

🌟 Roadmap to Becoming an AI & Data Engineer

For software engineers starting to specialize in AI + data (with GenAI focus)

☑ PHASE 1: Core Machine Learning Foundations (2–3 weeks)

Start here before going deep into GenAI — it's essential to understand how models work.

◇ What to Learn:

- Supervised learning: regression, classification
- Overfitting, underfitting
- Loss functions, metrics (MSE, accuracy, precision/recall)
- Model evaluation & validation techniques

📖 Course:

- [Supervised Machine Learning: Regression & Classification \(Andrew Ng – Coursera\)](#)
- ☑ FREE if audited

🔗 Practice:

- Kaggle [Intro to ML](#) & [Intermediate ML](#)
-

☑ PHASE 2: Data Engineering & Pipeline Skills (3–4 weeks)

These are must-have skills for any AI or data engineer.

◇ What to Learn:

- SQL (joins, CTEs, window functions, optimization)
- Data cleaning & preprocessing (Pandas, missing values)
- ETL/ELT pipelines
- Data modeling: star/snowflake schema
- Workflow orchestration (Airflow or Prefect)

📖 Courses:

- [Data Engineering Zoomcamp \(FREE\)](#)
- [Kaggle SQL Course](#)

🔗 Tools:

- PostgreSQL or BigQuery (for SQL)
- Airflow (workflow orchestration)
- dbt (data modeling, optional)

☑ PHASE 3: Deep Learning & Neural Networks (3 weeks)

To move beyond APIs and fine-tune or train your own models.

◊ What to Learn:

- Neural networks (forward/backpropagation)
- CNNs (for vision), RNNs (for sequence), Transformers (for NLP)
- Loss functions, optimizers
- Overfitting, dropout, batch norm

▣ Courses:

- [DeepLearning.AI Specialization \(Andrew Ng, Coursera\)](#) (*Free audit available*)
- [Practical Deep Learning for Coders \(fast.ai\)](#)

🔧 Tools:

- PyTorch (or TensorFlow if preferred)

☑ PHASE 4: Generative AI for Developers (3 weeks)

This is where your recent interest lies — and it's a great "builder" phase.

▣ Course:

- [Generative AI Full Course for Developers – @dswithbappy](#)

◊ What You'll Learn:

- LLMs, embeddings, transformers
- LangChain, Hugging Face Transformers
- RAG (Retrieval-Augmented Generation)
- Vector databases (Chroma, FAISS)
- Deployment: Google Vertex AI, AWS Bedrock
- Prompt engineering & agent tools

🔧 Project Ideas:

- RAG-based chatbot for PDF/CSV data
- Text summarizer with custom prompts
- Code doc assistant powered by OpenAI + LangChain
- Deploy an AI assistant via Streamlit + GCP

☑ PHASE 5: Portfolio & Positioning (Ongoing)

Start sharing & documenting your work publicly.

What to Do:

- GitHub repos with clean README, setup, and examples
- Medium/Hashnode posts explaining what you built
- LinkedIn posts with images, use cases, and takeaways

Suggested Projects to Showcase:

- AI Data Pipeline (ETL + BI Dashboard + KPIs)
- GenAI Search over docs (LLM + Chroma + LangChain)
- Fine-tuned LLM assistant for a niche (e.g. legal docs, resumes)
- Real-time KPI dashboard from CSV/SQL to Power BI or Streamlit

Claim Titles with Confidence:

Title	Can you claim it after this roadmap?	Notes
AI App Developer	<input checked="" type="checkbox"/> Yes	Build 2–3 apps
GenAI Engineer	<input checked="" type="checkbox"/> Yes	This course helps directly
ML Engineer	<input checked="" type="checkbox"/> Almost	Add a few custom training projects
Data Engineer	<input checked="" type="checkbox"/> Yes	Zoomcamp + SQL + modeling <input checked="" type="checkbox"/>
AI & Data Engineer	<input checked="" type="checkbox"/> YES	With this roadmap, you've earned it

Certifications to Boost Your Profile (Choose 1–2)

☒ 1. **IBM AI Engineering Professional Certificate**

Best for: Building full-stack AI understanding, with hands-on projects (PyTorch, ML, DL)

☒ 2. **DeepLearning.AI GenAI with LLMs**

Best for: GenAI workflows, LLM fine-tuning, enterprise-level tools (AWS, RAG)

☒ 3. **TensorFlow Developer Certificate**

Best for: Proving hands-on DL implementation skills in production

☒ Any of these certificates would be **recruiter-approved signals** on LinkedIn, CV, and GitHub.

Key Mentions:

- **@dswithbappy** – for his incredible GenAI course
- **Andrew Ng** – foundational ML & DL guidance

- **DataTalksClub Zoomcamp** – for world-class free DE training