# HARSHIL SHARMA

St. Louis, MO | +1(314)393-5686 | harshilsharma808@gmail.com |
https://www.linkedin.com/in/harshil-sharma-727a04202/ | https://github.com/xxender13 | Portfolio

## PROFESSIONAL SUMMARY

Machine Learning Engineer with experience in developing AI systems and optimizing Python workflows for data processing and automation. Led the redesign of a data visualization system, enhancing user engagement by 50%, and engineered Python workflows to reduce manual processing time. Skilled in Python, cloud deployment, and AI technologies, aiming to leverage expertise in Python development to contribute to innovative engineering solutions.

## TECHNICAL SKILLS

**Programming Languages**: Python, SQL, JavaScript, ReactJS
**Machine Learning**: PyTorch, TensorFlow, Keras, Scikit-learn, Hugging Face Transformers
**AI/NLP Technologies**: LangChain, OpenAI API, Retrieval Augmented Generation, FAISS, Vector Databases
**Cloud Platforms**: AWS (Lambda, EC2, S3), Azure, Docker, Kubernetes, Microservices
**Data Engineering**: Pandas, NumPy, OpenCV, FastAPI, REST APIs, ETL Pipelines
**Development Tools**: Git, Streamlit, Gradio, WCAG Compliance, MLOps, CI/CD

## EDUCATION

**Saint Louis University** — St. Louis, MO
*Master of Science in Computer Science | GPA: 3.61/4.00* — *Aug 2023 – May 2025*

**SRM Institute of Science and Technology** — Chennai, India
*Bachelor of Technology in Computer Science Engineering* — *Aug 2019 – Jun 2023*

## PROFESSIONAL EXPERIENCE

**PRiME Center, Saint Louis University** — St. Louis, MO
*Tech & Research Assistant* — *Oct 2024 – May 2025*

- Led React portal redesign for research data visualization system used by 500+ faculty members, increasing user engagement by 50%
- Developed automated Python workflows using TensorFlow and SQL, which reduced manual data processing time and improved efficiency

**Taylor Geospatial Institute, Saint Louis University** — St. Louis, MO
*Software Developer* — *Feb 2024 – Aug 2024*

- Optimized geospatial processing scripts handling 500GB datasets, achieving 25% performance improvement for mapping research
- Architected automated image validation pipelines processing 1,000+ daily satellite inputs with 95% quality assurance rate

**Neem Consulting Ltd** — Chennai, India
*AI Engineering Intern* — *Dec 2022 – Jun 2023*

- Designed deep learning image classification system achieving 92% accuracy for manufacturing quality control with 500+ daily images
- Delivered 18% improvement over baseline models through hyperparameter optimization and advanced model tuning techniques

**Agniforge Industries** — Chennai, India
*Machine Learning Intern* — *Jun 2021 – Dec 2021*

- Architected 6 production ML pipelines for predictive analytics, maintaining 85%+ accuracy across supply chain and customer behavior models
- Launched neural networks processing 10,000+ samples for demand forecasting, supporting business decisions for 3 client companies

## Key Projects

**AI Medical Diagnostic System - Chest X-ray Analysis** [Live Demo]                    Apr 2025 – Jul 2025
- Created AI diagnostic application detecting COVID-19, Pneumonia, and Lung Opacity from chest X-rays with 94% accuracy.
- Launched production system on Hugging Face Spaces with multilingual explanations, achieving 500+ user interactions.

**Autonomous Vehicle SLAM System**                    Mar 2025 – Apr 2025
- Established monocular SLAM pipeline with pose estimation and trajectory smoothing, reducing tracking drift by 35%   https://github.com/xxender13/AV-SLAM-SYSTEM
- Implemented Savitzky-Golay filtering and pose graph optimization across 10 KITTI sequences for improved positioning accuracy.

**Enterprise Contract Intelligence System**                    Jul 2024 – Aug 2024
- Spearheaded RAG-based QA system with LangChain and FAISS, achieving 93% answer accuracy and sub-800ms response time.
- Processed 100+ contract documents serving 50+ daily queries through AWS EC2 deployment with Docker containers.

**Edge AI Model Optimization Platform**                    Oct 2024 – Nov 2024
- Applied quantization techniques on 3 computer vision models, reducing inference time by 45% and memory usage by 60%
- Deployed optimized models on AWS Lambda serving 1,000+ daily inference requests with maintained accuracy.

## Publications

**Publication**: IEEE Paper - "Automated Attendance System Using AdaBoost Algorithm" - Improved system efficiency by 60% for 1,000+ users. https://ieeexplore.ieee.org/document/10046982

## Certification

**AI ML Foundations**: 2025
**Generative AI for Business**: 2025
**TensorFlow Developer**: 2023
**AWS AI Practitioner**: 2023