Training Course on Transformer Architectures in Natural Language Processing

Professional Development Training Course Outline

Category	Data Science	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

The landscape of Natural Language Processing (NLP) has been revolutionized by the advent of Transformer architectures, fundamentally transforming how machines understand, generate, and interact with human language. This advanced training course delves into the intricate mechanisms and practical applications of leading Transformer models such as BERT, GPT, and T5, equipping participants with the expertise to design, implement, and fine-tune state-of-the-art NLP solutions. Through a blend of theoretical understanding and hands-on coding, attendees will master the nuances of contextual embeddings, attention mechanisms, and transfer learning, enabling them to tackle complex linguistic challenges and drive innovation in AI-driven language technologies.

Training Course on Transformer Architectures in NLP (Advanced) is designed to empower data scientists, machine learning engineers, and AI developers to harness the full potential of Transformer models. From bidirectional contextual understanding with BERT to advanced text generation with GPT and the unified text-to-text framework of T5, the course provides a deep dive into model architectures, pre-training strategies, and fine-tuning techniques. Participants will gain practical skills in leveraging Hugging Face Transformers, optimizing model performance, and deploying robust NLP applications across diverse domains, fostering a competitive edge in the rapidly evolving field of artificial intelligence.

Programme Curriculum

Training Course on Transformer Architectures in NLP (Advanced)

Introduction

The landscape of Natural Language Processing (NLP) has been revolutionized by the advent of Transformer architectures, fundamentally transforming how machines understand, generate, and interact with human language. This advanced training course delves into the intricate mechanisms and practical applications of

leading Transformer models such as BERT, GPT, and T5, equipping participants with the expertise to design, implement, and fine-tune state-of-the-art NLP solutions. Through a blend of theoretical understanding and hands-on coding, attendees will master the nuances of contextual embeddings, attention mechanisms, and transfer learning, enabling them to tackle complex linguistic challenges and drive innovation in AI-driven language technologies.

Training Course on Transformer Architectures in NLP (Advanced) is designed to empower data scientists, machine learning engineers, and AI developers to harness the full potential of Transformer models. From bidirectional contextual understanding with BERT to advanced text generation with GPT and the unified text-to-text framework of T5, the course provides a deep dive into model architectures, pre-training strategies, and fine-tuning techniques. Participants will gain practical skills in leveraging Hugging Face Transformers, optimizing model performance, and deploying robust NLP applications across diverse domains, fostering a competitive edge in the rapidly evolving field of artificial intelligence.

Course Duration

10 days

Course Objectives

1. Comprehend the core principles of the **Transformer architecture**, including **self-attention**, positional encoding, and encoder-decoder mechanisms.
2. Understand the **bidirectional contextual embeddings** of BERT, its pre-training objectives (**Masked Language Modeling, Next Sentence Prediction**), and its applications in **text classification** and **question answering**.
3. Grasp the **autoregressive nature** of GPT models, their capabilities in **text generation**, **dialogue systems**, and **creative writing**, and the concept of **few-shot learning**.
4. Learn how T5's **text-to-text framework** enables a unified approach to diverse NLP problems, including **summarization**, **translation**, and **text simplification**.
5. Gain proficiency in utilizing the **Hugging Face Transformers library** for loading, fine-tuning, and deploying pre-trained models efficiently.
6. Explore techniques for **model optimization**, including **quantization**, pruning, and knowledge distillation for efficient inference and deployment.
7. Understand and mitigate **AI bias**, ethical considerations, and responsible deployment practices for large language models (LLMs).
8. **Develop Custom Transformer Models:** Acquire the skills to adapt and extend existing Transformer architectures for specific domain-specific or niche NLP tasks.
9. Master advanced **transfer learning strategies** and fine-tuning methodologies for achieving state-of-the-art results with limited labeled data.
10. Learn comprehensive **evaluation metrics** for Transformer models across various NLP tasks, including BLEU, ROUGE, and perplexity.
11. Gain practical experience in deploying Transformer-based NLP models using **Docker** and cloud platforms for scalable solutions.
12. Investigate and implement advanced attention variants beyond the original Transformer, such as **Longformer** for longer sequences.
13. Understand how to integrate Transformer model development and deployment into robust **MLOps workflows** for continuous integration and delivery.

