

# Shubhang Dhawan

📞 9335219193    ✉ shubhangdhawan2000@gmail.com     LinkedIn     GitHub     Google Scholar

●● Medium    📅 29 September 2000

## Profile

- Familiarity in **Python** and **OOPs** concepts, also hands on experience in libraries such as **NumPy and Pandas**. Build **APIs** with **Flask** and implementing **POST** and **GET** requests, used **Postman** for testing, and integrating APIs with databases for data handling.
- **Worked with RAG and vector embeddings** for **similarity search**, using **OpenAI API** for AI text generation. Used **Prompt Engineering** to optimize prompts for accurate, domain-specific outputs.
- Basic knowledge in building and training **Machine Learning** models, performed **Hyperparameter tuning**, **Feature engineering**, and utilizing algorithms such as **Regression and Classification**. Familiarity with model evaluation metrics such as **Precision, Recall, F1 score**.
- Familiarity in **NLP** with **Spacy** and **NLTK**, knowledge in **Text preprocessing**, **Tokenization**, **Stemming**, **Lemmatization**, **POS** tagging, **NER**, and Syntactic parsing.

## Education

Master's in Data Science, Chandigarh University 2022 – 2024

Bachelor's in Science, CSJM Kanpur University 2019 – 2022

## Work History

AI/ML Intern | Times Now | Bennett, Coleman & Co. Ltd Nov 2024 – Feb 2025  
Noida

- **AI-Powered Article Generation Based on Editor's Persona** Implemented **Azure GPT-4o**, **Azure Cosmos DB**, and **Azure AI Search** with **vector embeddings** to generate articles matching the **editor's persona** and **5W1H** approach.
- **AI-Driven News Video Translation, Transcription, and Hindi AI Voiceover** using OpenAI API, GenAI voiceover, and transcription models.
- **Car Specifications Chatbot on WhatsApp** using **Meta API**, **GPT-4o Mini**, and **RAG** with vector-based similarity search to provide car specifications from **Times Drive data**.

LLM Engineer | Frontera Health · Internship Mar 2024 – Oct 2024  
Remote

- Worked with the team, building a **GenAI Chatbot** for the Health & Wellness sector using **RAG**, **Semantic search** concepts
- Used **MongoDB Atlas** and **LlamaIndex** for knowledge base to store embeddings & vectors, improving query handling for **RAG** search, also used **LLMs API** like **Mistral AI 7b Instruct** and **GPT 3.5 Turbo**.

## Skills

Machine Learning | Python | Natural Language Processing | MySQL | Flask | Google Cloud


## Personal Projects


Retrieval Augmented Generation for the book '48 Laws of Power' 

NewsFeed Harmony 

Collaborative-Filter-Approach 

## Certificates & Publications

Machine Learning -Analytics  
Vidhya 

An Enhancement of online Drug  
Recommendation with BGFT-DBI-  
LSTM- IEEE 

Ethical Implications of AI in  
Healthcare - IJSREM 