Shreya Pandey

+918451002702 | shreyap2702@gmail.com | linkedin.com/in/shreyarpandey | github.com/shreyap2702

Summary

AI and Data Science student with hands-on experience in Python, NLP, and early-stage Generative AI workflows. Have built classification systems, recommender engines, and data preprocessing pipelines across diverse projects. Actively exploring LangChain and RAG-based retrieval methods. Strong interest in applying ML to real-world domains through both technical and collaborative roles.

Education

K. K. Wagh Institute of Engineering, Education and Research — Nashik, India
Bachelor of Technology in Artificial Intelligence and Data Science - Semester 6: 9.09 SGPA — *Nov. 2022 – Present*

Rao Junior College of Science — Thane, India
Higher Secondary Education - 95% — *2021*

Projects

github-repo-preprocessing-pipeline | *Python, Docker (in progress), GitHub API* — 2025

- Building a data pipeline to parse and structure GitHub repositories by detecting file types, directory structure, and code dependencies.
- Designed to support downstream tasks in a broader GitRAG system aimed at contextual retrieval over repository data.
- Currently integrating Docker for containerized cloning and exploring vector DB compatibility for scalable retrieval.

stack-own | *FastAPI, spaCy, React, SQLite* — June 2025

- Built an ML-powered recommender that suggests tech stacks from natural language project descriptions.
- Used spaCy's TextCat model for multi-label classification; integrated with a FastAPI backend.
- Showcased end-to-end ML workflow with real-time suggestions and persistent project storage.

last-resort | *FastAPI, SQLModel, React, Vite* — May 2025

- Developed a note-sharing platform for paired users to exchange daily messages (morning/night) securely.
- Implemented FastAPI-based REST APIs for user pairing, CRUD operations, and date-wise note filtering.
- Designed a responsive UI using React + Vite and deployed the frontend on Vercel.

Technical Skills

Languages: Python, SQL, C/C++, JavaScript
ML & NLP: PyTorch, TensorFlow, spaCy, Scikit-learn
Data Handling: pandas, NumPy, SQLModel, PySpark
Backend & APIs: FastAPI, Django, Streamlit, React
Tools: Git, Docker (in progress), Jupyter, VS Code, MongoDB, Vercel
Current Explorations: LangChain, Vector Database (in progress)

Additional Learning

Machine Learning Specialization — January 2025
Coursera — *Online*

- Explored Supervised, Unsupervised, and advanced ML algorithms with structured data use cases.

Build Basic Generative Adversarial Networks (GANs) — 2025
Coursera — *Online*

- Built and trained simple GANs using PyTorch for image transformation and generation.