Madhumitha Saravanan

+ 1 315-561-5150 | msaravan@syr.edu | https://www.linkedin.com/in/mithasaravanan7/ | https://github.com/madhusaravanan27

EDUCATION:

Syracuse University, School of Information Studies, Syracuse, NY **Dec 2024**
M.S. Applied Data Science
Relevant Coursework: Introduction to Data Science | Applied Machine Learning | Big Data Analytics | Quantitative Reasoning for Data Science | Cloud Management | Business Analytics | Natural Language Processing | Data warehouse | Advanced Big Data Management

Anna University, Meenakshi Sundararajan Engineering College, Chennai, India **Apr 2020**
BE Electronics and Communication Engineering
Relevant Coursework: Computer Programming | Object-Oriented Programming and Data Structures

SKILLS:

Programming Languages: Python, R, SQL, Java, Scala, MATLAB
Software/Tools: Microsoft Excel, Tableau, Power BI, Git, Docker, Google Analytics, JIRA, Confluence
Libraries: Pandas, Scikit-Learn, TensorFlow, NumPy, Keras, Stats Models, Matplotlib
ETL/Big Data Technologies: Apache Spark, Hadoop, Apache Airflow
Databases: MSSQL, MySQL, PostgreSQL, MongoDB, Cassandra, Elasticsearch, Kafka, Minio, Neo4j
Cloud Platforms & DevOps: AWS, Azure, Jenkins, Databricks

WORK EXPERIENCE:

Data Engineer – Infosys Ltd **Apr 2021 – Dec 2022**
- Led the design and implementation of an orchestration layer (REST APIs) using Java, JSON, and Apache Camel for three microservices, ensuring seamless integration and supporting cross-channel performance.
- Performed API testing across multiple environments using Postman, reducing code defects by validating JSON responses and enhancing data accuracy and integration, aligned with client business and marketing goals.
- Built and managed CI/CD pipelines using Jenkins, AWS EC2, and Kubernetes, streamlining deployment processes and delivering efficient, scalable, and automated workflows.
- Delivered client-focused solutions in an agile environment, adapting to evolving business needs, fostering strong client relationships, and ensuring high-quality results aligned with complex project requirements.
- Mentored team members on best practices and new technologies, fostering knowledge sharing, collaboration, and continuous improvement.

PROJECTS:

Zillow House Value Trend Analysis **Feb 2024 - Apr 2024**
- Analyzed Zillow pre-sales data (2018-2024) using statistical models and Matplotlib, improving trend forecasting accuracy by 30% through ARIMA and KNN methods.
- Developed and deployed a CNN-based time series model using Keras and TensorFlow, achieving 99.09% prediction accuracy on test data, which optimized forecast reliability and supported scalable ML solutions.
- Visualized Zillow house value trends using advanced EDA techniques, facilitating the identification of key factors influencing regional price fluctuations by 15%.

A Visual Exploration of Natural Disasters and Economic Impact in the USA **Feb 2024 - Apr 2024**
- Supported data-driven decision-making for disaster preparedness and recovery by creating a consolidated Tableau dashboard with interconnected views of economic and human impacts, improving data quality and reinforcing user-centered insights.
- Constructed visual narratives to optimize data storytelling, enabling stakeholders to better interpret trends and insights related to disaster impacts.

Crime Classification Based on Geographical Area **Oct 2023 - Dec 2023**
- Translated business objectives into actionable analyses by developing a crime classification system, increasing prediction accuracy to 72% using Random Forest, XGBoost, and Logistic Regression.
- Designed interactive heatmaps and applied KMeans clustering, enhancing crime pattern analysis by 20% and delivering clear data communication to stakeholders, thereby supporting data-driven decision-making.
- Optimized crime impact analysis by engineering features such as Crime Severity and Victim Demographics, facilitating stakeholder engagement and presenting actionable insights through user-centered visualizations.

VBAY Data Warehouse and ETL Optimization Project **Oct 2023 – Dec 2023**
- Designed a Snowflake-based data warehouse with DBT, applying data architecture principles to improve compliance and data governance for VBAY analytics, in line with enterprise data management best practices.
- Architected dimensional data models for key domains: Ratings & Review, Sales, Location, and Bidding, optimizing data modeling and cataloging to support product analytics and market insights.
- Increased query performance by 25% through SQL optimization techniques, ensuring efficient data retrieval and validation for business and strategic reporting.