

# Ethan Lu

ethanlu0929@gmail.com | 973-997-3240 | Richmond Hill, GA | [ethanlu.netlify.app](https://ethanlu.netlify.app) |  elu139 |  ethan-lu

## EDUCATION

### Georgia Institute of Technology

Bachelor of Science in Computer Science

Atlanta, GA

Expected May 2027

- **GPA:** 4.00 | **Concentrations:** Information Internetworks and Intelligence
- **Relevant Coursework:** Data Structures and Advanced Algorithms, Systems and Networks, Computer Architecture, Database Systems, HCI Research Methods, Discrete Mathematics, Linear Algebra, Java OOP

## PROJECTS

### factr

Founder

Atlanta, GA

May 2025 – Present

- Reduced misinformation detection false negatives by ~20% over unimodal baselines by developing a multimodal AI pipeline cross-validating PyTorch CNN, OpenCLIP, and LLaVA image-text-vision models
- Deployed a scalable FastAPI service on AWS Lambda, integrating Docker for modular experimentation and preparing for future public API access
- Trained models on LIAR, DFDC, and other benchmark datasets to ensure cross-modal robustness for real-world misinformation scenarios

### The Climate Reality Project

Web Co-Developer

Sparta, NJ

Mar 2022 – Oct 2022

- Expanded access to 100+ climate education resources integrated into the curricula of 15 local schools and libraries upon live, in-person demonstrations, by developing a searchable HTML/CSS/JavaScript website
- Enabled real-time content publishing by integrating the TinyMCE rich text editor, allowing multiple teammates to upload and update educational material directly to the site

### Senior Home Sign-In Application

Full-Stack Co-Developer

Springfield, NJ

Dec 2021 – Dec 2022

- Improved accuracy and efficiency of daily resident and visitor log management at a local senior community, by developing a Python-based web application with an accessible Tkinter interface tailored for elderly users
- Implemented JSON-based data storage to enable efficient read/write operations and offline admin access

## EXPERIENCE

### Data-Driven Education, Vertically Integrated Project

Lead Machine Learning/Data Engineer

Atlanta, GA

Jan 2025 – Present

- Increased scalability of model training by generating a synthetic dataset with ~90% statistical similarity to the institution's course forum data, using OpenAI's o1 model and validating similarity through Latent-Dirichlet Allocation (LDA) topic modeling and personalized natural language processing tokenization
- Improved instructor insight by building a React dashboard with pyLDAvis visualizations for topic modeling, with planned integration of sentiment analysis
- Developing a BERT-based classifier to categorize forum posts into cognitive presence categories, increasing accuracy and interpretability of decision-making and discussion analytics

### MathWorks Math Modeling Challenge

Lead Data Analyst and Scientist

Philadelphia, PA

Mar 2024

- Achieved a predictive accuracy of 0.86 AUC in forecasting homelessness trends in U.S. cities by developing a binomial logistic growth model to analyze complex socioeconomic datasets using Pandas and MATLAB

## LEADERSHIP/COMMUNITY ENGAGEMENT

### Investment Management Club

Tech+Commodity Sector Researcher and Analyst

Atlanta, GA

Aug 2024 – Present

- Delivered interpretable 0–1 scaled growth potential and market impact scores for companies in a simulated \$100K investment portfolio, by designing a scoring framework using Gaussian Mixture Models and log-probability transformations
- Built data pipelines to collect and normalize real-time and historical financial metrics from the Financial-ModelingPrep API, enabling unsupervised sector-based performance analysis with Python and PyTorch

### Data Science @ Georgia Tech

Sports Analysis Project Data Scientist

Atlanta, GA

Aug 2024 – Dec 2024

- Determined that investing in fewer, high-impact rather than many, low-impact players correlated with higher league survival rates, based on analysis of ~500 Premier League transfer records (2008–2024) filtered from a dataset of 75K+ entries using BeautifulSoup and visualizing results in Seaborn and Matplotlib

## PROFESSIONAL SKILLS

- **Programming Languages:** Python, Java, JavaScript, TypeScript, SQL, C, C++, MATLAB, HTML, CSS
- **Frameworks, Libraries & Tools:** FastAPI, Django, React.js, Node.js, Pandas, NumPy, Seaborn, Tkinter
- **Databases & Platforms:** MySQL, PostgreSQL, AWS, Docker, CI/CD
- **AI & Machine Learning:** BERT, OpenCLIP, LDA, NLTK, LLaVA, Hugging Face, PyTorch