John Huang

john_huang@brown.edu | linkedin.com/in/john-huang | github.com/johnlhuangP

EDUCATION

Brown University   Providence, RI
Bachelor of Science in Computer Science and Applied Mathematics   *Expected May 2027*

EXPERIENCE

Bioinformatics and Data Science Intern   May 2024 - August 2024
AbbVie   *South San Francisco, CA*

- Developed a robust **scRNA sequencing computational workflow** in **R**, leveraging industry standard bioinformatic tools and libraries, including **Seurat** and **SingleR**, for clustering analysis using **k-nearest neighbor** and **shared-nearest neighbor algorithms**.
- **Processed, performed quality control, and structuralized** scRNA and TCR sequencing datasets in **R**, enhancing accuracy of data analysis by **15**% and validated pre-clinical study results with **93**% **accuracy**.

Machine Learning Researcher   May 2024 - Present
Brown University   *Providence, RI*

- Leading a team of 3 researchers to **fine-tune Foundation Models** on high-quality, internally curated **medical datasets**, improving the efficiency and accuracy of the clinical note-writing process by incorporating advanced **machine learning techniques**.
- Implementing **qLoRA fine-tuning**, enhancing computational efficiency by **50**%, and utilizing **Fully Sharded Data Parallel (FSDP)** to enable **multi-GPU training**, optimizing large-scale model training.
- Designed and built an end-to-end **transcription-diarization pipeline** using **WhisperX** and **Llama 3.1** in Python, significantly improving the quality of clinical notes.

Full Stack Developer   September 2023 - Present
Fullstack@Brown   *Providence, RI*

- Developed the frontend of pro-bono web applications using **React** and the backend with **Node.js**, integrating **AWS** for data storage and cloud services.
- Led a team of developers in creating software solutions for university clubs and organizations, enhancing user experience by **30**%.
- Utilized **Google Cloud Platform (GCP)** for deployment, efficient data storage and role management.

PROJECTS

Bag-of-Words Model for Sentiment Classification | *Python*   September 2024

- Built a **Bag-of-Words (BOW) model** for sentiment classification on Twitter data using **spaCy** for NLP preprocessing (tokenization, lemmatization) and **scikit-learn** for feature extraction, train-test splits, and model training.
- Applied **n-grams (n=1 to 5)** to enhance model accuracy and robustness, optimizing feature extraction and achieving reliable sentiment predictions and **stratified k-fold cross-validation** to assess model performance.

24cast.org | *NextJS, React, Express, AWS*   April 2024 - Present

- Worked as a **full stack developer** in a fast paced, collaborative team to deliver an election prediction website that **outperformed major competitors** like FiveThirtyEight and The Economist in accuracy.
- Designed and implemented **AWS Lambda functions** for data processing, orchestrating daily data fetching with **AWS EventBridge**, and utilizing **Amazon DynamoDB** for efficient data storage and retrieval.

TECHNICAL SKILLS

Languages: Python, Java, JavaScript, TypeScript, R, SQL, HTML/CSS
Frameworks: React, Node.js, Next.js, Express, PyTorch
Tools: Git, Maven, AWS, Google Cloud Platform (GCP), MongoDB

AWARDS

AbbVie Possibilities Scholarship | *AbbVie*   June 2024
Undergraduate Teaching and Research Awards | *Brown University*   May 2024