Yash Thakare

+91-9309926529 | yashthakare4280@gmail.com | linkedin.com/in/thakare | github.com/TangoCharlieYT | Pune, India

Summary

Data-driven and analytical student with a strong foundation in machine learning, feature engineering, and statistical analysis. Proven ability to build high-accuracy predictive models (over 90%) for stock price and agricultural forecasting. Adept at generating actionable insights from large, complex datasets using Python, Scikit-learn, and TensorFlow. Seeking to apply my skills in developing robust analytics systems and contributing to high-impact projects as an Associate Data Scientist at Motorq.

Technical Skills

Languages: Python, SQL, MATLAB, C++, Java

ML/DS Libraries: TensorFlow, Scikit-learn, Pandas, NumPy, Matplotlib, NLTK, XGBoost

Tools & Platforms: FastAPI, MongoDB, Git, Microsoft Power BI, Google Cloud Platform (Vertex AI)

Techniques: Information Theory, Shannon Entropy, Feature Engineering (PCA, LDA), Time-Series Analysis

Experience

AI Intern Mar 2025 – Present
amasQIS.ai *Remote*

- Analyzed outputs from an Isolation Forest model to identify optimal green hydrogen production sites.
- Designed and integrated interactive map layers using the OpenWeatherMap API to visualize real-time data.
- Processed and structured complex time-series data from fuel cells to prepare for predictive modeling.

Treasurer & Board Member Sep 2024 – Present
AWS Cloud Club, VIT Bhopal *Bhopal, India*

- Managed a $4,000 budget for club operations and led coordination for technical workshops and events.
- Directed recruitment initiatives that increased active membership by over 15% in one semester.

Projects

Entropy Driven Analysis of Misinformation in Crisis Communication Jan 2025
Research Paper *Python, NLTK, Pandas, Scikit-learn*

- Analyzed 5,613 distinct misinformation stories from the COVID-19 pandemic by integrating Shannon entropy with NLP models.
- Quantified information disorder patterns, improving nuanced misinformation detection by 15%.

Crop Yield Prediction Model Inter-Sem 2024
Data Science Project *Python, Scikit-learn, Pandas*

- Implemented and fine-tuned 10+ ML algorithms (e.g., SVM, XGBoost) on 50,000+ agricultural records.
- Applied dimensionality reduction (PCA, LDA), reducing dataset features by 60% while retaining 95% of variance.
- Developed a final ensemble model that achieved 90% predictive accuracy, boosting the F1-score by 15%.

Education

VIT Bhopal University Madhya Pradesh, India
B.Tech in Computer Science (AI & ML) *2023 – 2027*

- CGPA: 9.00/10.00
- Relevant Coursework: Machine Learning, Data Structures & Algorithms, Database Systems

Achievements

- **Qualified** for Round 2 of **Flipkart GRiD 7.0**, Software Development Track (2025)
- **Qualified** for Round 2 of **Unstop Talent Park** (2025)
- Ranked in **Top 60** out of 160+ teams in a National Capture The Flag (CTF) competition (2024)

Certifications

- **Google**: Cloud & Generative AI Study Jam, 15+ Skill Badges (Vertex AI, API Gateway, Cloud Storage)
- **Data Science**: IU Indianapolis (Computational Data Science), Tata Group (Data Visualization)
- **Domain Specific**: ISRO (Geodata Processing), IIRS (Deep Learning in Ecology)
- **Technical**: NPTEL (Cloud Computing), MathWorks (MATLAB Onramp)