

Sahil Ghevariya

Aspiring Machine Learning and Data Science Engineer

8799094066 | sahilghevariya001@gmail.com | [LinkedIn](#) | [GitHub](#)

Professional Summary

Highly motivated and detail-oriented Computer Science student with a strong interest in Artificial Intelligence and Machine Learning. Skilled in Python, data preprocessing, and model development using frameworks like TensorFlow and Scikit-Learn. Familiar with deploying ML models using Flask and Docker, and version controlling projects with Git/GitHub. Passionate about learning new technologies and applying data-driven solutions to solve real-world challenges.

Technical Skills

Programming Languages: Python, C, Java, HTML, CSS, JS

Data Analysis & Machine Learning: Pandas, NumPy, Scikit-learn, Seaborn, Matplotlib, Plotly, CatBoost, XGBoost, LightGBM

NLP & Deep Learning: NLTK, SpaCy, Transformers, Hugging Face, TensorFlow, PyTorch, Keras, OpenCV Data

Analysis Tools: Tableau, Power BI, Streamlit, IBM Cognos

Development Environments: Git, VS Code, Spyder, Docker, PyCharm, Anaconda, Jupyter Notebook, Eclipse, Visual Studio, GitHub Actions

Database Management Systems (DBMS): MySQL, Oracle DB, PostgreSQL, MongoDB

MLOps: MLflow, DVC, GitHub Actions, Docker, FastAPI, Flask

Soft Skills: Analytical Thinking, Problem-Solving, Communication, Team Collaboration

Other Skills: Business Analysis, Data Analysis, Data Pipeline Development, Model Deployment

Certification

[AI/ML Internship Certificate](#), [Google Cloud Agent AI Day](#), [Microsoft Azure Fundamentals \(AZ-900\) training at Parul University](#), [AI/ML Hackathon 2.0, TechSpark Hackathon 2025](#), [TCS iON Career Edge - Young Professional certification](#),

Projects

Customer Churn Prediction | Python – Scikit-Learn – Pandas – Matplotlib | [GitHub](#)

- Developed a machine learning model to predict customer churn using classification algorithms such as Logistic Regression, Random Forest, and XGBoost.
- Performed extensive data preprocessing including handling missing values, encoding categorical features, and feature scaling for improved model performance.
- Achieved **~86% accuracy** on test data, with precision-recall tradeoff optimization to handle class imbalance.

House Price Prediction | Python – Scikit-Learn – Pandas – Matplotlib | [GitHub](#)

- Built a prediction model for house pricing using a synthetic dataset with features including size, type, location, room count, and age.
- Developed preprocessing pipelines (handling missing values, encoding, scaling), implemented Linear Regression & Random Forest algorithms, and compared their performance using visualizations and evaluation metrics.

Titanic Survival Prediction | Python – Pandas – Scikit-Learn | [GitHub](#)

- Developed classification models to predict passenger survival on the Titanic by preprocessing dataset features (missing value imputation, encoding, scaling) and exploring feature importance.
- Trained and compared models such as Logistic Regression and Decision Tree, evaluated using accuracy, precision, recall, and visualised insights to highlight key survival predictors.

Experience

OutriX – AI/ML Intern (Remote)

Aug 2025 – Sep 2025

- Completed a 1-month remote internship as an AI/ML Intern at OutriX, gaining hands-on experience in machine learning workflows and practical model development.

Education

Parul University

Vadodara, Gujarat

B.Tech in Computer science engineering – Artificial intelligence (4-Year)

June 2023 – May 2027

Sanskartirth Gyanpeeth Vidhyalaya

Surat, Gujarat

Higher secondary Education

June 2022 – Mar 2023

Achievements

- Participated in **AI ML Hackathon 2.0**, developed a *Voice-Enabled Patient Intake System* using React.js and Flask.
- Completed **Microsoft Azure Fundamentals (AZ-900) training** at Parul University and **currently preparing for the official certification exam**.

Languages

- English(Professional fluency)
- Hindi
- Gujarati(Native)