

**Introduction to Generative AI Engineering for Data Scientists and ML Engineers**

**Course Number:** AI-130WA
**Duration:** 3 days

**Overview**

This Introduction to Generative AI (GenAI) training is tailored for Machine Learning (ML) and Data Science professionals who want to gain a practical understanding of Generative AI and large language models (LLMs).

**Prerequisites**

* Practical experience (6+ months) minimum in Python - functions, loops, control flow
	+ Data Science basics - NumPy, pandas, scikit-learn
* Solid understanding of machine learning concepts and algorithms
	+ Regression, Classification, Unsupervised learning (clustering, Neural Networks)
* Strong foundations in probability, statistics, and linear algebra
* Practical experience with at least one deep learning framework (e.g., TensorFlow or PyTorch) recommended
* Familiarity with natural language processing (NLP) concepts and techniques, such as text preprocessing, word embeddings, and language models

**Materials**

All Generative AI training students receive comprehensive courseware.

**Software Needed on Each Student PC**

All attendees must have a modern web browser and an Internet connection.

**Objectives**

* Gain a solid understanding of Large Language Models (LLMs) and their foundational concepts, including generative AI and transformer architecture
* Master prompt engineering techniques to effectively communicate with LLMs and achieve desired outcomes across various NLP tasks
* Evaluate and compare different LLMs to select the most suitable model for specific natural language processing tasks
* Fine-tune and adapt open-source LLMs using domain-specific datasets to optimize performance for specialized applications

**Outline**

* Introduction
* LLM Foundations for ML and Data Science
	+ Overview of Generative AI and LLMs
	+ LLM Architecture and Training Techniques
		- Deep dive into the transformer architecture and its components
		- Exploring pre-training, fine-tuning, and transfer learning techniques
* Prompt Engineering for LLMs
	+ Introduction to Prompt Engineering
		- Techniques for creating effective prompts
		- Best practices for prompt design and optimization
	+ Developing prompts for various NLP tasks
		- Text classification, sentiment analysis, named entity recognition
* LLM Evaluation and Comparison
	+ Overview of metrics and benchmarks for evaluating LLM performance
	+ Techniques for comparing LLMs and selecting the best model for a given task
	+ Evaluating and comparing LLMs for a specific NLP task
* Fine-Tuning and Domain Adaptation
	+ Introduction to Open-Source LLMs
		- Advantages and limitations in ML and data science projects
	+ Preparing domain-specific datasets for fine-tuning LLMs
	+ Techniques for adapting LLMs to new domains and tasks using transfer learning
	+ Fine-tuning and adapting an open-source LLM for a specific domain
* Conclusion