



BAJA
CALIFORNIA
GOBIERNO DEL ESTADO

ECONOMÍA
Secretaría de Economía e Innovación



Programa de formación en **Ingeniería de Inteligencia Artificial Generativa para docentes de Baja California**

Temario 2025

TEMARIO

Módulo 1: La repentina prevalencia de la Inteligencia Artificial (IA)

Objetivos de Aprendizaje: Comprender el impacto global y regional del avance de la IA y sus implicaciones para la educación superior y el desarrollo de la fuerza laboral.

Temas Cubiertos:

- 1.1 La evolución de la Ciencia de Datos a Machine Learning (ML).
- 1.2 La evolución de ML a Inteligencia Artificial Generativa (GenAI).
- 1.3 Arquitectura de Transformers para Grandes Modelos de Lenguaje (LLMs).
- 1.4 Más allá de la generación de texto: GenAI para audio, video e imágenes.
- 1.5 La promesa de la Inteligencia Artificial General (AGI). 1.6 La IA en el contexto educativo de México y América Latina.

Módulo 2: Datos Potenciados por IA

Objetivos de Aprendizaje: Repasar el marco de Modelos de Datos (ej: Variable de Resultado/Dependiente). Comprender los matices de los conceptos de Machine Learning, Deep Learning y Redes Neuronales, como subcampos de la IA, junto con su experiencia, fronteras y desafíos (ej: Explicabilidad).

Temas Cubiertos:

- 2.1 Fases del Modelado de Datos.
- 2.2 Aprendizaje Supervisado/No Supervisado versus Aprendizaje Semi Supervisado.
- 2.3 Aprendizaje por Refuerzo con enfoque en Agentes.
- 2.4 Redes Neuronales Recurrentes para Procesamiento de Lenguaje Natural, Redes Neuronales Convolucionales para Procesamiento de Imagen/Video.

Módulo 3: Python para el Desarrollo de IA

Objetivos de Aprendizaje: Aprender scripts de Python que manejen y preprocesen eficazmente varios formatos de datos (ej., JSON, CSV) comúnmente utilizados en flujos de trabajo de IA. Comprender las librerías esenciales de Python como Pandas y NumPy para una manipulación y análisis de datos eficientes. Aprender a construir programas en Python que interactúen con APIs externas para obtener datos o integrarse con endpoints de modelos de GenAI.

Temas Cubiertos:

- 3.1 Análisis Exploratorio de Datos y Flujos de Trabajo de IA.
- 3.2 Librerías de Python imprescindibles: Pandas, NumPy y Matplotlib.
- 3.3 Servicios Web: APIs.

Módulo 4: Transformers como Modelos Fundacionales

Objetivos de Aprendizaje: Comprender la arquitectura de los Transformers y su capacidad para generar contenido (texto, imagen, audio, video), y su relevancia como base para la construcción de Grandes Modelos de Lenguaje (LLMs).

Temas Cubiertos:

- 4.1 Procesamiento Secuencial y Paralelo.
- 4.2 Atención y Auto-Atención.
- 4.3 Aprendizaje de Contexto y Significado.
- 4.4 Procesos de Tokenización.

4.5 LLMs: Una aplicación práctica de los Transformers.

Módulo 5: Grandes Modelos de Lenguaje (LLMs)

Objetivos de Aprendizaje: Comprender qué son los LLMs y cómo se diferencian entre sí, su estado actual, fronteras y desafíos: pre-entrenamiento, alucinaciones, protección de información personal.

Temas Cubiertos:

- 5.1 Componentes fundamentales de la mayoría de los LLMs modernos: mecanismos de auto-atención, embeddings y codificaciones posicionales.
- 5.2 Pre-entrenamiento, fine-tuning (ajuste fino) y aprendizaje por refuerzo a partir de retroalimentación humana (RLHF).
- 5.3 Ingeniería de Prompts y Aprendizaje en Contexto (In-Context Learning).

Módulo 6: Modelos de Código Abierto e Inferencia Local (On-Premise)

Objetivos de Aprendizaje: Navegar por el ecosistema de IA de código abierto para descubrir, evaluar y seleccionar modelos apropiados para tareas específicas. Articular las ventajas estratégicas y los desafíos del despliegue local (on-premise), incluyendo la privacidad, el costo y la personalización.

Temas Cubiertos:

- 6.1 Una introducción a las plataformas clave (ej., Hugging Face).
- 6.2 Modelos fundacionales de peso abierto (ej., Llama, Mistral).
- 6.3 Una revisión de las consideraciones de hardware (CPU vs. GPU, VRAM) y la configuración de un entorno robusto utilizando herramientas como [aquí iría el nombre de las herramientas si se especifican, como Ollama, LM Studio, etc., si no, se puede dejar genérico].

