Pavan Subhash Chandrabose Nara

Proficient in **Python | SQL | AI/ML Research | Data & ETL Pipelines**

| npsc1601@gmail.com | (314)-319-8129 | GitHub: github.com/Subhashnara | LinkedIn: Pavan Subhash Chandrabose

EDUCATION

A results-driven Computer Science professional with a recent Master's degree (4.0 GPA) and multiple peer-reviewed publications in AI and Cybersecurity. Proven expertise in developing end-to-end data solutions, from building automated ETL pipelines in AWS to deploying machine learning models and fine-tuning Large Language Models (LLMs). Combines advanced skills in Python, SQL, and cloud technologies with a strong research background to translate complex data into actionable insights and robust software.

SKILLS

AI & Machine Learning: Generative AI & LLMs, Retrieval-Augmented Generation (RAG), Prompt Engineering, Natural Language Processing (NLP), Vectorization & Embedding, LLM Finetuning, Scikit-learn, PyTorch, TensorFlow, Keras

Data Analysis & Visualization: Statistical Analysis, Data Mining, Data Modeling, ETL Development, Data Warehousing, Reporting, Tableau, Power BI, Matplotlib, Seaborn

Programming & Databases: Python (Pandas, NumPy), SQL, NoSQL, Object-Oriented Programming (OOP), MySQL, PostgreSQL, SQLite, SQL Server

Cloud & Big Data: AWS (S3, Lambda, EC2, RDS, Athena, CodeBuild, CloudWatch), Snowflake, Docker, CI/CD, API Integration

Development & QA: Agile Methodologies, Version Control (Git/GitHub), Test Case Design, Functional & Regression Testing, Selenium, Postman, Jira

Professional Skills: Stakeholder Communication, Project Coordination, Process Optimization, Technical Writing, Mentoring, Problem-Solving

EXPERIENCE

Southeast Missouri State University **Missouri, USA**

Graduate Teaching & Research Assistant *01/2024 – 05/2025*

- Led the design of a specialized LLM Q&A system for cybersecurity threat analysis, implementing a novel framework using Python and the Google Gemini API that resulted in a peer-reviewed ACM conference publication.
- Designed and implemented robust data pipelines using Python (Pandas) and AWS services to collect, clean, and standardize large datasets, enhancing data quality and integrity for advanced analytical reporting.
- Developed comprehensive data visualizations and interactive dashboards using Power BI, Tableau, and Seaborn to communicate complex findings, leading to improved data-driven decision-making by faculty.
- Applied advanced statistical and data mining techniques to analyze complex datasets, identifying key trends in student performance that informed targeted academic interventions.
- Enhanced operational efficiency by developing Python automation scripts for routine administrative tasks, reducing manual processing time by over 50%.
- Mentored and instructed over 120 undergraduate students on topics in Data Analysis, Cybersecurity, and AI, achieving 90% positive feedback on teaching effectiveness.

PROJECTS

Prompt Tree: Enhancing guided interactions for Accessible Cybersecurity Training **Missouri, USA**

Master's Thesis *01/2024 – 12/2024*

- Developed a specialized LLM for cybersecurity training using Python, Google Gemini API, and a Retrieval-Augmented Generation (RAG) architecture to educate users on cybersecurity concepts.
- Introduced the "Prompt Tree" framework to enhance LLM interaction, drawing on credible sources like NVD, CVE, and OWASP to improve prompt clarity, relevance, and user satisfaction.

Local Healthcare Access and Wait Time Analyzer **Missouri, USA**

Data Analyst *07/2024 – 01/2025*

- Reduced manual data preparation time by 70% by engineering automated ETL pipelines in Python (Pandas, Geopandas) to process and standardize disparate public health datasets.
- Architected a unified SQLite database from multiple public sources (HRSA, Census API) to analyze over 3,000 healthcare facilities, identifying 150+ high-risk "healthcare desert" counties.
- Created interactive dashboards with Streamlit and Folium, including choropleth maps, to visualize geographic inequities in care and influence policy proposals.

Course Gradebook Analysis and Insights **Missouri, USA**
Data Analyst *05/2024 – 12/2024*

- Architected a secure, serverless data pipeline using AWS (S3, Lambda, RDS) to automate the ingestion and analysis of 700+ student grade entries, reducing manual processing time by 80%.
- Monitored pipeline health with Amazon CloudWatch and implemented CI/CD with AWS CodeBuild, achieving 99.9% reliability and reducing new feature deployment time by 50%.

EDUCATION

Southeast Missouri State University **Missouri, USA**
Master of Science in Computer Science *Graduated, 05/2025*

- **GPA**: 4.0/4.0
- **Relevant Coursework**: Data Analytics, Data Mining, Advanced Artificial Intelligence, Research Methods, Advanced Database Management Systems, Database Management, Data Architecture

Kakatiya University **Telangana, India**
Bachelor of Technology in Computer Science *Graduated, 06/2022*

- **GPA:** 8.39/10.00, *Dean's List*
- **Related Coursework:** Data Analytics, Data Science, Mathematics, Statistics & Applications, Database Management, Machine Learning, Design and Analysis of Algorithms

AWARDS AND PUBLICATIONS

[1] *Outstanding Computer Science Graduate Student*, 2025.

[2] Subhash Chandrabose Nara, P., Mitra, R., Roy, I., Robin Cole III, T. (2025). Accessible Cybersecurity Education Using Prompt Tree. In: Hu, G., Kambhampaty, K.K., Roy, I. (eds) Computer Applications in Industry and Engineering. CAINE 2024. Communications in Computer and Information Science, vol 2242. Springer, Cham. https://doi.org/10.1007/978-3-031-76273-4_9

[3] Kavya Nikhita Meda, Pavan Subhash Chandrabose Nara, Svoboda Bozenka, Tarek Zormati, Seth Turner, Wayne Worley, and Reshmi Mitra. 2025. Integrating Prompt Structures Using LLM Embeddings for Cybersecurity Threats. In Proceedings of the 2025 ACM Southeast Conference (ACMSE 2025). Association for Computing Machinery, New York, NY, USA, 180–187. https://doi.org/10.1145/3696673.3723069