

Manav Bhuta

9820107893 | bhutamanav@gmail.com | linkedin.com/in/manav-bhuta | github.com/bhutamanav11

SUMMARY

3rd-year Computer Engineering student with 2 published ML/AI research papers (DACS 2025, IEEE WCSC 2026). Datathon finalist with experience in end-to-end ML pipelines and model evaluation. Proficient in Python, C++, and SQL with 100+ LeetCode problems solved. Seeking ML/AI internship for real-world impact.

EDUCATION

Mukesh Patel School Of Technology and Engineering (NMIMS), Mumbai *Expected May 2027*
B.Tech in Computer Engineering | **CGPA: 3.68/4.0**
Relevant Coursework: Data Structures & Algorithms, Artificial Intelligence, Data Analytics

TECHNICAL SKILLS

Languages: Python, C++, SQL
ML/AI Frameworks: Scikit-learn, LangChain, TensorFlow/Keras, NLTK, Hugging Face Transformers
Data Tools: Pandas, NumPy, Matplotlib, Seaborn, Streamlit
Development Tools: Git, Jupyter Notebook, VS Code, Google Colab, LaTeX
Databases: MySQL | **APIs:** OpenAI, Anthropic, Gemini, Spotify

EXPERIENCE

Research Intern, Suvidha Foundation *Aug 2024 – Oct 2024*
– Worked on **EcoFinanceAI**, applying **NLP and transformer-based models** to link financial investment data with environmental impact reports
– Designed an entity linking system to connect financial decisions with sustainability outcomes, enabling **automated transparency and responsible investment analysis**

PROJECTS

Multi-Document RAG System | Python, LangChain, Scikit-learn | [GitHub](#) *Aug – Dec 2024*
– Published in **IEEE WCSC 2026**: Comparative analysis of **GPT-3.5, Claude-3-Haiku, Gemini-Pro** in RAG framework with **172-page** enterprise corpus
– Engineered novel metrics achieving **statistical significance** ($p < 0.05$, **Cohen's $d > 0.96$**) in model differentiation
– Key finding: Claude highest faithfulness (**0.904**) with **40%** honest refusals; Gemini/GPT **100%** coverage (**0.846/0.890**)

Box Office Revenue Prediction | Python, Scikit-learn, XGBoost | [GitHub](#) *Jan – Mar 2025*
– Published in **DACS 2025**: Built end-to-end ML pipeline for pre-release box office revenue forecasting using real-world TMDB data
– Benchmarked **3 regression** and **3 ensemble models** with feature engineering and hyperparameter tuning; tuned **XGBoost** achieved best performance (**$R^2=0.77$, MAE=\$43.9M**)

EmoSound: Mood-Based Music Recommender | Python, Librosa, Streamlit | [GitHub](#) *Jun – Aug 2024*
– Real-time music recommendation via voice emotion detection with **87%** accuracy across **7** emotion classes
– Integrated collaborative filtering + content-based algorithms with Spotify API; deployed on Streamlit

ACHIEVEMENTS & LEADERSHIP

Achievements: Datathon Finalist – Analytika NMIMS (2024) | SQL 50 – LeetCode (2025)
Sub-Head, Tech, DataMavericks | NMIMS University, Mumbai *Aug 2024 – Present*
– Led **2** ML/data events; mentored **5** peers; built industry webinar partnerships (**+30%** participation)