Naveen K

Chittoor, Andhra Pradesh | +91 93925 46035 | sknaveen148@gmail.com
linkedin.com/in/naveen840 | github.com/NaveenKumar1822 | *Open to relocation*

Professional Summary

BCA graduate with project-based experience in Python, SQL, machine learning, and statistical analysis across logistics, financial risk, and credit domains. Built end-to-end data science pipelines covering EDA, feature engineering, predictive modeling, and interactive Streamlit dashboards. Seeking an entry-level Data Scientist or Data Analyst role where I can apply practical ML skills to real business problems.

Education & Certification

Vignana Sudha Degree & P.G College Chittoor, Andhra Pradesh
Bachelor of Computer Applications (BCA) *2022 – 2025*
Relevant Coursework: Statistics, Database Management Systems, Programming in Python, Mathematics

Advanced Data Science & AI Certification Feb 2025 – Aug 2026
Intellipaat in collaboration with IIT Roorkee *Verify on LinkedIn*

Projects

Nassau Candy Logistics Intelligence *Python, Pandas, Streamlit, Plotly, Parquet*
Candy Manufacturer Supply Chain Analytics | github.com/NaveenKumar1822

- Analyzed 10,194 shipment records across 196 lanes and 5 factories to diagnose logistics performance gaps and SLA compliance failures.
- Identified a 46.5% SLA breach rate; surfaced operational patterns across factories, routes, and shipment characteristics to support corrective decision-making.
- Built an interactive Streamlit application with Plotly visualizations to monitor shipment performance, delivery trends, and SLA compliance; delivered reproducible pipeline for ongoing logistics monitoring.

Realized Volatility Forecasting for Options Risk Management *Python, Pandas, NumPy, Statsmodels, LightGBM*
Financial Time Series Forecasting | github.com/NaveenKumar1822

- Developed a time-series forecasting pipeline to estimate near-term realized volatility of the XLE ETF using 14 years of daily OHLCV data (3,522 trading days, 2010–2023); engineered volatility-based features and analyzed temporal patterns including autocorrelation and rolling statistics.
- Benchmarked persistence, ARIMA, and LightGBM models using time-aware cross-validation; LightGBM achieved the lowest RMSE across out-of-sample evaluation windows.
- Framed modeling as a risk-estimation problem (realized volatility) rather than price direction; identified the Garman-Klass range-based volatility estimator as the top predictive feature, ahead of lagged RV terms, producing a practical framework for options-side risk assessment.

Horizon – Probability of Default Modeling *Python, Pandas, Scikit-learn, SQL, Streamlit*
Consumer Credit Risk – ML Classification Pipeline | github.com/NaveenKumar1822

- Built a machine learning pipeline to estimate borrower Probability of Default (PD) for consumer credit risk using application and bureau data, achieving a calibrated Logistic Regression model with ROC-AUC of 0.7415 (Gini 0.4831, KS 0.3778) that outperformed a tuned XGBoost benchmark.
- Performed data preprocessing, EDA, feature engineering, and hyperparameter tuning; evaluated models using AUC-ROC, precision, recall, and F1-score across stratified cross-validation folds, and calibrated probability outputs to reduce the Brier Score by 55.6% (0.2082 → 0.0924).
- Trained and compared Logistic Regression and Random Forest classifiers to select the best-performing model for point-of-origination credit decisioning.
- Deployed an interactive Streamlit decision engine mapping predictions into a 10-tier Basel risk framework (1.42%–34.12% observed default rate) to demonstrate real-time borrower risk scoring, expected-loss estimation, and model prediction explanations.

Technical Skills

Programming & Data Analysis: Python, SQL, Pandas, NumPy, EDA, Feature Engineering, Hypothesis Testing, Statistical Analysis
Machine Learning: Scikit-learn, Regression, Classification, Random Forest, LightGBM, Cross-Validation, Hyperparameter Tuning, Time Series Forecasting, Model Evaluation (AUC, RMSE, F1)
Visualization & BI: Matplotlib, Seaborn, Plotly, Power BI **Databases:** MySQL **Tools:** Git/GitHub, Streamlit, Jupyter Notebook