AI Fine Tuning and Data Preparation for Pre-Trained Models

- 3-Day Class
- Hands-on labs

Course Overview

Train Your AI: Fine-Tuning and Data Preparation for Pre-Trained Models is a comprehensive 3-day course that prepares students for obtaining and organizing data for fine-tuning pre-trained LLMs and Generative AI! Select your own Transformer model GenerativeAI, and fine-tune it with datasets you've prepared, or gathered. This course offers a combination of lectures and hands-on labs, providing participants with a solid understanding of AI concepts and the skills to design and implement AI solutions.

Throughout the course, you will learn about AI transformer-based architectures, the fundamentals of Python programming for AI models, and the deployment of open source Transformer models. You will begin with data preparation by gathering, cleaning, labeling, and organizing data for the purpose of fine-tuning an LLM. The course delves into training techniques, including back propagation, gradient descent, and various AI tasks such as classification, regression, and clustering.

You will gain practical experience through hands-on exercises with open source LLM (Language Learning Model) frameworks, allowing you to work with fine-tuned models and run workloads on different models to understand their strengths and weaknesses. Additionally, the course covers the conversion of model formats and provides in-depth exploration of AI programming environments like PyTorch+transformers and transformers' low-level interactive inspection.

Towards the end of the course, you will delve into advanced topics such as context extension through fine-tuning and quantization for specific application target environments. By the completion of the course, you will have the opportunity to earn an AI certification from Alta3 Research, further enhancing your credentials in the field of Artificial Intelligence. This course is ideal for Python Developers, DevSecOps Engineers, and Managers or Directors seeking a overview of AI and its practical application in the enterprise.

Note: Students are expected to already have basic Python skills.

- Access the classroom from anywhere via browser and internet
- Each participant will have access to a fully configured, GPU-accelerated server.
- Obtain hands-on experience with the most widely used, industry-standard software, tools, and frameworks.
- Learn to build deep learning and accelerated computing applications Review this course online at https://www.alta3.com/courses/ai-tune

Who Should Attend

- Project Managers
- Architects
- Developers
- Data Acquisition Specialists

What You'll Learn

- Go from Drowning in Data to maintaining Clean and Curated Data for AI Fine-Tuning and Training!

- Establish guidelines for obtaining RAW Data for AI Fine-Tuning.
- Clean and Curate Data for AI Fine-Tuning.
- Understand AI architecture, specifically the Transformer model.
- Describe the role of tokenization and word embeddings in AI processing.
- Install and use AI frameworks like Llama-3.
- Explore model quantization and fine-tuning.
- Fine-Tune AI Models with PyTorch.
- Explore Low-Rank Adaptation (LoRA) and Quantized Low-Rank Adaptation (QLoRA).
- Perform LoRA and QLoRA Fine Tuning against Llama3 Models.
- Deploy and Maximize AI Model Performance.

Outline

Data Curation for AI

- 🗫 Lecture: Curating Data for AI
- 🖥 Lecture + Lab: Gathering Raw Data
- 🖥 Lecture + Lab: Data Cleaning and Preparation
- 🖥 Lecture + Lab: Data Labeling
- 🖥 Lecture + Lab: Data Organization
- 🗫 Lecture: Premade Datasets for Fine Tuning
- 🖥 Lecture + Lab: Obtain and Prepare Premade Datasets

Deep Learning

- 🗫 Lecture: What is Intelligence?
- 🗫 Lecture: Generative AI
- 🗫 Lecture: The Transformer Model
- 🗫 Lecture: Feed Forward Neural Networks
- 🖥 Lecture + Lab: Tokenization
- 🖥 Lecture + Lab: Word Embeddings
- 🖥 Lecture + Lab: Positional Encoding

Pre-trained LLM

- 🗫 Lecture: A History of Neural Network Architectures
- 🗫 Lecture: Introduction to the LLaMa.cpp Interface
- 🗫 Lecture: Preparing A100 for Server Operations
- 🖥 Lecture + Lab: Operate LLaMa3 Models with LLaMa.cpp
- 🖥 Lecture + Lab: Selecting Quantization Level to Meet Performance and Perplexity Requirements

Fine Tuning

- 🗫 Lecture: Fine-Tuning a Pre-Trained LLM
- 🗫 Lecture: PyTorch
- 🖥 Lecture + Lab: Basic Fine Tuning with PyTorch
- 🖥 Lecture + Lab: LoRA Fine-Tuning LLaMa3 8B
- 🖥 Lecture + Lab: QLoRA Fine-Tuning LLaMa3 8B

Operating Fine-Tuned Model

- 🗫 Lecture: Running the llama.cpp Package
- 🖥 Lecture + Lab: Deploy Llama API Server
- 🖥 Lecture + Lab: Develop LLaMa Client Application
- 🖥 Lecture + Lab: Write a Real-World AI Application using the Llama API

Prerequisites

- Python - PCEP Certification or Equivalent Experience
- Familiarity with Linux

Next Courses

Course 1 Course 2