
EDUCATION**Harvard University, T.H. Chan School of Public Health**

Boston, MA

*Master of Science in Health Data Science**Aug. 2022 – Mar. 2024 (Expected)*

Relevant courses: Data science programming in R, Statistical inference, Artificial Intelligence (MIT)

University of North Carolina at Chapel Hill

Chapel Hill, NC

*Bachelor of Science in Computer Science & Physics – Astrophysics Option; GPA: 3.99**Aug. 2018 – May 2022*

Relevant courses: Algorithms, Data structures, Computer organization, Databases, Object-oriented programming, Computer security, Distributed system, Programming languages

PROGRAMMING SKILLS

- **Languages:** Python, Java, MySQL, R, C, C++, TypeScript, Prolog, SML, Lisp
- **Technologies:** Git, LaTeX, Pandas Data Frame, Scikit-Learn, Keras, TensorFlow, Huggingface, OpenAI, BERT, GPT-3

RELEVANT EXPERIENCE**University of North Carolina at Chapel Hill**

Chapel Hill, NC

*Research Assistant**Aug 2020 - May 2022***Project: Towards a Comprehensive AI Teaching Assistant Based on Course Forums:**

- Used neural networks, word embedding, IBM Watson natural language understanding API, and Keras deep learning API to build the kernel of the AITA that helps classify forum posts and detect forum post questions that are duplicate or incomplete.
- Achieved an average accuracy of 95% on the model, collaborated on team of 3, led 95% of work, composed an honor thesis, held a thesis defense, and published the work on the Carolina Digital Repository.

Project: Automating Testing of Visual Observed Concurrency:

- Developed a new testing-based framework using Java to provide both a grading management and automation system for evaluating the concurrency requirements of assignments implemented in Java.
- Collaborated on team of 6, and published a paper on The 3rd Workshop on Education for High Performance Computing.

Project: Broad Awareness of Unseen Work on a Concurrency-based Assignment:

- Used different technologies to record events related to work on a Java assignment that exercised threads, synchronization, and coordination and provided preliminary answers to questions about the unseen work behind the concurrency aspects of the assignment.
- Collaborated on team of 6 and published a paper on The Workshop on Education for High Performance Computing.

Project: The Role of Activity & Youth on the M_KS-M_* Relation:

- Measured rotation periods and S indices of 143 binaries, fitted the position of the binaries using Monte Carlo Markov Chain (MCMC), and colored the HR diagram using rotation periods, H-alpha lines, and S indices using Python.
- Collaborated on team of 7, participated in writing the paper about this work which will be published on American Astronomical Society Journal and presented at the UNC research symposium by the end of the semester.

Project: Structure in The APOGEE/Cannon Stars:

- Used Python unsupervised learning techniques to find groupings of the Cannon dataset in chemical abundance space and determine if the same groupings also occur in physical space.
- Collaborated on team of 3, composed a thesis, and gave a presentation at the end of the semester.

Project: Computer Simulation of Nemesis's Effect on Earth:

- Used various numerical techniques in Python to determine if Nemesis, a hypothetical brown dwarf companion to the Sun, was responsible for the 26 million-year periodicity in the extinction of Earth's living species due to perturbing comets.
- Collaborated on team of 3, composed a thesis, and gave a presentation at the end of the semester.

PUBLICATIONS

- **Bowen Gu**, Hao Wang, and Kaizhuo Chen, “Towards a Comprehensive AI Teaching Assistant Based on Course Forum”, Published on Carolina Digital Repository.
- Prasun Dewan, Samuel George, **Bowen Gu**, Zhizhou Liu, Hao Wang, and Andrew Wortas, “Broad Awareness of Unseen Work on a Concurrency-based Assignment”, Published on The 3rd Workshop on Education for High Performance Computing (EduHiPC 2021).
- Prasun Dewan, Andrew Wortas, Ken Liu, Sam George, **Bowen Gu**, and Hao Wang, “Automating Testing of Visual Observed Concurrency”, Published on Workshop on Education for High Performance Computing.