Data Scientist/Data Analyst

Harshdip S. Ghaterao
Contact no. 9604761828
Email: harshdip1804@gmail.com

CARRIER OBJECTIVE:

Summary:

Highly skilled and dedicated Data Scientist with a strong background in statistical analysis, machine learning, and data visualization. Proficient in programming languages like Python and R, with experience in handling large datasets and utilizing advanced analytics techniques. Strong problem-solving abilities, excellent communication skills, and a passion for deriving actionable insights from complex data sets.

Education:

- B.E. in CSE, Solapur University 2019
- CDAC in Big Data Analytics 2021

Skills:

- Statistical analysis and modeling
- Machine learning algorithms (supervised and unsupervised)
- Data preprocessing and cleansing
- Data visualization (Tableau, Matplotlib, etc.)
- Programming languages: Python
- Deep learning frameworks (TensorFlow)
- Database:SQL

Experience:

VINTECH ELECTRONIC SYSTEMS PVT.LTD, PROJECT ENGINEER [APR 2022 - NOV 2022]
- Developed predictive models using machine learning algorithms
 to improve customer   segmentation and enhance marketing strategies.   - Conducted exploratory data analysis and generated actionable insights.
- Collaborated with cross-functional teams to design.
- implement data-driven solutions for complex business problems.
- Working on DOCKER ,K8s,Github,Python .

INNOMATICS RESEARCH LABS, DATA SCIENCE INTERN [OCT 2021 - JAN 2022]
- Conducted statistical analysis on large datasets to identify patterns.
- Created interactive dashboards and visualizations.
- To communicate complex findings to stakeholders.

Projects:

Starbucks Customer Segmentation :

Use tableau for graphical data visualization,considering Starbucks   Dataset to predict customer purchasing habits to perform ETL operations & pySpark and MongoDB for storage of transformed
Data, multiple machine learning algorithms (PCA, K-means) to perform segmentation of customers.Use of tableau for graphical data visualization, considering Starbucks Dataset to predict customer purchasing habits to perform ETL operations & pySpark and MongoDB for storage of transformed data, multiple machine learning algorithms (PCA, K-means) to perform segmentation of customers.

WebServer using Docker :

Docker is a popular container technology for hosting services and applications that virtualizes operating systems and builds lightweight containers
Focused on automation and integration.
You will create a default base image for this project that other developers can use in their production environments.
Experience in orchestrating Docker containers.
Created docker files for easier replication of DEV Environments in local machines.

Certifications:

- edureka:AI/ML workshop [2021]
- IBM:Python for data science [2021]
- Noble Work Foundation:DvOps Tools For Beginners [2023]
- Udemy:Data Science and Machine Learning Basic to Advanced [2023]

DECLARATION:

I hereby declare that all information furnished above are true, complete and correct to the best of my knowledge and belief.