

# SARTHAK ARSUL

+91 9168342306 [sarthakarsul18@gmail.com](mailto:sarthakarsul18@gmail.com) [LinkedIn](#) [GitHub](#) Balaji Nagar, Pune, Maharashtra - 411043

## Profile Summary

Certified Data Scientist skilled in ML, DL, NLP, and GenAI with expertise in Python (Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow) and SQL. Experienced in regression, classification, clustering, neural networks, and LLM-based applications. Proficient in data visualization (Power BI, Matplotlib, Seaborn) and deploying ML apps using Flask, FastAPI, and Docker. Open to roles like Data Scientist, ML/NLP Engineer, GenAI Engineer, Python Developer, and Data Analyst.

## Technical Skills

**Programming & Frameworks:** Python, Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow, LangChain, LangGraph, SQL, FastAPI, Flask, Django(basic)

**Machine Learning & AI:** Regression, Classification, Clustering, Neural Networks, Deep Learning, NLP

**Tools & Visualization:** GitHub, Matplotlib, Seaborn, MySQL, PostgreSQL, Power BI, Jupyter Notebook, Docker

## Soft Skills

**Fast Learner & Problem Solver:** Quickly grasps concepts and finds effective solutions

**Good Communicator & Team-Oriented:** Works well in teams and communicates clearly

**Organized & Time-Conscious:** Manages tasks efficiently and meets deadlines

## Projects

### AskMyDocs AI – RAG-based Document Assistant (End-to-End Project)

[Live Demo](#) | [GitHub Link](#)

- Developed and deployed an end-to-end **RAG system using LangChain, Hugging Face, and Streamlit**.
- Enabled PDF ingestion, **embeddings**, and semantic search for real-time Q&A.

### Air Quality & Health Risk Prediction System (End-to-End Project)

[GitHub Link](#)

- Built a real-time air quality health risk predictor using **WAQI API and ML model**
- Created **Streamlit** app for city-wise predictions and actionable health insights

## Internship

### Data Science & Analytics Intern

*Zidio Development | Jul 2024 – Aug 2024*

Built speech emotion detection models using NLP & ML for real-time applications.

### Data Science Intern

*CodSoft | Jun 2024 – Jul 2024*

Executed ML projects in data analysis & prediction; maintained GitHub & LinkedIn presence.

## Certifications

**Data Science – Ethan's Tech**

**Python Programming – Ethan's Tech**

## Education

Deogiri College, Sambhajinagar - BCA – 73.43%

Jun 2020 – May 2023

Balbhim College, Beed - 12<sup>th</sup> (HSC) – 63.38%

Jun 2019 – May 2020