Jahnavi Nalla

(480)-605-8270 | jahnaviwork04@gmail.com |LinkedIn | GitHub | Tempe, Arizona

SUMMARY

Highly skilled and results-driven graduate with experience in designing and deploying robust, scalable data pipelines and advanced AI solutions. Proven expertise in building high-volume ETL/ELT workflows using PySpark and Airflow on AWS and GCP. Proficient in developing and orchestrating LLM-based AI agents, implementing Retrieval-Augmented Generation (RAG) frameworks, and optimizing AI performance. Eager to leverage a strong foundation in data engineer, data governance, cloud architecture, and MLOps to drive data-driven innovation.

EDUCATION

Arizona State University, Tempe, AZ GPA: **3.75/4.00**

Master of Science, Information Technology *Dec 2025*

Relevant courses: Advanced Data Structures & Algorithms, Database Systems & Design, Distributed Systems, Object-Oriented Programming, Software Engineering Principles, Prompt engineering, AI into Cybersecurity, Cloud Security and Ops

SKILLS

- **Programming:** Python, Java, SQL, NoSQL, Object-Oriented Programming (OOP), Design Patterns, Microservices, API Development, RESTful APIs
- **Data Engineering:** ETL/ELT, Data Cleansing, Data Quality, Data Warehousing, Data Lake Architecture, PySpark, Apache Airflow, Databricks
- **AI/ML:** Large Language Models (LLMs), RAG, Prompt Engineering, Agentic AI Platforms, NLP, Responsible AI, Decision Intelligence
- **Databases & Cloud:** MySQL, PostgreSQL, MongoDB, Delta Lake, Docker, Kubernetes, AWS, GCP (based on your experience)
- **Tools:** Git/GitHub, Jira, Agile, Scrum, Jenkins, Postman, Figma.

EXPERIENCE

Neuralseek *08/2025- present*

AI agent Developer -Internship *Miami, Florida*

- Engineered multi-agent workflows and deployed solutions using NeuralSeek's mAIstro orchestration tool and AWS, implementing data ingestion, Retrieval-Augmented Generation (RAG), and source explainability features to enhance contextual reasoning.
- Developed AI agents with LLMs, utilizing fast engineering and no-code tools to create agentic solutions for enterprise decision support and automation.
- Implemented RAG and Context-Augmented Generation (CAG) frameworks in practice to build robust AI agents, demonstrating proficiency in enhancing contextual reasoning and optimizing data retrieval from knowledge bases.
- Assessed and refined LLM performance, evaluating key metrics for scalability, correctness, and dependability, and contributed to the mAIstro platform by documenting agent designs and technical specifications.
- Integrated human-in-the-loop systems and ethical AI frameworks, gaining experience in the AI + data lineage ecosystem to ensure compliant AI deployments.

Colruyt Group *05/2023- 08/2024*

Data Engineer *Hyderabad, India*

- Designed and deployed robust, high-volume ETL/ELT data pipelines from scratch using PySpark, Airflow, and Databricks across AWS and GCP cloud environments, resulting in a 30% reduction in data processing latency for critical supply chain analytics.
- Implemented advanced data cleansing, standardization, and data quality checks using expert Python and raw SQL techniques on complex retail datasets, which reduced data-related errors by 45% and significantly improved business intelligence reliability.
- Designed and governed a secure, high-performance Delta Lake architecture, enabling concurrent batch and streaming data loads and achieving a 35% optimization in cloud compute costs through cluster tuning and efficient data partitioning.
- Partnered with data scientists and business analysts to translate complex requirements into scalable data models and solutions, accelerating the time-to-market for new reporting and advanced analytics use cases by 20%.
- Maintained data pipeline reliability through proactive troubleshooting and issue resolution, while participating in code reviews and mentoring junior team members, collectively leading to a 25% decrease in P1 production incidents and faster mean time to recovery.

Arizona State University *08/2024- 12/2024*

Teaching Assistant: Advanced Database management systems *Tempe, Arizona*

- Provided personalized feedback to over 60 students on their projects, focusing on conceptual modeling, relational schema design, and adv. SQL implementation.
- Guided students through the integration of Relational and NoSQL databases, helping them troubleshoot complex issues in their final projects.
- Assessed student understanding of database security and integrity principles, ensuring their projects met industry standards for data protection and reliability.

ACADEMIC DEVELOPMENT PROJECTS

Local AI (Link)

- Developed a full-stack web application using Python and the Flask framework to enable real-time, dynamic chat interactions with a locally hosted AI model.
- Engineered a real-time data streaming solution by implementing Server-Sent Events (SSE) on the backend to dynamically update the user interface with AI-generated responses as they were created.
- Designed and implemented a clean user interface using HTML and CSS, ensuring a seamless and intuitive user experience for interacting with the AI model.

Government scheme Bot (Link)

- Developed a conversational AI chatbot using Python and Streamlit, leveraging Google's Gemini Pro model to accurately interpret user queries and provide relevant information about government schemes.
- Engineered a vector-based search system by implementing ChromaDB with HNSW to perform efficient similarity searches, enabling the chatbot to quickly retrieve and process information from government websites.
- Built a user-centric application that streamlines access to complex government information, incorporating features for eligibility checking and natural language understanding to enhance the user experience.

ACHIEVEMENTS & ADDITIONAL INFORMATION

- Volunteered for Foster your Future in maintaining their website
- Received a 4.8/5 rating in end-of-semester feedback as a Teaching Assistant
- **Campus Partner for Perplexity**