Abhishek Mule
Aspiring Data Scientist
Pune Maharashtra
+91 9307208070 muleabhi553@gmail.com
Linkedin:- https://www.linkedin.com/in/abhishek-mule-ds/ **GitHub:-** https://github.com/AbhishekMule03

Profile Summary
Motivated individual with strong grasp of supervised and unsupervised machine learning models. Proficient in optimization algorithms and data science applications. Skilled in analyzing complex datasets to generate actionable insights. Eager to collaborate with teams in analyzing large datasets for meaningful insights.

Education
Maharashtra Mahavidyalaya, Nilanga
Bachelor of Computer Applications (BCA) 2021-2024 CGPA : 9.27
Commerce (12th)
Shree Chhatrapati Shivaji College , Omerga (82.83%) 2019-2021

Work history
Data Science Consultant, Intern | Rubixe (June 2024 – December 2024)
- Collaborated with the team members to understand client requirements and business challenges. Worked on a client project where machine learning algorithms like decision tree , xgboost , random forest algorithm were applied and achieved 95% test accuracy. Played a key role in the project team and ensured timely submission of the projects.
- Constructed predictive models using machine learning algorithms , including Logistics Regression and Multiple Linear Regression , achieving an accuracy rate 85% in attributing revenue and conversions to specific marketing touchpoint.

Data Science Intern | Elite Tech Intern(Jan 2025 – March 2025)
- Hands-on experience in data analysis, machine learning, and data visualization.
- Improved my problem-solving skills through practical project work.
- Earned outstanding remarks for my contribution and performance.

Skills
- **Power BI** - Reports , Dashboards , Extract Transform and load (ETL)
- **Machine Learning Algorithms** - Linear Regression , Logistics Regression , Decision Tree , SVM , Naïve Bayes , KNN , K-means , Random Forest , Algorithm , Gradient Boosting and Ada Boosting.
- **Python** – Functions , datatypes , Loop , Conditions
- **Pandas** - Pandas Functions
- **Numpy** - Numpy Functions
- **SQL** – CURD Operation
- **Excel** – Excel Formula And Pivot Table

Certificate
- **Datamites** - **Certified Data Science** :- Gained Comprehensive knowledge in Data Science , including Machine Learning , Data Analysis , and Python programming , through hands – on projects and industry-relevant curriculum.
- **IABAC** - **Certified Data Science** :- Demonstrated expertise in Data Science concepts , Statistical Modeling and data Visualization techniques , with a focus on practical applications and advanced methodologies.
- **Satish Dhawale** - **Power BI** :- Acquired advanced skills in Power BI for creating interactive dashboards , data storytelling and business intelligence solutions.
- **NASSCOM** - **Certified Data Science** :-This certification, supported by the Ministry of Electronics & IT, has enhanced my skills in data science, analytics, and Al-driven decision-making. Looking forward to applying these insights in real-world projects!
- **Infosys Springboard** - **Python for Data Science** :- course offered by Infosys Springboard. The course covered key concepts in Python programming, data analysis, and visualization techniques, providing a strong foundation for data science and machine learning applications.

- **WsCube Tech - 'Master SQL Basics in 90 mins'** :- Masterclass by WsCube Tech! Looking forward to applying these skills in real-world projects.

Projects | Rubixe , Intern (June 2024 – December 2024)

- **Auto Imports Predictions**
  Conducted an in-depth analysis of automobile import datasets, identifying key trends, seasonality, and anomalies. Developed and fine-tuned predictive models using machine learning algorithms like Random Forest and XGBoost to enhance the accuracy of forecasting import volumes. Improved import scheduling efficiency by 20% through optimized model predictions, leading to reduced delays and cost savings. R^2 Score: 93% (for the regression-based forecast model)

Skills applied: supervised learning algorithm | data pre-processing | Regression | predictive modelling | model evaluation.

- **Indian Liver Patients predictions**
  Analyzed clinical datasets of liver disease patients to understand factors influencing liver health. Developed classification models using machine learning algorithms (Logistic Regression, Decision Trees , Random Forest) to predict the risk of liver disease with improved accuracy of 79%. Employed Python libraries like Pandas, NumPy, and Scikit-learn for data preprocessing, feature engineering, and model building to ensure the robustness of predictions. Accuracy: 79% (for classification model)

Skills applied: supervised learning algorithm | data pre-processing | classification | predictive modelling | model evaluation.

- **Texas Salary predictions**
  Analyzed large datasets to uncover patterns in salary distribution across multiple sectors in Texas. Built interactive data visualizations and dashboards using Matplotlib, Seaborn, and Plotly, enabling stakeholders to explore the salary trends across industries. Generated actionable insights related to income disparities and sector-wise compensation trends, supporting HR and finance decision-making. R^2 Score: 90% (for regression-based salary prediction model)

Skills applied: supervised learning algorithm | data pre-processing | Regression | predictive modelling | model evaluation.

- **Heart Diseases predictions**
  Developed a heart disease prediction system based on clinical data, utilizing machine learning techniques such as Random Forest and Gradient Boosting Naive Bayes. Enhanced model accuracy by optimizing hyperparameters, achieving an overall prediction accuracy of 94%. Created comprehensive documentation of model performance and project workflow, aiding future improvements and potential deployment for healthcare applications. Accuracy: 94% (for classification model)

Skills applied: supervised learning algorithm | data pre-processing | classification | predictive modelling | model evaluation.

Projects | Elite Tech Intern , (Jan 2025 – March 2025)

- **Email Spam Detection**
  Machine Learning Approach for Spam Detection
  Data Preprocessing – Text cleaning and vectorization (TF-IDF, Count Vectorizer)
  Feature Engineering – Identifying important words and patterns
  Model Training – Using Logistic Regression, Naïve Bayes, Random Forest, and Deep Learning for classification
  Performance Optimization – Hyperparameter tuning and model ensembling for better accuracy

Skills applied: Machine learning · Python (Programming Language) · Pandas · NumPy

- **Iris Flower Classification**
  Performed Exploratory Data Analysis (EDA) to uncover patterns in the dataset.
  Used classification algorithms like Logistic Regression, KNN, Decision Tree, and Random Forest to classify iris species.
  Achieved an accuracy of over 95%, optimizing model performance through hyperparameter tuning.
  Data Preprocessing & Feature Engineering
  Supervised Learning Algorithms
  Model Evaluation & Performance Tuning
  Data Visualization with Matplotlib & Seaborn

Skills applied: Python (Programming Language) · Machine Learning · Pandas · NumPy · Matplotlib · Seaborn