SACHIN MAURYA

Data Scientist | Statistical Modeling | Time-Series & Predictive Analytics

sachin.mauryaa113@gmail.com | +91-8528187710 | GitHub | LinkedIn | Portfolio

Professional Summary

Data Scientist with hands-on experience building end-to-end predictive analytics solutions across time-series and medical imaging domains. Strong foundation in statistical modeling, hypothesis testing, and feature engineering, with proven ability to improve model performance through structured experimentation. Experienced in translating analytical findings into actionable insights to support predictive maintenance and healthcare decision-making.

Technical Skills

- **Programming & Data Manipulation:** Python, SQL, NumPy, Pandas
- **Statistics & Analytics:** Descriptive/Inferential Statistics, Hypothesis Testing, Probability, Correlation Analysis, A/B Testing Fundamentals
- **EDA & Visualization:** Data Cleaning, Feature Engineering, Matplotlib, Seaborn, Streamlit Dashboards
- **Machine Learning:** Regression, Classification, Clustering, Random Forest, XGBoost, SVM, Model Selection, Cross-Validation, Hyperparameter Tuning
- **Deep Learning:** CNN, LSTM (PyTorch, TensorFlow/Keras)
- **Time-Series & Signal Processing:** FFT, Statistical/Time/Frequency Features, Trend Analysis, Anomaly Detection
- **Model Evaluation:** Precision, Recall, F1-Score, ROC-AUC, Confusion Matrix, Bias-Variance Analysis
- **Tools:** Git, Jupyter Notebook, MLflow, FastAPI (Basics)

Professional Experience

AI Engineer Intern – Predictive Analytics & Maintenance Jan 2025 – Dec 2025
NSTL–DRDO | IIT Roorkee Collaboration

- Conducted exploratory data analysis on high-frequency vibration time-series data and engineered statistical and FFT-based features, improving fault classification ROC-AUC to above 0.90.
- Developed and compared ML/DL models (Random Forest, XGBoost, CNN, LSTM) and built interactive dashboards to support data-driven maintenance decisions and early fault intervention.

Data Science Intern – Medical Imaging Analytics Feb 2025 – Aug 2025
QuickCURD (Remote)

- Performed preprocessing and exploratory analysis on CT and X-ray datasets, identifying data inconsistencies and improving overall dataset quality.
- Built and validated deep learning classification models, enhancing F1-score and ROC-AUC through structured augmentation and cross-validation techniques.

Key Projects

Predictive Maintenance – Centrifugal Pump Vibration Analysis GitHub

- Designed a complete analytical workflow including signal processing (FFT), statistical feature engineering, and multi-model comparison (RF, XGBoost, SVM, LSTM) for equipment health prediction.
- Visualized degradation patterns and feature importance to enable interpretable insights for condition-based maintenance planning.

Human Activity Recognition – UCI HAR Dataset GitHub

- Conducted EDA and time-series feature extraction on accelerometer and gyroscope sensor data.
- Built classification models using cross-validation, achieving **96.7% accuracy** with strong generalization.

Publications

Vibration Fault Classification – ML Research (VETOMAC 2025) GitHub

- Developed an end-to-end vibration fault classification framework using statistical and frequency-domain feature engineering and benchmarked ML models to identify optimal predictive performance.

Education

B.Tech (Honors) – Computer Science & Engineering (Artificial Intelligence) CGPA: 7.6/10
Chhattisgarh Swami Vivekanand Technical University 2022 – 2026

Achievements

- Two Letters of Recommendation from IIT Roorkee faculty for analytical contributions in predictive maintenance.
- 5 ratings in Python and SQL on HackerRank.
- NPTEL Elite certifications in Computer Vision and Management Information System.