

Sarthak Doshi

Indore (M.P.) | +91 78987 70372 | sarthakdoshi231@gmail.com | github.com/SdSarthak | linkedin.com/in/sd-sarthak

SKILLS

Programming Languages:	Python, SQL, HTML/CSS
Frameworks:	TensorFlow, NLTK, Transformers, OpenAI, SciKit Learn, Keras, LangChain, Beautiful Soup, Flask, Django, PyTorch
Tools & Technologies:	Postman, Git/ Gitlab, JIRA, MongoDB MySQL Workbench, DBeaver, Tableau, PowerBI, Jupyter, Colab, MS Office
Soft Skills:	Problem Solving, Team Work

EDUCATION

B.TECH. in Computer Science (Data Science)	08/2022 - 05/2026
Acropolis Institute of Technology and Research - Indore, (M.P.)	CGPA: 7.9 / 10.0
Higher Secondary School PCM	07/2020 - 05/2022
Indore International School – Indore (M.P.)	Score: 82.4%

WORK EXPERIENCE

AI/ML Engineer Intern | GreenStory 01/2025 - 08/2025

- Developed an AI chatbot using a hybrid RAG (Retrieval-Augmented Generation) approach and prompt engineering with OpenAI models, increasing response accuracy by 35%.
- Engineered a bulk product matching system that automated SKU alignment across e-commerce platforms, reducing manual processing time dramatically.
- Optimized vector-based search with embeddings, boosting retrieval speed by 40% for high-volume product data queries.
- Developed an automated report generation system that aggregates and parses data from four APIs, leveraging GenAI and OpenAI to produce a 40-page report—reducing manual effort from 1–2 weeks to minutes.

Data Science Intern | Sabudh Foundation 07/2024 - 12/2024

- Built and deployed a deep learning-based recommendation system that increased user engagement by 30% for retail analytics.
- Designed and implemented machine learning models for fraud detection, reducing false positives by 25% using feature engineering and anomaly detection.
- Optimized complex SQL queries for financial data analysis, improving query performance by 50% and supporting real-time transaction monitoring.

NOTABLE PROJECTS

Sentiment Analysis | Python, BeautifulSoup, NLTK, GeminiAPI, Langchain 12/2024

- Developed a web scraping and sentiment analysis pipeline classifying sentiments into five levels: highly positive, mildly positive, neutral, negative, and highly negative.
- Utilized advanced text preprocessing (stopword removal, stemming, TF-IDF vectorization) to refine input data, achieving 85% accuracy on real-world datasets.
- Built and trained a classification model using both Naïve Bayes and Transformer-based approaches.

CERTIFICATIONS

- IBM Coursera - Introduction to Deep Learning & Neural Networks with Keras
- NPTEL - Python for Data Science, Introduction to Machine Learning, DBMS
- Hackerrank - Software Engineer Intern, SQL Intermediate
- Sabudh Foundation - Natural Language Processing and Recommender Systems, Deep learning, Machine learning
- Others - Dataiku Developer Series, AIML IIT Indore Fluxus, Python, Numpy from Kaggle

VOLUNTEERING

- Vice Chairperson and Ex-Trainer, ACM Student Chapter – Organized technical workshops, coding competitions, and mentoring sessions to foster a strong technology community.
- National NGO PRAWES – Contributed to community outreach programs, promoting social causes and supporting underprivileged initiatives.