

RAVINDRA SINGH KUSHWAHA

📞 6393344614 Noida, Uttar Pradesh

✉ ravindrakushwaha761@gmail.com

🌐 [linkedin.com/in/kushravindra](https://www.linkedin.com/in/kushravindra)

🐙 github.com/ravindrakush11

Education

Lovely Professional University

M.Tech (Machine Learning & Artificial Intelligence)

Sep. 2021 - May 2023

Jalandhar, Punjab

CCS University

B.Tech (Information Technology)

Sep. 2017 - May 2021

Meerut, Uttar Pradesh

Relevant Tools/Technology

Generative AI & LLMs	LLMs, OpenAI API, Hugging Face Transformers, LangChain, LlamaIndex, Prompt Engineering, Embedding Models, Mistral, LLaMA, Grok, Claude
LLM Fine-Tuning	PEFT, LoRA, QLoRA, Adapters, Prompt Tuning, RAG
AI Agents	smol-ai/smol-dev, Hugging Face Agents
Experiment Tracking	MLflow
Vector Databases	FAISS, ChromaDB, Weaviate
Dev Tools	Git, GitHub, GitHub Actions, Docker, CI/CD, Jupyter Notebooks, VSCode
APIs & Web Frameworks	Flask, FastAPI, REST API, API Integration, Streamlit, Gradio
Cloud Platforms	AWS Sagemaker
Databases & ORMs	MySQL, PostgreSQL, XAMPP
Python & Data Tools	Python, SQL, NumPy, Pandas
ML/DL Libraries	Scikit-learn, Matplotlib, Seaborn, PyTorch Lightning

Experience

Humanitics Dimensions Software PVT. LTD.

AI/ML Engineer

July-2023 to Aug-24

Noida, Uttar Pradesh

- Developed a **Text-to-SQL application** using both **open-source models from Hugging Face** and the **OpenAI API**, enabling users to query structured databases using natural language. Built user interfaces with **Streamlit** or **Gradio** for accessibility and ease of use.
- Integrated **LLMs and LangChain** to automate language-based tasks, including document summarization, NER, and information extraction.
- Automated OCR-based pipelines using **PyTesseract**, improving document processing efficiency in unstructured data scenarios.
- Built and deployed AI-powered web apps using **Flask, FastAPI, Streamlit, and Gradio**

Independent AI Researcher

August 2024 – Present

- Applied **prompt engineering** techniques to fine-tune model outputs for tasks like summarization, structured data querying, and generative reasoning.
- Conducted independent research and prototyping in **Natural Language Processing (NLP)** and **LLM-based systems** using open-source models from **Hugging Face** and local LLM deployment via the **OLLAMA framework**.
- Built advanced **Retrieval-Augmented Generation (RAG)** pipelines using **LangChain**, combining embedding models and vector databases (**FAISS, Chroma**) for document-grounded question answering.

Projects

Retrieval-Augmented Generation (RAG) Q&A System | *LLamaIndex, Hugging Face, FAISS, Streamlit* [GitHub](#)

- Built a modular **RAG pipeline** combining document retrieval via **FAISS** and **LLM-based** answer generation. Developed a **Streamlit app** for real-time querying, with custom text preprocessing, embedding generation, and optimized retrieval for accurate, context-aware responses.

Text-to-SQL: Natural Language Querying for Databases | *LangChain, OpenAI, Flask, PostgreSQL* [GitHub](#)

- Developed a **Text-to-SQL** system that translates natural language queries into executable SQL commands using **LangChain** and **OpenAI GPT-3.5**. Built a **Flask API** for real-time interaction with dynamic **PostgreSQL** connections, enabling users to query databases conversationally without writing SQL.

Multi-Agent Systems with Hugging Face Agents | *Hugging Face Agents, Transformers, LLMs Colab* [GitHub](#)

- Built modular Jupyter notebooks including `HF_Agentic_RAG`, `HF_Agent_text_to_sql`, and `hf_agent_orchestrate_a_multi_agent_system` to demonstrate real-world use-cases of autonomous agents and collaborative task execution.

Research Publications

- Kushwaha, Ravindra Singh, and Rajan Kakkar. "*A Systematic Study of Super-Resolution Generative Adversarial Networks.*" [Click here](#)
International Conference on Advanced Computing, Machine Learning, Robotics and Internet Technologies, Springer Nature Switzerland, 2023.
- Kushwaha, Ravindra Singh, et al. "*An overview: super-image resolution using generative adversarial network for image enhancement.*" [Click here](#)
5th International Conference on Contemporary Computing and Informatics (IC3I), IEEE, 2022.