

# ATUL PANDEY

Harla, India 815314 | +91 6287766540 | atul1226@hotmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

---

## Professional Summary

---

Detail-oriented Data Analytics and Data Science professional with expertise in ETL pipelines, predictive modeling, and data-driven decision making. Skilled in Python, SQL, machine learning, and full-stack development, with a track record of transforming complex datasets into actionable insights. Strong background in communication, problem-solving, and operational optimization.

## Technical Skills

---

- **Data Science & ML:** Python, scikit-learn, LightGBM, SQL, ETL Pipelines, NLP, TF-IDF, Feature Engineering, Predictive Modeling, Demand Forecasting
- **Web Development:** Flask, Flask-RESTful, Vue.js, JavaScript, HTML, CSS, Bootstrap, SQLAlchemy, REST APIs, Celery, Redis
- **Tools & Platforms:** Git, GitHub, Jupyter Notebook, Kaggle, MS Excel, MS Word
- **Soft Skills:** Analytical Research, Strategic Thinking, Report Writing, Communication, Mentoring

## Professional Experience

---

### Data Scraping Intern / ETL *Vanel Tech India Private Limited*

- Performed end-to-end ETL operations and data scraping from diverse sources to support analytics initiatives
- Automated workflows for data collection, transformation, integration, and validation
- Maintained strict data confidentiality and prepared analytical reports for stakeholders
- Improved processing efficiency and reduced manual intervention by 40%

### Data Analytics (Education Analytics) *EAT English Data Analytics*

- Tracked student engagement and response accuracy to optimize learning outcomes
- Generated customized learning paths using predictive analytics
- Analyzed voice interactions to improve educational delivery quality
- Created reports supporting curriculum enhancement and decision-making

### Data Analytics Intern *College Dunia*

- Developed predictive models to forecast trends and identify improvement opportunities
- Analyzed operational data to uncover efficiency gaps and optimization strategies
- Worked with modern analytics tools to improve analytical capabilities
- Delivered insights supporting strategic planning and resource allocation

### Data Analytics (Remote) *Al Shajarah, Kuwait*

- Collected and analyzed structured and unstructured datasets from multiple departments
- Prepared reports and presentations communicating business insights
- Collaborated with cross-functional teams to support data-driven decision making
- Recommended operational improvements and optimization strategies

## Projects

---

### Trekking Management System (V2) *Flask, Flask-RESTful, Vue.js, Celery, Redis, Flask-Security-Too, SQLAlchemy*

**GitHub:** [Trekking-Management-System-V2](#)

- Built a full-stack, role-based web application (Trekker, Guide, Admin) for managing trek bookings, guide assignments, payment verification, and reporting
- Designed a decoupled architecture with a Flask-RESTful backend and Vue.js frontend, including a secure payment-approval workflow with automated slot management
- Implemented Redis caching on trek-list queries (50s TTL) with automatic cache invalidation on database writes, improving read performance while preserving data integrity
- Used Celery and Celery Beat for asynchronous CSV exports, scheduled email reports, and Google Chat webhook notifications, keeping the web server non-blocking
- Enforced role-based access and admin-gated account approval using Flask-Security-Too

Comment Category Prediction (Kaggle Project)

Python, LightGBM, scikit-learn, pandas, NumPy

GitHub: Comment-Category-Prediction-using-ML

- Built an end-to-end ML pipeline for multi-class comment classification (4 classes) on a ~198K-row dataset combining text and numerical metadata, achieving a **0.8349 Macro F1 score** with LightGBM (outperforming Logistic Regression at 0.82 and SGD at 0.79)
- Conducted EDA to diagnose severe class imbalance (~55% vs. 4%) and ~73% missingness in demographic fields, dropping unreliable columns and identifying `if_1/if_2` and vote counts as the strongest predictive signals
- Engineered numerical features (log-transformed votes, vote ratio/difference) and temporal features (hour, day-of-week, month) alongside word-level (1–3 gram) and character-level (2–5 gram) TF-IDF text features after cleaning URLs, mentions, and hashtags
- Benchmarked LightGBM against Logistic Regression and SGD classifiers, selecting LightGBM for its ability to model non-linear interactions between text and numerical features while remaining robust to class imbalance and outliers
- Applied cross-validation and hyperparameter tuning for robustness; identified transformer-based models (BERT/RoBERTa) and model stacking as future improvements

Hospital Management System

Flask, Flask-SQLAlchemy, Bootstrap, Python

GitHub: hospital-management-system-v1

- Built a full-stack, role-based hospital management system with dedicated Admin, Doctor, and Patient panels for managing patients, doctors, appointments, and treatments
- Implemented time-slot conflict checking to prevent double-booking of appointments across doctors' schedules
- Developed an Admin panel to add/edit/delete doctors and patients, search appointments and treatments, and view hospital statistics via summary cards and tables
- Built a Doctor panel enabling appointment management (complete/reject), diagnosis and treatment entry, patient medical history lookup, and daily schedule management
- Built a Patient panel supporting self-registration/login, appointment booking, rescheduling and cancellation, visit history, and profile management
- Designed a modular architecture separating Flask views, HTML templates, and SQLAlchemy database models, with a clean, responsive Bootstrap interface

Operational Efficiency Optimization

Python, MS Excel

GitHub: Capstone\_project

- Analyzed sales and inventory data to identify profit stagnation causes
- Applied optimization strategies improving inventory utilization
- Presented actionable recommendations to stakeholders

Education

BS in Data Science and Applications

Indian Institute of Technology (IIT) Madras | 2022 – 2027

Bachelor of Science in Chemistry

Binod Bihari Mahto Koylanchal University, Dhanbad | 2021 – 2024

Intermediate

Kids Garden Senior Secondary School, Dhanbad | 2021

Certifications

- AWS Cloud Foundations | [View Credential](#)

Achievements

- Consistently maintained strong academic performance
- Secured 1st place in Inter-High School Debate Competition
- Successfully led collaborative projects with timely delivery
- Continuously upskilling in data science and analytics technologies