Athang Bachhav

+1 (716) 939-8152 | athangvi@buffalo.edu | linkedin

Aspiring Data Scientist with a focus on machine learning and predictive analytics, seeking an internship to apply advanced statistical models and data-driven solutions in real-world business scenarios.

EDUCATION

MPS (M.S), Data Science and Application — Buffalo, NY
University at Buffalo | SUNY — Class of 2025
Coursework : Machine Learning, Statistics, Predictive Modeling, NLP, Database Management

BE in Information Technology, Specialization in Data Science — India
Savitaribai Phule Pune University — June 2018 - June 2022
Coursework : Big Data Analytics, Business Intelligence, Human-Computer Interaction

CORE COMPETENCIES

Machine Learning | Statistical Modeling | Big Data Analytics | Time Series Analysis | Cloud Operations

TECHNICAL SKILLS

Languages : Python, SQL, R
Libraries/Frameworks : TensorFlow, PyMC-Marketing, Scikit-Learn, Pandas, NumPy, OpenCV
Cloud/Data Tools : AWS SageMaker, Google Cloud (BigQuery/GCS), Azure, Hadoop, Parquet Optimization
Visualization : Tableau, Plotly, Matplotlib/Seaborn, MS Excel
IDEs : Jupyter Notebook, RStudio, Git, Colab

INDUSTRY ALIGNMENT

Retail Analytics | E-commerce Systems | Financial Markets | ML Research & Development

PROJECTS

PyMC-Marketing Bayesian Modeling | Retail Analytics | CLV — Sept 2024

- Developed a Bayesian CLV forecasting system using PyMC-Marketing on 400K+ transactions from 2,778 customers, achieving 93% posterior certainty in purchase/dropout rates.
- Identified a 132% CLV gap between France and Belgium, enabling 15% budget reallocation to high-value markets, with top 5% customers generating $19,366 mean CLV (48x baseline).
- Accelerated insights by reducing inference latency 92% (47s → 3.2s) and automating stakeholder reports with Plotly dashboards, cutting analysis time by 65%.

Large-Scale Temporal Recommendation System Pipeline for Amazon Reviews (571M+ Records) — Jan 2025

- Achieved 89.4% validation AUC for next-purchase prediction (vs. 75.2% with random splits)
- Cut cloud compute costs by $2.34/M records through GCS lifecycle policies
- 92% pipeline success rate at 28.4M reviews/day throughput

Cryptocurrency Market Forecasting: Implementing AR/ARMA Time Series Models for Bitcoin Price Prediction — Feb 2025

- Utilized ARMA models in Python (statsmodels 0.14.0) to analyze 366 days of Bitcoin prices ($988-$10,178), achieving 89% stationarity through first-order differencing (Dickey-Fuller p=0.001).
- Designed temporal train-test splits (80-20) with walk-forward validation, testing on 73-day horizon (Feb 2018 peak volatility period).
- Grid-searched ARMA(p,q) parameters (p=2-5, q=1-3) using AIC minimization, identifying optimal (3,2) configuration.

PUBLICATIONS

- Sharma S., Bachhav A.V., et al., Facial Feature Extraction and Emotional Analysis Using Machine Learning | International Journal of All Research Education and Scientific Methods ISSN:2455-6211 | Vol.10(11), Nov 2022 | ML-driven emotion classification
- Saqlain F., Bachhav A.V., et al., Driver Drowsiness Detection Using Artificial Intelligence | International Journal of All Research Education and Scientific Methods ISSN:2455-6211 | Vol.10(12), Dec 2022 | Computer vision system for road safety