

# Adithya Singupati

(930) 333-2369, [adisingu@iu.edu](mailto:adisingu@iu.edu), <https://www.linkedin.com/in/adithyasingupati>  
<https://github.com/MrAdithya21>, <https://mradithya21.github.io/mywebs.github.io/>

## Education

**Indiana University Bloomington** — *Master of Science in Data Science*

Graduation: May 2025

Relevant Coursework: Applied Machine Learning, Applied Algorithms, Computer Vision, Usable AI, Data Visualization, Advanced Database Technologies, Intro to Statistics

**Gayatri Vidya Parishad College of Engineering** — *Bachelor of Engineering in Computer Science*

Graduation: June 2023 — GPA: 3.7/4.0

Relevant Coursework: Software Engineering, Big Data Analytics, Data Mining, Artificial Intelligence

## Skills

**Programming Languages:** Python, Java, R, C, SQL, JavaScript

**Machine Learning & AI:** Supervised Learning, Regression, Clustering, Dimensionality Reduction, Classification, Generative AI (LangChain, Hugging Face)

**Data Engineering:** Data Modeling, Data Pipeline Development, ETL Workflows, Data Validation, SQL Query Optimization, Data Quality Monitoring

**Tools & Frameworks:** NumPy, pandas, scikit-learn, SQLAlchemy, dbt, Apache Airflow, Git, AWS (SageMaker, Lambda), JIRA, Agile

**Data Visualization:** Tableau, Power BI, AWS QuickSight, Data-Driven Dashboards, Python (Matplotlib, Seaborn, D3.js)

**Collaboration:** Cross-functional Teamwork, Stakeholder Communication, Agile Methodologies, Task Management

## Projects

### Iris Classification Project

- Designed a classification model using Logistic Regression and Decision Trees to classify iris species with 97% accuracy.
- Implemented model evaluation and used cross-validation to ensure robust results.
- Visualized feature importance and model performance using Matplotlib and Seaborn.

### Database Management System for Retail Inventory

- Developed a scalable SQL database and implemented ETL routines to manage retail inventory data with real-time insights.
- Created data pipelines for seamless data extraction, cleaning, and transformation to prepare data for analysis.
- Built interactive Power BI dashboards to monitor sales and inventory trends.

### Market Data Modeling and Customer Segmentation Analysis

- Designed a pipeline using Python and SQL to ingest and validate over 1M market/customer records from diverse sources.
- Applied PCA and clustering techniques to model data groupings and uncover patterns in customer behavior and asset interaction.
- Automated monitoring processes to flag anomalies, ensuring data accuracy and improving downstream analytics reliability.

### Generative AI Chatbot with Hugging Face

- Built a chatbot using Hugging Face's transformers library, leveraging GPT-3 for natural language understanding.
- Developed and deployed the model using AWS Lambda for real-time inference and SageMaker for model training and monitoring.
- Integrated the chatbot into a customer service application, improving response time by 30%.

## Experience

**Data Analyst** — *O'Neill School of Public and Environmental Affairs*, Bloomington, IN, USA November 2024 – Present

- Analyzed and processed 500K+ data points, ensuring 99% data accuracy, and developed interactive visualizations using D3.js for 1,000+ stakeholders.
- Designed and built data pipelines using Python to automate data extraction, cleaning, and transformation for analysis.
- Implemented automated workflows to monitor and track data quality, improving reporting efficiency by 25%.

**Data Scientist (Research Assistant)** — *Kelley School of Business*, Bloomington, IN, USA March 2024 – Present

- Processed and analyzed 80M+ financial news articles using NLP to extract sentiment metrics aligned with market movements.
- Built model pipelines that informed business-level decisions using topic modeling, classification, and anomaly detection.
- Applied clustering models to detect shifts in public sentiment trends related to economic indicators and investment behaviors.

**Data Scientist Intern** — *MyEdMaster LLC*, Leesburg, Virginia, USA

May 2024 – August 2024

- Built regression models to predict SAT performance with 85% accuracy, leading to a 20% improvement in personalized recommendations.
- Created reproducible and maintainable code for machine learning workflows using Python.
- Collaborated with software developers to integrate machine learning models into a user-facing platform.

## Publications

**Accessing General Health Care Facilities**, IRJET, Vol. 9 Issue 10

Developed an e-health platform for real-time medical equipment availability.

<https://www.irjet.net/archives/V9/i10/IRJET-V9I10150.pdf>

**Tracking the Storm: Visualizing Trends in U.S. Hurricanes & Climate Impact**, IRJET, Vol. 12 Issue 1

Analyzed hurricane data to develop climate impact visualizations and trends using advanced analytics.

<https://www.irjet.net/archives/V12/i1/IRJET-V12I1138.pdf>

## Certifications

- AWS Certified Cloud Practitioner
- Tableau Desktop Specialist