

Naga Venkata Sai Ruthvik Kasi

DATA ANALYST – Data Visualization, Predictive Analysis & Project Management

✉ sairkasi792@gmail.com ☎ +1 (226) 935-7752 📍 [Canada](#) **in** [LinkedIn](#) 🌐 [GitHub](#) 📁 [Portfolio](#)

SKILLS

- **Programming Languages:** Python, R, SQL, Java, JavaScript for scripting, analytics, automation workflows.
- **Databases and Querying:** MySQL, PostgreSQL, MongoDB, Hive, Snowflake, Neo4j, and Spark SQL queries.
- **Business Analysis:** Market analysis, KPIs, risk assessment, ROI, pilot programming, Agile, traceability matrix.
- **Data Platforms and Tools:** ETL, Azure, SAS, Power BI, Tableau, Jira, Visio, Advanced Excel, ML models.
- **Project Management:** Timeline management, backlog control, budget analysis, stakeholder and QA reporting.

WORK EXPERIENCE

Data Analyst Intern

Sharaf DG Energy

February 2024 – August 2024

UAE

- Processed 50,000+ energy records using pandas, SQL & Excel, applying data cleaning techniques to reduce raw data errors by 35% & decrease preparation time by 40% for weekly performance & yield reports across regions.
- Developed over 12 automated ETL pipelines for data engineering to process sensor data from 200+ solar panels, decreasing manual data handling by 85% and enabling 24/7 live updates for energy monitoring dashboards.
- Designed 30+ Power BI dashboards for data visualization, used weekly by 6 departments, reducing report generation time by 60% and increasing insight visibility on panel efficiency, downtime, and predicted performance.
- Flagged 40+ system fault patterns with O&M through anomaly detection techniques, reducing downtime by 18% and earning recognition from the CFO for the impact of real-time diagnostics on performance metrics.

Research Intern

National University of Singapore

July 2023 – November 2023

Singapore

- Developed ML pipelines on 100,000+ image-caption pairs using tools & Transformers, leveraging transfer learning to achieve a 20% boost in output accuracy and contribute to a cross-team AI research collaboration.
- Created a CNN-based genre classifier trained on 5,000+ audio wave forms, reaching 91% test accuracy and enabling integration of audio features into a novel multimodal storytelling framework via supervised learning.
- Deployed 4 ML models on AWS EC2 and S3 with pipeline orchestration, using containerization and CI/CD to cut setup time by 60% and increase team deployment efficiency across iterations in the research lifecycle.
- Co-authored a research paper on multimodal AI storytelling, integrating 3 model types across 2 modalities and completing the project ahead of schedule with 100% productivity using cloud computing infrastructure.

Data Analyst Intern

SmartInternz - [GitHub Link](#)

May 2023 – July 2023

Hyderabad, India

- Engineered a full-stack analytics platform for NYC bike data using React.js and Flask, enabling live query processing, API routing, and client-side interactivity for multi-dimensional data exploration across 100K+ records.
- Analyzed trip logs using Python & Tableau to extract hourly usage trends, station-wise output, and demographic segmentation, improving visual precision by 38% via data preprocessing, filtering layers, and aggregate metrics.
- Integrated Tableau dashboards into the app interface for rendering, applying live data filtering and backend triggers to automate data refresh cycles, improving analytical response time by 45% during multi-slice queries.
- Deployed front-end components for station-wise analysis, connecting to Flask endpoints for asynchronous rendering and map-based metrics display, advancing mobility trend inputs by 40% using geographic data alignment.

PROJECT EXPERIENCE

Genre Classification System

Python Developer

[GitHub Link](#)

- Engineered a CNN using Mel spectrograms from 1,000+ audio files (GTZAN) via Librosa; achieved 92% accuracy in 10 genres by optimizing hyper parameters and dropout layers using TensorFlow for robust genre classification.
- Processed waveform audio data to generate normalized frequency-time features with Librosa; optimized signal input for CNN performance by experimenting with 5+ augmentation tactics including noise injection and shifting.
- Integrated Flask interface for real-time audio genre prediction with CNN, ensuring less than 2s latency on inputs.

Heart Disease Prediction System

Machine Learning Developer

[GitHub Link](#)

- Implemented Logistic Regression, KNN, and Random Forest models on 900+ patient records; maximized 89% model precision, optimizing recall and F1 score through k-fold validation and hyperparameter grid search.

- Leveraged feature engineering to 14 clinical indicators, including ECG data, cholesterol levels, and resting blood pressure using pandas; coordinated missing values and scaling through one-hot encoding and MinMaxScaler.
- Created 10+ ROC curves, heatmaps and matrices using seaborn to assess model performance and efficiency.

COVID-19 Analysis and Visualization

[GitHub Link](#)

Data Analyst

- Analyzed and normalized daily COVID-19 data from 190+ countries using pandas; applied time-series indexing, per system scaling & schema unification to ensure accurate longitudinal analysis across zones & reporting norms.
- Leveraged EDA using seaborn tools on 50K+ records to visualize surges, case fatality ratios & country wise spikes; extracted insights from moving averages and cumulative trends to reveal global outbreak patterns.
- Enabled live tracking via API-fed dashboards with less than 3s latency for filtering COVID rates by date, country.

EDUCATION

Masters in Applied Computing (AI Stream)

September 2024 – Present

University of Windsor, Windsor

Bachelor of Technology in Computer Science Engineering

September 2020 – August 2024

Vellore Institute of Technology, India

CERTIFICATIONS

- | | |
|---|-----------------------|
| • Microsoft Azure Certification | July, 2025 [Expected] |
| • Data Analytics Job Simulation – Deloitte | July, 2025 |
| • Requirements Elicitation for Business Analysis – IIBA | June, 2025 |
| • BA Business Process Modeling – IIBA | June, 2025 |
| • Business Benefits Realization Foundations – PMI | May, 2025 |
| • Agile Requirements Foundations – IIBA | May, 2025 |
| • Business Analysis Foundations – IIBA | May, 2025 |
| • Predictive Analysis and Data Mining – IIBA | November, 2024 |
| • Data Science 360 – Edoxi | August, 2024 |
| • AWS Certified: Data Analytics with Big Data – AWS | July, 2023 |
| • Data Analytics with Deep Learning – National University of Singapore (NUS) | July, 2023 |
| • Microsoft Certified: Security, Compliance and Identity Fundamentals – Microsoft | July, 2023 |

LEADERSHIP & VOLUNTEERING

- **Chairperson** – Uddeshya NGO
- **Design Head** – Uddeshya NGO
- **Volunteer** – Arduino Coding Event
- **Core Committee Member** – TEDx VIT
- **Core Committee Member** – RoboVITics
- **Core Committee Member** – Youth Red Cross