Aditya Raj Sinha

adityaharsh03@gmail.com | +91 9279314598 | Aditya/Portfolio | linkedin.com/in/Aditya | github.com/Aditya

PROFESSIONAL

Data analyst skilled in transforming raw data into actionable insights to support business decision-making. Hands-on experience in SQL, Python (Pandas, NumPy), Tableau, Power BI, statistical analysis, exploratory data analysis, and KPI reporting through real-world projects. Built interactive dashboards, performed sentiment analysis, forecasting, and financial comparison analytics. Strong ability to analyze large datasets, identify patterns, and clearly communicate insights to stakeholders.

EDUCATION

- **M.Sc. Data Science—CGPA: 8.13**

Vellore Institute of Technology–AP | 2024-2026

- **B.Sc. Physics —CGPA: 8.05**

Ranchi University, Ranchi | 2019–2022

COURSEWORK

Machine Learning, Time Series Analysis, Forecasting, Data Mining, Deep Learning, Statistical Inference, SQL, Visualization, NoSQL, Multivariate Analysis

TECHNICAL SKILLS

- **Programming & Analysis**: Python, SQL, R, Pandas, NumPy, EDA, Statistical Analysis, A/B Testing
- **Machine Learning & AI**: Scikit-learn, TensorFlow, PyTorch, Ensemble Models, Regression, Classification, Time Series Analysis
- **Natural Language Processing (NLP):** TF-IDF, BM25, Cosine Similarity, LSA, Sentiment Analysis (VADER, NRC Lexicon)
- **Visualization & BI:** Tableau, Power BI, Excel, Matplotlib
- **Tools & Platforms:** Git & GitHub, Jupyter Notebook, Streamlit, Google Colab
- **Data Handling & Analysis:** Data Cleaning, Preprocessing, Feature Engineering, Panel Data Construction, Exploratory Data Analysis (EDA), Hypothesis Testing, ETL
- **Databases & Data Storage**: MongoDB, Neo4j

SOFT SKILLS

Analytical Thinking, Problem-Solving & Critical Thinking, Attention to Details, Adaptability, Time Management, Interpersonal Skills, Communication & Collaboration Teamwork, Detail-oriented

PROJECTS

Plant Disease Detection – Image Analytics Project

Technologies: PyTorch, Transfer Learning, EfficientNet

- Applied transfer learning techniques to classify crop diseases using the PlantVillage dataset.
- Used data augmentation to improve generalization on limited datasets.
- Evaluated models using confusion matrix, classification report, and training history visualizations.
- Impact: Demonstrated strong capability in supervised learning and image-based data analysis.

News Research & Analytics Tool – Data & NLP Application

Technologies: Python, Streamlit, TF-IDF, BM25, Cosine Similarity, LSA

- Designed and developed an end-to-end text analytics system to extract, preprocess, rank, and summarize news articles.
- Implemented document similarity scoring using TF-IDF, BM25, and cosine similarity to improve information retrieval relevance.
- Built automated summarization pipelines to reduce manual article review time by approximately 40%.
- Deployed the application using Streamlit, enabling interactive search, filtering, and real-time relevance evaluation.
- Impact: Improved efficiency in analyzing large volumes of unstructured text data.

HR Analytics Dashboard – Workforce Insights (Tableau)

Technology: Tableau, Excel, Data Cleaning, KPI Tracking

- Built interactive dashboards for HR teams to track attrition rate, department-wise performance, and experience distribution.
- Developed dynamic filters and drill-down features to support data-driven workforce decisions. Cleaned and transformed raw HR data using Excel and Python for accurate reporting.
- Key Impact: Helped stakeholders identify patterns in turnover and performance variations across departments.

Renewable Energy Forecasting & Optimization

Technology: Python, Scikit-learn, Ensemble Models, Time Series Analysis

- Performed EDA on solar and wind datasets to identify the correlation between weather conditions and energy output.
- Built ensemble regression models (Random Forest, Gradient Boosting, XGBoost, LightGBM) to forecast energy production.
- Evaluated models using RMSE, MAE, and time-based splitting for reliable forecasting.
- Key Impact: Enabled predictive insights that can streamline energy planning and reduce variability in supply estimation.

Mental Health Survey Analytics – Sentiment Insight Reporting

Technology: Python, VADER, NRCLexicon, Data Visualization

- Conducted sentiment and emotion analysis across demographic groups (age, stress level, occupation).
- Created visual reports to identify trends in mental health concerns and emotional polarity.
- Key Impact: Presented insights useful for designing targeted mental health interventions.

HACKATHON

GenAI Hackathon - Winner

Built an enhanced News Research & Analytics Tool using a retrieval-based (RAG-style) approach with TF-IDF, BM25, and cosine similarity, deployed via Streamlit.

CERTIFICATIONS

- Deloitte-Data analytics job simulation - Forage
- JP Morgan Chase & Co. - Quantitative Research job simulation - Forage
- Python and Statistics for Financial Analysis - Coursera