# Dhruv Gorasiya

+1(925) 297-7110 | Boston, MA
dhruvgorasiya19@gmail.com | linkedin.com/in/dhruvgorasiya/ | github.com/DhruvGorasiya | dhruvgorasiya.netlify.app

## EDUCATION

**Northeastern University, Khoury College of Computer Science** — Boston, MA
*Master of Science in Computer Science,* ***GPA: 3.91 / 4.0*** — **Expected:** May 2026

**Relevant coursework:** Algorithms, Artificial Intelligence, Web Development, Program Design Paradigm(Java)

**California State University, Long Beach** — Long Beach, CA
*Bachelor of Science in Computer Science;* ***GPA: 3.6 / 4.0*** — January 2020 – May 2024

**Relevant coursework:** Data Structures & Algorithms, Machine Learning, Database Fundamentals, OOP, Probability & Statistics

## EXPERIENCE

**California State University, Long Beach** — *Long Beach, CA*

***Research Assistant – Machine Learning*** — August 2023 – June 2024

- Accelerated dataset discovery across **5,000+ heterogeneous tables** by **22%** (average query latency), by implementing a containment-optimized variant of **LSH Ensemble**
- Boosted attribute-matching precision by **14%** through hybrid similarity modeling using **Minhash**, **q-grams**, and **word embeddings** across schema-agnostic data sources
- Reduced false positives by **27%** in similarity search by designing **adaptive partitioning** and applying **cost-based rebalancing** for skewed data distributions
- Enabled scalable benchmarking by building an end-to-end **Python pipeline** integrating **FastText**, **Scikit-learn**, and custom **hash-based indexing modules**

***Computer Science Tutor*** — August 2023 – December 2023

- Tutored 20+ students across topics like **Data Structures**, **Algorithms**, **OOP**, and **Machine Learning**, adapting explanations to different learning styles
- Guided debugging and code review for **10+** diverse student projects, helping reinforce clean coding practices and secure implementation techniques

## PROJECTS

**Fraudulent Transaction Detection** — February 2025 – April 2025

- Developed a fraud detection pipeline on 6M+ transactions using **Random Forest**, **Naive Bayes**, and **Logistic Regression**
- Performed **data cleaning**, **normalization**, and **feature selection**, reducing feature space by 53% and improving training speed
- Tackled class imbalance with **stratified sampling** and **SMOTE**, achieving **F1-score of 0.92** and **AUC-ROC of 0.98**
- Containerized the model with **Docker** and deployed as a RESTful **Flask API** on **Heroku**

**Genetic Algorithm-based Course Schedule Planner** — January 2025 – March 2025

- Engineered a **Genetic Algorithm** to generate optimal course schedules minimizing cognitive load and time conflicts
- Modeled **burnout risk** using **Random Forest** on workload and GPA-based features, achieving **91%** cross-validated accuracy
- Deployed as a **FastAPI** service with **PostgreSQL** for state persistence and **Pinecone** for personalized goal matching
- Reduced projected burnout by **40%** based on **workload scores** and **course feedback**

**Adaptive Tutor using Reinforcement Learning and NLP** — December 2024 – March 2025

- Designed an intelligent **AI tutor** using **Reinforcement Learning** and **RAG-based NLP** to personalize feedback
- Integrated **GPT-4** via **LangChain** with **Pinecone** vector search for **semantic retrieval** and few-shot prompting
- Reduced hallucinations by **35%** through **A/B prompt tuning**, reward-based ranking, and **instruction tuning**
- Implemented real-time inference backend with **FastAPI**, maintaining latency under **250ms** with **GPU batch prediction**

**Smart Event Scheduler with LSTM and XGBoost** — August 2024 – December 2024

- Built a **delay forecasting system** using **LSTM** for temporal data and **XGBoost** for structured features
- Created a full **ML pipeline**: data ingestion → training → **FastAPI** API → dashboards with **Plotly**
- Provisioned via **Azure ML** for managed **CI/CD** and lifecycle integration
- Reduced venue mismatch by **22%** through post-deployment A/B testing with user feedback loops

## SKILLS

| | |
|---|---|
| **Languages:** | Python, Java, Golang, JavaScript, C++, SQL |
| **Frameworks/Libraries:** | Flask, FastAPI, TensorFlow, PyTorch, LangChain, scikit-learn |
| **Databases:** | PostgreSQL, MongoDB, Pinecone, Firebase |
| **DevOps/Cloud:** | AWS, Azure, Docker, Git |
| **Concepts:** | LLMs, Transformers, RAG, NER, RL, XGBoost |