Módulo 7: Personalización de Modelos con Parameter-Efficient Fine-Tuning (PEFT)

Objetivos de Aprendizaje: Aprender las limitaciones computacionales y financieras del fine-tuning completo y articular la propuesta de valor de PEFT. Diferenciar entre los métodos clave de PEFT, con un enfoque principal en la Adaptación de Bajo Rango (LoRA). Problema central que PEFT resuelve: personalizar grandes modelos de manera eficiente. Comprender cómo fusionar un adaptador LoRA entrenado con el modelo base para crear un modelo personalizado independiente listo para inferencia.

Temas Cubiertos:

- 7.1 PEFT (Parameter-Efficient Fine-Tuning).
- 7.2 LoRA (Low-Rank Adaptation).
- 7.3 LoRA Cuantizado (QLoRA).

Módulo 8: Integración de IA con APIs de LLMs

Objetivos de Aprendizaje: Autenticarse y ejecutar llamadas a la API de los principales proveedores comerciales de LLMs (ej.: OpenAI, Google, Anthropic). Integrar la llamada a funciones basada en API (uso de herramientas) para conceder a los modelos acceso a herramientas externas y fuentes de datos.

Temas Cubiertos:

- 8.1 Librerías cliente de Python para las APIs respectivas (ej., openai, google-generativeai, anthropic).
- 8.2 Profundización en el paradigma REST de la API.
- 8.3 Orquestación de flujos de trabajo de IA con plataformas low-code/no-code.

Módulo 9: Fundamentos de la Ingeniería de Prompts

Objetivos de Aprendizaje: Desglosar los prompts en sus componentes centrales (ej., rol, instrucción, contexto, formato) para analizar su efectividad. Aplicar patrones de prompting fundamentales, incluyendo Zero-Shot, Few-Shot y Chain-of-Thought (CoT) prompting. Estructurar prompts para controlar formatos de salida, tonos y personas específicos. Construir prompts de alta calidad y evaluar el rendimiento de los mismos. Distinguir el prompting Zero-Shot del Few-Shot.

Temas Cubiertos:

- 9.1 Principios de optimización en la comunicación con LLMs.
- 9.2 "Cookbook" de la API de OpenAI.
- 9.3 Guías de "Diseño de Prompts" de Google.
- 9.4 "Guía de Ingeniería de Prompts" por [DAIR.AI](#).

Módulo 10: Ingeniería de Prompts Avanzada y Diseño de Agentes

Objetivos de Aprendizaje: Ir más allá de los prompts básicos utilizando técnicas avanzadas como Chain-of-Thought (Cadena de Pensamiento) y ReAct para hacer que los modelos piensen paso a paso. Comprender la Llamada a Funciones (Function Calling) para otorgar a tus agentes acceso a cualquier API, desde la búsqueda web hasta bases de datos personalizadas.

Temas Cubiertos:

- 10.1 Patrones y Principios de Diseño de Agentes (Agentic Design).
- 10.2 Trazado, Depuración y Evaluación de Agentes.
- 10.3 Chain-of-Thought (Cadena de Pensamiento).
- 10.4 Librería LangChain.
- 10.5 Llamada a Funciones (Function Calling).

Módulo 11: IA Ética y Confiable

Objetivos de Aprendizaje: Explorar los riesgos, sesgos y la necesidad de transparencia de la IA. Examinar las consideraciones éticas, los impactos sociales, la privacidad y la toma de decisiones autónoma. Investigar los sesgos de datos, algorítmicos y humanos, discutiendo estrategias de mitigación para la equidad y el despliegue ético de la IA.

Temas Cubiertos:

- 11.1 Marcos de Ética de la IA.
- 11.2 Equidad: Cómo evitar sesgos en la práctica.
- 11.3 El humano en el bucle (Human-in-the-loop): Aumentando la precisión general del modelo.
- 11.4 Inteligencia Artificial Explicable (XAI) en NeMo.

Módulo 12: Generación Aumentada por Recuperación (RAG) - Parte 1

Objetivos de Aprendizaje: Construir bases de datos vectoriales y utilizar la búsqueda por similitud.

Temas Cubiertos:

- 12.1 Embeddings: Representaciones numéricas de objetos del mundo real que los sistemas de ML e IA utilizan para comprender dominios de conocimiento complejos como los humanos.
- 12.2 FAISS: Búsqueda de Similitud de IA de Facebook.
- 12.3 Chroma: Base de datos vectorial nativa de IA diseñada para construir aplicaciones de IA.

Módulo 13: RAG - Parte 2 (Construcción de Aplicaciones)

Objetivos de Aprendizaje: Desarrollar aplicaciones RAG completas utilizando LangChain.

Temas Cubiertos:

- 13.1 Encadenamiento de almacenes vectoriales (Vector Store Chaining).
- 13.2 Memoria: ¿Un incentivo o un obstáculo?
- 13.3 Formatos de respuesta a consultas (Query Response Formats). 13.4 RAG en despliegue.

Módulo 14: Evaluación de Errores, Manejo y Explicabilidad

Objetivos de Aprendizaje: Evaluar, explicar y depurar el comportamiento del modelo.

