MD SHAHID AFRIDI

+91-9570268175 | mdshahidafridi9060@gmail.com
LinkedIn | GitHub | New Delhi

Summary

Results-driven **Machine Learning and Data Analytics** with expertise in **SQL, Python, Power BI, MS Excel, and Machine Learning**. Skilled in **Data Cleaning**, **Exploratory Data Analysis (EDA)**, and building **Predictive Models** to drive **data-driven decisions**. Passionate about transforming complex **data into actionable insights** to optimize **business performance**.

Education

- **Maulana Azad National Urdu University, Hyderabad** *2022 - 2024*
  Master's in Computer Applications (M.C.A) Hyderabad , India
  - CGPA: 8.16

Projects

- **Diabetes Prediction Using Machine Learning** *Jan 2024 - Apr 2024*
  Tools: ***Python, Pandas, NumPy, Streamlit, Seaborn, Matplotlib*** [GitHub]
  - Collected and processed **diabetes datasets** to develop an **interactive prediction system**.
  - Implemented **data preprocessing techniques** including **missing value imputation, feature scaling, and outlier detection** to enhance model performance.
  - Built a **Random Forest Classifier** to predict diabetes with a **testing accuracy of 97% and training accuracy of 99%**.
  - Developed an **interactive web application** using **Streamlit** for real-time diabetes prediction based on user inputs.
- **Supervised Regression Price Prediction** *Aug 2024 - Oct 2024*
  Tools: ***Python, Pandas, NumPy, Streamlit, Matplotlib*** [GitHub]
  - Designed and implemented an end-to-end solution for predicting prices across multiple categories: **Electronics** and **Vehicles**.
  - Developed separate **Machine Learning** models (**Linear Regression** and **Random Forest Regression**) for six sub-projects, including **Laptop**, **Mobile**, **Television**, **Camera**, **Car**, and **Bike** price prediction.
  - Utilized diverse datasets from **Kaggle**, performing extensive **data preprocessing** to handle missing values, scale features, and ensure data quality.
  - Built an interactive **Streamlit** web application enabling real-time price prediction based on user-selected specifications.
  - Deployed the project on **Render**, providing seamless access for users to explore price predictions online.
- **IPL Dashboard Using Power BI** *Dec 2024 - Jan 2025*
  Tools: ***Power BI, DAX, Data Visualization, Data Cleaning*** [GitHub]
  - Collected and processed **IPL datasets from Kaggle** to create an **interactive and visually appealing dashboard**.
  - Built key insights such as **team and player performance, stadium rankings, and match statistics** across IPL seasons.
  - Developed **custom visualizations** to display metrics like **average runs per match, top scorers, and top wicket-takers**.
  - Used **Power BI's DAX** to calculate **complex measures** and enhance data interactivity.

Skills

- **Programming Languages :** Python, SQL (MySQL)
- **Libraries :** Numpy, Pandas, Seaborn, Matplotlib
- **Tools :** Power BI, MS Excel
- **Coursework :** Data Science, Machine Learning, Statistical Analysis, Database Management
- **Miscellaneous :** Github, Data Analysis, Data Cleaning
- **Soft Skills :** Critical Thinking, Problem-Solving Ability, Interpersonal skill, Collaboration

Certifications

- **Python professional Certificate (Mindluster)** *Jan 2024*
- **Data Science with Python professional Certificate (simplilearn)** *Feb 2024*
- **Machine Learning for Beginners Using Python (Mind Luster)** *Sep 2024*

Additional Information

- **C.C.F.A. [Certificate in computer Financial Accounting] Tally (ASCENT COMPUTER)** *Oct 2019*