

Rahul Shewatkar

 +91-9881658802  rshevatkar@gmail.com

 Portfolio: rshevatkar.github.io/portfolio-website

 GitHub: github.com/rshevatkar

 LinkedIn: linkedin.com/in/rahul-shewatkar-ml-engineer

Profile

Data professional with a background in logistics operations and 1+ years of hands-on ML internship and project work. Built end-to-end machine learning projects — from EDA and feature engineering to model deployment — achieving 97.5% classification accuracy and live-deployed apps used for restaurant demand analysis and shipment delay prediction. Skilled in Python, Pandas, Scikit-learn, Matplotlib, and Seaborn. Transitioning into a data analyst or junior data scientist role to combine domain knowledge with predictive modelling capability.

Relevant transition: Completed 500+ hours of self-directed ML training (2023–2024) covering Python, statistics, supervised learning, and model deployment

EDUCATION

Shree H.V.P. M's COET Amravati, Amravati University | Jun 2013 – May 2016

Bachelor of Engineering in Electronics and Telecommunication

P.R. Pote Patil Institute Polytechnic, Amravati | Jun 2010 – May 2013

Diploma in Engineering

SKILLS

Languages: Python, SQL (basic)

ML Libraries: Scikit-learn, XGBoost, TensorFlow (basic), H2O AutoML, LangChain, Retrieval-Augmented Generation (RAG), HuggingFace Transformers, Sentence Transformers

Data & Visualization: Pandas, NumPy, Matplotlib, Seaborn

MLOps & Deployment: MLflow, Docker, Docker Compose, GitHub Actions, CI/CD Pipelines, FastAPI, Streamlit, Render, Hugging Face Spaces

Vector Databases: ChromaDB

Tools: Git, GitHub, Jupyter Notebook, VS Code

Familiar with: AWS, Azure ML, GCP (exposure, no production experience)

EXPERIENCE

Codec Technologies Pvt. Ltd.

Mumbai, India

Data Science Intern

Aug/2025 – Nov/2025

- Worked on end-to-end machine learning pipelines, from data preprocessing and feature engineering to model training and evaluation.
- Built and tested supervised learning models (regression & classification) using Python, focusing on performance, accuracy, and business relevance.
- Conducted exploratory data analysis (EDA) to identify patterns, correlations, and data quality issues that informed model design.
- Applied feature selection and transformation techniques to improve model stability and predictive performance.
- Documented workflows, results, and insights to support reproducibility and stakeholder understanding.

Unified Mentor Pvt. Ltd.

Gurugram, India

Machine Learning Engineer Intern

Jan/2025 – Jul/2025

- Contributed to price prediction and classification pipelines; performed EDA using Pandas and Seaborn to identify key correlations across structured datasets of 10K+ records.
- Contributed to data cleaning, preprocessing, and feature engineering tasks using Python (Pandas, NumPy).
- Supported exploratory data analysis and visual storytelling using Matplotlib and Seaborn.
- Assisted in model training and evaluation for real-world use cases such as price prediction and classification.
- Documented findings, created dashboards, and presented results to mentors and project stakeholders.

Shipco IT Pvt. Ltd.

Pune, India

Export Documentation Executive

Sep/2019 – Feb/2023

- Prepared and managed all export documentation, including invoices, packing lists, and bills of lading, ensuring compliance with international regulations.
- Coordinated shipments to international customers across the US, Europe, and Asia, achieving 98% on-time delivery rate.
- Led the migration of the Manifest process for major ports like Keelung, Rotterdam, and Shanghai, improving operational accuracy.
- Handled Master Bill of Lading (MBL) queries and collaborated with international offices to resolve discrepancies promptly.

WNS Global Services Pvt. Ltd.

Pune, India

Associate Operations (Air Freight)

Mar 2018 – Apr 2019

- Managed customer onboarding and generated credentials for air freight tools, ensuring smooth system access.
- Compiled monthly performance reports and presented insights to clients, driving data-driven decision-making.
- Calculated and allocated reward points based on revenue data, enhancing customer engagement and retention.

GTL Pvt. Ltd. (Micro-Internship)

Pune, India

Engineer

Feb/2018

- Monitored and resolved network issues by collaborating with escalation and incident management teams.
- Delivered satisfactory services by managing customer-end devices and troubleshooting network problems efficiently.

