Wenyao Yu

w6yu@ucsd.edu | (1) 909-536-3815

EDUCATION

University California, San Diego   La Jolla, California, USA
B.S. in Mathematics – Computer Science, Minor in Data Science   Sept 2022 - Present
- Current GPA: 3.9; Provost Honors for 2022-2025 academic year
- Coursework Highlights: Game Theory, AI: Probabilistic Models, Representation Learning, Discrete Math & Graph Theory, Modern Algebra, Computer Organization and System Programming, Data Mining, Advanced Data Structures, Programming/Data Structure for Data Science, Mathematical Reasoning, Vector Calculus, Linear Algebra

PROFESSIONAL EXPERIENCE

Hangzhou Dianzi University   Hangzhou, Zhejiang, China
Researcher, Medical Research Center   July 2023 - Jan 2024
- Downsampled high-resolution pathological slides (100k*100k pixels) in the Pancreatic Cancer Pathology Image (PCPI) dataset, extracted 329 image patches (224×224 pixels) with rescaling and cropping
- Applied Macenko method to normalize H&E staining variations to ensure consistency and reliability for subsequent analysis
- Used PILLOW and tkinter to develop a GUI (Graphical User Interface) that offered user-friendly visual experience for surgeons for enhanced diagnostic capabilities with step-by-step image comparisons
- Implemented a feature in the GUI that allowed surgeons to upload image directories for a classified display of each image

The First Affiliated Hospital of Wenzhou Medical University   Wenzhou, Zhejiang, China
Researcher, Medical Research Center   July 2023 – December 2023
- Conducted comprehensive data collection by gathering datasets related to chronic obstructive pulmonary disease (COPD), idiopathic pulmonary fibrosis (IPF) and pulmonary hypertension (PH) sequences of RNA
- Utilized tools including GATK and Illumina for screening, quality checking, and evaluating process on the acquired 5 open datasets in FASTQ format, reducing the data collection timeframe from 5 months to 3 months
- Drafted the paper as a co-first author, and the paper is ready for submission in Dec. 2023, the paper title: Dissecting the Role of Gut Microbiota Heterogeneity in the Onset of Chronic Lung Diseases via Mendelian Randomization Analysis and In-silico In-situ Microbiota Resequencing

Data Analyst Intern, Information Dept   June 2022 - Aug 2022
- Achieved a 56% reduction in discrepancies between medicine names and their corresponding ids MySQL and PostgreSQL, contributed to a cleaner and more consistent information system for the hospital
- Pioneered the efforts for improved workflow by introducing a new code structure, which enhanced efficiency for all 8 members in the team

PROJECT EXPERIENCE

Python Programming, a Tool for Data Scientist
Participant under the supervision of Assistant Professor Mark Vogelsberger from MIT
- Developed a CNN classifier that could distinguish between cats and dogs with 80% precision rate

Kaggle Competition: House Price Prediction Challenge Competition
Participant in a competitive machine learning challenge
- Engineered a predictive model using XGBoost to estimate rental prices with a robust feature set derived from extensive data exploration
- Achieved a root mean square error (RMSE) of approximately 84, indicating high accuracy and model reliability

EXTRACURRICULA ACTIVITIES
- Host and Participant, Honors Society at Warren College: discussed worldviews, philosophies (epistemology, metaphysics, axiology, etc.) with other fellow students
- Member of the Leaderboard, Chinese Engineering Society (CES): organized events including grad panels, lab tours, research showcase etc, and managed the Department of Academic Development (DAD)

SKILLS & OTHERS
- **Computer Tools:** Excel, MySQL, PostgreSQL, Python (pytorch), Java, MATLAB
- **Language:** English (professional), Chinese Mandarin (native)
- **Hobbit:** Cucurbit Flute (Grade Ten Certificate), Choir (Gold Award in WorldStrides choir competition, 2022)