Organizational Benefits

- Equip teams with the ability to build sophisticated NLP applications that accurately interpret and process human language, leading to better customer insights, improved content analysis, and intelligent automation.
- Leverage pre-trained Transformer models and transfer learning to significantly reduce development time and computational resources required for new NLP solutions, fostering rapid innovation.
- Extract deeper insights from unstructured text data, enabling data-driven decision-making in areas like market intelligence, sentiment analysis, and risk assessment.
- Cultivate in-house expertise in cutting-edge NLP technologies, positioning the organization at the forefront of AI innovation and enabling the development of unique, high-value language-centric products and services.
- Learn to fine-tune and optimize Transformer models for efficient deployment, minimizing computational costs and maximizing the return on investment in AI infrastructure.
- Develop the capability to build and deploy scalable NLP solutions that can handle large volumes of text data and diverse linguistic tasks, supporting business growth and expanding operational capabilities.

Target Audience

1. Data Scientists.
2. Machine Learning Engineers
3. AI Developers.
4. Researchers in NLP/AI.
5. Product Managers (AI/Data Products).
6. Software Engineers with ML Interest.
7. Ph.D. Students in Computer Science/AI
8. Technical Leads & Architects

Course Outline

Module 1: Introduction to Transformer Architectures and NLP Landscape

- Evolution of NLP: From statistical models to neural networks and the Transformer revolution.
- The "Attention Is All You Need" paper: Core concepts and why Transformers are superior.
- Overview of the Transformer Encoder-Decoder architecture.
- Setting up your advanced NLP development environment (Python, PyTorch/TensorFlow, Hugging Face).
- **Case Study:** Analyzing the shift from RNNs/LSTMs to Transformers in Google Translate's architecture.

Module 2: The Core Components of Transformers

- Deep dive into Multi-Head Self-Attention: Query, Key, Value matrices.
- Positional Encoding: Capturing sequence order information.
- Feed-Forward Networks and Layer Normalization.
- Residual Connections and Skip Connections for stable training.
- **Case Study:** Deconstructing the attention patterns in a simple sequence-to-sequence task using visualizations.

Module 3: Bidirectional Encoder Representations from Transformers (BERT)

- BERT's pre-training tasks: Masked Language Modeling (MLM) and Next Sentence Prediction (NSP).
- Fine-tuning BERT for downstream NLP tasks.
- Tokenization strategies for BERT (WordPiece).
- Understanding BERT's contextual embeddings.

- **Case Study:** Implementing sentiment analysis on customer reviews using a fine-tuned BERT model and analyzing its performance against traditional methods.

Module 4: Generative Pre-trained Transformers (GPT Family)

- Autoregressive language modeling and its implications for text generation.
- GPT-1, GPT-2, GPT-3, and their scaling laws.
- Few-shot, one-shot, and zero-shot learning with large GPT models.
- Prompt engineering techniques for controlled text generation.
- **Case Study:** Generating creative marketing copy or product descriptions using a fine-tuned GPT-2/3 model and evaluating coherence and relevance.

Module 5: Text-to-Text Transfer Transformer (T5)

- The unified text-to-text framework: treating all NLP tasks as text generation.
- Pre-training objectives of T5.
- Exploring T5's encoder-decoder architecture.
- Fine-tuning T5 for specific tasks like summarization and question answering.
- **Case Study:** Building a text summarization system for news articles using T5 and comparing its output with human-generated summaries.

Module 6: Advanced Transformer Variants and Adaptations

- RoBERTa: Robustly optimized BERT approach.
- ALBERT: Lite BERT for self-supervised learning.
- XLNet: Generalized autoregressive pretraining.
- Longformer: Handling long sequences with sparse attention.
- **Case Study:** Comparing the performance of BERT, RoBERTa, and ALBERT on a complex text classification task.

