KISHAN YADAV

Faridabad, Haryana

📞 +91 8826495998   ✉ kishanyadav.29102002@gmail.com   in   ⊙   🌐 Portfolio

Objective

- Innovative Data Scientist with a full-stack AI development mindset and experience in designing end-to-end intelligent systems.
- Proficient in building **LLM-powered applications** using **Python**, **LangChain**, **Streamlit**, and **Gradio**.
- Hands-on with large models including **Mistral-7B**, **GPT-4**, and **LLaMA 3**.
- Skilled in developing **conversational agents**, **real-time computer vision solutions**, and **automation pipelines**.
- Built impactful AI projects like **MockMate**, **EcoVision**, **DreamSpaceAI**, and **StyloGenie** across domains including **LLMs**, **Generative AI**, **Computer Vision**, and **MLOps**.
- Passionate about crafting **user-centric, data-driven products** and driving impactful **AI innovation** in fast-paced environments.

Education

Aravali College of Engineering & Management   **Aug 2020 – Aug 2024**
B.Tech in Computer Science (AI & ML)   *Faridabad, Haryana*

Shree Ram Model School   **Apr 2019 – Apr 2020**
Senior Secondary (12th)   *Faridabad, Haryana*

Experience

Data Science Intern   **Jan 2025 – Jun 2025**
Global Source Technology   *Nashik, India*

- Performed data wrangling, EDA, and feature engineering on structured datasets using **Pandas** and **NumPy**.
- Developed and fine-tuned supervised models like **Logistic Regression**, **Random Forest**, and **XGBoost** using **Scikit-learn**.
- Visualized model performance and extracted insights using **Matplotlib** and **Seaborn**.

Technical Skills

Languages: Python

Tools/Platforms: VS Code, Jupyter Notebook, PyCharm, Spyder, Git, GitHub, MS Excel, Tableau, Streamlit Cloud, Ollama

Cloud & Big Data: Microsoft Azure ML, Azure ML Pipelines, Databricks, PySpark

Database: MySQL

Libraries/Frameworks:

- **Data Science & ML**: NumPy, Pandas, Seaborn, Matplotlib, Scikit-learn, XGBoost, LightGBM, SciPy
- **Deep Learning**: TensorFlow, PyTorch, Keras, Theano
- **Computer Vision**: OpenCV, MediaPipe, YOLO, Haar Cascade Classifier
- **NLP**: SpaCy, NLTK, Gensim, Hugging Face Transformers
- **LLMs & Generative AI**: LangChain, LangGraph, Hugging Face Transformers, GPT-3/4, Mistral-7B (Ollama), LLaMA 3, Gemini, Prompt Engineering, BERT, ChatGPT, AutoGen, Transformers Agents
- **Web & GUI**: Streamlit, Gradio, Tkinter
- **Automation & Others**: PyAutoGUI, Pycaw, SpeechRecognition, BeautifulSoup
- **MLOps**: MLflow, GitHub Actions, Azure ML Pipelines, CI/CD Pipelines — Model versioning, experiment tracking, training pipeline automation, model registry, lifecycle management
- **LLMOps**: LangSmith, LangGraph — Prompt evaluation, LLM observability, agentic flow management, chain debugging, multi-agent orchestration

Concepts & Techniques:

- **Machine Learning (Supervised)**: Linear & Logistic Regression, KNN, SVM, SVR, Naive Bayes, Decision Trees, Random Forest, Gradient Boosting, PCA

- **Unsupervised Learning**: K-Means, Hierarchical Clustering, DBSCAN

- **Deep Learning Architectures**: CNN, RNN, ANN, LSTM

- **NLP Techniques**: Tokenization, Lemmatization, NER, Word Embeddings, Sentiment Analysis, Text Classification

- **Transformers & Agentic AI**: Prompt Engineering, Context-aware Chains, Retrieval-Augmented Generation (RAG), LLM Agents, Tool Use with LangChain

- **Generative Models**: GANs, VAEs

- **Data Science Workflow**: EDA, Feature Engineering, Data Preprocessing, Hyperparameter Tuning, Regularization (L1/L2)

Soft Skills: Analytical Thinking, Problem Solving, Adaptability, Collaboration, Communication, Research-Oriented

Projects

MockMate: Virtual Mock Interview Assistant | *Python, Streamlit, LangChain, LangGraph, Ollama* **GitHub**
- Built an AI-powered mock interview platform generating contextual questions using Mistral-7B via LangChain.
- Integrated voice input, transcription, and real-time scoring through prompt-based feedback pipelines.
- Implemented structured feedback, result export (PDF/JSON), and session-wise progress analysis.
- **Client Approach:** Designed to help job seekers practice interviews with AI-generated coaching and evaluation.

StyloGenie – AI Wardrobe Assistant | *LangChain, Streamlit, Mistral, Pexels API* **GitHub**
- Developed a LangChain agent recommending outfits based on item, color, and occasion using memory-aware prompts.
- Applied color theory for contextual styling and fetched visuals dynamically using the Pexels API.
- Deployed in a responsive Streamlit interface with real-time outfit rendering.
- **Client Approach:** Personalized styling aid for fashion platforms, stylists, and e-commerce wardrobe curation.

EcoVision – Intelligent Trash Sorting | *TensorFlow, Deep Learning, OpenCV* **GitHub**
- Trained MobileNetV2-based classifier on TrashNet dataset to detect 5+ waste categories with 90%+ accuracy.
- Applied image augmentation and optimization techniques to improve generalization and robustness.
- Built a real-time video classification pipeline using OpenCV for smart bin applications.
- **Client Approach:** Positioned as a solution for smart city projects, IoT-driven waste management, and sustainability tech.

DreamSpaceAI – Room Redesign with GenAI | *SDXL, Hugging Face, T2I Adapter, Python* **GitHub**
- Implemented a depth-aware image-to-image pipeline using Stable Diffusion XL with T2I Adapter for style transfer.
- Used Hugging Face diffusers for GPU-accelerated rendering of room redesigns based on user-defined themes.
- Enabled multimodal input (image + voice/text) and exportable outputs with real-time preview.
- **Client Approach:** Interior design previews, architectural mockups, and creative visualization for client engagement.

Certifications

- **NPTEL** — Soft Skills

- **Naresh i Technologies** — Full Stack Data Science and AI

Languages

- Hindi – Native

- English – Proficient