แผนการเรียนหลักสูตร AI & Machine Learning Engineering (15 สัปดาห์)

โครงสร้างเวลา: สัปดาห์ละ 6 ชั่วโมง (ทฤษฎี 2 ชม. + 4 เวิร์กชอปย่อย รวม 4 ชม.)

อ้างอิงจาก: Hands-on ML, Designing ML Systems, GenAI Design Patterns, Software Eng. for DS, AI Engineering และ FastAI/PyTorch

Phase 1: Foundations & Classical Machine Learning

(สัปดาห์ที่ 1-4 เน้นปูพื้นฐานการเขียนโค้ดที่ดี และโมเดล ML ดั้งเดิม)

สัปดาห์ที่ 1: Introduction to ML & Software Engineering Basics

ทฤษฎี (2 ชม.): ภาพรวมของ ML, กระบวนการทำโปรเจกต์ ML แบบ End-to-End, แนวปฏิบัติที่ดีในการเขียนโค้ด (Git, Environment)

- **Workshop 1:** การตั้งค่า Environment (Conda/Venv) และ Git Version Control
- **Workshop 2:** Data Manipulation พื้นฐานด้วย Pandas & NumPy
- **Workshop 3:** การทำ Data Cleaning และ EDA (Exploratory Data Analysis) เบื้องต้น
- **Workshop 4:** สร้างโมเดลแรกด้วย Scikit-Learn (Linear Regression) และดูผลลัพธ์

สัปดาห์ที่ 2: Classification, Metrics & Model Selection

ทฤษฎี (2 ชม.): การแยกประเภทข้อมูล, การวัดผล (Precision, Recall, ROC Curve), ปัญหา Overfitting/Underfitting

- **Workshop 1:** การสร้างโมเดล Binary Classification (Logistic Regression / SGD)
- **Workshop 2:** การจัดการข้อมูลแบบ Multiclass และสร้าง Confusion Matrix
- **Workshop 3:** การทำ Cross-validation และ Hyperparameter Tuning (GridSearch)
- **Workshop 4:** การใช้ Support Vector Machines (SVM) และเปรียบเทียบประสิทธิภาพ

สัปดาห์ที่ 3: Trees, Ensembles & Unsupervised Learning

ทฤษฎี (2 ชม.): หลักการทำงานของ Decision Trees, Ensemble Learning (Bagging, Boosting), K-Means

- **Workshop 1:** สร้างและแสดงภาพ (Visualize) Decision Tree
- **Workshop 2:** การใช้ Random Forest และดู Feature Importance
- **Workshop 3:** การใช้ Gradient Boosting (XGBoost หรือ LightGBM)
- **Workshop 4:** การจัดกลุ่มข้อมูล (Clustering) ด้วย K-Means และ PCA ลดมิติข้อมูล

สัปดาห์ที่ 4: Software Engineering for Data Scientists

ทฤษฎี (2 ชม.): Object-Oriented Programming (OOP) สำหรับ Data Science, การเขียนโค้ดให้ทำงานเร็วขึ้นและใช้หน่วยความจำน้อยลง

- **Workshop 1:** การเขียน Python แบบ OOP (สร้าง Class สำหรับ Data Pipeline)
- **Workshop 2:** Exception Handling และ Logging ในระบบ ML
- **Workshop 3:** การจัดการ Memory (Memory profiling) และการทำ Multiprocessing เบื้องต้น
- **Workshop 4:** สร้าง RESTful API พื้นฐานด้วย FastAPI เพื่อรองรับ Model Inference

Phase 2: Deep Learning & Unstructured Data

(สัปดาห์ที่ 5-8 เน้น Deep Learning ด้วย PyTorch/FastAI และ Transformers)

สัปดาห์ที่ 5: Introduction to Deep Learning (PyTorch & FastAI)

ทฤษฎี (2 ชม.): พื้นฐาน Neural Networks, Gradient Descent, Backpropagation, แนะนำ PyTorch

