

# Manish Tiwari

AI Engineer | Automation LLM Systems | Open Source Contributor

+91 9779405863 | manish.tiwari.09@zohomail.in | LinkedIn | GitHub | Portfolio | Blogs

## Projects

---

### SafeRun AI – Secure Sandbox Execution Platform [Package] [GitHub] [Live]

May 2026

- Built Docker-based sandbox executing untrusted AI-generated Python code in isolated containers with CPU, memory, and timeout limits.
- Implemented AST scanner and YAML policy engine detecting dangerous imports, blocked calls, and infinite loops; classifies execution into **4 risk levels**.
- Developed FastAPI backend with SQLite audit logging, optional LLM-powered explanations, and CLI + browser UI; published as **saferun-ai** on PyPI .

### Trimurti-LM: 4.2M Parameter Multilingual GPT [Model] [Dataset] [Demo] [GitHub]

Jan 2026

- Trained **4.2M** parameter GPT-style language model from scratch on **450K** multilingual (EN/HI/PA) sentences using custom SentencePiece tokenizer with byte fallback support.
- Achieved perplexity **3.32** with training completed in **2.38 hours** on GTX 1650 4GB using FP16, gradient checkpointing, and gradient accumulation.
- Published multilingual model, dataset, and Dockerized Streamlit frontend on HuggingFace.

### Production Model Drift Monitoring System [GitHub] [Live]

Feb 2026

- Built MLOps monitoring system detecting concept drift using statistical(KS) performance comparison with configurable alert thresholds and real-time scoring.
- Built FastAPI async logging pipeline, Plotly dashboard, and Discord webhook alerting system.
- Containerized deployment with Docker and automated CI/CD testing via Github Actions; achieved **<500ms** inference latency.

## Experience

---

### Smartivity Labs Pvt. Ltd.

New Delhi

*Automation Intern*

*June 2026 – August 2026*

- Automated bulk certificate generation, reducing processing time from ~1 minute per certificate to **10–12 seconds for 200+ certificates** through a single automated workflow.
- Automated PPT → Excel product-data extraction, scaling processing from **4–5 minutes per product to 20+ products in 4–5 minutes**, reducing processing time by over 90% at 20-product scale.
- Built a PDF → Excel extraction pipeline, reducing a **2–3 hour manual workflow to a few minutes**, with only ~10 minutes of human review required for final validation.
- Built an end-to-end Slack-integrated project automation platform used by **100+ designers and managers** across **200+ projects**: implemented Slack OAuth (bot install + OIDC login with PKCE), automated per-project channel creation, and a timezone-aware daily reminder scheduler — eliminating **hundreds of minutes each week spent scanning a 100-user spreadsheet** for individual project updates and deadlines.
- Engineered a Slack-driven approval workflow where designers mark stages complete via chat and managers approve in-channel to **auto-advance project state**; added analytics dashboards tracking **overdue projects, designer performance comparisons, and monthly completion trends**.

### Open Source Software Contributor

Global

*Hugging Face · CHAOSS/Augur · Moss*

*Jan 2024 – Present*

- Contributed 32+ merged PRs across 10+ open-source organizations including Hugging Face, Moss, and CHAOSS/Augur, shipping SDK improvements, automated tests, developer tooling, documentation, and AI application features.
- Built Bharat Benefits, a voice-first RAG assistant merged into Moss, integrating Sarvam AI speech models and Moss retrieval implementing an end-to-end STT → Retrieval → LLM → TTS pipeline for Indian government welfare scheme discovery.

### Travintop.com

Remote

*Data Operations Intern*

*Dec 2024 – Feb 2025*

- Built Python automation scripts to process and normalize 500+ travel package records for CMS upload workflows.
- Analyzed influencer engagement metrics and structured operational datasets supporting marketing and listing operations.

## Technical Skills

---

**Automation:** Python, Workflow Automation, Slack Integrations, Document Processing, Excel/PPT Automation  
**AI/ML:** PyTorch, HuggingFace, Transformers, LangChain, OpenAI SDK, RAG, LLM Workflows, LoRA, Fine-Tuning  
**Backend/Infra:** FastAPI, REST APIs, Docker, Linux, GitHub Actions  
**Tools:** Streamlit, Git&Github, Ollama, Plotly, SentencePiece, Scikit-learn

## Education

---

|                                                                                                                                |                                                          |
|--------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|
| <b>IK Gujral Punjab Technical University</b><br><i>Bachelor of Technology in Computer Science &amp; Engineering; CGPA: 8.0</i> | Jalandhar, Punjab, India<br><i>July 2023 – June 2027</i> |
|--------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|

## Achievements

---

- Built and published Ollama libraries for Sarvam 30B and 105B multilingual LLMs with 100+ combined community pulls.
- Won Industrial Automation hackathon category for building “EnergiX Copilot,” an AI assistant for energy optimization and operational monitoring.
- Fine-tuned Mistral-7B and Qwen-14B using LoRA for dual-agent workflows at AMD Hackathon, IIT Delhi.