CHANDRAKANT CHOUDHARY

11, Patrakaar Colony — Ashok Nagar Sarkanda Bilaspur (C.G.), 495006

📞 7974473704 ✉ 🔗 🐙

Objective

- Data Scientist with strong practical experience in building AI solutions using real-world datasets and end-to-end system design.

- Proficient in Python-based ML/DL development with hands-on experience in tools like **Gradio**, **Streamlit**, **OpenCV**, and **XGBoost**.

- Built intelligent systems such as *Duplicate Question Detection (Quora)* using NLP + **XGBoost** and *Fight/Anomaly Detection in CCTV Videos* using **YOLOv8** + **ConvLSTM**.

- Experienced with advanced LLMs and frameworks including **GPT-4**, **Mistral-7B (Ollama)**, **LLaMA 3**, **LangChain**, and **LangGraph**.

- Skilled in agentic AI flows, computer vision, automation tools, and full **MLOps**/**LLMOps** pipelines for scalable deployment.

- Passionate about designing intelligent, user-centric, and data-driven applications in fast-paced, innovation-driven environments.

Education

Guru Ghasidas Central Unversity
Masters of Computer Application
November. 2022 – Dec 2024
Bilaspur, Chhattisgarh

Mats Unversity
Bachelor of Computer Application
Sep. 2016 – dec 2019
Raipur, Chhatisgarh

Sanskar Public School
Senior Secondary
May 2016
Raigarh, Chhattisgarh

Relevant Coursework

- Machine Learning
- Deep Learning
- Generative AI
- Agentic AI Systems
- Natural Language Processing
- Computer Vision
- Time Series Forecasting
- Data Visualization Storytelling

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

Duplicate Question Detection | *Python, NLP, XGBoost, Scikit-learn*   **April 2025**
- Built a machine learning model to detect whether two questions from Quora have the same intent or meaning.
- Preprocessed text using techniques like lowercase conversion, stopword removal, and token normalization.
- Created pairwise features such as word overlap ratio, fuzzy matching scores, token count differences, and length differences.
- Used CountVectorizer to convert text into numerical features and engineered additional metrics such as fuzzy matching scores, word overlap, and length difference.
- Trained and evaluated the model using XGBoost for high accuracy in classification tasks related to semantic similarity.

Anomaly Event Detection in CCTV Footage | *Python, OpenCV, YOLOv8, ConvLSTM*   **May 2025**
- Designed and implemented a deep learning-based video analysis system to detect fight or abnormal behavior in surveillance footage.
- Used YOLOv8 for real-time person detection and overlaid bounding boxes (green for normal, red for fight) on the video frames
- Developed a ConvLSTM model to learn spatial-temporal patterns across video sequences for fight detection.
- Processed videos into frame sequences and integrated the system into a functional pipeline capable of handling local video inputs.

AI Resume Screening System | *Python, BERT, Groq LLM, Gradio*   **July 2025**
- Engineered an end-to-end resume screening system using BERT embeddings for semantic matching between job descriptions and resumes
- Implemented intelligent text chunking and embedding pipeline using SentenceTransformers, achieving efficient document comparison
- Integrated Groq's LLM API for automated candidate scoring (1-10) with reasoning, processing multiple resumes simultaneously
- Developed a user-friendly Gradio web interface supporting both text input and file uploads (.txt, .pdf) with real-time processing

Certifications

- **Full Stack Data Science Program** – Naresh i Technologies *Jan 2025 – July 2025*
  Completed a comprehensive end-to-end training in data science covering machine learning, deep learning, NLP, computer vision, and model deployment using Python-based tools.

- **Intermediate Machine Learning Certificate** – Kaggle Learn *May 2025*
  Completed hands-on projects on real-world datasets covering feature engineering, model validation, handling missing values, and overfitting control.