Hitesh Kumar

Gurgoan, Haryana, India

📞 +91-7206139480 ✉ hiteshofficial0001@gmail.com 🔗 https://www.linkedin.com/in/hitesh-kumar-aiml/
🐙 https://github.com/Hiteshydv001 🌐 https://hiteshydv001.github.io/Hiteshydv001/

Summary

Results-driven Data Scientist with expertise in machine learning, deep learning, and statistical analysis. Proficient in Python, SQL, and TensorFlow, with hands-on experience in predictive modeling, NLP, and big data technologies. Skilled in data wrangling, feature engineering, and optimizing ML algorithms to drive business growth. Passionate about solving real-world challenges through AI and collaborating with cross-functional teams.

Education

Chandigarh University   **Jun 2022 - Jun 2026**
Mohali, Punjab, India
Bachelor of Engineering in CSE (Artificial Intelligence and Machine Learning) (BE-CSE-AIML)   (CGPA: 7.98)
- **Coursework:** Machine Learning, Deep Learning, Data Structures & Algorithms, Natural Language Processing (NLP), Statistical Analysis, Data Mining, Database Management Systems (SQL, NoSQL), Big Data Analytics, Cloud Computing, Computer Vision, Image Processing, Reinforcement Learning, AI Optimization, Information Retrieval, Time Series Analysis.

Technical Skills

- **Languages:** Python, SQL, R, C++
- **Machine Learning & AI:** Supervised Learning, Unsupervised Learning, Reinforcement Learning, Deep Learning, Generative AI, NLP, Computer Vision, Time Series Analysis, Recommendation Systems, Anomaly Detection
- **Libraries & Frameworks:** TensorFlow, PyTorch, Scikit-learn, Keras, OpenCV, Hugging Face, Pandas, NumPy, Matplotlib, Seaborn, Plotly, Dask
- **Data Science & Analytics:** Statistical Analysis, A/B Testing, Hypothesis Testing, Feature Engineering, Data Wrangling, Exploratory Data Analysis (EDA), Predictive Modeling, Data Visualization
- **Big Data & Cloud Computing:** Apache Spark, Hadoop, Google BigQuery, AWS, Azure, Databricks
- **Databases & Data Storage:** MySQL, PostgreSQL, SQLite, MongoDB, NoSQL, Snowflake
- **MLOps & Model Deployment:** Model Training & Optimization, ML Pipeline Automation, Model Deployment using TensorFlow Serving, Flask, FastAPI, Docker, Kubernetes, CI/CD for ML
- **Development Tools & IDEs:** Jupyter Notebook, VS Code, PyCharm, Git, GitHub, GitLab, Tableau, Power BI

Certifications:

- **Salesforce Certified AI Specialist:** Salesforce
- **Salesforce Certified AI Associate:** Salesforce
- **AI-900: Microsoft Azure AI Fundamentals:** Microsoft
- **GitHub Foundations:** GitHub
- **Postman API Fundamentals Student Expert:** Postman

- **Predictive Analytics using IBM SPSS Modeler:** IBM
- **Advanced Credit Program** - AI Models: Chandigarh University
- **Data Analyst Associate:** DataCamp
- **Multicloud Network Associate:** Aviatrix
- **Generative AI Fundamentals Accreditation:** Databricks
- **Oracle Cloud Infrastructure 2024 Generative AI Certified Professional:** Oracle
- **Oracle Cloud Data Management Foundation Associate:** Oracle

Projects

- **Multi-RAG-Agent Tools:** *(Selected in Open Source Program: JGEC Winter of Code)*

Engineered a Multi-Agent AI System for automated text summarization, research article generation, and data sanitization, improving processing efficiency by 40%. Designed a modular agent-based architecture with 5+ specialized agents, automating workflows with 95% accuracy. Optimized the system to extract, refine, and sanitize large datasets, reducing manual effort by 60%.

Tech Stack: Python, Streamlit, FAISS, Hugging Face, Gemini API, GitHub, NextJS, Flask
🔗 GitHub | 🔗 Live Demo

- **LinkedIn Automate Comment** *(Selected in Open Source Program: JGEC Winter of Code And DWOC)*

Built an AI automation system for LinkedIn interactions, increasing engagement efficiency by 50%. Leveraged web scraping, sentiment analysis, and LLM-based text generation to process 1,000+ posts daily. The system analyzes sentiment, generates AI-driven comments with 90% relevance, and automates posting, enabling professionals to engage effectively with minimal effort.

