John Huang

650-455-7938 | 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, Data Structures and Algorithms, Discrete Structures and Probability, Probability Theory, Stochastic Calculus, Linear Algebra

Experience

Software Engineer September 2023 - Present

Fullstack at Brown *Providence, RI*

- Developed full-stack web applications for 3 university organizations using **React, Node.js** and **Django**, implementing features such as user authentication and interactive data visualizations.
- Integrated **AWS S3** and **DynamoDB** for scalable storage, increasing system efficiency by **25**%.
- Collaborated in **Agile** teams for sprint planning, code reviews, and issue tracking with **Git**, **implementing CI/CD practices** to enhance deployment efficiency and contribute to a streamlined development process.
- Applied **TDD** practices, ensuring code quality and reducing bug occurrences by **30**%.

Machine Learning Research Intern May 2024 - Present

Brown University *Providence, RI*

- Led a team of 3 researchers to fine-tune **Large Language Models (LLMs)** on medical datasets, **improving clinical note accuracy by 15**%.
- **Designed and developed an end-to-end transcription and diarization pipeline** in **Python**, leveraging WhisperX and Llama 3.1, which resulted in a **20**% **improvement** in clinical note transcription quality.
- Implemented **Retrieval Augmented Generation (RAG)** using **LlamaIndex**, optimizing model performance and retrieving relevant medical information.

Data Science Intern May 2024 - August 2024

AbbVie *South San Francisco, CA*

- **Engineered a scalable analysis pipeline** using **Python** and **R** for single-cell RNA sequencing data, improving data processing efficiency by **15**%.
- Processed and analyzed large-scale scRNA and TCR sequencing datasets, conducting data wrangling, quality control, and dimensionality reduction, **improving data integrity by 15**%.
- Automated data preprocessing and pipeline execution of T-Cell Exhaustion workflow, **reducing computational runtime by 30**%.

Projects

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

- **Developed both frontend and backend** of the Brown Political Review's innovative election prediction website, reaching over **10,000 active users** and **over 100,000 views**.
- Engineered responsive user interfaces using **Next.js** and **React**, implementing advanced frontend features such as **interactive Highcharts** for data visualization, integrating **GeoJSON** data for dynamic election maps.
- Designed and implemented **RESTful APIs** to fetch local prediction data using **Express.js**, achieving **99**% **uptime** during peak election periods and orchestrated serverless architecture web scraping with **AWS Lambda**.

Neural Machine Translation Using Transformer Architecture | *Python, PyTorch* September 2024

- Developed a custom **neural machine translation model** from scratch using the **transformer encoder-decoder architecture**, implementing core components such as **self-attention** and **cross-attention sublayers**.
- Constructed a **6-layer encoder and decoder**, integrating **feed-forward networks**, **layer normalization**, token/positional embeddings, and efficient **boolean masking** operations.
- **Implemented beam search decoding**, achieving **34**% **BLEU score**, and compared it against **greedy decoding** to optimize translation performance.

Skills

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