John Huang

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

Education

Brown University Providence, RI

B.S. Computer Science and Applied Mathematics (4.0 GPA) *Expected May 2027*

Relevant Coursework: Computational Linguistics, Software Engineering, Probability Theory, Data Structures and Algorithms, Discrete Structures and Probability, Stochastic Calculus, Operations Research: Deterministic Models

Experience

Machine Learning Research Intern May 2024 - Present

Brown University *Providence, RI*

- **Led a team of 3 researchers** to fine-tune Large Language Models on internally curated high-quality medical datasets, **improving clinical note accuracy by 15**% **and enhancing efficiency by 20**%.
- **Designed and developed an end-to-end transcription and diarization pipeline** in **Python**, leveraging WhisperX and Llama 3.1, which resulted in a **30**% **improvement** in clinical note transcription quality.

Data Science Intern May 2024 - August 2024

AbbVie *South San Francisco, CA*

- **Engineered a scalable scRNA sequencing analysis pipeline** in **Python** and **R**, utilizing bioinformatics tools such as **Seurat**, **SingleR**, and **Scanpy** to perform clustering analysis with k-nearest neighbor (k-NN) and shared-nearest neighbor (SNN) algorithms.
- **Processed and analyzed large-scale scRNA and TCR sequencing datasets**, conducting data wrangling, quality control, and dimensionality reduction with PCA and UMAP, which **improved data integrity by 15**% and contributed to a **93**% **accuracy** in validating pre-clinical study outcomes.
- Automated data preprocessing and pipeline execution of T-Cell Exhaustion workflow, **reducing runtime by 30**%.

Software Engineer September 2023 - Present

Fullstack at Brown *Providence, RI*

- **Built full-stack web applications** for 3 university organizations using **React** and **Node.js**, implementing features like user authentication and data visualization dashboards.
- Integrated **AWS S3** for file storage and **DynamoDB** for scalable **NoSQL data storage**, resulting in a **25**% improvement in system efficiency.
- Collaborated in **Agile teams**, conducting sprint planning and regular code reviews, using **Git** for version control and issue tracking.

Projects

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

- **Developed both frontend and backend** of an election prediction website, reaching over **7,800 active users** and **76,000 views**.
- Engineered responsive user interfaces with **Next.js** and **React**, enhancing SEO performance by **30**% and increasing traffic by **25**%.
- Designed and implemented **RESTful APIs** to fetch local prediction data using **Express.js**, achieving **99.99**% **uptime** during peak election periods and orchestrated serverless architecture web scraping with **AWS Lambda**.

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

- **Created a sentiment classification model** for financial data using a **Bag-of-Words (BOW)** approach, leveraging **spaCy** for natural language preprocessing and **scikit-learn** for feature extraction, model training, and evaluation.
- Increased features using **n-grams (n=1 to 5)** and implemented **TF-IDF weighting** and **stopword removal** to optimize feature representation and improve model accuracy, achieving **63.8**% **accuracy**.

Skills

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