

# Gurnaj Singh Mann

✉ gurnajsinghmann@gmail.com | 📞 +91 771-735-2608 | 🌐 Gurnaj1Singh | 🌐 LinkedIn | 🌐 Portfolio

## EDUCATION

2022 – 2026 **B.Tech in Computer Science and Engineering** **CGPA: 7.14**  
Dr. B.R. Ambedkar National Institute of Technology (NIT), Jalandhar

## TECHNICAL SKILLS

<b>Data Science &amp; Machine Learning</b>	Data Analysis, Supervised/Unsupervised Learning, Feature Engineering, Model Evaluation.
<b>Languages</b>	Python (Primary), SQL (MySQL, PostgreSQL), C++, JavaScript.
<b>Full-Stack</b>	FastAPI, React.js, MongoDB, REST APIs, Tailwind CSS, Node.js.
<b>Libraries &amp; Tools</b>	Git/GitHub, Pandas, NumPy, Scikit-learn, NLTK, OpenCV.

## EXPERIENCE

**Softtrix Tech Solutions** | *Industrial Trainee (Backend & AI Integration)* June 2024 – Aug 2024

- **Architected** backend logic for a Lead Management System (LMS) using **FastAPI**, implementing rigorous validation rules to ensure 100% integrity of incoming marketing lead data.
- **Optimized** high-volume SQL queries for analytics dashboards, reducing report generation latency by **40%** through strategic indexing and query restructuring.
- **Developed** automated scripts to aggregate marketing funnel metrics, enabling real-time ROI tracking for client-facing reports.

## PROJECTS

**Assignment Master: Agentic RAG Academic Integrity Engine** Jan 2026 – Present

- **Engineered** an Agentic AI workflow using **FastAPI** and **Qdrant** Vector Store to store and retrieve 768-dimensional sentence embeddings for semantic plagiarism detection.
- **Implemented** a Retrieval-Augmented Generation (RAG) pipeline with **SBERT**, achieving sub-millisecond query latency for contextual similarity analysis across multi-page PDF documents.
- **Developed** a graph-based collusion detection agent using **NetworkX** to identify "Cheating Rings" by clustering pairwise similarity matrices with **95% precision**.
- **Integrated** secure **JWT** authentication and SMTP-based OTP verification to restrict access to institutional domains (@nitj.ac.in).

**Sales & Operations Data Analysis Pipeline (SQL + Python)** Jan 2025 – May 2025

- Built an end-to-end data analysis pipeline using SQL and Python to process large-scale operational datasets.
- Performed data cleaning, transformation, and **exploratory data analysis (EDA)** using Pandas and NumPy.
- Designed optimized SQL queries to **extract KPIs** such as revenue trends, customer segmentation, and demand patterns.
- Generated insights to support data-driven decision-making and improve operational efficiency.

**SMS Spam Classification Pipeline** Oct 2024

- **Developed** an end-to-end NLP pipeline for real-time text classification using **NLTK** for tokenization and lemmatization.
- **Trained** a Multinomial Naive Bayes model, achieving **99.07% Precision** and **96.22% Accuracy** through hyperparameter tuning on the UCI dataset.

## CERTIFICATIONS & ACHIEVEMENTS

- **Machine Learning Specialization:** Andrew Ng, Stanford University (Coursera) — Expertise in Supervised/Unsupervised Learning.
- **Meta Professional Certificates:** Certified in **Frontend Development (React.js)** and **Version Control (Git)**.