

YASH MALVIYA

Boston, MA | yash.malviya0416@gmail.com | (774)-232-5158 | <http://www.linkedin.com/in/yashmalviya16> | <https://github.com/YashMalviya16>

SUMMARY

- I'm a 3+ years experienced Data Scientist and MS in Data Science student at Worcester Polytechnic Institute, graduating in May 2025, actively seeking full-time roles in Data Science, Machine Learning, MLOps, and Generative AI.
- Currently working as a Machine Learning Engineer with the Commonwealth of Massachusetts and as a Research Assistant at WPI, driving synthetic data research using GANs and Diffusion models.
- Experienced in building AI-powered applications, scalable pipelines, and research-backed solutions deployed using AWS, Docker, and Kubernetes.

SKILLS

- **Programming & ML:** Python, R, SQL, PyTorch, TensorFlow, Scikit-learn
- **Data Engineering & MLOps:** AWS SageMaker, Azure ML, Databricks, Docker, Kubernetes, CI/CD, Github Actions
- **Big Data & Visualization:** Spark, Snowflake, Power BI, Tableau
- **Certifications:** NVIDIA Certified in Building RAG with LLMs, AWS Certified in Machine Learning

PROFESSIONAL EXPERIENCE

Machine Learning Engineer | Commonwealth of Massachusetts | Boston, MA Jan 2025 – Present

"Introduced public sector AI strategy—built generative models that redefined infrastructure planning that transformed public sector AI"

- Conducted research on AI innovation strategies, contributing to 2 proposals impacting over 3 million Massachusetts residents statewide.
- Built and evaluated 3 full-stack ML pipelines using GANs and Diffusion, generating and processing over 10 million synthetic records.
- Led a 5-member cross-functional team to prioritize and evaluate 20+ AI use cases spanning data governance, infrastructure, and policy.

Research Assistant | Worcester Polytechnic Institute | Worcester, MA Dec 2023 - Present

"Set a new gold standard in synthetic healthcare data—achieved 99.89% error reduction with a breakthrough matrix encoding technique"

- Conducted a literature review of 90+ papers and synthesized insights from 12 EHR models to benchmark synthetic data generation.
- Developed a binary matrix technique to simulate 50,000+ EHR time-series, achieving 0.9989 SSIM and industry-leading fidelity metrics.
- Presented findings to 4 WPI research labs and 2 industry partners, influencing real-world adoption of synthetic data methods.

Machine Learning Researcher | Availity Clinical Solutions | Jacksonville, FL Dec 2023 – Jan 2025

"Engineered enterprise-ready GANs that crushed latency, supercharged privacy by 18%, and unlocked \$25K+ in AI-driven revenue streams"

- Engineered synthetic GAN pipelines generating over 1 million patient records, improving training efficiency and HIPAA compliance.
- Deployed 4 modular ML pipelines using AWS SageMaker, reducing resource usage and saving \$30K+ in cloud compute costs.
- Enhanced synthetic data accuracy across 6 KPIs, validated by FID improvements and supporting a \$25K boost in revenue.

Data Scientist | Tata Consultancy Services | India Jun 2021 – Aug 2023

"Orchestrated a \$17K+ turnaround at TCS by exposing sales leaks and deploying predictive ML, achieving 22% forecast precision boost"

- Diagnosed revenue drop using 10K+ transaction records, driving targeted sales strategy interventions that restored business flow.
- Applied SGD on 12 product categories, increasing forecast precision and streamlining inventory cycles across multiple markets.
- Delivered 24+ strategic insights to the COE, driving executive-level decision-making on operations and growth optimization.

LEADERSHIP EXPERIENCE

Massachusetts Institute of Technology (MIT) | Event Co-Ordinator | MIT Bitcoin Hackathon & Expo | Apr 2025

- Hosted 100+ hackers, MIT Expo Director, and Former U.S. Assistant Treasury Secretary at a high-profile blockchain event. [🔗](#)

Commonwealth of Massachusetts | Speaker | Mar 2025

- Presented on generative AI to state leadership, with commonwealth leadership team during a government innovation session. [🔗](#)

EDUCATION

• **Master of Science in Data Science, Worcester polytechnic Institute-Worcester, MA, (CGPA 4/4)** Aug 2023 – May 2025

• **Bachelor of Technology, Computer Science Engineering, State Technical University (RGPV), India (CGPA 3.4/4)** Aug 2021 – Jun 2023

RELATED PROJECTS

Legacy Mental Health Counselor | WPI Research | Advisor: Dr.Randy Paffenroth [🔗](#)

- Deployed a production-ready sentiment chatbot on AWS, serving 10K+ users monthly with 35% latency reduction and optimized deployment for production traffic exceeding 10,000 monthly requests, slashing API latency by 35% and boosting user throughput.

Synthetic Data generation (EHR) | WPI Research | Advisor: Dr.Chun-kit Ngan [🔗](#)

- Developed 2M+ synthetic EHR records using GANs/Diffusion models, resolving 7+ AWS automation issues. Built GAN and Diffusion architectures that generated over 2 million synthetic EHR entries.

Optum Insight Business Intelligence | WPI Major Project | Advisor: Dr.Nima Kordzedah [🔗](#)

- Built real-time BI system for 500+ analysts at Optum, reducing downtime by 20% and accelerating 300+ incident resolutions and designed a BI framework for Optum's ITSM used by 500+ analysts, integrating real-time analytics to cut downtime by 20%.