

KSHITIJ TYAGI

☎ +91 9149357724 ✉ kshitijtyagi281@gmail.com [in LinkedIn](#) [Github](#)

Education

Graphic Era University, Dehradun

B.Tech in Computer Science & Engineering

2022 - present

8.95/10 CGPA

Raja Rammohan Roy Academy, Dehradun

I.S.C.

2020 - 2022

85.6%

Projects

RAG Chatbot for ISP Customer Care

[Source Code](#)

- Built a **Retrieval-Augmented Generation (RAG)** chatbot using **LangChain** and the **LLaMA 3** model to resolve customer queries efficiently.
- Scraped customer care Q&A from the JioFiber website using **Selenium**, storing the data in **ChromaDB** with nomic-embed-text embeddings.
- Integrated a lightweight local deployment with **Ollama** to run all models, ensuring low latency and offline functionality.
- Achieved contextually accurate and business-specific query resolution by fetching relevant data from the database and generating dynamic responses on the fly.

Sentiment Analysis of YouTube Comments

[Source Code](#)

- Developed a **Python** scraper using **Selenium** to collect and analyze **10,000+ YouTube comments** for sentiment analysis.
- Utilized **Natural Language Processing (NLP)** techniques, leveraging **Word2Vec** for feature extraction and **LSTM** for sentiment classification, achieving **79.58% accuracy** on the test set.
- Trained the model using the **Sentiment140 dataset** from Kaggle, consisting of **1.6 million tweets**, for robust sentiment analysis.
- Generated sentiment insights for video content, categorizing comments into positive, negative, and neutral sentiments, and analyzing trends across **50+ videos**.

Reddit and Twitter Scrapers

[Source Code](#)

- Created a **Reddit Scraper** using **Python**, **requests**, and **Reddit API** to extract data from **500+ posts per hour**.
- Integrated **SQLAlchemy** for seamless database operations, enabling efficient storage and querying of **thousands of records** for further analysis.
- Boosted data retrieval speed by **20%** through enhanced handling of **API rate limits**.
- Developed a **Twitter Scraper** using **Selenium** to capture real-time tweets, processing **100+ tweets per minute** based on hashtags and keywords.

Times of India Scraper and REST API

[Source Code](#)

- Designed a **Times Scraper** using **BeautifulSoup** to extract real-time news articles across multiple categories.
- Set up a **REST API** using **Django** and **Django REST Framework** to serve the scraped data in **JSON format**, enabling on-demand retrieval of news articles.
- Enhanced the scraper for efficient data extraction, processing **100+ articles per day**, and integrated a **SQLite database** for real-time access.

Technical Skills

Languages: Python, Java, C/C++

Frontend: HTML, CSS, Tailwind CSS

Backend: Django, Django REST framework

Databases: SQLite, MySQL

Data Science & AI: Pandas, Numpy, Tensorflow, NLP

Developer Tools: Postman, Insomnia, VS Code, GitHub

Coding Profiles

- Leetcode: [pewpewnowdie](#) (**1727** peak rating)