

Sai Shashank Paidipelly

Miami, FL | shashanksai0708@gmail.com | (786) 893-2390 |

LinkedIn: [Sai Shashank Paidipelly](#)

Education

Master of Science, Computer Science (GPA: 3.8/4.0)

Florida International University

Jan 2023 – May 2024

Miami, United States

Bachelors in technology, Information Technology (GPA: 3.1/4.0)

Keshav Memorial Institute of Technology

Jul 2017 - Aug 2021

Hyderabad, India

Professional Experience

Data Science Intern, Smart Infratech, Wellington, Florida

Nov 2024 – Present

- Developed a Domain-specific Retrieval QA chatbot with Langchain, FAISS, OpenAI Embeddings to facilitate accurate and context-aware responses to user queries, streamline customer support operations, and reduce response times by 60%.
- Developed a Flask application for automated PDF data extraction and analysis by training a custom extraction model in Azure Document Intelligence. Engineered backend workflows to handle and process file uploads via Azure Cognitive Services API.
- Designed a dynamic frontend using JavaScript to render real-time PDF previews and structured JSON outputs with collapsible data views, facilitating the conversion of unstructured data into a structured JSON format and optimizing processing for seamless user interaction.

Research Assistant, ABC-UTC(FIU), Miami, Florida

April 2023 – April 2024

- Collaborated with cross-functional teams, including web developers, to develop and optimize the dynamic search functionality for the ABC-UTC WordPress website using React, WordPress REST API, and Relevanssi plugin, implementing real-time results, relevance-based backend processing, and a custom JavaScript front-end with advanced filters and sorting capabilities for an enhanced user experience.
- Performed comprehensive data cleaning on ABC Research database by removing NULL values, eliminating duplicate records, ensuring consistent data types, and standardizing column names to maintain data integrity and enhance the accuracy of research database queries and reporting.
- Organized and structured data for each UHPC mix, including material ratios and test results recorded across various curing periods. Ensured the data was properly formatted and stored systematically to enable efficient analysis and reporting of performance trends.
- Utilized NLTK to analyze and enhance the semantic relevance of keywords by integrating it into the backend to provide users with more precise, context-based keyword suggestions, resulting in improving overall user experience.

Data Analyst, Nalanda College, Telangana

October 2021 – December 2022

- Automated the process of importing student admission data from Excel sheets into the college database by leveraging Pandas and SQLAlchemy for data cleaning and transformation, ensuring efficiency and accuracy in data management.
- Performed data preprocessing and validation to prepare datasets for scatter plot visualizations, enabling effective interpretation of negative or weak correlations and improving the accuracy of insights.
- Conducted department-wise analysis of attendance data using SQL GROUP BY and aggregate functions, reducing reporting workload by 50% and enhancing efficiency in presenting key metrics.
- Implemented SQL CASE statements and conditional expressions to classify attendance patterns, flag low-attendance students, and facilitate timely interventions, driving actionable insights through trend identification and visualization.
- Performed predictive modeling to analyze alumni employment rates, leveraging GPA and subject proficiency as key predictors to identify trends and factors influencing career outcomes, supporting data-driven decision-making.

Projects

Livestock Portfolio (PostgreSQL, Excel)

Built a comprehensive stock portfolio by analyzing livestock data and comparing bought prices with current market prices, utilizing PostgreSQL to query and manage large datasets efficiently. Implemented dynamic tables and stock history formulas to track past trends and facilitate informed investment decisions with real-time data updates.

Retail Sales Performance and Trend Analysis (Python, SQL)

Developed a retail data analysis pipeline to clean, transform, and load data into SQL Server for efficient querying. Utilized Python for preprocessing and SQL to derive actionable insights, including top-performing products, regional sales trends, year-over-year growth comparisons, and sub-category profit growth. Enabled data-driven decision-making by providing insights into sales performance and profitability.

Technical Skills

Programming: Python, Pandas, NumPy, Matplotlib, LangChain. **Database:** PostgreSQL, MySQL.

Data Analysis & Visualization: Power BI, Excel, PowerPoint. **Data Management:** WordPress, ETL processing.