

DataScience With GenAI Project Details

There will be 9 mini projects and 2 major projects and for each concept we will be having a Real time Use Case scenario for better understanding of the Concept

Major Projects

Major Project 1: AI-Powered Profile Onboarder

Overview:

A Django-based Gen AI web app that automates resume parsing and onboarding for job seekers. Users upload PDF resumes, which are parsed using LLMs (OpenAI or LLaMA) to extract structured data. The system creates user accounts, calculates profile scores, and enables job applications.

Core Features:

- Resume parsing via Gen AI (PyMuPDF, Regex, LLM)
- Auto account creation & email notifications
- Profile status lifecycle (Created → Reviewed → Open to Work → Not Interested)
- Editable parsed profile dashboard
- Profile scoring system (0–100) with suggestions
- Job listing & application module
- Admin dashboard with filters & manual overrides
- Security (encrypted PII, role-based access)
- Tech Stack: Django, DRF, MySQL, Bootstrap/Tailwind, Celery, Redis

Major Project 2: AI-Driven Bank Statement Analyzer

Overview:

Automates financial data extraction from bank statements (PDF/scanned) using OCR

and LLMs. Classifies transactions into categories, identifies recurring payments, and detects anomalies for lending, auditing, and financial analysis.

Core Features:

- OCR-based text extraction
- AI classification of transactions (salary, rent, utilities)
- Recurring payment detection
- Fraud and anomaly detection
- Financial health summary
- Adaptability to multiple formats via Gen AI

Mini Projects

Mini Project 1: Analysis of Capitals and Universities Data

Overview:

Analyzes CSV data of countries and universities to explore geographical and educational trends.

Core Features:

- Load & merge two CSV datasets
- Correlate national rank, cost, and location
- Provide insights for global education patterns

Mini Project 2: Online Shopping System

Overview:

Simulates a basic shopping system using Python OOP.

Core Features:

- Add, search, remove, update products in cart

- GST calculation & invoice generation
- Final order summary with subtotal and tax

Mini Project 3: Analytic American Association

Overview:

A law firm management system to track clients, lawyers, tasks, and billing.

Core Features:

- Client & employee (lawyer) profiles
- Billable task tracking with time sheets
- Invoice generation and 8-hour/day validation
- Efficient legal workflow automation

Mini Project 4: Sales Data Insights

Overview:

Performs end-to-end analysis on sales data using Pandas and NumPy.

Core Features:

- Load CSV sales data
- Sales summary, product trends, and customer behavior
- Country & time-based analysis
- Useful for business insights and forecasting

Mini Project 5: Melbourne Housing Snapshot

Overview:

Analyzes real estate trends from Melbourne's property sale dataset.

Core Features:

- Load & process real estate data
- Analyze price trends, location impacts
- Identify best-value suburbs
- Great use case for real estate data science

Remaining Mini Projects will be given as Assignments

Mini Project 6. Decision Tree Classifier on Moon Dataset.

Mini Project 7. Ensemble on MNIST Dataset

Mini Project 8. Titanic Dataset - Classification Task

Mini Project 9. SVM Classifier on MNIST Dataset