# Yash Gondkar

936 N 15th Street, Milwaukee, WI 53233 | +1 (602) 539-1439 | yashgondkar@gmail.com | [linkedin.com/in/yash-gondkar](linkedin.com/in/yash-gondkar) | [github.com/Yash-G](github.com/Yash-G)

## SUMMARY

Machine Learning Engineer and researcher with 4+ years of industry experience. Expertise in transformer architectures, MLOps, and production AI systems. Published researcher in group activity recognition with linear-time transformer optimization. Experienced in building end-to-end ML pipelines, RAG systems, and scalable backend services.

## TECHNICAL SKILLS

- **Languages** – Python (3 yrs), Java (4 yrs), JavaScript, TypeScript, C, C++, Bash, PowerShell, Golang (1 yr), SQL (4 yrs)
- **ML/AI (2.5 yrs)** – PyTorch (2 yrs), scikit-learn, Transformers, CUDA, LlamaIndex, LangChain, MCP, RAG, MLOps
- **Frameworks/Tools** – Spring Boot (4 yrs), Spring Data JPA (4 yrs), Angular (2 yrs), AWS (Lambda, ECS), PostgreSQL (2 yrs), MySQL (2 yrs), BigQuery, Docker, Git, CI/CD
- **Specialized** – Vector DBs (Pinecone), GitHub Actions, Pandas, NumPy, Matplotlib, Google Colab, ELK Stack

## WORK EXPERIENCE

**Model Earth / DreamStudio**, Machine Learning Engineer                                    **July 2025 – Present**
- Built scikit-learn-compatible Random Bits Forest wrapper achieving **top performance on Blinks-2016**: 95.36% accuracy, 98.92% ROC-AUC, 95.00% F1 score – endorsed as 'major' update by supervisor.
- Shipped first working Vector DB sync for CodeChat RAG system using **Pinecone + GitHub Actions + Tree-sitter**, handling add/modify/delete/rename operations with robust error handling and retry logic.
- Built AWS Lambda deployment automation with boto3 for one-step backend deployment including Function URL permissions; integrated dynamic parameter handling in UI with graceful edit-conflict resolution.

**Mobiuso Technologies**, Backend Engineer, Mumbai                                    **April 2021 – August 2023**
- Developed dynamic service generating/modifying JPA entities from JSON via Java reflection, reducing **new entity setup from 7 days to under 1 hour** – saving 40+ hours per entity type.
- Built filtering engine for on-the-fly JPQL query generation supporting nested attributes, pagination, and custom response types; created Angular frontend library improving UI development time by similar margins.
- Deployed AWS ECS stack and managed PostgreSQL migrations for B2B hospital management CMS serving multiple healthcare facilities.

**Nvizion Solutions**, Backend Engineer, Mumbai                                    **September 2018 – March 2021**
- Implemented efficient BigQuery solutions achieving **10x cost reduction** for queries scanning 5TB/minute using advanced techniques (aggregate functions, cross joins, nested queries).
- Built recursive algorithms for Firestore data retrieval reducing GET call time from **20s to 0.3s** while eliminating per-chart storage costs through on-the-fly result generation.
- Configured and managed **ELK stack analytics engine** using Elasticsearch REST API, created watches for alerting with Painless DSL, built Kibana dashboards with KQL/Lucene queries, and deployed Java/JavaScript APM agents for application monitoring.

## RESEARCH EXPERIENCE

**SUGAR: Sequence Unfolding Transformer for Group Activity Recognition** | [Published Paper](Published Paper)
**First Author, University of Massachusetts Boston ([Master's Thesis](Master's Thesis))**                **January – August 2025**
- Developed novel technique **linearizing transformer performance** for Group Activity Recognition, outperforming state-of-the-art Bi-Causal model and competing methods (Lightning Attention-2, Mamba SSM).
- Validated on AnimalKingdom dataset achieving superior accuracy; published in **Transactions on Artificial Intelligence**. Evaluated using precision, recall, precision@k, AUROC, and PR curves.

**Additional Research Projects** – UMass Boston
- **Ensemble Learning with Std Dev as Signal:** Implemented CUDA-optimized training pipeline with test-time adaptation (TENT framework), achieving high accuracy on ImageNet with dynamic logging of ensemble diversity metrics.
- **LLM Watermarking:** Modified beam search with temperature scaling and top-k/top-p filtering for StarCoder/LLaMA; implemented KV caching and GPU-optimized batching with AUROC and Pass@K evaluation on HumanEval/MBPP.

**Teaching Assistant**, UMass Boston – ENGIN 241, ENGIN 246, CS 109                **September 2024 – May 2025**

## EDUCATION

**University of Massachusetts Boston**, Boston, MA                                    **September 2023 – May 2025**
Master of Science in Computer Science, Thesis Track | **GPA: 3.97/4.0**

**Thadomal Shahani Engineering College**, Mumbai University                                    **June 2014 – June 2018**
Bachelor of Computer Engineering

## HONORS AND AWARDS

- **First Prize** in SEED MAHA IT IDOL 2018 (statewide competition across Maharashtra) – awarded laptop
- **First Prize** in SEED IT Idol Mumbai 2018 (17,000 participants) – awarded smartphone
- **3rd in India** in HackerRank Week of Code 36 (ranked 77 out of 12,759 participants worldwide)
- **Top 200 Finalist** in Code Gladiators 2017 out of 200,000 participants nationwide
- Cleared both Week 1 and Week 2 of the IBM Quantum Computing Qiskit Challenge