

SURAJ CHAUDHARY

Siwan, Bihar 841236

7488498933 skc818527@gmail.com [Linkedin](#) [GitHub](#)

PROJECTS

Retrieval-Augmented Generation | [GitHub](#)

Feb 2026 - Present

- **Objective:** Built an end-to-end Retrieval-Augmented Generation (RAG) system to answer financial queries strictly from documents, reducing hallucination and enforcing safety-aware responses.
- **Outcome:** Delivered a modular, production-style assistant that retrieves relevant context from a vector database and generates grounded, human-readable answers suitable for compliance-sensitive financial environments.
- **Tech Stack:** Python, LangChain, ChromaDB, Sentence-Transformers, Google Gemini API, CLI, Git.

SMS Spam Classifier NLP ML Pipeline (DVC, AWS S3) | [GitHub](#)

Jan 2026 - Present

- **Objective:** Developed an end-to-end NLP pipeline for SMS spam detection using text preprocessing, TF-IDF vectorization, and supervised classification models.
- **Outcome:** Engineered a reproducible ML workflow with DVC-based experiment tracking, data version control, and AWS S3 artifact storage to support scalable experimentation and reliable model deployment.
- **Tech Stack:** Python, scikit-learn, DVC, AWS S3, NLTK, TF-IDF, Pandas.

Ads Optimization Engine | [GitHub](#) | [Live](#)

Dec 2025 - Present

- **Objective:** Developed an explainable machine learning-based system to select and rank advertisements under budget constraints using both ML predictions and business rules.
- **Outcome:** Delivered a production-ready API that generates ranked, human-readable ad **decisions** (rank, label, reason), enabling transparent and reliable decision-making for real-world use.
- **Tech Stack:** Python, FastAPI, scikit-learn, Pandas, Streamlit, Docker, Render.

EXPERIENCE & TRAINING

Data Structures and Algorithms - Self Paced | GeeksForGeeks

Jul 2024 - Oct 2024

- **Completed** the Self-Paced DSA Course by GeeksforGeeks through the **90 Days DSA** Challenge, strengthening foundations in problem-solving and core computer science principles.
- **Covered** key Data Structures arrays, **linked lists**, stacks, queues, **trees**, **graphs**, heaps and Algorithms **recursion**, sorting searching, backtracking, dynamic programming, greedy, etc.
- **Solved 250+** coding problems, with consistent daily practice focused on logic building, code **optimization**, and efficient time and space **complexity**.
- **Enhanced proficiency** in **C++, java** gaining confidence for coding rounds, technical interviews, and competitive programming.

CERTIFICATION

- Generative AI Professional | **Credential:** 102976190OCI25GAIOPC | [Oracle University](#)
- AI Foundations Associate | **Credential:** 102976190OCI25AICFA | [Oracle University](#)

Nov 2025

Oct 2025

SKILLS

- **Languages and DataBase:** Python, SQL, Java, C++
- **ML & DL :** LLM, GenAI , Supervised and Unsupervised , Regression, Classification, Tree Models, Model Evaluation , Feature Engineering
- **Library & Frameworks:** NumPy, Pandas, Matplotlib, Scikit-learn, TensorFlow
- **Deployment & Tools:** FastAPI, REST Api Streamlit, Git, GitHub, Docker, DockerHub, MLOps, Render, DVC, **AWS S3**

ACHIEVEMENT

- **Participated in SIH (SMART INDIA HACKATHON)**
- **50 Days Badge** on LeetCode after solving **100+** problems.
- **90 Days DSA** Challenge on GeeksforGeeks, successfully working through 250+ problems.

Sep 2025

Jul 2025

Oct 2024

EDUCATION

Lovely Professional University

Aug 2022 - Present

- B.Tech. Computer Science and Engineering — 7th Sem CGPA: 7.83

Phagwara, Punjab

Daroga Prasad Ray College, Siwan

Jul 2018 - 2020

- Intermediate PCM Percentage — 68.8

Siwan, Bihar

CHANDRA Badan UCH High School

Jul 2017 - 2018

- Matriculation Percentage — 60.0

Siwan, Bihar