Arush Verma

linkedin.com/in/arushv04/| github.com/ArushVe| (858) 282-7026| averma343@gatech.edu

U.S. Citizen

Education

Georgia Institute of Technology | Atlanta, GA Expected Graduation **-** May 2025

Degree: BS in Computer Science – GPA: 3.88

Threads: Intelligence and Information/Info networks

Coursework: Data Structures & Algorithms, Design & Analysis of Algorithms, Machine Learning, Intro to Probability and Statistics

Experience

Software Engineer Intern - *Readyly* Summer 2024

- Developed a scalable content ingestion module capable of extracting and organizing information from various web sources into a neo4j instance with 92% accuracy
- Engineered a multi-modal Chat assist feature utilizing knowledge graphs and Retrieval Augmented Generation (RAG) and implementing a neural re-ranking system that improved response relevance by 37% and reduced query response time by 25%
- Integrated OpenAI models and implemented a unified API supporting both text and voice interactions enhancing user compatibility over across multiple platforms

Undergraduate Research Assistant – *Automated Algorithm Design* Fall 2023 - Spring 2024

- Utilized an in-house generative AI model (EMADE) to detect occurrences of 14 different chest diseases based on X-ray images
- Analyzed the results of the transfer learning model with an area under curve (AUC) of the Pareto Optimal points being 0.07
- Created a more balanced version of the dataset that was used in the original paper consisting of over 100,000 images

Data Science Intern - *John Deere Sustainability Solutions Team* Summer 2023

- Scaled a Python script on an EC2 server to read 50% more data points on sustainability from farms across 15 states
- Queried the data from the S3 bucket over the past five years and increased sustainability customers by 31%
- Leveraged satellite image data to predict tillage of farms without recorded data with a 76% accuracy
- Applied a linear regression model to estimate CO2 emission impact based on changing different data points in farms

Projects

Detecting Disease from Respiratory Audio Spring 2024

- Preprocessed the data with all the patient information and developed Mel spectrograms for each breathing cycle
- Conducted Principal Component Analysis to better illustrate the trends present in the data
- Developed the CNN model architecture which predicted the appropriate disease based on the spectrogram at 82% accuracy

Footprint Fall 2023

- Developed a custom algorithm to create a sustainability score by extracting and analyzing keywords from Amazon product pages through web scraping
- Designed an interactive website that features a comprehensive dashboard, enabling users to track their sustainability scores over time
- Integrated the OpenAI API to provide GPT-generated suggestions within the dashboard, offering potential improvements for user sustainability scores

Dungeon Crawler Fall 2023

- Constructed interactions between the player, environment, and enemies in Android Studio with an in house made game screen
- Led weekly meetings using Agile Development and extreme programming framework
- Charted Sequence Diagrams, System Sequence Diagrams, Design Class Diagrams to map out the project structure

Orbital Dynamics Spring *2022*

- Queried through 250 data points to train a linear regression model to determine position of the comet over time
- Established algorithms to determine the velocity and acceleration of the comet throughout its orbit
- Scaled this to a different comet and verified that the model can be applied to other comets given enough data points
- Mapped new orbit with different speed and position data using matplotlib

Skills

Languages: Java, Python, C, JavaScript, HTML, CSS, SQL, R, PySpark, Cypher, Swift

Technologies: RDBMS, AWS, Databricks, Docker, Jupyter Notebook, Neo4j, Git, FastAPI, GraphQL

Libraries: Pandas, NumPy, MatPlotLib, Tensorflow, Pinecone, BeautifulSoup, Playwright, Selenium