Module 7: Tokenization and Subword Segmentation

- Byte-Pair Encoding (BPE), WordPiece, and SentencePiece.
- Handling out-of-vocabulary (OOV) words.
- Impact of tokenization on model performance.
- Building custom tokenizers.
- **Case Study:** Analyzing tokenization strategies for a specialized domain (e.g., medical texts) and their impact on model training.

Module 8: Fine-tuning Strategies and Transfer Learning

- Effective fine-tuning techniques for pre-trained Transformers.
- Adapter layers and LoRA for efficient fine-tuning.
- Multi-task learning with Transformers.
- Domain adaptation and continued pre-training.
- **Case Study:** Adapting a pre-trained Transformer for a low-resource language or a highly specialized industry dataset.

Module 9: Evaluation Metrics for NLP Tasks

- Classification metrics: Accuracy, Precision, Recall, F1-score.
- Generation metrics: BLEU, ROUGE, METEOR.

- Perplexity for language models.
- Human evaluation and qualitative assessment.
- **Case Study:** Evaluating the performance of different Transformer models on a question-answering dataset using relevant metrics.

Module 10: Model Optimization and Efficiency

- Quantization: Reducing model size and inference time.
- Pruning and weight sharing techniques.
- Knowledge distillation for creating smaller, faster models.
- Optimizing Transformer models for edge devices.
- **Case Study:** Applying quantization techniques to a BERT model and observing its impact on inference speed and accuracy.

Module 11: Responsible AI and Ethical Considerations in LLMs

- Understanding and mitigating bias in large language models.
- Fairness, accountability, and transparency (FAT) in NLP.
- Privacy concerns and data anonymization.
- Ethical guidelines for deploying generative AI.
- **Case Study:** Identifying and addressing potential biases in text generated by a GPT model for a specific application.

Module 12: Deployment and MLOps for Transformers

- Serving Transformer models with Flask/FastAPI.
- Containerization with Docker.
- Deployment on cloud platforms (AWS, GCP, Azure).
- Monitoring and maintaining NLP models in production.
- **Case Study:** Deploying a fine-tuned Transformer model as a web service for real-time text classification.

Module 13: Advanced Applications of Transformers

- Information Extraction: Named Entity Recognition (NER), Relation Extraction.
- Semantic Search and Information Retrieval.
- Machine Translation and Cross-lingual NLP.
- Dialogue Systems and Conversational AI.
- **Case Study:** Building a custom NER model using a Transformer architecture to extract specific entities from legal or medical documents.

Module 14: Vision-Language Transformers (Multimodal NLP)

- Introduction to multimodal AI.
- CLIP, DALL-E, and their underlying Transformer principles.
- Applications in image captioning and visual question answering.
- Challenges and future directions in multimodal Transformers.
- **Case Study:** Exploring an open-source multimodal Transformer for generating image captions or performing visual question answering.

Module 15: The Future of Transformers and Research Directions

- Transformers beyond NLP: Vision Transformers, Audio Transformers.

- Sparse attention mechanisms and efficient Transformers.
- The promise and challenges of Artificial General Intelligence (AGI) with LLMs.
- Open problems and cutting-edge research in Transformer architectures.
- **Case Study:** Discussing the potential impact of future Transformer developments on a specific industry or research area.

Training Methodology

This course will employ a highly interactive and hands-on training methodology, blending theoretical lectures with practical coding sessions, real-world case studies, and collaborative problem-solving.

- **Interactive Lectures:** Engaging presentations explaining complex concepts with clear examples and analogies.
- **Live Coding Demonstrations:** Step-by-step demonstrations of implementing Transformer models using Python and Hugging Face.
- **Hands-on Labs/Practicals:** Participants will work on coding exercises and mini-projects to solidify their understanding and build practical skills.
- **Case Study Analysis:** In-depth examination of how Transformer models are applied in real-world scenarios across various industries.
- **Group Discussions and Q&A:** Fostering a collaborative learning environment where participants can share insights and address challenges.
- **Project-Based Learning:** A culminating project where participants apply their learned skills to build an end-to-end NLP solution using Transformer models.
- **Expert-Led Sessions:** Taught by experienced practitioners and researchers in the field of NLP and AI.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call *+254769199797*

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Online	$2,000
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com