

# ADITYA PACHPANDE

✉ [adityapdev13@gmail.com](mailto:adityapdev13@gmail.com) | [in aditya-pachpande30](https://www.linkedin.com/in/aditya-pachpande30) | [G Adi2p30](https://github.com/Adi2p30) | US Citizen | [Scholar](#)

## Education

### Purdue University

3.3 GPA Bachelor of Science in Data Science, Entrepreneurship Certificate

May 2027

West Lafayette, Indiana

## Technical Skills

**Languages:** Python, R, PostgreSQL, Julia, Java, C++

**Software:** Tableau, QGIS, Panoply, MATLAB, Unity Microsoft Suite

**ML/AI & Data:** PyTorch, TensorFlow, Scikit-learn, Keras, (Agentic AI) Langchain, Google-ADK, Autogen, HuggingFace, Pandas, Polars, NumPy, SciPy, Matplotlib, NLTK, XArray, Spark, BigQuery, FastMCP

**Software Engineering:** Flask, FastAPI, NextJS, Redis, MongoDB, PostgreSQL, Neo4j, Docker

## Experience

### National Aeronautics and Space Administration, Goddard

June 2025 – August 2025

*Data Science Intern – CmCt & Cryosphere Modeling*

*Greenbelt, Maryland*

- Built CryoSphere Model Comparison tool with **Xarray**, **NumPy**, **Geopandas**; processes **40M+ data points**, **reduced analysis time from 3 min to 20 sec** per region. Designed with **Shapely** geospatial processing for multiple data formats.
- Optimized ice sheet modeling with **Numba JIT** and **NVIDIA CUDA** parallel processing; enabled simultaneous validation of **15+ climate models** against satellite data with 40% faster processing. These automated pipelines helped researchers iterate on their models **4x faster**

### Intelligent Systems, Assistive Technologies Lab - Purdue University

November 2024 – Present

*Machine Learning Undergraduate Researcher*

*Purdue University, West Lafayette*

- Co-authored and published research papers in journals including **Nature Journal (Scientific Data)**
- Trained **YOLOv11** computer vision models achieving **96% mAP@0.5:0.95**; applied **mixed-precision inference** to optimize throughput for real-time medical applications.
- Built **custom mixed reality visualization software** on the Microsoft HoloLens with real-time data overlay in **C#** and **Unity**; leveraged system to generate additional training data and iteratively refine CV model.
- **Fine-tuned LoRA adapters (Computer Vision using LLMs)** on multiple open-source LLMs; built training pipelines that reduced setup, deployment, training and eval times for lab members.

### Daniels School of Business

March 2025 – Present

*Marketing Data Science & Analytics Intern*

*West Lafayette*

- Built end-to-end **NLP pipeline** using **Sentence Transformers**, **BERT**, **LDA**, and **LSA** for topic modeling across media articles and graduate surveys; reduced dimensionality with **PCA** and tested at scale on **AWS SageMaker**.
- Engineered **admit-decline funnel analysis** over **129,147 rows** (2018–2023); built **Random Forest** and **Decision Tree** models to predict yield likelihood, identifying top factors: Student Experience (21.8%), Rankings (12.3%), and Financial Aid (11.2%).
- Developed **GPA normalization system** using **Gemini API** with web-search grounding and heuristic fallback across 17 countries; optimized geocoding pipeline reducing runtime from **15+ minutes to 10 seconds**.
- Automated survey summarization across **100+ column datasets** using **OpenAI** and **Gemini APIs**; integrated results into interactive **Tableau** dashboards with demographic filters and presented findings to stakeholders.

### Polistock.app

March 2025 – March 2026

*CEO*

*Remote*

- Built **multi-agent (Agentic AI) architectures** using **Google ADK**, **MCP Servers**, and **Supabase Realtime** to process live news into **Postgres pgvector** for low-latency market analysis.
- Developed **Natural Language Orders** using **Neo4j graph RAG** and vectorization to enable automated, event-driven trade execution on policy signals.
- Integrated **NewsAPI**, **Alpha Vantage**, and **SEC EDGAR** for real-time sentiment scoring with LLM reasoning across **5,000+ stocks**.
- Scaled full-stack platform (**FastAPI**, **Next.js**, **Railway**) to **300+ active users** and **45,000+ events analyzed**, driving over **\$250,000** in platform investments.

## Publications

---

- An Egocentric Life-Saving Interventional Procedure Dataset of Actions,  
Medical Questions, Maneuvers and Tools | *Nature Scientific Data*** **2025**
- Trauma THOMPSON: A Dataset and Realistic Generative Framework  
for AI Copilots in Emergency Care** **2025**