MUSTAFA SHAIK

Data Scientist

+91 7780110507

shaikmustafa572@gmail.com

mustafashaikmd.github.io

Proddatur

EDUCATION

Bachelor of Technology, CSE
Chaitanya Bharathi Institute of Technology
2014 - 2018

Intermediate
Sree Vidya Junior College
2012 - 2014

EXPERTISE

Python, Pandas, NumPy, Flask

Data Analytics,Data Visualization

HTML, CSS

Java, C, C++

SQL

LANGUAGE

English

Hindi

Telugu

Experience

Jan 2022- Present
Innomatics Research Labs
Data Science Intern

- **Data Collection and Preparation**: Gather, clean, and preprocess data from various sources, including databases, APIs, and web scraping, to make it suitable for analysis.
- **Exploratory Data Analysis (EDA)**: Perform EDA to gain insights into the dataset, visualize data distributions, identify patterns, outliers, and correlations using statistical methods and data visualization tools like matplotlib, seaborn, or Tableau
- **Model Development**: Assist in building machine learning models for predictive analytics, classification, regression, clustering, or other data-driven tasks. This involves selecting appropriate algorithms, feature engineering, model training, and evaluation
- **Model Evaluation and Validation**: Evaluate model performance using appropriate metrics and cross-validation techniques. Perform hyperparameter tuning and model selection to optimize model performance.
- **Tool and Technology Proficiency**: Gain proficiency in data science tools and libraries such as Python (NumPy, Pandas, Scikit-learn), R, SQL, and data visualization tools (Matplotlib, Seaborn, Plotly). Familiarize yourself with version control systems like Git and collaboration platforms like GitHub.

2020 - 2021
Sai Rajeswari Institute of Technology
Programmer

- Instrumentation Integration: Interface laboratory instruments and equipment with software systems to automate data acquisition, instrument control, and experiment execution. This may involve working with hardware interfaces, communication protocols, and device drivers.
- Database Management: Design and maintain databases to store experimental data, metadata, and research findings. Implement data management systems for organizing, querying, and retrieving laboratory data efficiently.
- Collaboration and Support: Collaborate with researchers, scientists, and other laboratory staff to understand their requirements and provide technical support. Assist laboratory personnel in troubleshooting software-related issues, optimizing performance, and integrating new technologies
- Training and Education: Provide training sessions and workshops to laboratory staff on the use of software tools, programming languages, and data analysis techniques. Share best practices, tips, and tutorials to empower users to leverage software effectively.