MD ARSHAD HUSSAIN

Bangalore

📞 +91 8877020394

✉ arshadattari007@gmail.com 💼 LinkedIn 💻 Github

Education

Ramgarh Engineering College | **2021 – 2024**
B.Tech in Computer Science and Engineering(Lateral Entry) | *Ramgarh, Jharkhand*

AL-Kabir Polytechnic, Jamshedpur | **2018 – 2021**
Diploma in Computer Science Engineering | *Jamshedpur,Jharkhand*

Government High School,Koderma | **2016 – 2017**
Matriculation(State Board) | *Koderma,Jharkhand*

Relevant Coursework

- Data Science
- Python Programming
- Database Management
- Machine Learning
- Data Visualization
- SQL NoSQL Databases
- Statistics for Data Science

Experience/INTERNSHIP

RUBIXE | **Feb 2025 – Aug 2025**
Data Science Intern | *Bangalore, Karnataka*

- Currently undergoing hands-on training in real-world data science projects with a focus on Python, machine learning, and analytics.
- Working on data preprocessing, model building, and performance evaluation under industry mentors.
- Gaining exposure to tools like Pandas, Scikit-learn, and Power BI for data-driven decision making. .

Projects

Insurance Cost Prediction | *Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn* **GitHub**

- Developed a regression model to predict individual medical insurance costs based on features like age, BMI, smoking status, and region.
- Performed data preprocessing, feature encoding, and outlier detection using Pandas and NumPy.
- Visualized feature relationships using Seaborn and Matplotlib to identify key cost drivers.
- Implemented and compared multiple models including Linear Regression, Ridge, Lasso, and Decision Tree Regressor using Scikit-learn.
- Evaluated models using RMSE, MAE, and R^2 score to select the best-performing algorithm.

Flight Fare Prediction | *Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn* **GitHub**

- Developed a regression model to predict flight ticket prices based on features like airline, source, destination, and duration.
- Performed data cleaning and exploratory data analysis using Pandas, NumPy, Matplotlib, and Seaborn.
- Trained and evaluated models including Random Forest and XGBoost using Scikit-learn for optimal accuracy.

INX Future Inc (IABAC Project) | *Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn* **GitHub**

- Developed classification models to predict employee performance levels based on factors such as department, education, experience, and training hours.
- Performed data cleaning and exploratory data analysis using Pandas, NumPy, Matplotlib, and Seaborn.
- Trained and evaluated models including Decision Tree and Random Forest using Scikit-learn for optimal accuracy.

Technical Skills

Languages: Python, SQL
Tools/Visualization: Matplotlib, Seaborn, Plotly, PowerBI, Google Colab, Jupyter Notebook,Pycharm, VS Code, Git/GitHub
Libraries for Data Science / Machine Learning: Pandas, Numpy, Scikit-learn,XGBoost,tensorflow

Certifications

- Certified Data Scientist- Datamites
- Certified Data Science Foundation - IABAC
- Python Programming - Simplilearn
- Certified Data Scientist- IABAC
- Certified Data Scientist- NASSCOM