Harsh Tomar

+91-8817969476 tomarharsh28303@gmail.com linkedin.com/in/harsh-tomar-a96a38256
github.com/HarshTomar1234 kernel-crush.netlify.app

Professional Summary

AI/ML Engineer with 8+ months of hands-on experience in Computer Vision, Deep Learning, and Generative AI. Currently developing intelligent research profiling systems using NLP and RAG architectures at i3 Digital Health. Proven track record building production-ready AI applications with 78% accuracy improvements. Strong expertise in PyTorch, OpenCV, LangChain, and modern ML frameworks. Open-source contributor with 41+ GitHub repositories, 25+ stars, and active CNCF community member.

Experience

AI Intern **May 2025 – Present**

i3 Digital Health *Remote*

- Architected intelligent research profiling system aggregating 10,000+ research papers from PubMed, ResearchGate, Google Scholar, and ORCID APIs, reducing manual research time by 85%
- Built NLP pipelines using LangChain and transformer models to extract research themes and collaboration patterns, achieving 92% accuracy in topic classification
- Developed RAG-powered search agents providing contextual recommendations for research collaborators, improving match relevance by 78%
- Deployed scalable AI solutions in production healthcare environments serving 500+ researchers using FastAPI, Docker, and AWS infrastructure
- Collaborated with healthcare professionals to optimize search algorithms, resulting in 40% faster query response times and enhanced user satisfaction

Community Contributor **January 2023 – Present**

CNCF and Google Developer Groups *Remote*

- Active member of Cloud Native Computing Foundation participating in 15+ cloud-native technology discussions and contributing to open-source projects
- Engaged in Google Developer Groups collaborating on machine learning initiatives, presenting at 2 tech talks on AI/ML best practices
- Mentored 10+ junior developers in AI/ML concepts through community workshops and open-source contributions

Projects

Tennis Vision | *PyTorch, OpenCV, YOLOv8, Supervision* | GitHub Link **2025**

- Developed comprehensive computer vision system for tennis match analysis achieving 95% player detection accuracy and 88% ball tracking precision using YOLOv8 and custom-trained models
- Implemented real-time player tracking with position heatmaps and ball trajectory analysis, processing 30 FPS video streams with ¡100ms latency
- Built automated shot classification system categorizing 12 tennis stroke types with 91% accuracy, enabling detailed match analytics
- Open-sourced project gaining 21 GitHub stars and 1 fork, demonstrating community impact and code quality

MolecuQuest | *Python, FastAPI, MongoDB, React* | GitHub Link **2025**

- Platform connecting 200+ biologists and medical researchers for molecular research collaboration, facilitating 50+ successful partnerships
- Implemented ML-powered research matching algorithms using cosine similarity and collaborative filtering, achieving 84% match satisfaction rate
- Developed full-stack application with React frontend and FastAPI backend, supporting real-time messaging and document sharing for 500+ users
- Integrated with 5 major research databases (PubMed, NCBI, UniProt) for comprehensive researcher profiling and discovery

Field Fusion | *YOLOv8, OpenCV, K-means, Optical Flow* | GitHub Link **2024**

- Built unified computer vision platform for sports analysis supporting 5 sports with real-time tracking, processing 4K video at 25 FPS
- Implemented object detection pipeline using YOLOv8 achieving 93% player detection accuracy and 87% ball tracking precision in various lighting conditions
- Developed K-means clustering for dynamic field segmentation and optical flow for camera movement analysis, reducing tracking errors by 35%

- Created comprehensive analytics dashboard providing team formation analysis, player movement patterns, and performance metrics

AgentForge | *CrewAI, LangChain, Python, Multi-Agent Systems* | GitHub Link **2025**

- Comprehensive guide for building AI agents using CrewAI, LangGraph, and smolagents with 15+ hands-on examples and practical implementations
- Created orchestrator, explorer, and coder agents working collaboratively, reducing development time by 60% for complex coding tasks
- Open-sourced frameworks experimentation gaining 2 GitHub stars and 1 fork, with detailed documentation and tutorials for community adoption

Breast Cancer Predictor | *Scikit-learn, Streamlit, Ensemble Methods* | GitHub Link **2024**

- Developed ML-powered diagnostic tool achieving 97% accuracy in breast cancer prediction using ensemble methods (Random Forest, XGBoost, SVM)
- Created interactive Streamlit dashboard with real-time predictions and SHAP explanations, serving 100+ medical professionals
- Implemented comprehensive data preprocessing pipeline handling missing values and feature scaling, improving model robustness by 15%
- Deployed production-ready application with automated model retraining and performance monitoring using MLflow

Technical Skills

Programming Languages: Python, Bash, Javascript
ML/AI Frameworks: PyTorch, TensorFlow, Scikit-learn, HuggingFace Transformers, OpenCV, Keras, XGBoost
Computer Vision: YOLOv5-v8, Object Detection, Image Segmentation, Supervision, Optical Flow, OpenCV, PIL
Generative AI & LLM: LangChain, LlamaIndex, CrewAI, LangGraph, AG2 (AutoGen), RAG, Prompt Engineering, OpenAI API
Data Science: NumPy, Pandas, Matplotlib, Seaborn, Statistical Analysis, Feature Engineering, A/B Testing
Development Tools: Git, Docker, Jupyter Notebooks, VS Code, Streamlit, FastAPI, Flask
Cloud & Deployment: AWS (EC2, S3, Lambda), Google Cloud Platform, Firebase, Netlify, CI/CD, MLflow, Model Deployment
Databases:MongoDB, Vector Databases (Pinecone, Chroma)

Education

Bachelor of Technology in Artificial Intelligence and Data Science **Nov 2022 – May 2026**
Lakshmi Narain College of Technology, Bhopal *CGPA: 7.2/10*

- Relevant Coursework: Machine Learning, Computer Vision, Deep Learning, NLP, Data Structures & Algorithms, Reinforcement Learning, Statistical Analysis, Neural Networks, Database Management Systems
- Academic Projects: Implemented 10+ ML algorithms from scratch, developed 2 computer vision applications, completed 2 NLP projects

Certifications & Achievements

Machine Learning Specialization - Udemy (November 2024)
Generative AI Certification - Google Cloud Skills (December 2024)
Open Source Impact - 41+ GitHub Repositories, 25+ Stars, Active Contributor (2023 – Present)
Community Recognition - CNCF Member, Google Developer Groups Contributor
Research Publications - 2 Technical Blog Posts on AI/ML