Piyush Choudhari

Pune, Maharashtra, India 411033

+91 9168088565 | GitHub | LinkedIn | Portfolio | choudhari.piyush@gmail.com

Technical Skills

- **Languages**: Python, Java, SQL
- **API Development**: FastAPI, Flask, Spring Boot, Postman, Insomnia
- **Data Analytics**: Excel, PowerBI, Tableau
- **Databases**: Sqlite3, PostgreSQL, MySql, MongoDB, ChromaDB
- **CI/CD**: GitHub Actions, Docker, Render, Vercel
- **Version Control**: Git, GitHub

Projects

Stock News Sentiment Analysis Platform - GitHub
- **Streamlined** real-time data ingestion by scraping over **60 news websites** using Selenium, extracting and parsing relevant content with the **Llama 3.2 LLM** for higher accuracy.
- Designed a robust inference pipeline to classify stock news sentiment (**positive**, **negative**, or **neutral**) using the **finBERT text classification model**, ensuring precise insights for users.
- **Engineered** a RAG-based AI Chatbot using Langchain, ChromaDB, and the **Gemini-1.5-Flash LLM**, enabling interactive, query-based access to stock news.

Mars Weather Data ETL Pipeline - GitHub | Video Link
- Built an automated **ETL pipeline in Java** to fetch and process **Mars weather data from NASA's InSight API**, storing it in **MongoDB with 99.9% uptime**.
- Created an interactive **Dash-powered dashboard** to visualise critical metrics like **pressure, temperature, and wind patterns** for data-driven insights.
- Automated the entire workflow with **Apache Airflow DAGs** and containerized the application using **Docker Compose**, integrating services like **PostgreSQL and Web Servers GUI** for seamless operations.
- Transformed **MongoDB data into structured SQL tables** with dynamic schema creation in **MySQL**, improving data analysis and accessibility.

Pokedex API - GitHub | Video Link
- **Developed** a scalable **RESTful API** to query data on **1,015 Pokemon**, enabling query by attributes like name, species, type, size, and abilities.
- **Scraped data** from www.pokemondb.net, exporting to CSV and storing it in a **PostgreSQL database** hosted on **Neon DB** on **AWS**.
- Published the dataset on **Kaggle**, receiving **1,200+ downloads and 5,000+ views** as an open source dataset.
- Utilised **FastAPI** as the backend server. **Containerized the API with Docker** and optimised to get less than **100ms response times.**

Achievements
- **1st Rank** In Software Category At **SIH Internal College Hackathon**.
- **2nd Rank** At **Data Visualization Hackathon**- Organised by ISA Club, D.Y.Patil College Of Engineering.
- **4th Rank** At **Visionary Hackathon**- Organised by Vishwakarma University And Binghamton University, New York.

Leadership
- **AI&ML Lead At IASC Club -** Organized a bootcamp for 70+ peers, covering topics like VCS, Git, Github, Open Source. Link
- **AI&ML Coordinator At IASC Club -** Conducted a 3-day GenAI Workshop for 120+ students, receiving 100+ positive student feedbacks. Link

Certifications
- **IBM Data Science** *(Coursera)* **-** Link
- **Complete Data Science,Machine Learning,DL,NLP Bootcamp** *(Udemy)* - Link

Education

D.Y.Patil College Of Engineering, Akurdi   *2022-2026*

Bachelor of Engineering in A.I.D.S(CGPA of 8.37)   *Pune, Maharashtra, India*