AI DEVELOPER ROADMAP

2024 – 2026

From Fundamentals to Production-Ready AI Engineering

OVERVIEW

This roadmap provides a structured, phase-by-phase guide for developers aiming to build, deploy, and scale AI-powered applications. Whether you are transitioning from software engineering or deepening existing ML skills, this path covers the full stack — from Python fundamentals to production LLM systems.

4	18–24	100+
Learning Phases	Months Total	Skills Covered

PHASE 1 — FOUNDATIONS

PHASE 1 Core Foundations □ 0–3 Months

▸ **Python for AI**
- Advanced Python: generators, decorators, async/await
- NumPy, Pandas for data manipulation
- Data visualization: Matplotlib, Seaborn, Plotly

▸ **Math Essentials**
- Linear algebra: vectors, matrices, eigenvalues
- Calculus: derivatives, gradients, chain rule
- Probability & statistics: distributions, Bayes theorem

▸ **Developer Environment**
- Git & GitHub for version control
- Jupyter Notebooks / VS Code setup
- Virtual environments, Docker basics

PHASE 2 — MACHINE LEARNING

PHASE 2 Machine Learning Core □ 3–9 Months

▸ Classical ML

Supervised: regression, decision trees, SVMs, random forests
Unsupervised: k-means, PCA, DBSCAN
Model evaluation: cross-validation, precision/recall, ROC-AUC

▸ Deep Learning

Neural network architecture fundamentals
CNNs for computer vision tasks
RNNs and LSTMs for sequence modeling
PyTorch & TensorFlow / Keras workflows

▸ Data Engineering

ETL pipelines and feature engineering
Handling imbalanced datasets
Data augmentation strategies

PHASE 3 — LARGE LANGUAGE MODELS & GENERATIVE AI

PHASE 3 LLMs & Generative AI □ 9–15 Months

▸ **LLM Fundamentals**
- Transformer architecture: attention mechanisms, positional encoding
- Pre-training vs fine-tuning vs RAG approaches
- Prompt engineering: few-shot, chain-of-thought, ReAct

▸ **Working with APIs**
- OpenAI, Anthropic Claude, Google Gemini APIs
- LangChain & LlamaIndex orchestration frameworks
- Embeddings, vector stores: Pinecone, Weaviate, ChromaDB

▸ **Fine-Tuning & Alignment**
- LoRA and QLoRA parameter-efficient fine-tuning
- RLHF and DPO preference alignment
- Hugging Face Transformers and PEFT libraries

▸ **Multimodal AI**
- Vision-language models (GPT-4V, LLaVA, CLIP)
- Text-to-image: Stable Diffusion, DALL-E, Midjourney API
- Audio AI: Whisper, ElevenLabs, speech synthesis

PHASE 4 — PRODUCTION & MLOPS

PHASE 4 Production AI Engineering □ 15–24 Months

▸ **MLOps & Deployment**
- Model serving: FastAPI, BentoML, TorchServe
- Containerization: Docker, Kubernetes for ML workloads
- CI/CD pipelines for ML: GitHub Actions, MLflow

▸ **AI System Design**
- Retrieval-Augmented Generation (RAG) architectures
- Agentic AI systems and autonomous workflows
- Guardrails, content filtering, responsible AI patterns

▸ **Monitoring & Scaling**
- Model performance monitoring: drift detection, data quality
- Cost optimization: quantization, caching, batching
- Cloud AI services: AWS SageMaker, GCP Vertex AI, Azure ML

▸ **Security & Ethics**
- Prompt injection and adversarial attack defense
- Privacy-preserving ML: federated learning, differential privacy
- AI governance frameworks and compliance (EU AI Act)

SKILLS & TOOLS MATRIX

Core Languages & Frameworks
- Python 3.11+
- PyTorch / TensorFlow
- FastAPI / Flask
- SQL & NoSQL databases
- JavaScript (for AI demos)

AI/ML Platforms & APIs
- OpenAI / Anthropic / Gemini
- Hugging Face Hub
- LangChain / LlamaIndex
- Weights & Biases (W&B)
- Replicate / Modal

Data & Vector Infrastructure
- Pinecone / Weaviate / Qdrant
- Apache Spark (big data)
- Airflow / Prefect (pipelines)
- DVC (data versioning)
- Great Expectations (quality)

Cloud & DevOps
- AWS SageMaker / Lambda
- GCP Vertex AI / BigQuery
- Azure ML / Cognitive Services
- Docker + Kubernetes
- Terraform (infrastructure)

LEARNING RESOURCES

Courses: fast.ai Practical Deep Learning • DeepLearning.AI Specializations • Stanford CS229 / CS224N • Hugging Face NLP Course

Books: Hands-On Machine Learning (Géron) • Deep Learning (Goodfellow) • Designing ML Systems (Huyen) • Build a Large Language Model (Raschka)

Practice: Kaggle Competitions • Papers With Code • LeetCode (ML interviews) • Hugging Face Spaces (demos)

Communities: Hugging Face Discord • r/MachineLearning • AI Twitter/X — Andrej Karpathy, Yann LeCun • Local AI meetups & NeurIPS / ICML

CAREER PATHS

ML Engineer
Build & deploy ML pipelines and model infrastructure

AI Product Engineer
Integrate LLMs into consumer and enterprise products

Research Scientist
Publish novel architectures and training techniques

MLOps Engineer
Scale, monitor and govern AI systems in production

Produced by Techtrep • thetechtrep.com • Your source for the AI developer ecosystem