AMAN KUMAR (DATA SCIENCE)

+91 8377898127 ◇ Greater Noida, India

atomcode2002@gmail.com ◇ linkedin.com/in/aman-kumar-ml ◇ github.com/AtomCode01

PROFESSIONAL SUMMARY

Detail-oriented Data Scientist with a passion for leveraging machine learning, statistical analysis, and data visualization to solve real-world problems. Proficient in building end-to-end pipelines, automating workflows, and extracting actionable insights from large datasets. Skilled in Python, SQL, and AWS, with proven success in improving model performance and reducing deployment time. Strong collaborator committed to driving data-driven decision-making.

CORE COMPETENCIES

- **Programming:** Python (NumPy, pandas, scikit-learn), SQL
- **Data Analysis:** Data Visualization (Matplotlib, Seaborn), ELT Pipelines, Statistical Analysis
- **Machine Learning:** Model Development, Evaluation, Deployment
- **Cloud Technologies:** AWS (S3, EC2), Docker, CI/CD Pipelines, MLflow
- **Tools:** Flask, Git/GitHub, Jupyter Notebooks
- **Soft Skills:** Problem-Solving, Communication, Collaboration, Critical Thinking

PROFESSIONAL EXPERIENCE

Data Science Intern   *CODTECH Solutions (Remote)*
Dec 2024 – Present

- Analyzed large datasets and developed predictive models to enhance data-driven decision-making.
- Automated data preprocessing pipelines, reducing manual effort by 25% and improving efficiency.
- Collaborated with cross-functional teams to implement machine learning solutions for real-world business problems.

Data Science Intern   *Cognify Technologies (Remote)*
Dec 2024 – Jan 2025

- Built machine learning models to predict user behavior, achieving a 12% improvement in prediction accuracy.
- Conducted exploratory data analysis (EDA) and feature engineering to optimize model performance.
- Integrated machine learning models into production systems within an Agile team environment.

PROJECTS

Student Performance Predictive Tool →

- Built a machine learning model with 81% accuracy to predict student performance metrics.
- Developed and deployed an end-to-end pipeline using Flask, reducing data processing time by 20%.
- Implemented cross-validation techniques to minimize overfitting and enhance model reliability.

ML Pipeline Automation with MLflow →

- Automated ML pipelines on AWS, integrating MLflow for real-time tracking and monitoring.
- Streamlined deployment using CI/CD workflows, reducing deployment time by 15%.
- Conducted rigorous model testing to ensure stability and accuracy.

NYC Airbnb Data Analysis →

• Preprocessed and analyzed Airbnb data using ELT pipelines, improving data efficiency by 20%.

• Delivered actionable insights on pricing and occupancy trends to support data-driven business strategies.

• Visualized data insights with Python libraries, enhancing stakeholder understanding.

EDUCATION

Bachelor of Computer Applications, Swami Vivekananda Subharti University, India   2024
Relevant Coursework: Data Structures, Machine Learning, Data Science, Algorithms

CERTIFICATIONS

Machine Learning Engineer Certification – Udemy   Nov 2024
Certified Data Analysis Specialist – Udemy   Oct 2024

PORTFOLIO AND LINKS

Portfolio: Portfolio