Tina Yao

PhD student specialising in data science and machine learning using Python and SQL for health data. Strong track record for research, statistical analysis, data visualisation and scientific communication.

a: 5 Poplar House
London, NW9 4DR
e: yaotina98@gmail.com
t: +44 7874350783
w: github.com/Ti-Yao

EDUCATION

UCL — *MRes + PhD AI-enabled Healthcare Systems CDT*
September 2020 - present

- Projects:
    - AFFIRM: predictions using time-series EHR data
    - EyeBERT: developing ophthalmology-specific NLP model
    - Segmenting Cardiac MRI images: machine vision for heart measurements
- Included in the Dean's List for outstanding students 2020-2021
- Training and evaluating machine learning models on high-performance computing clusters
- Relevant modules: Information Retrieval & Data Mining (Python, 96%), Data Methods for Health Research (MySQL, 96%), Software Development for Health Data Science (Python, 92%)

UCL — *MSci Physics, 1st Class (81%)*
September 2016 - July 2020

- Awarded Brian Duff Memorial Prize for the best 4th-year project in the department
- Relevant modules: Practical Skills and Computing 1 & 2 (Python), Introduction to Programming (Java), Statistical Data Analysis (C++), Developing Scientific Communication

Rushcliffe School
September 2008 - July 2016

- A-Levels: 3A* 2A in Physics, Maths, Chemistry, Further Maths, Biology

EXPERIENCE

UCL — *Postgraduate Teaching Assistant*
October 2021 - present

- Teaching three MSc health data science courses in Python, machine learning and SQL
- Guest lecturer for Machine Learning in Healthcare course

Nuffield Research Placement — *Project Supervisor*
June 2021

- Created and supervised a health data science research project for a year 12 student

UCL — *Institute for the Physics of Living Systems Summer Internship*
June 2020 - August 2020

- Conducted independent research to model cell division using molecular dynamics software
- Contributed to a paper published in PNAS (impact factor = 11.2)

HACKATHONS

- **UCL Code Club** — *2nd (out of 7),* Created an interactive website visualising the correlations between air quality and respiratory health: us-airhealth.herokuapp.com
- **Women in Data** — *8th (out of 40),* Two-week international collaboration. Presented COVID-19 data insights using Tableau.

PUBLICATIONS

- Harker-Kirschneck, Lena, Hafner, Anne, **Yao, Tina.** et al. "Physical mechanisms of ESCRT-III–driven cell division". *Proceedings of the National Academy of Sciences* 119.1 (2022).

TECHNICAL SKILLS

<u>Advanced</u>

Python (NumPy, SciPy, pandas, matplotlib)

Machine learning (scikit-learn, TensorFlow)

SQL (MySQL, PostgreSQL)

Machine Vision

Natural language processing

Linux

<u>Basic</u>

C++

Java

AWARDS

Dean's List Award

Brian Duff Memorial Prize

CREST Gold Award

UKMT Intermediate Maths Challenge Gold Award

ONLINE CERTIFICATIONS

Natural Language Processing with Classification and Vector Spaces (Coursera)

Natural Language Processing with Probabilistic Models (Coursera)

Learn Natural Language Processing Course (Codecademy)

LANGUAGES

English (Native)

Mandarin (Intermediate)