

Utkarsh Tripathi *Data Science and Analyst(Fresher)*

✉ shreytripathi2004@gmail.com

☎ +91-7905-20-2121

📍 Prayagraj, Uttar Pradesh

🌐 LinkedIn

🐙 Github

🏆 Kaggle

Education

2019 – 2020 Prayagraj, India	10th Standard <i>Army Public School New Cantt</i> Standard Mathematics
2021 – 2022 Prayagraj, India	12th Standard <i>Dwarka Prasad Public School</i> PCM+Physical Education
2022/06 – 2025/06 Prayagraj, India	Bachelors of Science <i>Maa Saraswati Sita Degree College</i> Mathematics major
2024 Online	Master's Program in Data Science <i>Simplilearn, in collaboration with IBM</i> 🔗

Skills

Programming Languages

- **Python:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Tensorflow
- **SQL:** Writing queries, Joins, Subqueries

Machine Learning

- Regression
- Classification
- Clustering

Soft Skills

- Strong analytical thinking and problem-solving skills
- Effective communication and presentation skills
- Collaborative team player

Data Analysis & Visualization

- Data Cleaning and Wrangling
- Exploratory Data Analysis (EDA)
- Tools: Tableau, Power BI, Excel

Deployment

- **Web Frameworks:** Flask, Streamlit for building and deploying ML models
- **Containerization:** Basic knowledge of Docker (container creation and management)
- **AWS:** Deploying models using EC2, S3, and Lambda

Certificates

- Python for Data Science [🔗](#)
- Data Science with Python [🔗](#)
- SQL Certification Course [🔗](#)
- Industry Master Class – Data Science [🔗](#)

Projects

Sales Insights [🔗](#)

- Designed a PowerBI Dashboard to understand AtliQ hardware goods sales trend.
- The Final Dashboard was effective at displaying the sales trend of AtliQ hardware, allowing users to understand the data and make informed decision.
- This dashboard could help in increasing the revenue atleast by 7% in the next Quarter

Employee Turnover Analytics [🔗](#)

- **Data Analysis & Visualization:** Used exploratory data analysis and visualization to identify key drivers of turnover like satisfaction levels, evaluation scores, and workload.

- **Clustering:** Applied clustering techniques to group employees who left based on satisfaction and evaluation to uncover behavioral patterns.
- **Machine Learning:** Developed and trained models (Logistic Regression, Random Forest, Gradient Boosting) to predict employee turnover with high accuracy.
- **Risk Classification:** Categorized employees into risk zones (Safe, Low-Risk, Medium-Risk, High-Risk) based on turnover probabilities.

Family Financial Score Web Application [↗](#)

- **Financial Data Input:** Enabled users to input key financial details such as income, savings, expenses, loan payments, and credit card spending.
- **Score Calculation:** Built a scoring model based on savings-to-income, expenses-to-income, loan-to-income ratios, credit card spending, and discretionary spending.
- **Personalized Recommendations:** Provided targeted advice for improving financial stability based on calculated scores.
- **Data Persistence:** Stored financial data and scores in an Excel file for future reference.
- **User Interface:** Created an intuitive, responsive interface using Bootstrap for data input and viewing.
- **Streamlit Integration:** Developed a Streamlit app to visualize stored financial data and simulate changes in spending to improve financial health.

Online Gaming Behavior Classifier [↗](#)

- Developed a predictive model to forecast engagement levels based on specific features.

Extra Curricular Activities

Tata Data Visualisation: Empowering Business with Effective Insights Job Simulation on Forage - November 2024 [↗](#)

- Completed a simulation involving creating data visualizations for Tata Consultancy Services
- Prepared questions for a meeting with client senior leadership
- Created visuals for data analysis to help executives with effective decision making
- Prepared questions for a meeting with client senior leadership
- Created visuals for data analysis to help executives with effective decision making

BCG Data Science Job Simulation on Forage - November 2024 [↗](#)

- Completed a customer churn analysis simulation for XYZ Analytics, demonstrating advanced data analytics skills, identifying essential client data and outlining a strategic investigation approach.
- Conducted efficient data analysis using Python, including Pandas and NumPy. Employed data visualization techniques for insightful trend interpretation.
- Completed the engineering and optimization of a random forest model, achieving an 85% accuracy rate in predicting customer churn.
- Completed a concise executive summary for the Associate Director, delivering actionable insights for informed decision-making based on the analysis.

AWS APAC Solutions Architecture virtual experience program on Forage - December 2024 [↗](#)

- Designed a simple and scalable hosting architecture based on Elastic Beanstalk for a client experiencing significant growth and slow response times.
- Described my proposed architecture in plain language ensuring my client understood how it works and how costs will be calculated for it.