Venkata Pavithra Chaganti

Phone: 856-244-4197| Email:chaganva@mail.uc.edu

PROFESSIONAL SUMMARY

Information Technology graduate from the University of Cincinnati with hands-on project experience in data engineering and machine learning. Skilled in designing, implementing, and optimizing scalable data pipelines, ensuring data security compliance, and building efficient data models for analytics. Passionate about leveraging data to drive business insights and innovative solutions. Actively seeking full-time Data Engineer opportunities, available to start immediately.

EDUCATION

Master of Science in Information Technology **August 2023-December 2024**

University Of Cincinnati, Cincinnati, Ohio

Bachelor of Technology in Information Technology

Bapatla Engineering College, Bapatla, India

TECHNICAL SKILLS

Programming Languages: Python, Java, C, Linux.

Data Engineering Tools: Azure Data Factory, Databricks, Azure Data Lake, Azure Synapse Analytics, Azure Key Vault, Azure Blob Storage, ADLS, Snowflake Cloud, Data Pipelines, Spark SQL, Delta Lake, Data Dictionary, Data Modeling, PowerBI, Apache Kafka, PySpark.

Web Development&Design: HTML, CSS, JavaScript, React.js, Java Servlets, Support Analysts.

CI\CD: GitHub, Jenkins (basic familiarity).

Software&OS: Linux, Microsoft Office (Word, Excel, PowerPoint, Access) & Windows (All versions).

Database&Networking: MySQL& Knowledgeable with computer networking concepts and functionality.

PROFESSIONAL EXPERIENCE

Tata Consultancy Services (TCS): Assistant System Engineer and Trainee: **July 2022-July 2023**

- Led comprehensive data management tasks, ensuring dataset integrity across multiple projects.
- Proficient in HTML, CSS, JavaScript, React.js, and Node.js. Conducted code reviews, ensured quality assurance, and maintained coding standards.
- Implemented agile methodologies via Jira, accelerating projects by 25% and reducing post-launch bug fixes by 15%.
- Worked on system engineering tasks, focusing on seamless project execution and delivery.

CIOS Digital: Data Engineer Intern: **July 2021-June 2022**

Description:

Developed dashboards for MBA INSITE Product teams across all markets where Shell owns/operates fuel sales data. The business uses a data-driven approach to manage fuel sales across markets, leveraging Power BI for insights.

Responsibilities:

- Created fully parameterized pipelines to ingest data from SAP BW to Azure Gen2 Data Lake.
- Participated in business requirements discussions, backlog refinement, sprint planning, and review meetings.
- Split source data by country and organize it into year/month folders in ADLS Gen2.
- Developed Azure Databricks notebooks for data transformation and ingestion into Azure SQL Database.
- Created Delta tables in Azure Gen2 using Databricks. Wrote SQL queries using Spark SQL to process data.
- Validated data in higher environments and obtained project team sign-off before production deployment.
- Deployed and scheduled pipelines using GitHub. Involved in both development and production environments.

Environment: Spark, Spark SQL, SQL Server, Azure Databricks, ADLS, Python

Exposes Data Labs: Software Developer Intern **Jan 2021-May 2021**

- Leveraged cybersecurity best practices, such as Python, HTML, CSS, and JavaScript, to develop MSG Secure, ensuring robust encryption and protection against cyber threats.
- Reduced processing duration by 20% and increased user engagement by 25% by creating interactive web interfaces.
- Led a team in developing MSG Secure, showcasing leadership in cybersecurity and technical proficiency.

ACADEMIC PROJECTS

Student Performance Prediction System

- Collected and integrated multiple data sources (academic and personal) to build a predictive model for student performance.
- Built a web-based interface using HTML, CSS, Bootstrap, JavaScript, and React.js for seamless user interaction.
- Selected machine learning algorithms in Python, achieving 80% accuracy in predicting academic outcomes.
- Analyzed prediction results and provided actionable insights to enhance academic performance and educational outcomes.

Real-time Traffic Data Processing System

- I have built a data pipeline for real-time traffic data ingestion, processing, and analysis.
- Streamed data using Apache Kafka and processed it with Apache Spark and Azure Databricks.
- Stored data in Azure Data Lake and visualized insights with Power BI.
- Applied machine learning to predict traffic patterns and suggest optimal routes.

Customer Sentiment Analysis System

- Collected and preprocessed customer feedback data from various sources to build a sentiment analysis model.
- Developed a user-friendly dashboard using Flask and JavaScript to display real-time sentiment analysis results.
- Applied natural language processing (NLP) techniques in Python, achieving 85% accuracy in sentiment classification.
- Analyzed sentiment trends to provide insights for customer satisfaction improvements and marketing strategies.

Technologies: Apache Kafka, Apache Spark, Azure Databricks, Azure Data Lake, Power BI, Python.

CERTIFICATIONS

- Microsoft AWS Cloud Practitioner, Python, Linux, Machine Learning.