Afnitha Abdul Rahman

Aspiring Data Scientist

Kerala, India 8113871649 | afnitha.abdulrahman@gmail.com
LinkedIn – https://www.linkedin.com/in/afnitha-abdul-rahman-b689b727a

Analytical and detail-driven **entry-level Data Scientist** with strong proficiency in machine learning, data preprocessing, and model deployment. Hands-on experience through industry-level capstone projects with tools like **Python, Scikit-learn, CatBoost, and LightGBM**. Demonstrated ability to deliver accurate predictive models with strong business impact. Passionate about data storytelling and using analytics to solve real-world problems.

Education
• Bachelor of Arts in Economics – Mahatma Gandhi University, Kerala (2020–2023), CGPA: 7.9/10
• Higher Secondary – BHSS Perumbavoor, Kerala (2018–2020)
• SSLC – GGHS Perumbavoor, Kerala (2018)

Technical Skills
• **Languages & Libraries:** Python, NumPy, Pandas, Scikit-learn, XGBoost, LightGBM, CatBoost
• **Machine Learning:** Regression, Classification, Clustering, Feature Engineering, Model Tuning
• **Visualization Tools:** Matplotlib, Seaborn
• **Data Handling:** SQL, Excel, Data Preprocessing, SMOTE, Time Series Analysis
• **Tools & Platforms:** Jupyter Notebook, Git, VS Code

Projects & Experience
Data Science Intern
Rubixe AI Solutions | May 2024 – Nov 2024
• Built supervised ML models for real-world datasets (healthcare, finance, gaming).
• Improved model accuracy using feature engineering and hyperparameter tuning.
• Delivered client-ready dashboards and data-driven insights.
  R^2 = 0.93 (Game prediction), Accuracy = 100% (Skin classification)

Projects
Bike Rental Demand Prediction (May 2024 – Jun 2024)
• Created regression model using LightGBM to forecast daily bike rentals.
• Integrated seasonal and weather-based features for improved accuracy.
  Tools: Python, LightGBM, EDA, Regression, Visualization
  R^2: 0.89

Game Winner Prediction (PUBG Dataset) (Jun 2024 – Jul 2024)
• Processed 4.4M+ rows to predict player win percentages using CatBoost.
• Engineered features like kill count, walk distance, and survival rank.
  Tools: CatBoost, Regression, Feature Engineering
  R^2: 0.93 | RMSE: 0.08

Hospital Stay Duration Prediction (Jul 2024 – Aug 2024)
• Developed classification models to predict patient stay lengths.
• Used GridSearchCV and handled class imbalance with SMOTE.
  Tools: Gradient Boosting, Random Forest, Logistic Regression, Healthcare Data
  Accuracy: 83%

Skin Disorder Classification (Aug 2024 – Sep 2024)
• Built multiclass classifiers for early disease detection.
• Applied SMOTE to fix imbalance and achieved near-perfect performance.
  Tools: Logistic Regression, Random Forest, SMOTE, Classification
  F1 Score: 1.00 | Accuracy: 100%

Credit Score Prediction (Oct 2024 – Nov 2024)
• Developed predictive models for creditworthiness analysis.
• Achieved a Gini Coefficient of 1.0 using Gradient Boosting.
  Tools: SMOTE, Gradient Boosting, Financial Analytics
  Outcome: Accurate credit risk insight

Certifications
• Certified Data Scientist - DataMites (2024)
• Introduction to Generative AI - LinkedIn Learning (2025)
• Business Analysis Foundations – IIBA® (Jun 2025)
• Internship Certificate - Rubixe AI Solutions (2024)
• Career Skills in Software Development – LinkedIn Learning, 2024

Soft Skills
Communication | Critical Thinking | Problem Solving | Time Management | Collaboration