Eric E. Nunes

eric27n@me.com | (512) 350-5377 | https://linkedin.com/in/eric27n | https://eric27n.github.io

Education

University of Massachusetts Amherst *Amherst, MA*
Master of Science in Computer Science *May 2026*

- **GPA:** 3.82
- **Selected Coursework:** Algorithms in Data Science, Information Retrieval, Machine Learning, Natural Language Processing, Research Methods in Empirical Computer Science, Simulation Methods, Systems for Data Science

Texas A&M University *College Station, TX*
Bachelor of Science in Computer Science; Minors in Mathematics and Statistics *May 2024*

- **GPA:** 3.83, Magna Cum Laude
- **Awards:** Texas A&M CSCE Undergraduate Leadership Award (2024), Peer Teaching Excellence Award (2023), Thigbe Hardworking Endowed Scholarship (2022), Dr. Richard Volz Memorial Scholarship (2022)
- **Selected Coursework:** Artificial Intelligence, Bayesian Statistics, Computer Security, Data Analytics in Cybersecurity, Data Structures, Databases, Information Retrieval, Machine Learning, Software Engineering, Statistical Computing

Experience

Walmart Global Tech *Sunnyvale, CA*
Data Scientist III *Starting Jul 2026*

- Accepted full-time return offer following Summer 2025 Data Scientist internship.

Walmart Global Tech *Sunnyvale, CA*
Data Scientist Intern *Jun 2025 - Aug 2025*

- Built a scalable attribute value extraction data pipeline for Walmart product titles, improving missing Google Merchant Center attribute coverage resulting in a projected 3% increase in organic CTR and $10M incremental GMV.
- Fine-tuned and benchmarked transformer models for product attribute extraction, improving F1 over the production baseline model by up to 25% and achieving 84% F1-based performance.

Microsoft *Redmond, WA*
Software Engineering Intern *May 2023 - Aug 2024; May 2024 - Aug 2024*

- Engineered a workflow for Windows Autopilot deployments extracting summary statistics on hundreds of thousands of entries per week, saving an estimated 500 hours among on-call Microsoft engineers per year.
- Developed a pipeline for an updated setting recommendation system in collaboration with the machine learning team on Microsoft Intune, increasing scope to over 10k settings with 16% reduction in user error.

Texas A&M University Department of Computer Science and Engineering *College Station, TX*
Peer Teacher, Program Student Coordinator *Jan 2022 - May 2024*

- Tutored over 50 students in computer science courses per semester with a 95% approval rate.
- Uploaded weekly review sessions to YouTube, netting over 4,000 cumulative views in the first year and over 8,000 lifetime views, including increasing a course's review viewership by over 150% from a previous semester.
- Awarded the 2023 Peer Teaching Excellence Award, given to one peer teacher per year for outstanding achievement.

Projects

SciEncoder: A Tiny Scientific Domain Encoder [Python] *Oct 2025 - May 2026*

- Designed a compact 54M-parameter BERT-style scientific encoder pre-trained on ~1.5B tokens of multi-domain scientific text, creating a lightweight alternative with 50% fewer parameters than BERT Base.
- Led junior team members through their first NLP project, assigning data collection/preprocessing work with GitHub.

Identifying Clusters in NYT Connections [Python] *Dec 2024 - May 2025*

- Developed a multi-model inference pipeline combining NLP word embedding techniques, a web-based word search API, and a knowledge graph, with LLM-based semantic reasoning for answer validation and re-ranking.
- Achieved a 40% improvement in solve accuracy and 25% improvement in first-round accuracy over baseline LLMs (GPT-4o) by leveraging structured knowledge sources and LLM inferences.
- Won 1st place out of 30+ teams at the university's Machine Learning Club project showcase.

Texas A&M Album of the Week Website [JavaScript, React, PostgreSQL] *Dec 2022 - May 2024*

- Performed full-stack development on a website for a student organization to streamline and centralize activities for a club with over 125 active members per semester.
- Automated club tasks among members and offers with API functions and data pipelines, saving an estimated 90 hours among officers every year from manual tasks.

Skills

- **Programming Languages:** Python, R, SQL, C++, JavaScript, Scala, Java, HTML, CSS, C#, TypeScript
- **Python Libraries:** Scikit-Learn, Pandas, Matplotlib, NLTK, PyTorch, Transformers, TensorFlow, Gymnasium
- **Technologies:** Apache Spark, Postgres, React.js, Node.js, dplyr, tidyverse, Git, Jira, Docker, LaTeX
- **Languages:** English (Fluent, U.S. Citizen), Portuguese (Advanced), Spanish (Intermediate), French (Beginner)