



VATSAL NEGI

Data Analyst

CONTACT

- 9528618806
- Deputy Ganj, Civil Lines, Moradabad
- vatsal.055757@tmu.ac.in

EDUCATION

TEERTHANKER MAHAVEER UNIVERSITY

Master of Computer Applications 2024 - 2026

MAHATMA JYOTIBHA PHULE ROHILHAND UNIVERSITY

Bachelor of Science (Maths) 2020 - 2023

SKILLS

Proficient in Python and SQL for data analysis, visualization, and automation. Experienced with Power BI, Excel, Matplotlib, and Seaborn for creating dashboards and analytical reports. Skilled in applying machine learning algorithms like Random Forest and XGBoost for predictive modeling and insights.

Knowledgeable in MySQL and Git/GitHub, with hands-on experience using Jupyter Notebook, Google Colab, and VS Code. Adaptable to new technologies, with strong problem-solving skills and a focus on delivering accurate, data-driven solutions.

LANGUAGES

Hard Skills: Python, SQL, Power BI, MS Excel, MySQL, Data Visualization (Matplotlib, Seaborn), Machine Learning (Random Forest, XGBoost), Data Cleaning & Preprocessing, Statistical Analysis, Git/GitHub, Jupyter Notebook, Google Colab, VS Code

Soft Skills: Analytical Thinking, Problem-Solving, Team Collaboration, Adaptability, Attention to Detail, Time Management, Effective Communication

ABOUT ME

I am a dedicated MCA student and aspiring Data Analyst with a strong foundation in Python, SQL, Excel, and Power BI. I have hands-on experience in data preprocessing, visualization, and predictive modeling through projects like CareerSphere AI, AirPulse, and AquaSure. I am passionate about leveraging analytical skills to uncover insights, support data-driven decisions, and contribute effectively to organizational growth.

PROJECTS

AQUASAFE: WATER POTABILITY PREDICTION

Built a Streamlit dashboard using Random Forest (68% accuracy) to predict water safety from 9 parameters, providing confidence scores, visual insights, and actionable suggestions. Skills: Python, Data Cleaning, ML, Seaborn, Matplotlib, Streamlit.

AIRPULSE: AIR QUALITY MONITORING (BEIJING)

Forecasted PM2.5 levels using ARIMA and created a Plotly dashboard with 30-day predictions highlighting extreme pollution periods. Skills: Python, Time Series Forecasting, Pandas, NumPy, Plotly, Matplotlib, Seaborn.

ROAD ACCIDENT SEVERITY ANALYSIS

Analyzed 128K+ UK accident records to predict severity and identify high-risk zones using Random Forest (77.79% accuracy) and Power BI dashboards. Skills: Python, ML, Data Cleaning, Power BI, Matplotlib, Seaborn, Git/GitHub.

INTERESTS

- Data Analysis & Visualization:** Exploring datasets, creating dashboards, and deriving insights.
- Machine Learning & AI:** Building predictive models and improving analytics solutions.
- Programming & Automation:** Python, SQL, and scripting to streamline data workflows.
- Problem Solving & Analytics Challenges:** Participating in hackathons, competitions, and real-world data projects.