

Aditya Raj Sinha

adityaharsh03@gmail.com | +91 9279314598 | [Aditya/Portfolio](#) | [linkedin.com/in/Aditya](#) | [github.com/Aditya](#)

PROFESSIONAL SUMMARY

Postgraduate in Data Science with solid grounding in data analytics, business intelligence, and machine learning. Skilled in Power BI, Python, SQL, and NLP. Practical experience includes developing interactive dashboards, automating text analysis, and creating predictive models to derive insights from data. Looking for a starting position in Data Analytics, Business Intelligence, or Machine Learning.

EDUCATION

- **M.Sc. Data Science—CGPA: 8.19**

Vellore Institute of Technology-AP | 2024 - 2026

- **B.Sc. Physics —CGPA: 8.05**

Ranchi University, Ranchi | 2019 - 2022

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, DAX, Matplotlib
- **Tools & Platforms:** 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

EXPERIENCE

Data Science Intern (Remote)

NullClass | Dec 2025 – Present

- Designed and implemented a conditional GAN-based text-to-image generation pipeline using PyTorch
- Built and trained custom Text Encoder, Generator, and Discriminator models, integrating text embeddings for multimodal learning

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.

GDP Analysis Dashboard

Skills: Power BI, DAX, Data Modeling, Data Visualization, World Bank & OECD Data

- Built an interactive dashboard to analyze global GDP trends (2017–2024) by country, region, and income group
- Created DAX measures for Total GDP, YoY Growth %, and top GDP country identification
- Key Impact: Enabled clear economic comparisons and data-driven insights through interactive BI reporting

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.

CERTIFICATIONS

- [Deloitte-Data analytics job simulation - Forage](#)

- [JP Morgan Chase & Co. - Quantitative Research job simulation - Forage](#)

- [Python and Statistics for Financial Analysis - Coursera](#)