

Yash Sharma

9461572488 | syash2297@gmail.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

AI/ML Engineer and MCA student focused on building practical NLP, machine learning, and GenAI systems. Skilled in Python, SQL, scikit-learn, FastAPI, and end-to-end ML workflows, with hands-on projects in text classification, document clustering, topic modeling, and retrieval-augmented generation. Strong foundation in data analysis, model evaluation, API development, and deployment-ready ML systems.

WORK EXPERIENCE

Celebal Technologies

Data Science Intern

May 2025 - Jul 2025

Jaipur, Rajasthan

- Built an unsupervised NLP pipeline for document clustering and topic modeling using TF-IDF, K-Means, and LDA to extract themes from unstructured text data.
- Developed text preprocessing, feature extraction, model evaluation, and Streamlit visualization workflows to support faster document analysis and insight generation.

TaskUs

Teammate

Feb 2022 - Sep 2023

Gurugram, Haryana

- Delivered 95%+ CSAT by analyzing recurring customer issues and improving support workflows, reducing repeat cases by 15%.

Concentrix

Customer Service Representative

Apr 2021 - Feb 2022

Gurugram, Haryana

- Reduced repeat contacts by 15% through root-cause analysis of high-frequency customer issue patterns.

EDUCATION

Poddar Management and Technical Campus

Master of Computer Application

Aug 2024 - Present

Jaipur, Rajasthan

Singhania University

Bachelor of Computer Application

May 2020 - Jul 2023

Jhunjhunu, Rajasthan

SKILLS

Programming: Python, SQL, JavaScript, TypeScript

Machine Learning: Scikit-learn, XGBoost, LightGBM, Feature Engineering, Cross-Validation, Model Evaluation

NLP & GenAI: Text Classification, TF-IDF, Topic Modeling, Document Clustering, LDA, RAG, Prompt Engineering

Data Analysis: Pandas, NumPy, EDA, Matplotlib, Streamlit

Frameworks & Tools: FastAPI, React, Next.js, Node.js, Git, GitHub, GitHub Actions, Docker

Databases: SQL, PostgreSQL, SQLite, Firebase/Firestore

PERSONAL PROJECTS

Customer Inquiry Classifier - NLP Routing System

Python | Scikit-learn | FastAPI | Streamlit | GitHub Actions

- Engineered an end-to-end NLP classifier using TF-IDF, calibrated LinearSVC, and Logistic Regression to route customer inquiries across support categories.
- Built FastAPI prediction endpoints with confidence-based routing, explainable TF-IDF keywords, and low-confidence human-review fallback.
- Added Streamlit demo, Docker-ready setup, and GitHub Actions testing to make the ML system reliable, testable, and deployment-ready.

TraceRAG - Retrieval-Augmented Generation System

FastAPI | Next.js | PostgreSQL/pgvector | Hybrid Retrieval | Docker

- Built a full-stack RAG system with document ingestion, vector + keyword search, and citation-grounded AI response generation.
- Designed query tracing, access-control foundations, health-ready API endpoints, and evaluation workflows for faithfulness and relevance.

ACHIEVEMENTS & INTERESTS

- Prime AI/ML Batch, Apna College: Python, ML, DL, GenAI, and AI Engineering Stack.
- Completed GenAI Mastermind, Outskill: LLM workflows, AI agents, and automation.
- Built GenAI Trip Planner with Gemini/Vertex AI at Google Gen AI Exchange Hackathon 2025.