

Ankit Tiwari

Python / Data Science & ML Developer

unkit.vercel.app | GitHub/anktw | Kaggle/ankitfzd | +91 7897895718 | at792226@gmail.com
Ayodhya, India

PROFILE

Machine Learning and Data Science developer with hands-on experience building end-to-end AI systems using Python, PyTorch, TensorFlow, Scikit-learn, FastAPI, and Generative AI frameworks. Developed projects in customer analytics, RAG systems, reinforcement learning, and real-time edge AI applications. Strong foundation in machine learning, deep learning, optimization, and model deployment with a focus on building scalable, production-ready solutions.

PROJECTS

Customer Segmentation & Churn Prediction System [Link] 08/2025 – 09/2025
FastAPI, Scikit-learn, XGBoost, SHAP, MLflow

- Engineered an XGBoost-based customer churn prediction system achieving ROC-AUC of 0.836 and PR-AUC of 0.633, with SHAP-driven insights for customer retention analysis.
- Deployed real-time churn scoring and customer segmentation APIs with sub-100ms inference latency for scalable prediction serving.
- Achieved 68.5% recall in identifying churn customers using optimized threshold tuning and imbalance-aware training.
- Integrated SHAP explainability to identify key churn drivers such as tenure, contract type, and monthly charges, improving business interpretability.
- Developed data-driven retention targeting workflows by identifying high-risk customer segments for proactive churn mitigation strategies.

Healthcare RAG Intelligence System [Link] 10/2025 – 11/2025
Generative AI + LangChain + Gradio

- An AI system that helps hospitals instantly analyze patient reviews and answer operational queries using retrieved real-world feedback.
- Embeds hospital reviews, performs vector similarity search, and uses Gemini to generate context-aware answers with a Gradio chatbot interface.
- Achieved 92%+ relevant response accuracy, reduced review-analysis time by 85%.

Real-Time Sign Language Translation System [Link] 03/2026 – 04/2026
Scikit-learn, FastAPI, Streamlit, Arduino (flex sensors)

- Achieved 25ms inference + 120ms E2E response on embedded edge hardware; built full Arduino → RandomForestClassifier → TTS pipeline with gesture-to-sentence mapping and confidence-filtered smoothing.
- Built a streaming pipeline (Arduino → sliding window → RandomForestClassifier → smoothing + confidence filtering → TTS + UI) with real-time API endpoints and gesture-to-sentence mapping.
- Achieved 25ms inference latency and 120ms end-to-end response time with stable real-time predictions using smoothing and confidence thresholds.

Model-Based RL Agent with World Model Planning [Link] 08/2025 – Present
PyTorch, Gymnasium, VAE, RSSM, LSTM/GRU, MCTS, CEM, TensorBoard

- Reduce sample inefficiency in reinforcement learning by enabling agents to learn environment dynamics and plan without excessive real-world interactions.
- Built an end-to-end MBRL pipeline combining VAE-based latent encoding, RSSM dynamics prediction, reward modeling, and planning (MCTS/CEM) to select optimal actions via imagined trajectories.
- Improved sample efficiency over model-free baselines (PPO/SAC) with reduced environment interactions and stable policy learning in continuous control tasks.

EDUCATION

Bachelor of Technology (Computer Engineering) 2022 – 2026
S.I.E.T Nilokheri, Haryana afw kurukshetra University, Kurukshetra
CGPA: 7.4

EXPERIENCE

Data Structures Intern

YBI Foundation
New Delhi, India

- Developed proficiency in advanced Python programming and problem-solving techniques.
- Implemented and analyzed fundamental data structures and algorithms to improve computational efficiency.
- Worked with NumPy and Pandas and other libraries for data processing and learned core machine learning fundamentals.

Technical Exploration

- Reproduced machine learning models from research papers and open-source repositories to understand practical implementation details.
- Implemented classical machine learning algorithms from scratch, including Linear Regression, Logistic Regression, Decision Trees, and K-Means Clustering.
- Participated in code reviews and debugging sessions to improve code quality, readability, and performance.
- Explored reinforcement learning concepts including MDPs, Q-Learning, DQN, and policy optimization methods.

SKILLS

Languages: Python, C, JavaScript, SQL

Libraries & Frameworks: Scikit-learn, XGBoost, NumPy, Pandas, Matplotlib, Seaborn, Streamlit, TensorFlow, PyTorch, nltk, FastAPI, Next.js.

Tools & Systems: Git, Docker, Linux, Jupyter Notebook, Redis, PostgreSQL, GitHub Actions, Kubernetes

Technical Skills: LLMs, NLP, RAG pipelines, Vector Databases, Transformers (HuggingFace), CI/CD, REST APIs, model deployment, Elasticsearch, LangChain, LangGraph

COURSES

Data Science Classroom

GeeksForGeeks
08/2025 – 11/2025

Noida, India

CERTIFICATES

- Machine Learning - Classroom (GeeksForGeeks) ↗
- React Foundations (Vercel) ↗
- Python Programming with DS&A (YBI Foundations) ↗
- Next.js App Router Fundamentals (Vercel) ↗

INTERESTS

Quantitative Research, Systems Programming, Open source, Ethical Hacking (Nmap, BurpSuite)