- **Workshop 1:** การจัดการ Tensors และ Autograd พื้นฐานใน PyTorch
- **Workshop 2:** สร้าง Neural Network แบบง่าย (จากศูนย์) โดยไม่ใช้ Framework สำเร็จรูป
- **Workshop 3:** สร้าง Image Classifier อย่างง่ายโดยใช้ FastAI Library
- **Workshop 4:** นำโมเดลจาก Workshop 3 ไปสร้าง Web App บน Hugging Face Spaces (Gradio)

สัปดาห์ที่ 6: Computer Vision Advanced

ทฤษฎี (2 ชม.): สถาปัตยกรรม CNN, ResNet, Transfer Learning, Data Augmentation

- **Workshop 1:** การใช้ Pre-trained CNN Model ทำ Transfer Learning
- **Workshop 2:** การทำ Data Augmentation เพื่อแก้ปัญหาข้อมูลภาพน้อย
- **Workshop 3:** สร้างโมเดล Image Segmentation (แยกวัตถุตามพิกเซล)
- **Workshop 4:** เทคนิคการตีความหมายโมเดลภาพ (Visualizing CNNs - Grad-CAM)

สัปดาห์ที่ 7: Natural Language Processing (NLP) Foundations

ทฤษฎี (2 ชม.): Text Preprocessing, Tokenization, Word Embeddings, RNNs พื้นฐาน

- **Workshop 1:** การทำ Text Tokenization และ Text Cleaning
- **Workshop 2:** การใช้งาน Word Embeddings และดูความหมายของคำในรูปแบบ Vector
- **Workshop 3:** สร้างโมเดล Sentiment Analysis วิเคราะห์อารมณ์ข้อความ
- **Workshop 4:** การทำ Sequence Classification แบบง่ายด้วย PyTorch

สัปดาห์ที่ 8: The Era of Transformers

ทฤษฎี (2 ชม.): Attention Mechanism, สถาปัตยกรรม Transformer, แนะนำ Hugging Face

- **Workshop 1:** สร้างบล็อก Self-Attention แบบง่ายเพื่อให้เข้าใจการทำงาน
- **Workshop 2:** การดึงโมเดล pre-trained จาก Hugging Face (Transformers library) มาใช้งาน
- **Workshop 3:** Fine-tuning โมเดล BERT สำหรับงาน Text Classification
- **Workshop 4:** การทำ Text Generation เบื้องต้นโดยใช้โมเดลกลุ่ม GPT-2

Phase 3: Generative AI & Application Building

(สัปดาห์ที่ 9-11 เน้น LLMs, Prompt Engineering, RAG และ AI Agents)

สัปดาห์ที่ 9: Generative AI & Prompt Engineering Masterclass

ทฤษฎี (2 ชม.): หลักการทำงานของ LLMs, รูปแบบ Prompt Engineering, ค่า API พารามิเตอร์ (Temperature, Top-P)

- **Workshop 1:** การเรียกใช้ LLM ผ่าน API (OpenAI / Anthropic / Gemini API)
- **Workshop 2:** ฝึกทักษะ Zero-Shot, Few-Shot และ Chain-of-Thought (CoT) Prompting
- **Workshop 3:** การบังคับให้ LLM ตอบกลับเป็น Structured Data (JSON format)
- **Workshop 4:** สร้าง Chatbot UI ส่วนตัวด้วย Streamlit

สัปดาห์ที่ 10: Retrieval-Augmented Generation (RAG)

ทฤษฎี (2 ชม.): ข้อจำกัดของ LLM (Hallucination), คอนเซปต์ RAG, Vector Databases

- **Workshop 1:** การสร้าง Text Embeddings และการทำ Chunking เอกสาร PDF
- **Workshop 2:** การบันทึกและค้นหาข้อมูลใน Vector Database (ChromaDB หรือ Pinecone)
- **Workshop 3:** การต่อจิ๊กซอว์สร้างระบบ Basic RAG (ดึงข้อมูล -> ส่งให้ LLM ตอบ)
- **Workshop 4:** การเพิ่มประสิทธิภาพ RAG (Advanced RAG: Re-ranking & Query Expansion)

