

Vikash Kumawat

Jaipur, Rajasthan, India +91-9079165646 kumawatvicky2006@gmail.com
linkedin.com/in/vikashkumawat-data github.com/vikash-2006 vikash-2006.github.io

Summary

BCA (Data Science) student and Data Science Intern at Regex Software with 9+ months of hands-on experience building scalable backend systems, end-to-end ETL pipelines, and ML-ready datasets. Proficient in Python, SQL, Flask, and Pandas. Experienced in relational database design, REST API development, and interactive data visualization. Seeking a fresher role in Data Engineering, Backend Development, or Data Analytics at a high-impact organization.

Technical Skills

Languages: Python, SQL, HTML5, CSS3, JavaScript

Backend / Web: Flask, REST API Design, CRUD Operations, Component-Based Architecture

Databases: MySQL, Relational Database Design, JOIN Queries (INNER, LEFT, RIGHT), Foreign Keys, ETL Pipelines, Aggregations (GROUP BY, AVG, COUNT)

Data & Analysis: Pandas, NumPy, Matplotlib, Seaborn, Plotly, EDA, Data Preprocessing, Feature Engineering, Data Visualization

Machine Learning: scikit-learn, Linear Regression, Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), StandardScaler, MinMaxScaler, Label Encoding, SMOTE, Supervised Learning

Tools & Platforms: Git, GitHub, VS Code, Jupyter Notebook, Google Colab, Prompt Engineering, **Linux (Ubuntu/Bash)**

Concepts: Object-Oriented Programming (OOP), Database Management Systems (DBMS), System Design

Experience

Data Science Intern

REGex Software Services

Jaipur, Rajasthan

Sep 2025 – Present (9 months)

- Engineered ML-ready datasets from raw CSVs using Pandas — performing null imputation, label encoding, and categorical feature mapping — delivering clean data pipelines for downstream machine learning workflows.
- Conducted end-to-end EDA on COVID-19 and Titanic datasets using Pandas and NumPy, identifying demographic and geographic patterns through univariate and bivariate analysis across 1,000+ records.
- Developed interactive dashboards using Matplotlib, Seaborn, and Plotly — heatmaps, distribution plots, and donut charts — to present actionable data insights to stakeholders.
- Designed prompt-engineering workflows to automate repetitive data reporting tasks, reducing manual effort by 60% through structured AI instruction pipelines.

Projects

Mitti Cafe — Full-Stack Web Application / Python · Flask · JavaScript · HTML5 · CSS3

2025

- Architected a complete multi-page cafe web application with a scalable flat-file structure and component-based integration system spanning 6 pages (menu, contact, login, experience, subscribe, index).
- Engineered a reusable `components.js` module that auto-injects navigation and footer elements across all pages at runtime — eliminating code duplication and enabling single-point UI updates.
- Built a Flask (Python) backend to handle server-side routing, form submissions, and dynamic data serving with clean separation of concerns between frontend and backend.
- Implemented accessibility-first design with skip-links, semantic HTML5 landmark elements, and responsive CSS3 layouts ensuring cross-browser compatibility.

Booking Management System / Python · Flask · MySQL · Docker

2025

- Designed a normalized relational MySQL database schema with multiple tables linked via foreign key constraints to manage bookings, users, and resources at scale.
- Built RESTful backend API endpoints using Flask to handle full CRUD operations for booking records with server-side input validation and error handling.

- Wrote optimized SQL JOIN queries (INNER JOIN, LEFT JOIN) to retrieve booking details with associated user and resource data across related tables.
- Containerized the full-stack application using **Docker** — configured multi-container networking (Flask + MySQL) resolving localhost-to-container hostname issues for production-ready deployment.

User Feedback System / *Python · Flask · MySQL*

2025

- Built a feedback collection web application with referential integrity enforced through foreign key constraints between users and feedback tables.
- Automated feedback summarization using SQL aggregation functions (GROUP BY, AVG, COUNT) to generate structured rating and response-count reports.
- Exposed REST API endpoints via Flask to accept, store, and retrieve structured feedback data, handling JSON request/response cycles end-to-end.

