

Dedicated and enthusiastic Bachelor of Technology (B.Tech) student in Computer Science (AI/ML). Proficient in Data Preprocessing, Machine Learning, Statistical Modeling, and Visualization using Python, SQL, Pandas, Scikit-learn, and TensorFlow. Passionate about applying AI/ML solutions to real-world problems. Eager to leverage technical expertise in dynamic environments.

## SKILLS

**Programming Languages:** Python, MySQL

**Libraries/Frameworks:** NumPy, Pandas, Scikit-learn, XGBoost, TensorFlow, Keras, Matplotlib, Seaborn, Streamlit, LangChain

**Statistical Modeling:** Hypothesis Testing, Linear & Logistic Regression, Probability Theory, Statistical Inference

**Machine Learning:** Supervised Learning, Unsupervised Learning, Model Building, Model Evaluation, Cross-Validation, Hyperparameter Tuning

**Deep Learning:** Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN)

**Data Analysis Skills:** Data Cleaning, Exploratory Data Analysis (EDA), Feature Engineering, Handling Imbalanced Data, Web Scraping

## EXPERIENCE

### Data Science Intern

Jun 2024 — Jul 2024

*Snapdeal (via Mentormind)*

Remote

- Segmented 5 customer groups using K-means to improve marketing focus.
- Segmented customer groups and identified key behavioral patterns using K-means clustering and RFM analysis to enhance marketing strategies.
- Applied RFM (Recency, Frequency, Monetary) analysis for effective customer profiling.
- Used Python (Pandas, scikit-learn) to clean data and build models, cutting processing time by 30%

### Data Analytics Intern

Jun 2024 — Aug 2024

*IBM SkillsBuild (CSRBOX Foundation)*

Remote

- Completed a 6-week program, applying data preprocessing, analysis, and visualization techniques to real-world datasets.
- Developed end-to-end data analysis workflows to identify key business insights and trends.
- Worked on tasks involving data cleaning, trend identification, and business insights generation.
- Utilized Python and visualization libraries to create and present actionable insights from large datasets.

## PROJECTS

### Vidsage – Smart YouTube Video Assistant

- Made an AI-powered assistant that extracts YouTube video transcripts and enables natural language Q&A over video content.
- Implemented RAG (Retrieval Augmented Generation) pipeline using LangChain, FAISS vector database, and Google Generative AI.
- Built an interactive UI with Streamlit to process videos, store embeddings, and answer context-aware queries in Hindi/English.
- Overcame transcript limitations by designing robust preprocessing and chunking with LangChain's text splitters.
- Code:** <https://github.com/Abhipratapsingh123/VidSage>

### Movie Recommender System

- Built a movie recommender system using cosine similarity and content-based filtering in Python.
- Created an interactive frontend using Streamlit and deployed the app online.
- Code:** [github.com/Abhipratapsingh123/ML\\_projects-Movie\\_Recommender](https://github.com/Abhipratapsingh123/ML_projects-Movie_Recommender)

### WhatsApp Chat Analyzer

- Created a tool to analyze WhatsApp group chats for message frequency, media sharing and user activity.
- Used Python with pandas, regex, and matplotlib for data parsing and visualization.
- Deployed the tool with Streamlit and published source on GitHub.
- Code:** [github.com/Abhipratapsingh123/ML-project-Whatsapp\\_chat\\_analysis](https://github.com/Abhipratapsingh123/ML-project-Whatsapp_chat_analysis)

### Spam Classifier

- This project is an Email/SMS Spam Classifier built using machine learning techniques (Naive Bayes).
- Leveraged Natural Language Processing (NLP) to preprocess and transform text data.
- Achieved 100% precision score using the Multinomial Naive Bayes model, ensuring no ham messages were misclassified as spam.
- Deployed the model as a Streamlit web application.
- Code:** [github.com/Abhipratapsingh123/ML-project-Spam-classifier](https://github.com/Abhipratapsingh123/ML-project-Spam-classifier)

## EDUCATION

**Bachelor of Technology in Computer Science (AI/ML)** Sushant University, Gurugram

2022 – 2026

**Activities:** Coding competitions, hackathons, and tech club participation

CGPA: 8.71 / 10