Victor Verma

vpverm@bu.edu | (617) 838-4092 | victorverma.com | linkedin.com/in/victorverma | github.com/victorverma3

EDUCATION

Boston University   **Boston, Massachusetts**
B.A. in Mathematics and Computer Science, Minor in Data Science   *May 2025*
- **Honors/Activities:** 3.89 GPA, Dean's List, UROP Student Research Award, Kappa Theta Pi, Men's Rugby.
- **Relevant Courses:** Machine Learning and AI, Natural Language Processing, Algorithms, Database Systems, Distributed Systems, Cloud Computing, Probability, Linear Algebra, Stochastic Algorithms, Software Engineering.

SKILLS

Languages/Tools: Python, SQL, JavaScript/TypeScript, Java, C, Git, Docker.
Machine Learning: Scikit-learn, PyTorch, TensorFlow, LlamaIndex, Hugging Face, DSPy, OpenAI, LLMs, pandas, NumPy.
Cloud/Web: AWS/Azure/GCP, Firebase, React, Node.js, Flask, FastAPI, Gradio, PostgreSQL, MongoDB, Pinecone.

EXPERIENCE

Questrom School of Business, Boston University   **Boston, Massachusetts**
Research Assistant   *Jan 2023 - Present*
Large-Scale Mining and Classification of State Legislator Demographics
- Integrated the **Google Search API**, the **OpenAI API**, **pandas**, and **BeautifulSoup** to build an automated data processing pipeline to extract legislator biodata from **600,000+** web pages and PDFs.
- Engineered **neural network**, **random forest**, **XGBoost**, and **k-nearest neighbors** models achieving a **72% accuracy** rate and **0.80 F1** score in biodata classification.
- Created a **novel dataset** documenting the education and work history of **150,000** U.S. State Legislator candidates from 1967 to 2017, discovering only 40% of candidates have biodata available online.

Analysis of Medical Malpractice Trends with NLP
- Designed **DSPy** programs to summarize **70,000+** cases of medical malpractice and identify doctor violations.
- Enhanced program accuracy using DSPy optimizers, a custom evaluation metric, and 50+ curated examples.

Savvas Learning Company   **Boston, Massachusetts**
Software Development Engineering Intern   *Jun 2024 – Aug 2024*
- Architected a **ReAct agent** utilizing **Python**, **Amazon Bedrock**, and the **LlamaIndex API**, enabling **400+** employees to efficiently query enterprise documentation and the GitHub codebase.
- Incorporated the **Confluence and GitHub APIs** to automate the upload of **500+** Confluence documents and GitHub README files to **S3 buckets**, constructing a comprehensive knowledge base for the ReAct pipeline.
- Launched the ReAct agent as a chatbot service using **AWS Lambda**, featuring a **Gradio** frontend and **FastAPI** backend, and configured role-based access for employees through **Google OAuth**.

PROJECTS

Letterboxd Movie Recommendations
- Developed an AI-powered movie recommendation website using **content-based filtering** with **random forests** and **58,000+ data points**, serving **2,000+ users** across **50+ countries**.
- Built with **React**, **TypeScript**, **Tailwind CSS**, **Flask**, and **Supabase**, featuring **10+** real-time user stats, downloadable visualizations, and automated data updates using **3 GitHub Actions**.

Kappa Theta Pi Lambda Chapter Website
- Engineered the **RAG agent chatbot** on the fraternity website leveraging **HuggingFace** for inference, **Pinecone** for context retrieval, and **Google Cloud Functions** for scalable deployment.
- Led a **6-member team** to build the fraternity database with **React**, **Node.js**, **Supabase**, and **MongoDB** with **Google OAuth** using **Firebase**, hosting academic and professional resources for **111** fraternity members, and integrated **Dockerized shell scripts** to streamline data updates.

StatSense AI
- Designed **LSTM** and **RNN** models with **TensorFlow** to predict NFL statistics for QBs, RBs, WRs, and TEs using weekly time-series data, achieving a **75% success rate** against player props from sportsbooks.