Arav Mehta

📱 917-886-1035 ✉️ anmehta@umass.edu 💼 linkedin.com/in/aravmehta 💼 Portfolio �ongithub.com/arav16112004

Education

University of Massachusetts Amherst   **Expected Graduation - May 2027**

B.S. in Computer Science   *Amherst, MA*

- **Relevant Coursework:** Data Structures & Algorithms, Object Oriented Programming ,Computer Systems, Web Programming, Linear Algebra, Multivariable Calculus, Programming Methodologies, Reasoning Under Uncertainty, Discrete Mathematics.
- **Certifications:** CS50x (Harvard), Introduction to Programming Using JavaScript, JuniorMBA Technology with Infinity Cars.

Experience

Outamation   **May 2025 – Jul 2025**

AI and Automation Extern

- Processed **20+** documents using PyMuPDF and Tesseract OCR, reducing manual review time by approximately **80%.**
- Built and deployed a Retrieval-Augmented Generation (RAG) pipeline using LlamaIndex to enable NLP querying.
- Reduced manual document processing time by **80%** using OCR and Python-based automation.
- Enhanced retrieval accuracy by **40–60%** in RAG pipelines by implementing query rewriting, prompt engineering, semantic chunking techniques, vector indexing, hybrid and fusion retrieval, optimizing performance on complex mortgage documents.

University of Massachusetts Amherst   **Nov 2023 – May 2025**

Tech Assistant

- Diagnosed and repaired hardware/software issues for **40+** computers improving computer uptime by **50%**.

Goldman Sachs   **Dec 2024 – Jan 2025**

Software Engineering Virtual Experience Program - Forage

- Performed a job simulation as a governance analyst assessing IT security and proposing improvements.
- Identified use of outdated password hashing algorithms by cracking passwords with Hashcat.
- Authored a memo recommending security uplifts, including extending minimum password length and adopting stronger hashing algorithms, increasing password cracking resistance exponentially.

Kohinoor Electronics   **Dec 2021 – Jan 2022**

Software Engineering Intern — E-Commerce & Inventory Systems

- Automated backend workflows to update **1500+** product listings, reducing update time.
- Constructed a system with **10+** inventory tools to flag low-stock items and prevent fulfillment delays.
- Analyzed **4000+** sales datasets using Python to uncover trends that shaped marketing strategies.

Projects

LLM-Powered Document QA Chatbot | (Github repository)

- Developed a chatbot for Outamation using Google Gemini API that answers user questions over uploaded documents using a RAG pipeline, achieving a **40–60% improvement** in factual answer accuracy.
- Deployed semantic chunking algorithms within the pipeline, along with document classification for mortgage documents, decreasing irrelevant information injection by **45%** and boosting the accuracy of document retrieval by **23%**, dramatically improving efficiency.
- Implemented HuggingFace embeddings and ChromaDB to generate and store dense vector representations of document chunks, reducing data retrieval time by **50% (90% in larger documents)** for LLM-based question answering.

Assignment Organizer Web Application | (Github repository)

- Engineered a Flask-based assignment tracker using Python, HTML and CSS with SQL database integration, enabling students to digitally centralize **50+** assignments; resulting in reduce average completion time per assignment.

Technical Skills

Languages: Python, Java, C, SQL, TypeScript, HTML, CSS.
Tools/Frameworks: Flask, React Native, Node.js, LlamaIndex, Tesseract OCR, PyMuPDF, OpenCV, ChromaDb, HuggingFaceEmbedding, Pandas, NumPy, GitHub, MySQL, Canva, Data Extraction, RAG, API Integration