สัปดาห์ที่ 11: AI Agents & Model Customization

ทฤษฎี (2 ชม.): AI Agents คืออะไร, Tool/Function Calling, การทำ PEFT (Parameter-Efficient Fine-Tuning)

- **Workshop 1:** การทำ Function Calling เพื่อให้ LLM ใช้งานเครื่องมือภายนอกได้ (เช่น ดูสภาพอากาศ)
- **Workshop 2:** สร้าง AI Agent เบื้องต้นด้วย LangChain หรือ LlamaIndex
- **Workshop 3:** การเตรียม Dataset แบบ Instruction-tuning
- **Workshop 4:** จำลองการ Fine-tune LLM แบบประหยัดทรัพยากรด้วยเทคนิค LoRA/QLoRA

Phase 4: Production, MLOps & System Design

(สัปดาห์ที่ 12-14 นำระบบขึ้น Production, การทำ Monitoring และออกแบบระบบ)

สัปดาห์ที่ 12: Designing Machine Learning Systems

ทฤษฎี (2 ชม.): การตั้งเป้าหมาย Business Metrics vs ML Metrics, Data Engineering เบื้องต้น

- **Workshop 1:** วิเคราะห์และออกแบบ System Design สำหรับโปรเจกต์ ML (Case Study)
- **Workshop 2:** การจัดการ Feature Store พื้นฐาน
- **Workshop 3:** การทำ Data Pipeline ขนาดย่อม (Batch Processing)
- **Workshop 4:** การเขียน Test สำหรับ Machine Learning Data (Data Testing)

สัปดาห์ที่ 13: Model Deployment & Serving

ทฤษฎี (2 ชม.): รูปแบบการ Serve โมเดล (Batch vs Online), Model Optimization (Quantization)

- **Workshop 1:** การทำ Model Quantization เบื้องต้นเพื่อให้โมเดลเล็กลง
- **Workshop 2:** เขียน API ด้วย FastAPI รองรับการทำ Model Inference
- **Workshop 3:** การทำ Containerization ด้วย Docker (สร้าง Dockerfile สำหรับ ML API)
- **Workshop 4:** รัน Docker Container และทดสอบส่ง Request แบบจำลองโหลด (Load Test

เบื้องต้น)

สัปดาห์ที่ 14: MLOps, Monitoring & Maintenance

ทฤษฎี (2 ชม.): Data Shift, Concept Drift, CI/CD สำหรับ ML, การกำหนดรอบการ Retrain โมเดล

- **Workshop 1:** การจำลองเหตุการณ์ Data Drift และการใช้เครื่องมือตรวจจับ (เช่น Evidently AI)
- **Workshop 2:** สร้าง Dashboard สำหรับ Model Monitoring อย่างง่าย
- **Workshop 3:** จำลองการทำ Automated Retraining Pipeline
- **Workshop 4:** ทบทวน CI/CD ขั้นพื้นฐานผ่าน GitHub Actions สำหรับรันเทสต์ ML Code

Phase 5: Capstone

(สัปดาห์ที่ 15 ประมวลความรู้ทั้งหมด)

สัปดาห์ที่ 15: End-to-End Capstone Project

ทฤษฎี (2 ชม.): สรุปภาพรวมสิ่งที่เรียนมาทั้งหมด, Best Practices สำหรับการนำเสนอโปรเจกต์

- **Workshop 1:** Project Scoping & Data Preparation (เลือกใช้ ML ดั้งเดิม, CV, NLP หรือ GenAI)
- **Workshop 2:** Model Development & Experimentation
- **Workshop 3:** Deployment & API Integration (นำขึ้น Docker หรือ Cloud เบื้องต้น)
- **Workshop 4:** นำเสนอโปรเจกต์ (Demo) & Code Review รับฟีดแบค