

GOVINDRAJ KOTALWAR

govindrajkotalwar038@gmail.com | +91-7058039664 | [linkedin.com/in/govindraj-kotalwar-2004-297986257](#) |
[github.com/Starbuster2004](#)
Pune, India

PROFESSIONAL SUMMARY

Results-driven Data Scientist & Analyst with a strong foundation in statistical analysis and machine learning. Proven experience in optimising data workflows and enhancing data quality at Vodafone Idea Foundation. Proficient in Python, SQL, and Predictive Modelling, with a focus on translating complex data into actionable business insights.

TECHNICAL SKILLS

- Data Science & Analytics:** Statistical Analysis, Hypothesis Testing, A/B Testing, EDA, Predictive Modelling, Pandas, NumPy.
- Machine Learning:** Scikit-learn, XGBoost, Linear/Logistic Regression, Random Forest, K-Means Clustering.
- Deep Learning & AI:** TensorFlow, PyTorch, CNNs, YOLOv8, OpenCV, RAG, Llama 3.1.
- Data Engineering:** SQL (Advanced), ETL Pipelines, BigQuery, Data Cleaning, Airflow.
- Tools & Platforms:** Docker, Git, AWS (EC2, S3), Tableau, Power BI, Excel.
- Soft Skills:** Leadership, Team Collaboration, Problem Solving, Critical Thinking, Adaptability.

PROFESSIONAL EXPERIENCE

- Vodafone Idea Foundation** AI Data Analyst Intern
Pune, India
Oct 2024 – Nov 2024
- Process Optimization:** Developed and deployed predictive classification models using **Scikit-learn** and **SQL**, improving data processing efficiency by **30%** and automating 15+ hours of weekly manual workflows.
 - Statistical Analysis:** Conducted comprehensive Exploratory Data Analysis (EDA) on 500K+ row datasets; utilised **Confidence Intervals** and **p-values** to validate hypotheses, directly influencing 3 strategic product roadmap decisions.
 - Pipeline Integrity:** Integrated data quality validation into ETL pipelines, boosting accuracy by **25%** and implementing **Data Drift** detection frameworks using Airflow.
- MIT Aurangabad - ECESA** Technical Co-Head
Aurangabad, India
Feb 2024 – Present
- Leadership & Management:** Directed technical upskilling programs for 200+ students in AI/ML and Cloud Computing, resulting in a **45% increase** in workshop attendance through strategic planning and team coordination.

PROJECTS

- Cold Email Generator** | *Llama 3.1, Groq LPU, LangChain, ChromaDB*
- Engineered a high-performance email generation tool using **Llama 3.1** on Groq, achieving **sub-700ms inference latency** for real-time generation.
 - Implemented a **RAG pipeline** with ChromaDB to dynamically match user portfolios with job descriptions, significantly improving relevance.
- Legal Agent RAG** | *Python, OpenAI Embeddings, FAISS, NLP*
- Developed an AI system capable of querying complex legal documents via **FAISS-based semantic search**, ensuring high-accuracy retrieval of relevant legal precedents.
 - Integrated prompt orchestration to enable accurate extraction and summarisation of legal clauses.
- Deep Detective (Computer Vision)** | *TensorFlow, CNN, OpenCV, Transfer Learning*
- Built a transfer-learning CNN model to detect deepfake images with **92% precision** and **88% recall**, reducing false positives by 27% compared to baseline models.
 - Designed a frame-level preprocessing pipeline using **OpenCV** to ensure stable detection across varied lighting conditions.

EDUCATION

- Maharashtra Institute of Technology** **CGPA: 7.08**
Aurangabad, India
Nov 2022 – July 2026
- Coursework:** Operating Systems (OS), Computer Networks (CN), DSA, DBMS, Statistics, Artificial Intelligence.

CERTIFICATIONS

Oracle Certified Generative AI Professional (2025) — Oracle Certified AI Foundations Associate (2024)