Yu Hu

Gainesville, FL | 352-888-9044 | huyuaaaray@gmail.com | https://huyu.vercel.app

Education

University of Florida Gainesville, FL

Electrical and Computer Engineering | **GPA**:3.56/4.0 *Aug. 2021 – May 2023*

Honor: Herbert Wertheim College of Engineering Achievement Award Scholarship (Awarded for academic excellence and research contributions)

Hubei University of Technology Hubei, China

Electrical Engineering | **GPA**:3.3/4.0 *Jun. 2016 – Jun. 2020*

Honor: Undergraduate Scholarship awarded to top 5% of students for academic excellence in Electrical Engineering.

Experience

Application Programmer Mar. 2024 – Present

University of Florida *Gainesville, FL*

- Architect and optimize scalable pipelines for extracting and processing standardized U.S. healthcare data, enhancing data efficiency and integrity.
- Develop and refine advanced **machine learning** and **deep learning** frameworks to improve predictive performance in medical data analysis.
- Design and deploy interactive, model-driven visualization platforms to support clinical decision-making and facilitate healthcare data interpretation.

Intern Machine Learning Engineer Dec. 2022 – Feb. 2024

University of Florida *Gainesville, FL*

- Cleaned and preprocessed extensive healthcare datasets, generating insightful visualizations to support data-driven decision-making for research projects.
- Developed a **computer vision**-based object detection system, integrated with an automated data acquisition pipeline to enable efficient data collection and analysis.
- Built and deployed an interactive web platform for **3D modeling** and predictive analytics, enhancing the visualization of object detection outcomes and improving user experience.

Projects

Healthcare Risk Prediction Web Application Nov. 2023 – Present

- Engineer front-end components using **React.js**, implementing secure user authentication and patient selection interfaces to enhance user experience.
- Develop a back-end **Flask** API for model integration, leveraging causal inference models to assess patient risks within the prediction framework.
- Integrate **FHIR-based database** and **EPIC system** inquiry functionality to streamline patient data retrieval and improve interoperability with electronic health records, while containerizing and orchestrating the entire software stack using **Docker**.

OMOP-FHIR Database Integration and Deployment Platform Apr. 2024 – Present

- Design and develop a comprehensive **OMOP database** to support standardized healthcare data integration and advanced analytics.
- Deploy and evaluate a web-based platform integrating **OMOP** and **FHIR** query capabilities, enhancing data interoperability and enabling real-time access.
- Utilize **ontology concept codes** to query, debug, and maintain the database, ensuring data consistency and accuracy.

Graph-Based Modeling for Alzheimer's Disease Prediction Dec. 2023 – Present

- Construct a patient similarity graph from **Electronic Health Records (EHR)** using **K-Nearest Neighbors (KNN)** to identify the top 100 most similar patient encounters.
- Utilize **NetworkX** to build the graph structure and integrate it with a **Graph Neural Network (GNN)** for predictive analysis of Alzheimer's Disease progression.

- Identify disease progression subphenotypes to improve early detection of Alzheimer's Disease and support personalized treatment strategies.

Advanced EHR Continuity Integration for Robust T2D Risk Prediction Feb. 2024 – Present

- Engineered an automated, **PCRONet**-based pipeline to systematically extract disease phenotypes, integrate longitudinal patient encounters, and execute comprehensive data cleaning on linked EHR and claims datasets from the **OneFlorida+ Clinical Research Network**.
- Develop and validate a machine learning framework for **type 2 diabetes** risk prediction, leveraging the **Harmonized Encounter Proportion Score (HEPS)** to quantify EHR data-continuity and improve predictive accuracy.
- Conduct fairness assessments using advanced algorithmic fairness metrics to identify disparities in model performance across racial-ethnic groups, ensuring equitable risk prediction.

OCT-Driven Retinal Image Biomarker Analysis for ADRD Prediction Dec. 2023 – Present

- Develop an automated pipeline to process **OCT DICOM scans**, enabling precise extraction and quantification of retinal imaging data.
- Construct predictive models using retinal thickness metrics from OCT scans to assess **ADRD onset risk**, establishing a predictive link between ocular biomarkers and neurodegenerative progression.
- Integrate OCT-derived features into machine learning frameworks to evaluate the predictive relationship between retinal measurements and ADRD development, enhancing early detection capabilities.

Object Detection and 3D Visualization GUI Apr. 2023 – Oct. 2023

- Designed an automated image collection pipeline using **YOLOv8** object detection models and an Intel depth camera.
- Implemented **MaskRCNN** and **UNet** object detection models with adaptive threshold algorithms for object segmentation and accurate measurement extraction.
- Developed a graphical user interface (GUI) using **Tkinter** to integrate detection algorithms and provide interactive 3D visualizations.

Technical Skills

Programming Languages: Python, SQL, R, C/C++, Java, JavaScript, HTML, CSS, LaTeX, Shell, Matlab
Machine Learning & Data Analysis: PyTorch, TensorFlow, Scikit-learn, XGBoost, NetworkX, Causal Inference
Tools & Platforms: Docker, Linux, Tableau, Power BI, Cluster Computing, Cloud Computing