Shivam Sonawane

sasonawane2003@gmail.com 9359145675 Nashik , India https://github.com/ShivamSonawane2003
https://in.linkedin.com/in/shivam-sonawane-582b48346 https://www.kaggle.com/shivamsonawane96
https://shivamsonawane2003.github.io/ShivamSonawane2003_portfolio/

PROFESSIONAL SUMMARY

Aspiring Data Scientist with hands-on experience in supervised learning, model deployment (Streamlit,FastAPI), and data visualization (Tableau, Matplotlib). Skilled in Python, SQL, and EDA with real-world applications in healthcare, finance, and content analytics.

WORK EXPERIENCE

Data Science Intern, *AiVariant*
08/2024 – 05/2025
Pune, India

- **Improved** dataset **accuracy by 98%** through comprehensive **data preprocessing**
- **Monitored** data activities, Cleaned and validated **10,000+ data entries** weekly to improve model input quality
- Developed **dashboards and visual reports** using Python, Pandas, Matplotlib, and Tableau.

Junior Software Developer, *Humming Byte Technologies*
01/2024 – 06/2024
Nashik, India

- Developed **5+ mobile applications** using Android Studio, improving user **engagement by 20%**.
- Designed and maintained **3+ web applications** using HTML, CSS, **reducing** page load time by **15%**.
- **Collaborated** in team to debug and optimize software, **improving performance** by **25%**.

Junior Software Program Analyst, *Innovatus Technologies*
09/2023 – 01/2024
Pune, India

- Translated **10+ user requirements** into technical specifications, improving system efficiency by 18%.
- Debugged and resolved **15+ software issues**, reducing system downtime by **40%**.
- Managed client communications, leading to a **95%** customer **satisfaction rate**.

EDUCATION

Bachelor Of Computer Science and Engineering, *Sandip University*
10/2022 – 06/2025
Nashik, India

Diploma In Information Technology, *Amrutvahini Polytechnic*
08/2019 – 07/2022
Sangamner, India

Data Science Study Program, *Excelr Solutions*
08/2024 – 01/2025
Pune, India

SKILLS

Programming: Python, SQL
Machine Learning: Scikit-learn, Deep Learning,TensorFlow, PyTorch, XGBoost
GEN AI Skills: LLMs, RAG, Ollama, Transformers, Fine-tuning, Prompting, Embeddings, VectorDB, Inference, Automation, Agents, Chatbots, NLP, Deployment
Data Analysis : Pandas, NumPy, EDA, Feature Engineering,Data Preprocessing
Visualization: Matplotlib, Seaborn, Tableau
Problem-Solving : Analytical thinking ,Data-driven decision making,Creative solution design ,Debugging and optimization
Tools: Jupyter Notebook, Streamlit ,Excel, PowerPoint, GitHub,FastAPI

LANGUAGES

English — Professional Working Proficiency | **Marathi** — Native or Bilingual Proficiency | **Hindi** — Conversational Proficiency

CERTIFICATES

- ExcelR Data Science Certified | 2025 | Credential ID: 18994/07012025
- Data Science Intern Certificate| 2025 | Credential ID:AIV/24-25/Q1/03/16731
- Generative AI Foundations Certified | 2025 | Credential ID: plOwOFvoWscHe6tK

PROJECTS

Diabetic Retinopathy In Patients, *Demo App : https://diabetic-retinopathy-prediction-in-patients.streamlit.app/*

- Developed a **binary classification model** to **predict** diabetic retinopathy using **ML algorithms**.
- Processed **6,000+ medical records** using Python and creating an interactive prediction app using Streamlit.

Fraud Detection System Based on Data, *Demo App : https://fraud-detection-system-shivam-sonawane.streamlit.app/*

- Developed a fraud detection model using **Isolation Forest** with **89% recall** on **5,000+ transactions**, applying **SMOTE** for class imbalance.
- Deployed the solution via **Streamlit** for **real-time anomaly detection**.

Electric Motor Speed Prediction Module

- Implemented **ML algorithms** to **predict speed**, optimizing model choice based on **accuracy by 96%** and **computational efficiency**.
- **Reduced** training time by **20%** through **feature optimization** and deployed a **real-time prediction interface** using Streamlit.

Deep Learning-Based Automation and Object Detection Analysis, *Publish paper : https://ijmrset.com/upload/40_Deep.pdf*

- Designed an object detection system using advanced **CNN and K-NN techniques**, focused on **automation** and **real-time analysis**.
- **Published research** based on this deep learning project in **IJMRSET Journal** as part of final year work.

RAG-Based Document Question Answering System

- Developed a **Fast API**-based **Retrieval-Augmented Generation (RAG)** application to answer user queries from private documents with **reduced hallucinations**.
- Implemented semantic retrieval using **Hugging Face embeddings** and **FAISS Vector DB**, enabling accurate, source-grounded responses.