Malay Chhatbar

Ahmedabad, Gujarat, India | +91 91570 60922
machhatbar2005@gmail.com | GitHub | LinkedIn | Portfolio

Education

LJ Institute of Engineering and Technology, LJ University — Ahmedabad, India
Bachelor of Engineering in Computer Science and Technology, Running — *2024 - 2028*
- SPI: Semester 3 – 10.00/10, Semester 2 – 8.89/10, Semester 1 – 8.96/10

Indian Institute of Technology Madras — Online
BS in Data Science and Applications, Running — *2023 - Expected 2027*
- Completed Diploma in Programming and Diploma in Statistics

Skills

Languages: Python, Go, TypeScript, Rust, Java, SQL
Machine Learning / AI: TensorFlow, PyTorch, JAX, Hugging Face Transformers, DistilBERT Fine-tuning, RAG, Embeddings, LightGBM, Prophet
Data Science: Pandas, NumPy, Model Training, Feature Engineering, Model Evaluation, Large Dataset Processing
LLM / RAG Infrastructure: pgvector, Ollama, Vector Search, Multi-tenant RAG APIs, Document Ingestion, Prompt-based Query Systems
Backend / Deployment: FastAPI, Flask, REST APIs, Docker, Redis, Celery, PostgreSQL, SQLite, GitHub Actions
Frontend / Tools: React, Next.js, Vue, Tailwind CSS, Git, GitHub, Postman, Linux/WSL, AWS, Cloudflare

Projects

RAG Support API | *FastAPI, Python, pgvector, Ollama, PostgreSQL, Docker, TypeScript*
- Built a Docker-first RAG customer support API with FastAPI, pgvector, Ollama, PostgreSQL, and a separate Better Auth service for authenticated access.
- Implemented multi-tenant retrieval using tenant-aware request headers, document/source ingestion, vector search, and query endpoints for support-style knowledge retrieval.
- Added pluggable connectors for files, SQL, and Mongo-style sources, plus read-only live SQL querying and S3-compatible object storage through MinIO.
- Designed the system as a production-style AI backend with containerized services, API boundaries, authentication, embeddings, and local LLM-based chat/query workflows.

NLP Model Fine-tuning Experiments | *Python, TensorFlow, PyTorch, Hugging Face, DistilBERT*
- Trained and fine-tuned multiple ML/NLP models, including DistilBERT-based models, for text classification and language understanding experiments.
- Used Python, TensorFlow, PyTorch, and transformer-based workflows for model training, evaluation, and iterative performance improvement.
- Gained hands-on experience with tokenization, dataset preparation, model fine-tuning, validation metrics, and practical experimentation with neural models.
- Trained and fine-tuned an open source Gemma language model on custom instruction-style data for domain-specific text generation and response improvement.

Achievements

- Won a consolation prize at Hackovate LJ, competing individually among approximately 780 participants and 280 teams.
- Achieved perfect 10.00/10.00 SPI in Semester 3 of B.E. Computer Science and Technology.

Certifications

- Cloud Computing – IIT Kharagpur, NPTEL
- Getting Started with Competitive Programming – IIT Gandhinagar, NPTEL
- Data Analysis and Decision Making – IIT Kanpur, NPTEL