

NIKHIL GOKHALE

Buffalo, NY 14214 | +1-716-465-8994 | nikhilgokhale2232001@gmail.com | [LinkedIn](#) | [Github](#)

SUMMARY

Data Scientist with proven experience in predictive modeling, financial data analysis, and automation. Skilled in Python, SQL, and R to deliver accurate forecasts, support FP&A, and uncover actionable business insights. Strong track record of building reporting tools and scalable pipelines to drive data-informed decisions in cross-functional settings.

EDUCATION

- University at Buffalo, SUNY** | Buffalo, NY
Master of Science in Data Science | January 2024 - May 2025
Relevant Coursework: Python Programming, Databases, Statistics, Data Structures & Algorithms, Machine Learning, Data Intensive Computing
- Mumbai University** | Mumbai, India
Bachelor of Technology in Electronics & Telecommunications | August 2019 - June 2023

TECHNICAL SKILLS

- Languages** Python, SQL, R, Bash
- Libraries/ML:** Scikit-learn, XGBoost, TensorFlow, PySpark, ElasticNet, Random Forest
- Data & Reporting Tools:** Tableau, Power BI, Excel, Jupyter, Streamlit
- Big Data & Cloud:** AWS (S3, EC2), Hadoop, Spark, Docker, GCP
- Other:** REST APIs, Git, Linux, VS Code

CORE COMPETENCIES

Forecasting | Financial Modeling | Predictive Analytics | Python | SQL | R | Hive | Tableau | Financial Planning & Analysis | FP&A | Machine Learning | KPI Reporting | Ad Hoc Analytics | Data Wrangling | Risk Modeling | Statistical Analysis | Time-Series Forecasting | Profitability Analysis | Business Intelligence | Cloud Pipelines | Feature Engineering | Model Deployment

PROFESSIONAL EXPERIENCE

Junior Data Scientist

Maxgen Technologies Pvt Ltd., Mumbai, India | July 2023 – January 2024

- Constructed over 10 predictive dashboards in Tableau and Power BI, improving financial performance visibility by 35%.
- Streamlined ETL processes using Python and SQL, shortening data preparation cycles by 40%.
- Collaborated with finance and compliance units to resolve 100+ operational anomalies across monthly reporting flows.
- Published REST API endpoints for seamless delivery of model outputs and financial metrics to internal stakeholders.
- Aligned business reporting systems with strategic KPIs by coordinating across 5+ cross-functional teams.

PROJECTS

Financial Risk Forecasting | *Python, Flask, SQL, Streamlit*

- Modeled credit behavior for 1,000+ loan accounts, embedding business rules into forecasting logic.
- Enhanced model classification accuracy by 22% through parameter tuning of logistic and decision tree classifiers.
- Delivered executive-ready dashboards highlighting metrics like payment ratio, credit utilization, and default risk.

Biological Age Regression Predictor | *PySpark, ElasticNet, Random Forest*

- Processed 250,000+ health records to infer biological age using behavioral and physiological indicators.
- Elevated model accuracy (RMSE ↓18%) and accelerated runtimes by 40% via distributed computing with Spark ML.
- Interpreted feature contributions to prioritize impactful predictors, following explainable AI best practices.

Customer Segmentation for Media Recommendations | *Python, TensorFlow, AWS, Scikit-learn*

- Engineered a collaborative filtering system that improved user recommendation precision by 30%.
- Distinguished 8 audience personas via UMAP and PCA dimensionality reduction to drive targeting strategy.
- Presented findings with visual storytelling tools (Seaborn/Plotly) to support stakeholder decision-making

Socioeconomic Impact on Health Analytics | *SQL, PostgreSQL, Tableau*

- Built a normalized schema for 500K+ socioeconomic and public health records to enable faster reporting.
- Created 3 dashboards correlating financial access and obesity rates, revealing regional disparities across 5 clusters.

RESEARCH & EXTRACURRICULARS

- Published research paper on “Self-driving cars to drive autonomously using Deep Learning” in IJARESM (May 2023).
- Management Head of Journal Committee (May 2022 – May 2023).
- Google Analytics Certified (2021).
- Meta Kaggle Hackathon (2025) | Working on large-scale pattern mining from code metadata.
- Stanford/[DeepLearning.AI](#) | Supervised ML | Regression & Classification | (In Progress).