Surya Pavan Peruri

suryapavan7056@gmail.com | +1 (940) 843-9415 | USA | GitHub | LinkedIn | Portfolio

Education

Master of Science, *University of Houston* Jan 2023 – Dec 2024 | Houston, USA
Engineering Data Science

Bachelor of Technology, *Vellore Institute of Technology* July 2017 – June 2021 | Vellore, India
Computer Science and Engineering

Technical skills

- **Programming Languages:** Python, SQL, R, NoSQL, Shell, HTML, CSS, Java, Shell Scripting, VBA
- **Frameworks & Tools:** NumPy, Pandas, Matplotlib, Git, GitHub, Docker, CI/CD, Statistics, SAS, PySpark, Seaborn, KPI, ggplot2, AWS Cloud, Google Cloud Analytics, Microsoft Azure, Apache Kafka, Tensorflow
- **Databases:** MS-SQL, MySQL, Snowflake, Big Query, PostgreSQL, MongoDB, Spark SQL, Oracle
- **Data Analysis / ETL:** Power BI, Tableau, OpenCV, Orange, Data Bricks, MS Excel, Google Analytics, Machine Learning, Data Modeling, Data Mapping, Data Mining, Data Extraction, Transformation, Azure DevOps, ETL, Postman, Google Colab, Docker

Professional Experience

Data Analyst Intern, *NSF – Data Science for Energy Transition* May 2024 – Jun 2024 | USA

- Analyzed large-scale energy transition datasets using Python Pandas, NumPy and SQL, developing predictive models to optimize renewable energy efficiency and improve forecasting accuracy by 20%.
- Built interactive dashboards in Tableau and Power BI, streamlining energy usage monitoring and reducing reporting time by 30% through automated data workflows.
- Automated data preprocessing workflows using Python and SQL, reducing manual data cleaning time by 40% and ensuring high data accuracy

Research Assistant, *University of Houston* Jan 2024 – May 2024 | USA

- Conducted in-depth data analysis using Python and R, resulting improvement in the accuracy of predictive models for ongoing research projects in data analytics.
- Developed automated data cleaning processes leveraging Pandas and NumPy, reducing manual processing time by 40%, which significantly boosted research productivity.
- Partnered with research teams to analyze findings and contribute to peer-reviewed publications, strengthening the project's impact and academic recognition

Data Analyst, *Infosys Pvt. Ltd* Jul 2021 – Dec 2022 | India

- Developed an InfoArchive-based application using OpenText products, Core Java, and Python, ensuring compliant archiving and enhancing accessibility of structured and unstructured data.
- Automated project prioritization workflows by creating a data-driven tool in Microsoft Excel and SQL, reducing manual effort by 40% and improving operational efficiency.
- Designed and optimized SQL-based data pipelines and interactive Power BI dashboards, delivering real-time business insights and enabling a 20% reduction in decision-making time for stakeholders.
- Integrated REST APIs to enhance backend data accessibility for web applications, improving response time and optimizing data retrieval processes.
- Conducted data quality analysis and implemented data validation techniques, improving accuracy and consistency across reporting and analytics systems.

Projects

IMDb Top 1000 Movies Dataset **Aug 2023 – Dec 2023**

- Analyzed IMDb Top 1000 Movies dataset using Tableau to identify key trends and patterns, resulting in improved insights on box office performance across themes.
- Established interactive dashboards that visualized movie ratings, release years, and themes, increasing user engagement by 35%.

Seoul Bike Sharing Prediction **Jan 2023 – May 2023**

- Designed and implemented machine learning models, achieving improvement in bike count predictions by employing advanced algorithms and optimizing model parameters.
- Enhanced model performance by conducting hyperparameter optimization through grid search and random search, resulting in a 20% reduction in prediction errors.

Facial Reconstruction System Using Image Morphing Techniques **Jan 2023 – May 2023**

- Developed a facial reconstruction system using Python, OpenCV, and Dlib libraries, applying advanced image morphing techniques to restore incomplete or damaged images.
- Built a Flask-based backend and a responsive web interface, enabling users to upload images, process reconstructions in real time, and visualize transformations interactively.

Certifications & Publications

- Tableau Desktop Specialist Certification – Tableau, 2025
- Microsoft Certified : Azure Fundamentals
- Distributed and Parallelized Image Encryption Using DES Techniques, IRJMETS, June 2020 (Link)