

Shashank Yadav

☎+91-8817877625

✉shashankyadavriiii@gmail.com

🌐www.linkedin.com/in/shashank-yadav-a75a54302

Education

Indian Institute of Technology, Bhilai

CGPA: 8.34 (Current)

B.Tech. in Data Science and Artificial Intelligence

2023 - 2027

Relevant Coursework: Data Structures, Algorithms, Data Visualization, Statistical Programming

Projects

- **Personalized Cold Email Generator**

May 2025– May 2025

Technologies: LangChain, ChromaDB, Python, LLMs, Semantic Search, Embedding Models

[GitHub](#)

- Built an **AI system to auto-generate** personalized cold emails from company websites
- Used **ChromaDB** to match user portfolios with relevant web content via embeddings
- Created prompt templates** for effective and context-aware email generation
- Integrated **semantic search to retrieve and rank** the most relevant project links

- **NSSC Hackathon Solution – Asteroid Impact Prediction**

Oct 2024 – Nov 2024

conducted by iit khargpur ,selected for finals

[GitHub](#)

Technologies: Pandas, Random Forest, One-Hot Encoding, Cross-Validation, EDA

- Cleaned data, handled missing values, and performed **exploratory analysis**
- Applied **one-hot encoding and feature selection** for model efficiency
- Trained a Random Forest classifier with **cross-validation and hyperparameter tuning**
- Achieved **84% accuracy in impact prediction** on unseen test data
- Extracted the **most influential features**, improving interpretability of predictions

Mar 2025 – Apr 2025

- **Song Clustering and Similarity Retrieval**

Technologies: Librosa, scikit-learn, DTW, DBSCAN, Cosine Similarity

- Clustered 1,000+ songs using temporal audio features with unsupervised DBSCAN
- Identified genre-consistent groupings with over 85% purity based on known song categories
- Injected noise (mean shift) into 100+ samples to test retrieval robustness
- Used DTW and cosine similarity to retrieve original songs from noisy queries
- Achieved 92% retrieval accuracy on distorted audio feature sets

Technical Skills

Programming:C, C++, Python

Data Science & ML: NumPy, Pandas, Scikit-learn,Librosa, Matplotlib, Seaborn,PowerBI

GenAI: Tensorflow, RAG, LLMs, Langchain, ChromaDB, Prompt Engineering

Core CS: Data Structures, Algorithms, Computer Architecture, Algorithm Design

Certifications

By standford & Deeplearning.ai: Machine Learning Specialization

Supervised Machine Learning.

Advanced Learning Algorithms

Unsupervised Learning, Recommenders, Reinforcement Learning.

By IBM: Python for Data Science, AI & Development

Position of Responsibility

Core Member: Data Science and Artificial Intelligence Club | Space Exploration Society