

Sahil Singh

[Portfolio](#) | [LinkedIn](#) | [Github](#)

+91-7505085844 | sahilsingh82738235@gmail.com | Gurgaon

AI Engineer with experience building production-grade **Generative AI, Computer Vision, Retrieval-Augmented Generation (RAG), and multimodal AI systems.**

EXPERIENCE

Artificial Intelligence Engineer – Radiansys Inc. (Feb 2026- Present)

- Developed real-time computer vision pipelines for face detection, tracking, and recognition using OpenCV and YOLO models.
- Built AI-powered sports highlight generation systems using event detection and multimodal video understanding.
- Fine-tuned and evaluated multimodal models and LLMs including Qwen for domain-specific applications.
- **Tech:** Python, PyTorch, OpenCV, YOLO, FastAPI, Docker, Qwen, Generative AI

Machine Learning Engineer – Nexority (Aug 2025- Feb 2026)

- Developed ML solutions for predictive analytics and business intelligence workflows.
- Built and evaluated classification and NLP models with preprocessing, feature engineering, and optimization pipelines.
- Worked with SQL, Pandas, and Scikit-learn to improve analytical workflows and decision systems.
- **Tech:** Python, SQL, Pandas, NumPy, Scikit-learn

Data Science Intern – KultureHire (Feb 2025 – May 2025)

- Contributed to real-time projects using Python, SQL, and MySQL to deliver insights that supported strategic business decisions.
- Improved efficiency in data workflows and enhanced accuracy in analytical outputs.

PROJECTS

AI Sports Highlight Generation Pipeline

- Built an AI pipeline converting full-length sports videos into automated highlight reels using transcription analysis, LLM event detection, and FFmpeg clip generation.
- Designed workflows for event extraction, quarter detection, and distributed processing.
- Extended support across basketball, football, soccer, cricket, and ice hockey.
- Built orchestration workflows using AWS Lambda, SNS, S3, and NestJS.

Developer/Builder – InstaFuture ([Instafuture.in](#))

- Built InstaFuture, a production-grade AI productivity platform featuring PDF automation, document conversion, AI writing tools, and developer utilities with hybrid local/server processing.
- Architected a multi-tool platform spanning AI workflows, PDF processing, and browser-side utilities with mobile-first UX.
- Designed hybrid browser/backend processing pipelines balancing privacy, performance, and infrastructure cost.

AI-Powered Delivery Route Optimization System ([TransLogi](#))

- Engineered an AI-powered logistics system predicting delivery times and optimizing routes using Random Forest and Google Maps API.
- Built a FastAPI backend with dynamic route optimization and an interactive Streamlit GUI.

Domain-Specific SQL Assistant Using Fine-Tuning

- Fine-tuned an LLM for NL-to-SQL generation on business schemas.
- Built validation layers and schema-aware prompting pipelines.
- Developed APIs for querying and summarized responses.

Smart Search Tool for Analytics Vidhya Courses ([Smart Search](#))

- Developed an NLP-powered search engine using Sentence Transformer embeddings to enhance content discoverability.
- Implemented vector similarity search and data preprocessing to improve user engagement.

TECHNICAL SKILLS

- Languages: Python, SQL
- AI / ML: PyTorch, TensorFlow, Scikit-learn, XGBoost
- Generative AI: LLMs, RAG, LangChain, Prompt Engineering, Embeddings, Vector Search
- Computer Vision: OpenCV, YOLO, Video Analytics, Object Detection
- Backend / Cloud: FastAPI, Docker, AWS Lambda, S3, PostgreSQL, MongoDB
- Tools: Git, GitHub, Hugging Face, FFmpeg

EDUCATION

B.Tech. in Computer Science (Data Science) – Sharda University | 2021–2025

- CGPA: 8.08/10

CERTIFICATIONS & ACHIEVEMENTS

- 2nd Rank – Data Science Hackathon, Sharda University
- Published IEEE paper: “*A Comparative Review of Suspicious Activity Detection Methods in Surveillance Videos*”
- Generative AI with Large Language Models — DeepLearning.AI
- Google Data Analytics Professional Certificate – Coursera (Jul 2024)
- BCG Data Science Job Simulation – Forage (Aug 2024)