

Shuheng Chen

📍 Los Angeles ✉ shuhengc@usc.edu 🎓 Google Scholar 🔗 shuhengc.cv in LinkedIn

Education

-
- Tianjin University (China)** Aug 2020 – July 2024
B.E. in Industrial Design, GPA: 3.5/4.0
- **Core Courses:** Electrical Engineering, Computer Software, Computer Hardware, Python and Internet Programming Principles, Object-Oriented Programming, Mechanical Principles, Advanced Mathematics, Linear Algebra, Probability Theory and Mathematical Statistics, etc.
- University of Southern California** Aug 2024 – May 2026
M.S. in Data Analytics, GPA: 4.0/4.0
- **Core Courses:** Predictive Analytics, Optimization Methods, Text Analysis, Decision Analysis, etc.

Research Interest

My research focuses on artificial intelligence at the intersection with other domains, particularly the development and application of LLMs in medicine. I am currently working on RAG system and the integration of LLMs into clinical reasoning and decision support systems.

Experience

-
- Software Engineer Intern** Suzhou, Jiangsu, China
Guotai Epoint Software Co., Ltd July 2022 - Aug 2022
- Data analyst Intern** Shanghai, China
China Fortune Securities Co., Ltd May 2023 - Aug 2023
- Undergraduate Research Assistant** Peiking, China
Autonomous Driving Laboratory of Tsinghua University Mar 2023 - July 2023
- Collected scene images under extreme weather conditions, Data Preprocessing, and further trained pre-trained CNN models as the core algorithm for image recognition.
- Research Assistant** Los Angeles, CA, US
USC AI&Clinical Outcomes Group led by Dr.Maryam Pishgar Sep 2024 - Present
- Cooperated with Doctors to build interpretable Machine Learning models based on large-scale open-source clinical databases.
- Research Assistant** Los Angeles, CA, US
USC LIME Lab led by Dr.Jieyu Zhao Mar 2025 - Present
- Built a GraphRAG system from medical literature; Built a case-retrieval system based on keywords and embeddings comparison; Built a application demo website.
 - AI's ability to understand GUI videos; Built, filtered and preprocessed a large-size dataset including 100K videos.

Publications

-
- Machine Learning-Based Prediction of ICU Readmissions in Intracerebral Hemorrhage Patients: Insights from the MIMIC Databases** Jul 2025
First Author. Intelligent Computing [Springer Link](#) [🔗](#)
- Interpretable machine learning model for early prediction of 30-day mortality in ICU patients with coexisting hypertension and atrial fibrillation: A retrospective cohort study** Nov 2025
First Author. Informatics in Medicine Unlocked, Scopus Q1 in CS [Elsevier Link](#) [🔗](#)

Undergraduate Projects

Intelligent Express mini delivery vehicle

- **Arduino (C++), Solidworks.**
- Designed an unmanned autonomous tracking delivery car using Arduino (C++), capable of completing three trips along a specified route, accurately delivering packages to designated locations, and stopping in designated areas. Participated in market research, functional design, and preliminary studies, created 2D and 3D product models, designed circuits and software, fabricated car components, and completed assembly and program debugging.

Energy-Science LLM and its application website

- **LLM fine-tuning; full-stack Web development.**
- Collected data in Energy-Science by utilizing Python and other tools and then pre-processed them as training datasets. Trained large language model in the specific field of Energy Science based on the llama2-7b model. Designed and developed the corresponding application website.

Research Projects

Pneumonia Classification by Computer Vision

- **pre-trained models; Grad-CAM**
- Collected data from from MIMIC-IV-CXR database, and divided these CXR pictures into 3 types of pneumonia based on medical definitions. Further trained pre-trained models to classify different types of pneumonia and used Grad-CAM technique to detect the exact places where the models relied on most.

Medical Assitant Enhanced by GraphRAG and Similar Patients-Retrieval

- **This project is still ongoing. GraphRAG, self-consistency**
- Collected existing literature of Medical Knowledge, and used USC HPC to build a graphrag system. Collected large-scale patient cases and built a retrieval system to find the most similar patient as reference for the reasoning model.

Service

Course Producer: USC Viterbi School of Engineering, ISE 534: Data Analytics Consulting .

Reviewer: iScience, European Heart Journal - Digital Health.

Technologies

Programming Skills: C++, C; Python; html, css, JavaScript

Data Management: SQL; neo4j; Tableau

AI skills: Data analysis; LLM(experienced with RAG and ft); Computer Vision; AI+X