

RAHUL JAIN

Junior Data Scientist

Phone: +91 98765 32100 | Email:

rahul.jain.ds@gmail.com

LinkedIn: linkedin.com/in/rahuljain-ds | GitHub:

github.com/rahuljain-data

Location: Noida, Uttar Pradesh, India

PROFESSIONAL SUMMARY

Curious and motivated Data Scientist with 1.5 years of hands-on experience in applying machine learning, data analysis, and statistical techniques to solve business problems. Skilled in Python, SQL, and data visualization.

Demonstrated ability to work with real-world datasets, build predictive models, and support business teams with actionable insights. Passionate about learning and continuously improving data-driven decision-making skills.

KEY SKILLS

- ✓ Python (Pandas, NumPy, Scikit-learn)
- ✓ SQL (PostgreSQL, MySQL)
- ✓ Machine Learning (Regression, Classification, Clustering)
- ✓ Data Visualization (Matplotlib, Seaborn, Power BI)
- ✓ Jupyter Notebook, Git, Excel
- ✓ Web Scraping (BeautifulSoup, Selenium)
- ✓ Feature Engineering & EDA
- ✓ Model Evaluation & Tuning
- ✓ Basic NLP & Time Series Forecasting
- ✓ REST APIs (Basic usage with Python)

PROFESSIONAL EXPERIENCE

Junior Data Scientist

DataSolve Analytics Pvt. Ltd., Noida

Feb 2024 – Present

- Built a linear regression model to predict monthly customer spend for a retail client, improving accuracy by 22% over baseline.
- Cleaned and preprocessed large CSV and SQL-based datasets using Pandas and SQL queries.
- Supported senior data scientists in creating classification models for credit risk scoring using logistic regression and decision trees.
- Created dashboards in Power BI to visualize customer segmentation using KMeans clustering.
- Participated in weekly Agile sprint planning and model reviews with the analytics team.

Data Science Intern

InnoTech Labs, Remote Internship

Jul 2023 – Jan 2024

- Worked on a text classification project to categorize support tickets into issue categories using Naive Bayes and TF-IDF.
- Built a small Streamlit web app to demo the classifier to the product team.
- Conducted EDA and built visual summaries using Matplotlib and Seaborn.

PROJECTS

1. House Price Prediction (Kaggle + Real-World Dataset)

- Developed a regression model using Scikit-learn on house pricing data.
- Applied feature engineering, missing value treatment, and hyperparameter tuning.
- Achieved R^2 score of 0.87 on test data.

2. Twitter Sentiment Analyzer (NLP Mini Project)

- Built a sentiment analysis model using Vader and logistic regression.
- Scraped tweets using Tweepy and visualized sentiment trends with Seaborn.
- Deployed using Streamlit for live input demo.

3. Sales Forecasting with Time Series (Side Project)

- Used ARIMA and Prophet to forecast monthly sales for a small retail dataset.

- Evaluated models using MAPE and RMSE.

EDUCATION

Bachelor of Technology (B.Tech) – Computer Science

Dr. A.P.J. Abdul Kalam Technical University (AKTU), Uttar Pradesh

Graduated: 2022 | CGPA: 8.1 / 10

CERTIFICATIONS

- Python for Data Science – IBM (Coursera), 2023
- Machine Learning with Python – freeCodeCamp, 2023
- SQL for Data Analysis – Mode Analytics (Online), 2023
- Mathematics for Machine Learning – Coursera, 2022

TECH STACK

Languages: Python, SQL

Libraries: Pandas, NumPy, Scikit-learn, Seaborn, Matplotlib

Tools: Jupyter, Git, Power BI, Streamlit

Databases: MySQL, SQLite

Cloud: Google Colab, Kaggle Kernels

ACHIEVEMENTS

- ✓ Shortlisted in Analytics Vidhya's "DataSprint 88" hackathon
- ✓ Published a blog post on Medium: "Beginner's Guide to Regression Modeling" (5,000+ views)
- ✓ Consistently rated 4.5+/5 in internship feedback reviews

PERSONAL DETAILS

Date of Birth: 12th June 2000

Languages Known: English, Hindi

Nationality: Indian

Address: B-203, Galaxy Apartments, Sector 62, Noida – 201301

DECLARATION

I hereby declare that the information above is true to the best of my knowledge.

Date: 4th October 2025

Place: Noida