Long Bui

617-774-8278 | tonybui5503@gmail.com | linkedin.com/in/tonybui2003 | github.com/tonybuii2003 | tonybui.dev

EDUCATION

Stony Brook University — Stony Brook, NY
Bachelor of Science in Computer Science — *Aug 2021 – May 2025*
Specialization in Artificial Intelligence and Data Science — *GPA: 3.34 / 4.00*
Relevant Coursework: Data Structures, Algorithms, Software Engineering, System Fundamentals, Machine Learning, Natural Language Processing, Computer Networks, Offensive Security, Object-Oriented Programming

TECHNICAL SKILLS

Languages: Python, Java, C, HTML/CSS, JavaScript, TypeScript, SQL, Bash, PowerShell, R
Technologies/Libraries: React.js, AngularJS, Express.js, Spring Boot, PyTorch, NumPy, Pandas, nltk, scikit-learn
Tools/Platforms: Linux, CronJob, Windows, Jenkins, Git, Docker, AWS, GCP, Hugging Face, SPSS, Jira
Databases: MongoDB, MySQL

EXPERIENCE

Stanford Institute for Human-Centered Artificial Intelligence (HAI) — June 2025 – Present
Software Engineer — *Stanford, CA*
- Build agentic full-stack web apps to train new therapists utilizing NextJS, FastAPI, GCP, and OpenRouter API

SUNY Research Foundation — Aug 2024 – June 2025
Research Assistant (Data Scienctist) — *Stony Brook, NY*
- Developed machine learning data analysis pipelines utilizing DBSCAN clustering for behavioral feature extraction from 270 participants' phone sensor and fitbit data, enabling psychologists to gain deep behavioral insights
- Improved team productivity by 40% by creating a text reminder app using PyQt, Twilio API, and Jenkins CI/CD
- Preprocessed 1,000+ surveys to train a psychopathology AI model, enabling accurate behavioral diagnostics
- Boosted data reliability by 90% by analyzing behavioral data, calculating descriptive statistics, ICC, and cross-correlation in R and SPSS to uncover hidden patterns and validate results

The World Well-Being Project (WWBP NLP Lab) — May 2024 – Aug 2024
Software Engineering Internship — *Stony Brook, NY*
- Increased social media and Qualtrics pre/post surveys dataset by 180 users by monitoring AngularJS and Android data collection apps, integrating an ETL pipeline with MySQL for behavior analysis
- Secured additional funding by creating detailed reports on 480,000+ Facebook posts and SMS data with DLATK
- Produced weekly and final reports showing user pool growth from 12 to 204, helping researchers monitor progress, assess recruitment success, and perform root cause analysis on 30+ trial issues

PROJECTS

CuratAI | React Native, Spring Boot, scikit-learn, nltk, GPT4o, AWS, Hugging Face — GitHub | Project Demo
- Engineered a RAG React Native application to fine tune GPT4o model into an intelligent personal curator
- Boosted LLM precision by integrating 100 million tokens sourced from The Met's art collection and web data

Votifier | Python, Java, Pandas, Spring Boot, MGGG, PyEI, MongoDB — GitHub | Project Demo
- Aggregated census data using MGGG to generate district and precinct data for analysis and visualization
- Engineered backend APIs with Java Spring Boot and MongoDB to enable heatmaps and analysis results GUI

Fake StackOverFlow | JavaScript, HTML/CSS, React.js, Express.js, MongoDB — GitHub | Project Demo
- Built a StackOverflow clone enabling users to create, update, delete, and lookup questions, comments, and answers

LEADERSHIP

HackNYU, HopperHacks — Feb 2025
- Mentored a team of 4 underclassmen who won Best First Hack (HackNYU) and Best AI/ML (HopperHacks)

Loqi - Tech Businesses in Development VIP Team — Feb 2023 – Dec 2024
- Co-founded an AI social platform that allows 100+ students to network, and generate study materials from notes