VIS '18 (InfoVis). 2019, pp. 512-522.

RESEARCH EXPERIENCE

Northeastern University

Data Visualization Lab Research Assistant

Boston, MA

April 2019-Present

- Conduct between-subjects study on Amazon Mechanical Turk (MTurk), in-person interviews in NU lab
- Use statistic methods (t test, one-way Anova and etc.) to assess participants' performance for several temporal event sequence visualizations
- Write data analysis report on Jupyter Notebook
- Co-authored on papers and submit to conferences

INDUSTRY EXPERIENCE

Novartis Institutes for BioMedical Research

AI and Machine Learning Co-Op

Cambridge, MA

January 2020-Present

- Work with the Neuroscience team at NIBR on rodent EEG signals during rest status. The goal is to develop an automatic high throughput genotype classification model with feature selection.
- Owned, design and build end-to-end auto machine learning engines to automate data scientists' work at scale.
- Created python package can load MATLAB formatted data into pandas dataframe for better analysis and modeling.
- Frequently communicate and present data insights and model key metrics to senior data scientists and subject-matter experts.

Inspur Group Co., Ltd.

ERP Consultant

Beijing, China

October 2017-July 2018

- Tested and upgraded company financial data visualization software, identified problems and implemented solutions
- Ensured clear communication between clients and software developers
- Conducted **training to clients and provided relevant technical support**
- **Applied MYSQL language to import and process data for financial reports**
- **Streamed Oracle Database maintenance with daily imports and backup**

EXTRACURRICULAR EXPERIENCE

JNU Honors School of Innovation Quality Training Program HKU SPACE ICB 2015 Winter Class **Hong Kong, China**

Participant

March 2015

- Gained experience in marketing and brand management, influence and negotiation

- Completed a corporate R&D report, business strategic analysis, market plan, and financial analysis
- Assessed fund operations, analyzed investment strategies in school's Personal Investment and Stock Trading contest

ACADEMIC PROJECTS

Novartis – Academia Hackathon 2019

Cambridge, MA

Project: Reimagining Pathway Similarity Networks

August 2019

- Collected data from publicly 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, completed a case study involving pathway clustering

Gaussian based Active Learning Classification for Image Dataset

October-December 2019

- Created a Gaussian based binary Active Learning classifier for UCI image dataset
- Implemented the algorithm from paper 'Yan, Yan, et al. "Active learning from crowds." (2011)'
- Compared the performance between fully supervised and active learning approaches using F1 score and accuracy as criterions

Twitter Emoji Co-mention Network

September-December 2019

- Collected and preprocessed more than one million tweets with streamline Twitter API calls containing keywords "art"
- 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

SERVICE AND VOLUNTEERING

Volunteer at Bridge International Boston

May 2019- Present

Dining Service Worker at University of California Davis

Jan-June 2017

Exhibition Guide for Mobile World Congress Shanghai 2016

June -July 2016

Volunteer at Wuxi International Marathon.

March 2015, 2016

Vice President at Student Career Development Association at Jiangnan University

2013-2015

TECHNICAL SKILLS/ CERTIFICATIONS

Coding and Development: Developed tools in **Python**, SQL, MATLAB, HTML5, CSS3, D3.js, C, NetworkX, and LATEX

Visualizations and Analysis: Developed visualizations using Matplotlib, Seaborn and Gephi. Some experience with Sketch.

Report on **Jupyter Notebook**, R Markdown.

Techniques: **Machine Learning**, Active Learning

User Studies: Semi-structured **in-person evaluations**, Amazon **Mechanical Turk** Studies.

Language: Native in **Mandarin**, Fluent in **English**

Certificates: The Learning Python for Data Analysis and Visualization, Deep Learning and NLP, Neural Networks and Deep Learning