

CONTACT

📍 Pune, Maharashtra, India

☎ +91 92847 69827

✉ shantanushinde233@gmail.com

[LinkedIn](#)

[GitHub](#)

EDUCATION

2022 - 2026

B.Tech in Computer Science and Engineering (AI)

Nutan College of Engineering & Research, Pune

CGPA: 7.1/10 - Specialization in Artificial Intelligence with focus on Machine Learning, Deep Learning, and Data Science applications.

SHANTANU SHINDE

AI / ML Engineer

PROFILE

Machine Learning and Generative AI Engineer focused on designing intelligent systems that are scalable, secure, and production-ready. Experienced in building multi-agent AI workflows, retrieval-augmented generation (RAG) pipelines, and LLM guardrail architectures while leveraging strong foundations in data analysis, feature engineering, and predictive modeling. Skilled at bridging AI research concepts with real-world applications using Python, TensorFlow, LangChain, CrewAI, FAISS, and modern orchestration frameworks to create reliable, human-centered AI solutions.

ACHIEVEMENTS

Winner — PICT Impetus International Project Competition

2025

Recognized for outstanding project innovation and execution

Top 5 Finalist — IBM SkillsBuild Maharashtra Hackathon

2025

Selected among top finalists for innovative technical solution

Top 20 Winner — VOIS Tech Innovation Marathon 2.0

2026

Ranked in top 20 among 650 participating teams

Article Published in College Newsletter

2026

"Revolutionizing Game AI with NVIDIA ACE: Autonomous Game Characters for the Future of Gaming"

SKILLS

Programming / Tools

Python, Jupyter Notebook, Google Colab, Git, VS Code, Kaggle Notebooks

Frameworks

Pandas, NumPy, Scikit-Learn, TensorFlow, XGBoost, Matplotlib, Seaborn, Hugging Face Transformers, LangChain, Unsloth

Machine Learning

Data Analysis, Data Cleaning, EDA, Feature Engineering, Supervised Learning, Model Evaluation, Hyperparameter Tuning

Soft Skills

Leadership, Teamwork, Collaboration, Presentation, Public Speaking

CERTIFICATIONS

AI White Belt

Tech Mahindra

Advanced certification in AI fundamentals and applications

OCI Certified AI Associate

Oracle

Professional-level certification in Oracle Cloud AI services

NPTEL: Python for Data Science

Comprehensive online course in Python for data science applications

Fundamentals of Deep Learning

NVIDIA

Hands-on training in deep learning frameworks and techniques

Red Hat AI Foundations Executive Certificate

Executive-level AI and machine learning foundations certificate

WORK EXPERIENCE

Program Admin and Operations Intern

Feb 2026 - Apr 2026

Knowledge Networks

Managed programme operations maintaining dashboards, tracking milestones, and ensuring 100% accuracy in task and pipeline monitoring for fellowship cohorts.

- Monitored participant engagement and attendance across sessions, delivering structured weekly reports
- Organized and maintained internal databases with complete data accuracy
- Coordinated global scheduling across multiple time zones
- Identified operational bottlenecks and implemented workflow improvements

Technologies: Dashboards & Analytics, Data Management, Process Optimization

International Project Intern

Aug 2025 - Dec 2025

Tech Mahindra, Pune

Framed business automation challenges and translated them into structured solutions, integrating workflow tools to optimize operational efficiency.

- Built and validated an XGBoost predictive model for sustainability use cases
- Achieved 97% accuracy on production-ready outputs
- Applied structured data preprocessing and feature engineering pipelines

Technologies: XGBoost, Python, Data Preprocessing, Enterprise Reporting

Software Development Engineer Intern

June 2025 - July 2025

Bluestock Fintech, Pune

Collaborated with UX and engineering teams to build and deploy user-centric frontend components, integrating APIs into end-to-end fintech workflows.

- Built responsive frontend components
- Integrated APIs into fintech workflows
- Worked under agile delivery cycles

Technologies: React/Frontend Development, API Integration, Fintech Technologies

PROJECTS

1. AI-Powered Mathematics Learning Platform

[GitHub](#) | [Demo](#)

Problem: Traditional RAG systems face critical challenges: hallucination, inability to self-improve, generic approaches, and context disconnect.

Solution:

- Designed a guardrail-protected multi-agent orchestration system with a 70+ keyword LLM-based validation layer to control accuracy
- Built category-aware routing with intelligent agent orchestration for specialized solvers
- Leveraged DSPy with BootstrapFewShot to design a self-improving feedback loop from user ratings
- Integrated ChromaDB-backed hybrid RAG architecture with specialized solvers

Technologies: LangChain, DSPy, ChromaDB, LangGraph, CrewAI, Streamlit, Groq/Gemini, Sentence Transformers

2. AI-Powered RAG Document Chatbot

[GitHub](#)

Problem: Organizations struggle to efficiently extract specific information from large volumes of documents (PDFs, images, text files).

Solution:

- Built an intelligent RAG chatbot integrating LangChain, FAISS, and LLMs for natural language querying
- Designed and deployed FAISS vector database pipelines for semantic search
- Integrated Tesseract OCR and web scraping into the document processing workflow
- Built a citation system surfacing document ID, page, and paragraph references for explainability

Technologies: LangChain, FAISS, Google Gemini/Groq LLaMA, Tesseract OCR, Streamlit, Python

3. TradeGuide AI - Intelligent Trading Platform

[GitHub](#)

Problem: Retail traders struggle to develop and execute profitable trading strategies due to lack of technical expertise.

Solution:

- Built a multi-agent trading system using LangChain with specialized AI agents (Analyst, Risk Manager, Executor)
- Designed an AI-guided strategy developer translating user intent into custom entry/exit rules
- Integrated an automated backtesting engine to validate strategies across historical data

Technologies: Python, LangChain, Multi-Agent AI Systems, Streamlit, GCP, Trading APIs

4. Multi-Layer Guardrail Content Moderation API

 [GitHub](#)

Problem: Single-model moderation systems struggle with prompt injection attacks and lack defense-in-depth security.

Solution:

- Architected a 7-stage Model Chain Prompting (MCP) pipeline to protect LLMs from hallucination and jailbreak attacks
- Designed a fail-closed, multi-layer threat detection system with code security scanning and independent verification
- Built and deployed a FastAPI REST service with Prometheus monitoring and structured logging

Technologies: Python, FastAPI, Google Gemini API, Docker, Prometheus, JSON Logging, REST API

5. DSA Visualizer - Interactive Learning Platform

 [GitHub](#) |  [Demo](#)

Problem: Students struggle to understand complex data structures and algorithms through static diagrams.

Solution:

- Built an interactive web platform delivering real-time visualization of 15+ data structures and algorithms
- Designed step-by-step animation controls allowing users to pause, resume, and validate each operation
- Deployed a responsive, mobile-friendly UI with MDX-integrated documentation

Technologies: Next.js 16, React 19, TypeScript, Tailwind CSS, Framer Motion, React Flow, Radix UI, MDX