

THANG TRUONG

truongthoithang@gmail.com | (737) 341-1994 | [LinkedIn.com/in/truongthoithang](https://www.linkedin.com/in/truongthoithang) | truongthoithang.com

EDUCATION

Auburn University

Presidential Graduate Research Fellow | M.S. Student, Software Engineering & Computer Science

Expected May 2027

The University of Texas at Austin

Bachelor of Science in Informatics (GPA: 3.96/4.0)

Aug 2021 - May 2025

- **Certificates:** Elements of Computing (CS minor), Applied Statistical Modeling
- **Courses Taken:** Probability 1, Statistical Learning Interference, Discrete Math, Linear Algebra, Element of Database, Software Design, Data Engineering, Element of Data Science

EXPERIENCE

Natural Language Processing iSchool Research Group

NLP Developer | Python, Bash, SeamlessM4T

Austin, TX

May 2024 – May 2025

- Architected a multilingual NLP data pipeline translating 89.6GB of parallel corpora into 28 languages using SeamlessM4T
- Co-authored PolyverseVLM, a lightweight vision-language model optimized for GPUs using 80M SigLIP and 500M Qwen2
- Optimized batch parallelism for model fine-tuning, reducing training time by 40% on limited hardware.

Liberal Arts Instructional Technology Services

Junior Software Developer | Python, R, Laravel PHP, Node JS, Docker, Tableau, WordPress

Austin, TX

Aug 2022 – May 2025

- Developed 20+ full-stack web apps and dashboards applying modular codebases and REST API integration
- Migrated legacy Drupal systems to Laravel, implementing advanced filtering and user management features
- Revamped three WordPress websites with custom plugins and UI/UX upgrades, increasing user engagement by 40%
- Implemented PDF viewer for UT Instapoll, supporting 50,000+ users across UT Austin classrooms

AI Health Lab [\[website\]](#)

NLP Research Assistant | BERTopic, Python, Matplotlib, Google Colab

Austin, TX

Nov 2023 – March 2024

- Applied BERTopic with UMAP + HDBSCAN to uncover 20 thematic clusters in a corpus of 12,146 LLM paper abstracts
- Validated clustering integrity using 5 comparative techniques (bar chart, dendrogram, heatmap, topic graph, PCA plot)
- Conducted a survey of 12 foundational LLM papers, manually labeling to identify 3 previously unreported research themes

Claudius Legal Intelligence Inc.

Software Engineering Intern | Python, Django, Figma

Remote

June 2023 – Aug 2023

- Developed 4 production-ready web pages in Django, integrating interactive frontends based on Figma designs with Bootstrap
- Led UI/UX testing iterations with feedback from legal professionals, improving usability scores by 18% in internal testing

PROJECTS

Large Language Model Evaluation and Comparison | Python, PyTorch, NLP | [Link](#)

June 2024

- Evaluated and compared large language models (LLMs) like GPT-2 and BERT, focusing on accuracy and computation efficiency.
- Fine-tuned model achieved 43% faster evaluation runtime, 36% increase in evaluation loss and 163% increase in accuracy.
- Published the optimized model to Hugging Face Hub for open-source collaboration, supporting more efficient and accurate LLMs

RoboDog Therapy – HackTX Analog Devices Winner | Python, Llama 2, OpenCV, DeepFace | [Link](#)

Oct 2023

[HackTX Analog Devices First Place](#)

- Led the software development initiative to construct a simulated robotic embedded system therapy dog
- Implemented **OpenCV** for image processing, **Llama2** for text generation and **DeepFace** for emotion detection
- Engineered a resilient **HTTP server** for efficient data transfer of five input values, enabling real-time monitoring of the robotic dog

Personalized Internship Recommendations | Python, TF-IDF, Streamlit, NLTK, Selenium, BeautifulSoup | [Link](#)

May 2023

- Conducted web scraping to gather data from 1800+ job opportunities and used **GitHub CI/CD pipeline** to daily update data
- Applied **NLP** techniques, **TF-IDF** and **cosine similarity**, to align resumes with job descriptions with **80% accuracy rate**
- Connected **600+ users** with tailored internship opportunities, fostering career advancement and professional development

Technical Skills

Programming Languages: Python, Java, Swift, JavaScript, TypeScript, R, Go, Kotlin, SQL, Bash, HTML/CSS, PHP, Scala, Julia
Frameworks & Libraries: React.js, Node.js, Django, Flask, Express, Laravel, FastAPI, TensorFlow, PyTorch, Scikit-learn, Docker, GitHub, CI/CD, Firebase, AWS (EC2, S3, Lambda), GCP, Tableau, PowerBI, Excel, Vercel, Nelify, WordPress, Drupal, LaTeX
Databases & Others: MongoDB, PostgreSQL, MySQL, Big Query, Redis, Apache Spark, Hadoop, Tableau, Power BI, OpenCV