(Frank) Shuyu Lu

+86-15905695071
franklu3517@foxmail.com
Hefei, Anhui, China
github.com/Frank-LSY

Summary

I'm an energetic and passionate young man working towards further doctoral study in applying data mining and AI in healthcare and medical research, or anything related to that. I hope to use my data science knowledge to make medical, public health, and healthcare research more effective and efficient.

Work Experience

Software Development Engineer 2021.10 - Present

China Electronics Technology Group Corporation - 38th Research Institute

Hefei, Anhui, China

- I worked on the entire development process of the front-end part of two B/S applications: a data center management system, which has been delivered to the army along with the center, and a distributed data management platform, which is officially deployed and used throughout the institute.
- I supported the labor union of the institute handling the data related to staff attendance and hours of work, with the help of data cleaning, aggregating, and processing method which are closely related to feature engineering. I summarized a standard workflow for getting, cleaning, aggregating, and processing the data from the original data stored in the database in order to get the final results.

Research Experience

Understanding Heart Failure Patients EHR Clinical Features via SHAP Interpretation of Tree-Based Machine Learning Model Predictions 2020.10 - 2021.03 Pittsburgh, PA

Master thesis project ---- Under the supervision of Dr. Xinghua Lu

- We analyzed structured EHR data of heart failure patients from UPMC, 2014-2019 to explore the feasibility of interpreting clinical characteristics data using interpretable machine learning methods.
- Based on XGBoost Regression, a model was developed to predict EF scores of heart failure patients using structured EHR data; the practical implications of the model and prediction results were explained using the SHAP(a model-independent interpretable machine learning method); supervised clustering was performed using the generated SHAP values, and dimensionality reduction visualization was performed using t-SNE; analyzed the correlation between EHR features in an attempt to delineate different heart failure patient subgroups.
- A paper was accepted by AMIA 2021 annual symposium and I'm the first author:
  - *Lu S, Chen R, Wei W, Belovsky M, Lu X. Understanding Heart Failure Patients EHR Clinical Features via SHAP Interpretation of Tree-Based Machine Learning Model Predictions. AMIA Annu Symp Proc. 2022 Feb 21;2021:813-822. PMID: 35308970; PMCID: PMC8861751.*

Using Trajectories of Sequential Organ Failure Assessment (SOFA) Scores to Predict Outcomes after Cardiac Arrest 2019.10 - 2020.01 Pittsburgh, PA

A collaborative research project with the University of Pittsburgh Medical Center (UPMC) ER department physicians

- I created a program to calculate daily SOFA score sub-scales from raw daily assessments for a cohort of cardiac arrest patients admitted to UPMC Presbyterian Hospital and seen by the Post Cardiac Arrest Service from 2010-2019.
- We did a descriptive analysis of basic uni-variate Pearson's correlation between SOFA sub-scores and survival, mRS, and CPC.
- I applied the cluster method to study the correlation between sub-SOFA scores and patients' conditions(death, discharge, etc.)

Education

Master of Science in Biomedical Informatics

University of Pittsburgh --GPA 3.6/4.0

2019.08 - 2021.05

Pittsburgh, PA, US

Bachelor of Engineering in Computer Science

Central China Normal University -- GPA 3.45/4.0

2015.09 - 2019.06

Wuhan, Hubei, China

Bachelor of Art in French (Minor)

Wuhan University -- GPA 3.8/4.0

2017.03 - 2019.06

Wuhan, Hubei, China

Skills

- Tools/Frameworks:
  - Python, R, SQL, Scikit-Learn, Pytorch -- for data analysis;
  - Javascript, Typescript, Html, CSS, node, Vue, tailwindCSS -- for software development.
- Language:
  - Chinese (native), English (fluent), French (basic)