

Dhanesh B B

Data Scientist | Python, SQL & Machine Learning

Bengaluru, India, +91 80508 35879, dhaneshbb5@gmail.com

LinkedIn: [linkedin.com/in/dhaneshbb](https://www.linkedin.com/in/dhaneshbb), GitHub: github.com/dhaneshbb, Portfolio: dhaneshbb.github.io

SUMMARY

Data Scientist building ML models for classification, regression, and forecasting on large-scale datasets of up to 58.4M records, achieving up to 92.5% accuracy and 0.9756 ROC-AUC. Skilled in Python, SQL, EDA, and ETL pipeline development. Experience spans credit-risk modeling, LLM/GenAI workflows, and document data extraction, publishing Python tools with 11,200+ PyPI downloads and translating technical findings into actionable, client-facing insights.

SKILLS

Python, SQL, R, MySQL, PostgreSQL, Pandas, NumPy, EDA, ETL, Data Cleaning, Feature Engineering, Data Preprocessing, scikit-learn, XGBoost, LightGBM, CatBoost, Regression, Classification, Predictive Modeling, Model Evaluation & Tuning, TensorFlow, Keras, Statistical Analysis, Hypothesis Testing, Statsmodels, SciPy, SPSS, SHAP, ROC-AUC, F1 Score, RMSE, R^2 , Power BI, Tableau, Matplotlib, Seaborn, Excel, Git, GitHub, Docker, Agentic Coding, LangChain, OpenAI API, Prompt Engineering, LLM Integration, GCP, Forecasting, Credit Risk Modeling, Financial Data Analysis, Business Analytics

EXPERIENCE

Data Science Consultant Intern, Rubixe (Think Ahead Innovations Pvt. Ltd.)

Bengaluru, India

Sep 2024 - Mar 2025

- Built an XGBoost lead-quality classifier from MySQL-sourced FicZon sales data on 7,420 leads; achieved 84.74% recall and 0.8106 ROC-AUC for high-potential lead identification.
- Delivered POC and client assignments using Python, SQL, data cleaning, EDA, visualization, and machine learning; communicated findings in supervised client-facing team settings.

Coordinator, Pantaloons (Aditya Birla Fashion and Retail Ltd.)

Bengaluru, India

Apr 2022 - Mar 2023

- Maintained 99% transaction accuracy across 100+ daily records; used Excel reporting and checks, reducing errors by 15%.

PROJECTS

Classification Models, Machine Learning

Jun 2024 - Apr 2025

GitHub: [HomeLoanDef](#), [INX-Employee-Performance-Analysis](#)

- Processed 58.4M loan records across 7 datasets and 218 features, reducing memory usage by 68.5% and false positives by 20% for default-risk modeling.
- Achieved 83.0% accuracy and 0.785 ROC-AUC for loan default prediction; achieved 92.5% accuracy, 0.9242 F1, and 0.9756 ROC-AUC for INX performance analytics.

AutoCSV Data Profiling Tools, Python CLI / PyPI Libraries

Dec 2024 - Present

PyPI: [insightfulpy](#), [autocsv-profiler](#); GitHub: [insightfulpy](#), [autocsv-profiler](#), [autocsv-profiler-suite](#)

- Published reusable CSV profiling and EDA tools with 11,200+ combined PyPI downloads across AutoCSV Profiler and InsightfulPy.
- Built multi-engine profiling workflows with 4 isolated profiling engines, 17 statistical metrics, visual reports, chunked processing, and reusable pandas EDA utilities.

StatementForge, PDF Data Extraction & ETL Pipeline

Jun 2025 - Present

GitFront: [StatementForge-demo](#)

- Automated extraction and normalization for 16 bank, credit-card, UPI, and credit-line sources into one schema for cross-institution analysis.
- Handled encrypted PDFs, inconsistent formats, and multi-source consolidation for repeatable statement processing workflows.

Regression Modeling Projects, Machine Learning

Jun 2024 - Mar 2025

GitHub: [AutoPricePred](#), [BikeRental](#)

- Achieved 0.917 test R^2 and 1,987 RMSE for automobile price prediction, selecting an interpretable Lasso model with 29 non-zero coefficients.
- Achieved 0.832 test R^2 and 820 RMSE for bike rental demand forecasting, selecting a Ridge model with a 0.5% overfitting gap for practical forecasting.

JobSync, Next.js AI Workflow App

Dec 2025 - Present

Live: [jobsync.dbbflow.fun](#) (GCP); GitFront: [jobsync-demo](#)

- Built a full-stack AI job-search CRM centralizing applications, documents, interviews, tasks, notes, deadlines, job aggregation, and application review workflows.
- Developed a Manifest V3 job-extraction extension with 47 URL match patterns and AI workflows for resume review, parsing, job matching, cover-letter parsing, and reporting.

Python/Electron Automation Tools

Aug 2025 - Present

GitHub: [spectratract](#), [Multi-AI-Chat-Manager-demo](#)

- Built SpectraTact, a Python-Electron Windows desktop app that launches configured application profiles, detects windows, and arranges workspaces using priority-based grid, side-by-side, and multi-monitor layouts.
- Built a multi-AI workflow app for 12 configured AI platforms with YAML profiles, selective prompt targeting, window controls, and a browser-extension bridge for prompt/output automation.

EDUCATION

Master of Arts in Economics, Bangalore University

Bengaluru

Jan 2022 - Nov 2023

CGPA: 7.15/10 (Grade A) | Coursework: Econometrics, Statistical Methods, Research Methodology

- Masters Thesis:** *Job Satisfaction and Work-Life Balance among IT Professionals* -- conducted a mixed-methods study (205 respondents) using R for statistical analysis.

CERTIFICATIONS

Certified Data Scientist Certification Training, DataMites

Completed Mar 2025

Certified Data Scientist, IABAC (Valid Until Apr 2028) & NASSCOM FutureSkills Prime

Issued Apr 2025