Data Analytics & EDA Pipeline / *Python · Pandas · Plotly · scikit-learn*

2025

- Built end-to-end ETL pipelines to ingest, clean, and transform raw CSV datasets using Pandas — handling null values, encoding categoricals, and removing outliers.
- Performed EDA on Employee Attrition (1,400+ records), COVID-19, and Titanic datasets — uncovering patterns in age, department, and gender via univariate and bivariate analysis.
- Trained scikit-learn classification models on preprocessed data, applying StandardScaler and MinMaxScaler for feature normalization, achieving measurable accuracy improvements.

Supervised ML Models — Churn & Insurance Prediction / *Python · scikit-learn · Flask · Pandas · SMOTE*

2025

- Built a **Customer Churn Prediction API** using Random Forest and Decision Tree classifiers on Churn_Modelling dataset (10,000+ records) — handled class imbalance with SMOTE, achieving significant F1-score improvement over baseline.
- Developed an **Insurance Charge Prediction API** using Linear Regression on structured medical cost data — applied feature engineering, StandardScaler normalization, and served real-time predictions via Flask REST endpoint.
- Trained and compared Logistic Regression, Decision Tree, Random Forest, and SVM models — evaluated using accuracy, precision, recall, and F1-score to select the best-fit model per use case.
- Packaged ML models into Flask APIs and tested end-to-end via Postman — handling JSON input payloads, server-side validation, and structured prediction responses.

AmazonScout — AI Product Research Agent / *Python · Flask · Groq/LLaMA 3.3 70B · ScraperAPI · RainforestAPI · HTML5/CSS3*

2025

- Built a full-stack AI-powered Amazon product research agent that scrapes live product data via ScraperAPI and RainforestAPI, then feeds it to a Groq-hosted LLaMA 3.3 70B model to generate structured research summaries.
- Engineered a Flask backend with modular API routes that orchestrate multi-source data fetching, LLM prompt construction, and JSON response handling in a single pipeline.
- Designed and iteratively refined the HTML/CSS frontend with a clean product research UI — enabling users to enter a product query and receive AI-analyzed insights in real time.
- Integrated prompt engineering techniques to guide LLaMA 3.3 70B in producing consistent, structured outputs including price comparisons, review sentiment, and product recommendations.

DriveX — Car Booking & Feedback Web App / *Python · Flask · MySQL · Docker*

2025

- Developed a full-stack car booking web application during internship at Regex Software — handling end-to-end booking flows, user registration/login, and structured feedback collection.
- Designed a normalized MySQL relational schema with foreign key constraints linking users, bookings, and feedback tables, ensuring referential integrity across all operations.
- Built RESTful Flask API endpoints for full CRUD operations on bookings and feedback, with server-side validation and clean JSON request/response cycles.
- Containerized the application using Docker with multi-container networking (Flask + MySQL) — resolved localhost-to-container hostname issues, enabling consistent production-ready deployment across environments.

- Designed and deployed a fully responsive personal portfolio at vikash-2006.github.io — showcasing projects, technical skills, and certifications with zero external framework dependencies.

Education

Bachelor of Computer Applications (BCA) — Data Science

Commerce College Jaipur

Jaipur, Rajasthan

2023 – 2026 (Expected)

Relevant Coursework: Data Structures & Algorithms, DBMS, Python, Machine Learning, Web Technologies, Statistics

Higher Secondary Certificate (12th) — Commerce

Surendra Public School

Jaipur, Rajasthan

Completed 2023

Certifications & Achievements

- **SQL (Basic) Certification** — HackerRank 2024
- **SQL (Advanced) — Advanced Querying Certification** — HackerRank 2024
- **AI Tools Workshop** — Artificial Intelligence & Automation 2024
- **Google Data Analytics Certificate** — Google / Coursera (*In Progress*) 2025
- **Python, Pandas & Data Visualization** — Kaggle (*In Progress*) 2025
- Publicly documenting a structured day-by-day Machine Learning course on GitHub and LinkedIn — building a consistent, verifiable learning record.