SIDDHANTH VEERANNA ARKASALI

281-236-0192 | siddhanthva@gmail.com | linkedin.com/in/siddhanth-arkasali/

OBJECTIVE

Data Science and AI Engineering Graduate seeking a **Summer 2026 internship**, leveraging **Python**, **machine learning**, and scalable **AI platforms** to build data-driven solutions that improve product performance, and decision-making

EDUCATION

University of Houston, *MS in Engineering Data Science and AI, GPA: 3.23/4*   **Aug 2025 - Jul 2027**
Coursework: Data Science, Machine Learning, Probability and Statistics, Artificial Neural Networks, AI for Engineers

Amrita Vishwa Vidyapeetham, *Integrated Masters in Computer Applications, GPA: 3/4*   **Jul 2019 - Aug 2024**
Coursework: Database Management System, DSA, Software Engineering, Data Mining, Web services and Cloud Computing

TECHNICAL SKILLS

Programming: Python (NumPy, Pandas, Scikit-Learn), R, SQL
Data Science: Regression, Classification, Clustering, Random Forest, XGBoost, PCA, Feature Engineering, Predictive Modeling, Model Validation, Time Series Analysis
Tools: Excel, Power BI, Git, Jupyter Notebook, MySQL, n8n Workflow Automation, Databricks, Generative AI (Prompt Engineering)
Business Analytics: Hypothesis Testing, KPI Analysis, Cost Optimization, Financial and Operational Analytics, Forecast vs Actuals
Soft Skills: Data Storytelling, Problem Decomposition, Stakeholder Engagement, Decision Support

PROJECTS

The Anatomy of Livable Neighborhood   **Aug 2025 - Dec 2025**
- Developed an explainable **multivariate regression** model to quantify urban livability, analyzing how walkability and housing features impact affordability and neighborhood desirability across diverse residential areas
- Achieved R^2 = **0.806** (Adj. R^2 = **0.803**) with ~**58%** lower MAE vs. baseline on **2,080** housing records, enabling data-driven insights for urban planning and real-estate valuation decisions

Multitudinous Transfer Learning for Paddy Yield Prediction in Smart Agriculture | *IEEE Xplore*   **Jan 2024 - Jun 2024**
- Deployed a transfer learning–based time-series forecasting model for crop yield prediction using climate, groundwater, soil, and emissions data, achieving R^2 = **0.91** and outperforming Bi-LSTM baselines across multiple evaluation metrics
- Enabled early, data-driven yield forecasts to reduce uncertainty, improve planning accuracy, and support resource optimization in large-scale smart agriculture systems

Student Segmentation for Performance Analysis   **May 2023 - Jul 2023**
- Built an end-to-end unsupervised ML pipeline using K-Means and K-Medoids to segment student performance data, enabling early identification of at-risk groups without labeled outcomes and supporting proactive academic decision-making
- Enhanced clustering robustness by ~**5%** through stronger outlier resistance and cluster stability evaluation, leveraging distance-based metrics and validation techniques to deliver scalable, data-driven insights for academic intervention

Mountain Road Blind-Curve Alert System| *TIJER Research Journal*   **Jul 2021 - Dec 2021**
- Designed and engineered a real-time **IoT**-based blind-curve safety system integrating sensors, microcontroller logic, and cloud-backed alerting to proactively **mitigate collision risk** through real-time detection in mountain roads
- Implemented an end-to-end workflow with real-time alerts and cloud logging, testing demonstrated reliable detection and rapid response, enabling continuous monitoring and effective post-event safety analysis

WORK EXPERIENCE

Data Validation Intern | *JK Files and Engineering Limited | Bangalore, India*   **Dec 2024 - Jan 2025**
- Validated, audited, and standardized datasets across manufacturing, operations, and HR domains using SQL-based data quality checks, achieving **94%** data accuracy across audited employee and organizational records
- Implemented systematic data validation techniques (consistency, completeness, anomaly detection) to reduce data errors and improve analytics-ready data reliability for reporting and downstream analysis
- Streamlined and automated data verification workflows, reducing turnaround time by **15%** and enabling reliable dashboards, KPI tracking, and modeling-ready datasets for operational decision support

LEADERSHIP AND COMMUNITY SERVICE

Event Coordinator | *Amrita Graduation Ceremony Committee*   **Sep 2023 - Aug 2023**
- Coordinated hospitality operations for **100+** attendees, managing guest support, logistics, and on-site coordination

Organizing Committee Member | *Amrita CODATHON Competition*   **Apr 2023 - May 2023**
- Collaborated with faculty and peers to manage timelines and curate problem statements for a multi-team coding competition

Group Manager | *Inaugural Ceremony of Amrita Hospital*   **Jul 2022 - Aug 2022**
- Managed on-ground operations for a large-scale inaugural event, coordinating dignitaries and logistics for **500**+ attendees