

GAJENDRA SHARMA

Data Science & Machine Learning

+91-9258851367 • gajendra.sharma365777@gmail.com • [LinkedIn](#) • Mathura, India

PROFESSIONAL SUMMARY

Final-year BCA student (CGPA 8.1/10) with 400+ hours of structured training in Python, SQL, machine learning, and business analytics. Delivered two end-to-end ML projects — a customer churn predictor (82% accuracy, 0.78 AUC) and a K-Means segmentation model — complete with interactive Power BI dashboards and reproducible code on GitHub. Seeking a Data Science internship to apply analytical and modelling skills to real business problems.

TECHNICAL SKILLS

Languages	Python, SQL (MySQL, PostgreSQL)
ML / Data Science	Scikit-learn, XGBoost, Pandas, NumPy · Logistic Regression, Random Forest, Decision Trees, K-Means, Feature Engineering, EDA
Statistics	Hypothesis Testing, A/B Testing, Probability, Descriptive Statistics, Confidence Intervals, Statistical Inference
Visualisation	Power BI, Tableau, Matplotlib, Seaborn, Microsoft Excel
Tools	Git & GitHub, Jupyter Notebook, VS Code, Streamlit, Google Colab

PROJECTS

Customer Churn Prediction

Mar 2025 – Apr 2025

Tech: Python · SQL · Scikit-learn · Power BI · Pandas

- Engineered features on 7,043 customer records (Telco dataset); resolved class imbalance, performed null imputation and label encoding — eliminating all missing-value leakage prior to model training.
- Trained Logistic Regression and Random Forest classifiers; achieved **82% accuracy** and **0.78 AUC** — identified contract type and tenure as the top two churn drivers via feature importance analysis.
- Built an interactive Power BI dashboard with churn KPIs, contract-segment drill-down, and cohort analysis; derived 3 data-backed retention strategies for the high-risk month-to-month segment.
- Automated the EDA pipeline using Pandas Profiling, cutting manual exploration time by ~60% for future dataset iterations.

Customer Segmentation using K-Means Clustering

Feb 2025 – Mar 2025

Tech: Python · Scikit-learn · Seaborn · PCA

- Applied K-Means clustering on e-commerce behavioural data; selected optimal $k=4$ using the Elbow Method and validated with a **Silhouette Score of 0.62** — an 11% improvement over the $k=3$ baseline.
- Executed end-to-end preprocessing: MinMax scaling on 6 numeric features, one-hot encoding of 4 categorical variables, and PCA visualisation confirming well-separated cluster boundaries.
- Produced 4 actionable customer profiles (High-Value, At-Risk, New, Dormant) with targeted marketing recommendations; simulated acquisition cost savings of 15–20% for the Dormant segment under proposed strategy.
- Documented the full pipeline in a structured Jupyter Notebook with step-by-step commentary suitable for non-technical stakeholder review — hosted publicly on GitHub.

EDUCATION

Bachelor of Computer Applications (BCA)

Rajiv Academy for Technology and Management, Mathura

CGPA: 8.1 / 10

Aug 2023 – Jun 2026

Relevant Coursework: Data Structures · DBMS · Machine Learning · Statistics · Data Analytics · Business Intelligence

TRAINING

Data Science & Machine Learning

Jan 2024 – Present

Physics Wallah (PW Skills) · Full-Stack Industry-Grade DS Programme · 400+ Hours

- Completed 400+ hours of structured training covering Python, SQL, Statistics, Machine Learning, and Business Analytics with hands-on industry-style projects.
- Built end-to-end ML pipelines — data ingestion, cleaning, EDA, feature engineering, model training, evaluation, and insight communication — using Pandas, NumPy, Scikit-learn, Power BI, and Git.
- Solved 30+ SQL challenges covering window functions, CTEs, subqueries, and aggregations on real-world datasets.

CERTIFICATIONS & ACHIEVEMENTS

- Data Science and Machine Learning Certification — PW Skills (Physics Wallah)
- Business Email Writing Certification — HP Life
- Maintained CGPA 8.1/10 while concurrently completing 400+ hours of Data Science training — sustained high performance under dual academic load.
- Built and publicly documented 2 end-to-end ML/analytics projects with reproducible code and business-facing visualisations on GitHub.