Tech Stack: Python, Selenium, Gemini API, FastAPI, Streamlit, Vercel, GitHub, Playwright
🔗 GitHub | 🔗 Live Demo

• Waste Classification Model (Organic/Non-organic)
Designed an AI-based waste classification system with 92% accuracy using CNN and YOLO for real-time detection. Deployed via Streamlit for seamless interaction. Selected among the top 60 out of 5,000 projects for a Government of India conference, evaluated by industry experts (AICTE & SHELL).

Tech Stack: TensorFlow, OpenCV, Streamlit, Python, PIL

👓 [GitHub](#) | 👓 [Live Demo](#)

• Guard AI: Designing Remote Proctoring System *(Selected in Open Source Program: JWOC, DWOC, SWOC)*
Created an AI-powered remote proctoring system to secure online assessments with real-time facial recognition, gaze tracking, and lip movement detection. Enhanced exam integrity by monitoring screen activity, audio inputs, and unauthorized device usage. Successfully onboarded 15-20 users for testing and collaborated with 30-40 contributors for development with open source.

Tech Stack: Python, FastAPI, OpenCV, Machine Learning, Streamlit, Vercel

👓 [GitHub](#) | 👓 [Live Demo](#)

Professional Experience

Social Winter of Code | *Project Maintainer & Mentor*   **Dec 2024 – Present | Remote**

Guided contributors in AI-based open-source project, Worked on Flask, Next.js, and MongoDB for AI applications.
Skills: AI, Python, Flask, Next.js, MongoDB

Vodafone Idea Foundation | *Blockchain Intern*   **Jan 2025 – Feb 2025 | Remote**

Implemented a blockchain-based land registry system using Ethereum, Solidity, and Web3.py, ensuring secure, transparent, and immutable land ownership records.
Skills: Blockchain, Solidity, MetaMask, Ganache, Web3.py

African Association of Entrepreneurs | *Web Developer*   **Oct 2024 – Jan 2025 | Remote**

Developed the French version of their website, translating blogs and articles from English to French.
Skills: WordPress, CMS, Web Development

Edunet Foundation (AICTE & SHELL) | *AI/ML Intern*   **Nov 2024 – Dec 2024 | Remote**

Built a waste classification model using CNN & YOLO for real-time detection, One of 40 selected projects evaluated by industry experts.
Skills: CNN, YOLO, Streamlit, Python, Kaggle

Think Blue Data | *AI Research Volunteer*   **Sep 2024 – Oct 2024 | Remote**

Annotated datasets for urban planning and building topology models, Helped in developing AI-based solutions for societal planning in Thailand.
Skills: Label Studio, Data Annotation, AI, Machine Learning

Volunteering

- **Ministry of Social Development and Human Security of Thailand (MSDHS)**   **Accelerator-November 2023 (1 month)**
  Participated in the celebration of International Volunteer Day, promoting volunteerism to achieve Sustainable Development Goals (SDGs) in Thailand also helped them in promotion and inauguration of their website
- **African Association of Entrepreneurs (AAE)**   **Awareness Lead- June 2024 - August 2024 (3 months)**
  Collaborated with a team of 20 volunteers around world to raise awareness about economic empowerment for SMEs
- **African Association of Entrepreneurs (AAE)**   **Publication Lead**
  Led a team of 25 people around the world in producing articles and reports on the entrepreneurship landscape in Africa, enhancing knowledge of SME dynamics and leadership in collaborative settings.
- **African Association of Entrepreneurs (AAE)**   **Awareness Volunteer-July 2024 - August 2024 (2 months)**
  Supported initiatives to raise awareness about economic empowerment for SMEs, fostering collaboration among diverse stakeholders.

Research Work

Train Waitlisted Ticket Confirmation Prediction Using Machine Learning   **ICACCTech 2024 · Nov 17, 2024**
- Built an ML model to predict waitlisted ticket confirmation probability.
- Analyzed historical data for better travel planning.

Weather-Integrated Traffic Routing with Dynamic Speed Prediction   **ICRAAI-2025 · Feb 21-22, 2025**
- Integrated weather data for optimized traffic routing.
- Developed AI-driven speed prediction and path optimization models.

Hobbies

Listening to music
Staying with nature
Running