

MOHD SHAMI

Data Analyst & ML Engineer

✉ codexshami@gmail.com | 📞 +91 89235 91576 | 📍 Moradabad, UP, India
🌐 [linkedin.com/in/codexshami](https://www.linkedin.com/in/codexshami) | 🐙 github.com/codexshami | 🌐 codexshami.github.io

Professional Summary

Detail-oriented Data Analyst and ML Engineer with hands-on experience in machine learning model development, data analysis, and business intelligence. Proficient in Python, SQL, Power BI, and Tableau. Skilled in end-to-end ML pipelines including feature engineering, model evaluation, and Flask-based deployment. Adept at translating business requirements into data-driven solutions that enhance decision-making.

Skills

Programming: Python (Pandas, NumPy), R, SQL

Machine Learning: Scikit-learn, XGBoost, Random Forest, CNN, TF-IDF, SMOTE, Hyperparameter Tuning

Data Visualization & BI: Power BI, Tableau, Matplotlib, Seaborn

Tools: Jupyter, Git, Docker, Flask

Statistics & Core: Descriptive & Inferential Statistics, Hypothesis Testing, A/B Testing, Feature Engineering, EDA, Model Evaluation (AUC, Precision/Recall)

Soft Skills: Communication, Problem-Solving, Adaptability, Teamwork, Time Management, Critical Thinking, Leadership

Work Experience

Data Science Intern | *Codec Technologies India*

Aug 2025 – Oct 2025

- Developed an XGBoost diagnostic model improving accuracy from 72% to 85%; documented preprocessing and hyperparameter choices reducing misclassification rate.
- Conducted end-to-end EDA, feature engineering, and data-cleaning pipelines for clinical datasets; delivered stakeholder-facing reports.
- Implemented reproducible training scripts and versioned experiments in Git for team collaboration.

Python Developer Intern | *CodTech IT Solutions Pvt. Ltd.*

Jun 2025 – Aug 2025

- Built reusable Python ETL scripts (Pandas, NumPy) to clean and transform heterogeneous data sources for downstream analysis.
- Automated data validation and reporting templates, improving reproducibility and reducing manual preprocessing.

Projects

Revive — ML-Based Disease Prediction Web App 🌐 [codexshami/Revive](https://github.com/codexshami/Revive) | Python, Flask, XGBoost, Bootstrap | Jan 2026 – Mar 2026

- Improved diagnostic accuracy from 72% to 85% via feature engineering and hyperparameter tuning; deployed via Flask web UI for clinician review.

HamOrSpam-Classifer 🌐 [codexshami/HamOrSpam-Classifier](https://github.com/codexshami/HamOrSpam-Classifier) | Python, Scikit-learn, TF-IDF, SMOTE | May 2025 – Aug 2025

- Built email spam classifier using TF-IDF and Logistic Regression with SMOTE balancing, achieving >98% accuracy with robust recall on imbalanced datasets.

FarmAIQ — Smart Agriculture Recommendation System 🌐 [codexshami/FarmAIQ](https://github.com/codexshami/FarmAIQ) | Python, Random Forest, CNN | Nov 2025 – Dec 2025

- Designed crop recommendation and disease detection pipelines; applied cross-validation and tuning to improve generalization on public agricultural datasets.

Education

B.Tech in Data Science | *Teerthanker Mahaveer University, Moradabad, UP*

Expected 2027

CGPA: **8.5/10** | **1st Rank in Academic Cohort**

Certifications

IBM: Data Analysis & Problem Solving (2026) | **ISRO:** Scientific Observation from Space (2026) | **Simplilearn:** Deep Learning & ML with Python (2025)

HackerRank: Python, SQL, Problem Solving (2025) | **ITM EdTech:** Advanced Data Analysis — Excel & Power BI (2025) |

IIT Bombay: Intro to Computer Science (2024)

Key Achievements

Smart India Hackathon (SIH): Selected for college round | **India AI Impact Summit:** Presented project | Solved 850+ DSA problems in Python | **DUCAT School of AI:** Industrial visit in Data Science

Languages: Hindi, English, Urdu, Arabic