PROJECTS**RAGOps-AI – Production-Ready RAG & MLOps Pipeline**

Tech Stack: FastAPI, Streamlit, MLflow, Docker, GitHub Actions, ChromaDB, HuggingFace

- Developed a production-style AI Resume Assistant using Retrieval-Augmented Generation (RAG) architecture.
- Built REST APIs with FastAPI and interactive frontend with Streamlit.
- Designed vector database pipeline using ChromaDB and sentence-transformer embeddings.
- Tracked experiments and retrieval metrics with MLflow for prompt, chunk-size, k-value, and relevance evaluation.
- Implemented Docker-based microservice deployment and CI/CD automation using GitHub Actions.
- Deployed multi-service architecture on Render with separate API, UI, and MLflow tracking services.
- **Live app:** <https://ragops-streamlit-ui.onrender.com>
- **GitHub:** <https://github.com/rshewatkar/ragops-ai>

Shipment Delay Predictor

Tech Stack: Scikit-learn, XGBoost, Streamlit, Docker, GitHub Actions, HuggingFace

- Built an end-to-end ML system to predict shipment delays for logistics and courier companies before orders leave the warehouse.
- Performed EDA on 10,999 shipments, uncovering key delay drivers: high discounts, 2–4kg weight range, and 4+ customer care calls. Engineered 3 new features from these insights
- Trained and tuned 4 ML models. XGBoost achieved 71.51% F1 score and 75.29% AUC-ROC. Built an interactive Streamlit dashboard for operations teams to input shipment details and get instant delay risk predictions with actionable recommendations
- Deployed as a live Streamlit app on Hugging Face Spaces, allowing real-time delay risk assessment
- **Live app:** <https://rshewatkar-shipment-delay-predictor.hf.space/>
- **GitHub:** <https://github.com/rshewatkar/shipment-delay-predictor>

Ship CO₂ Emission Predictor

Tech Stack: Scikit-learn, Streamlit, GitHub Actions, HuggingFace

- Built an end-to-end machine learning pipeline using Python, Pandas, and Scikit-learn to predict ship CO₂ emissions
- Designed a pipeline-based architecture (Column Transformer + Pipeline) to ensure consistent preprocessing and avoid serialization issues
- Achieved R² score of 0.995 and RMSE of ~1013, indicating high prediction accuracy
- Developed an interactive web app using Streamlit with dynamic UI components (sliders, dropdowns) for real-time predictions
- Deployed the application on Hugging Face Spaces, handling environment and dependency management
- Debugged and resolved production issues including model serialization errors, dependency conflicts, and Git LFS artifact handling
- Implemented a fallback mechanism to retrain model dynamically when deployment artifacts were missing or corrupted
- **Live App :** https://huggingface.co/spaces/rshewatkar/ship_co2-emission-predictor
- **GitHub :** https://github.com/rshewatkar/Ship_CO2_Emission_Prediction

Restaurant Demand / Analytics ML Project

Tech Stack: Scikit-learn, Folium, Streamlit, Docker, GitHub Actions, HuggingFace

- Built end-to-end ML pipeline analyzing 9,551 restaurants across 15+ countries for rating prediction, cuisine classification, and personalized recommendations
- Trained and compared 3 ML models (Linear Regression, Decision Tree, Random Forest) achieving best R² score with Random Forest regressor
- Developed content-based restaurant recommendation system using TF-IDF vectorization and Cosine Similarity
- Created interactive geographical analysis with Folium maps showing restaurant density, ratings, and city-wise statistics
- Deployed full Streamlit web application on Hugging Face Spaces
- **Live App :** <https://rshewatkar-restaurant-ml-project.hf.space>
- **GitHub :** <https://github.com/rshewatkar/restaurant-ml-project>

Vehicle Price Prediction using Machine Learning

- Developed a regression model to predict used vehicle prices based on features like year, mileage, fuel type, and engine

size.

- Explored correlations using heatmaps and visualized key trends in pricing behavior.
- Trained and compared multiple models (Linear Regression, Random Forest, Decision Tree); achieved 85% R^2 on test data.
- Deployed insights to identify influential features and improve model accuracy.

Mobile Phone Price Category Prediction using Machine Learning

- Developed a multi-class classification model to predict mobile phone price categories (low/medium/high/very high) based on key specifications using supervised learning.
- Trained and evaluated five models (Logistic Regression, SVC, KNN, Random Forest, XGBoost), achieving 97.5% accuracy with Logistic Regression.
- Delivered actionable business insights, enabling data-driven pricing strategy for mobile retailers and e-commerce platforms.