Chandrakant Choudhary

11, Patrakaar Colony — Ashok Nagar Sarkanda, Bilaspur (C.G.) – 495006
7974473704 sonu.aps1998@gmail.com linkedin.com/in/chandrakant-choudhary-260818178 github.com/chanducn

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 University **Nov 2022 – Dec 2024**
Masters of Computer Application – Bilaspur, Chhattisgarh

Mats University **Sep 2016 – Dec 2019**
Bachelor of Computer Application – Raipur, Chhattisgarh

Sanskar Public School **May 2016**
Senior Secondary – 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**, **Pandas Profiling**
  - **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**, **FastAPI**, **Flask**
  - **Automation & Others: PyAutoGUI**, **Pycaw**, **SpeechRecognition**, **BeautifulSoup**
  - **MLOps: MLflow**, **GitHub Actions**, **Azure ML Pipelines**, **CI/CD Pipelines**, **DVC**, **Weights and Biases**
  - **LLMOps: LangSmith**, **LangGraph**
- **Concepts & Techniques:**
  - **Supervised Learning:** 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, RAG, LLM Agents
  - **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 on Quora Dataset | *Python, NLP, CountVectorizer, XGBoost* **April 2025**

- Built a binary classification model to detect semantically similar question pairs using Quora dataset.
- Preprocessed text using tokenization, stopword removal, and generated pairwise features like fuzzy ratios, word overlaps, and token count differences.
- Achieved **81% accuracy** using **XGBoost** classifier with CountVectorizer and manual feature engineering.
- Deployed the model using **Docker** and automated the pipeline with **GitHub Actions**, hosted on **AWS EC2** for real-time inference.

Anomaly Event Detection in CCTV Footage | *YOLOv8, OpenCV, ConvLSTM, Python* **May 2025**

- Designed an end-to-end anomaly detection system to identify fight/violent events from surveillance video feeds.
- Utilized **YOLOv8** for real-time person detection and **ConvLSTM** to capture spatio-temporal patterns across frame sequences.
- Achieved **92% F1-score** on validation set using RWF-2000 dataset with 2-class classification (fight / non-fight).
- Overlaid bounding boxes on detected persons in video frames — green (normal), red (fight) — for intuitive visualization.

Certifications

- **Full Stack Data Science Program** – Naresh i Technologies *Jan 2025 – July 2025*
- **Intermediate Machine Learning Certificate** – Kaggle Learn *May 2025*