

Shreyas Mahangade

+91-9322651611 | shreymahangade1@gmail.com | [linkedin.com/in/shreyas3011](https://www.linkedin.com/in/shreyas3011) | github.com/shreyas3011

PROFILE SUMMARY

Results-driven Data Scientist and Machine Learning Engineer with a strong foundation in Python, statistics, and modern Machine Learning techniques. Skilled in developing end-to-end machine learning pipelines, model evaluation, and deployment using tools like AWS, Streamlit, LangChain and github. Proficient in LLMs, Generative AI, and NLP, with hands-on experience in data preprocessing, feature engineering, and improving model performance through tuning and validation. Passionate about applying AI to real-world challenges in a scalable and explainable manner.

EDUCATION

- Bachelor of Engineering (B.E) Artificial Intelligence and Data Science** 2022 – 2026
Savitribai Phule Pune University (SPPU), Pune | CGPA: 9.20|
- Higher Secondary Certificate (H.S.C) Science** 2021 – 2022
Shivaji University, Kolhapur | 76%|

EXPERIENCE

Machine Learning Engineer Intern Feb 2025 – Jul 2025
Design Techno SERVICES PVT LTD Pune, Maharashtra

- Collaborated with cross-functional team members to build and tune classification models improving accuracy by 18% through effective feature engineering and hyperparameter tuning with GridSearchCV.
- Automated preprocessing workflows using Pandas and NumPy, reducing data preparation time by 25% and streamlining model evaluation through cross-validation.
- Deployed a Streamlit app for real-time inference, enhancing accessibility and showcasing an end-to-end ML pipeline.

AI/ML Intern Jan 2026 – Mar 2026
Ecocomitychain AI (ECC) Remote

- Collaborated with an AI/ML team to **design and implement knowledge graph-based systems** for supply chain disruption analysis, modeling entities and relationships to enable intelligent decision-making.
- Built and managed **graph data pipelines using Neo4j**, contributing to schema design, data transformation, and efficient querying of large-scale supply chain data
- Contributed to the development of **AI-driven applications** by performing data preprocessing and supporting machine learning workflows integrated with graph-based insights.

PROJECTS

AI-Powered Multi-Agent News Intelligence & Audio Summarization System | MCP, Multi-Agent Systems | [Project Link](#) |

- Designed a multi-agent AI system with **MCP** coordination for data ingestion, summarization, sentiment analysis & audio generation
- Built an **automated Python-based news crawling pipeline** with dynamic rendering, proxy rotation, and anti-bot handling.
- Implemented **LLM-powered reasoning and summarization agents** for extracting key insights from unstructured news and social data.
- Integrated **Reddit sentiment analysis with GenAI pipelines** using prompt engineering, RAG, and neural text-to-speech for contextual news summaries.
- Engineered a **scalable, modular AI system** with asynchronous processing, API-based communication, and cloud-ready deployment architecture.

Config-Driven Predictive Diagnostics Pipeline for APS Sensors | Scikit-Learn, Docker, AWS (EC2/S3), Linux | [Project Link](#) |

- Built a **modular, config-driven ML pipeline** for APS component failure prediction using high-dimensional sensor data.
- Applied **KNNImputer, RobustScaler, and SMOTE-Tomek** to handle missing data, outliers, and class imbalance.
- Optimized multiple models using a **cost-sensitive metric** to reduce high-risk false negatives.
- Implemented **logging, exception handling, and artifact tracking** for deployment-ready reliability.
- Integrated **CI/CD pipelines via GitHub Actions** to automate testing and streamline deployment.
- Containerized with **Docker** and deployed on **AWS** for scalable, cloud-based inference.
- Achieved **99.6%** accuracy using **XGBoost** while minimizing failure-related costs.

TECHNICAL SKILLS

- Languages:** Python, C++, SQL
- Machine Learning & AI:** TensorFlow, Scikit-Learn, Hugging Face, LangChain, LLMs, Generative AI, MCP, AI Agents, Phidata, Deep Learning, NLP, Streamlit, NLTK
- MLOps Tools:** MLflow, DVC, Docker, GitHub Actions
- Data & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, MySQL, neo4j
- Deployment & Tools:** Streamlit, Git, GitHub, FastAPI, AWS

CERTIFICATIONS

- Data Science [EduTECH Pune, 2025]
- Oracle Cloud Infrastructure 2025 Generative AI Professional (1Z0-1127-25) [Oracle University, 2025]