Sean McHale

seanryanmchale@gmail.com | (203) 312-4902 | github.com/mchales | linkedin.com/in/mchales/

EDUCATION

Georgia Institute of Technology, Atlanta, GA   **December 2024**
Master of Science in Computer Science | Machine Learning | GPA: 4.0

Vanderbilt University, Nashville, TN   **May 2023**
Bachelor of Science | Computer Science | Applied Mathematics | GPA: 3.92, *magnum cum laude*

WORK EXPERIENCE

Tacit Solutions, Founding Engineer and Lead Developer   **May 2023 – Present**
- Led the development of an AI-powered web application that uses LLMs, natural language processing, and vector databases for scalable one-on-one conversations, crowd-sourced knowledge exchange, and insight extraction.
- Engineered a full-stack solution with Django RESTful API, React, and PostgreSQL. Hosted on AWS and Heroku.
- Designed a flexible SQL database schema to accommodate new features; schema is updated with strategic database migrations. Created a Redis database to handle long and complex background tasks.
- Integrated feedback and suggestions with fast turnaround time from 5+ pilot customers, following Agile practices.

Vanderbilt HIDe Lab, Paid Research Assistant   **December 2021 – August 2023**
- Led a 2–3-person team for application development of a blockchain-based decentralized patient identity system.
- Onboarded new members to the lab, creating documentation and walkthroughs to streamline the process.
- Developed and authored a paper on an innovative patient identity framework using decentralized identifiers and the InterPlanetary File System, presented at the Int. Conference on Medical and Health Informatics 2023, Kyoto, Japan.

Hashed Health – Freelance Contractor, Full Stack Developer   **May 2022 – May 2023**
- Developed a full-stack platform to connect employees with students for micro-internships and contract work.
- Led frontend development using React and architected the backend with Spring Boot, integrating Spring Security, Hibernate, and a PostgreSQL database.

Duracell – Internship, R&D Data Analytics Intern   **May 2021 – August 2021**
- Leveraged web scraping and crawling to authenticate battery type data for thousands of products on Amazon. This data was utilized by Sales Analysts as input to sales forecast.
- Reviewed market research surveys for flow design, questions to reduce false responses, and general clarity.

LEADERSHIP

Vanderbilt Blockchain Club, Vice President and Co-Founder   **April 2021 – August 2023**
- Coordinated projects and guest speakers with companies such as Deloitte, Token Metrics, Blockchain Acceleration Foundation, and Reach. Led members to develop a blockchain-based Le Chat Noir application using Reach's platform.
- Outreached and communicated to grow the club to 200+ members with 30+ active members.

PROJECTS

Smart Summarization, LLM Text Summarizer   **August 2023 – December 2023**
- Developed length-insensitive summarization technique for LLMs. This approach enhances quality as measured by ROUGE scores and maintains high performance via parallel processing.
- Integrated two LangChain open-source methods to create a robust algorithm that balances the pros and cons of each.

Emergency Response Big Data Analytics   **February 2023 - May 2023**
- Executed Spark jobs on AWS EMR to merge a 26GB traffic dataset with traffic incidents from S3, processed geographic data with GeoPandas, and visualized slow-response roads using Plotly based on a weighted metric.
- Utilized BallTree machine learning algorithm for efficient proximity searches between traffic and weather data.

Named Entity Recognition Performance for LLMs   **February 2023 – April 2023**
- Utilized web scraping to gather 7,000+ forum discussion posts. Processed and cleaned with Pandas. Post-project, the professor used my dataset as the primary data in a paper accepted at the AMIA 2024 Informatics Summit.
- Analyzed performance of multiple LLMs using ChatGPT API on NER metrics and sentiment.

SKILLS & COMPETENCIES
- **Languages:** Python, C++, JavaScript, SQL, C, Rust, Java, Solidity
- **Technology:** Django, React, PostgreSQL, AWS, Spark, Pandas, PyTorch, LangChain, Hadoop, Pinecone