Temas Cubiertos:

- 14.1 Explicabilidad del Modelo usando SHAP (SHapley Additive exPlanations).
- 14.2 Explicaciones Locales Interpretables Ag-nósticas al Modelo (LIME).
- 14.3 Detección y prevención de alucinaciones.
- 14.4 Evaluación humana con bucles de retroalimentación (feedback loops).

Módulo 15-16: Prototipado de Aplicaciones con Streamlit (compatible con Python)

Objetivos de Aprendizaje: Construir interfaces de usuario (UIs) desplegables para flujos de trabajo de LLMs.

Temas Cubiertos: Gradio, Streamlit, integración básica de UI.

Recursos Sugeridos:

- 15.1 Streamlit y Gradio para prototipado rápido de UI.
- 15.2 Hugging Face Spaces para el despliegue de LLMs.

Módulo 17: Métricas Clave en MLOps

Objetivos de Aprendizaje: Diferenciar y seleccionar las métricas correctas entre el rendimiento del modelo, la eficiencia operativa y el impacto comercial. Rastrear y analizar métricas operacionales clave, incluyendo la latencia de inferencia, el rendimiento (throughput) y la utilización de recursos (CPU/GPU) para gestionar costos y asegurar una experiencia de usuario de calidad.

Temas Cubiertos:

- 17.1 Los Tres Pilares de las Métricas de MLOps: Calidad del Modelo, Salud Operacional e Impacto Comercial.
- 17.2 Weights & Biases y plataformas MLOps.
- 17.3 Microservicios de Inferencia NVIDIA (NIM) y Microservicios NVIDIA NeMo.

Módulo 18-19: Desarrollo de Proyecto Final (Capstone Project)

Objetivos de Aprendizaje: Finalizar un proyecto funcional basado en LLMs integrando múltiples habilidades.

Temas Cubiertos:

- 18.1 Planificación de tu proyecto: Documentación inteligente desde el inicio.
- 18.2 Entornos Sandbox, Staging y Producción.
- 18.3 Pruebas/ajustes iterativos para la mejora general del LLM. 18.4 Despliegue basado en la nube.



Módulo 20: Presentación del Proyecto Final (Capstone Showcase) y Preparación para la Certificación

Objetivos de Aprendizaje: Presentar y defender proyectos de IA; prepararse para el examen NVIDIA NCA-GENL.

Temas Cubiertos:

- 20.1 Presentación del proyecto.
- 20.2 Retroalimentación entre pares (Peer-sourced feedback).
- 20.3 Dominios del examen.
- 20.4 Ejemplos de ítems de prueba.

SYLLABUS (English version)

Module 1: The unsudden prevalence of Artificial Intelligence (AI)

Learning Objectives: Understand global and regional impact of the advancement of AI and its implications for higher education and workforce development.

Topics Covered:

- 1.1 Data Science evolving into Machine Learning (ML)
- 1.2 ML evolving into Generative Artificial Intelligence (GenAI).
- 1.3 Transformers architecture for Large Language Models (LLMs)
- 1.4 Beyond text generation, GenAI for audio, video and images.
- 1.5 The promise of Artificial General Intelligence.
- 1.6 AI the educational context of Mexico and Latin America.

Module 2: Data Powered by AI

Learning Objectives: Recap Data Models framework (i.e.: Outcome/Dependent variable). Understand the nuances of Machine Learning, Deep Learning and Neural Networks concepts, as AI's subfields, along with its expertise, frontiers and challenges (i.e.: Explicability).

Topics Covered:

- 2.1 Phases of Data Modeling.
- 2.2 Supervised/Unsupervised Learning versus Semi Supervised Learning.
- 2.3 Reinforcement Learning with focus on Agents.
- 2.4 Recurrent Neural Networks for Natural Language Processing, Convolutional Neural Network for Image/Video Processing.

Module 3: Python for AI Development

Learning Objectives: Learn Python scripts that effectively handle and preprocess various data formats (e.g., JSON, CSV) commonly used in AI workflows. Understand essential Python libraries like Pandas and NumPy for efficient data manipulation and analysis. Learn how to build Python programs that interact with external APIs to fetch data or integrate with GenAI model endpoints.

Topics Covered:

- 3.1 Exploratory Data Analysis and AI Workflows.
- 3.2 Python must-have libraries Pandas and NumPy, Matplotlib,
- 3.3 Web Services: APIs.

Module 4: Transformers as Foundation Models.

Learning Objectives: Understand the architecture of Transformers and its capability to generate content (text, image, audio, video), and their relevance as foundation for building Large Language Models (LLMs).

Topics Covered:

- 4.1 Sequential and Parallel Processing,
- 4.2 Attention and Self-Attention.
- 4.3 Context and Meaning Learning,
- 4.4 Tokenization Processes.
- 4.5 LLMs: A practical application of Transformers.

Module 5: Large Language Models (LLMs)

Learning Objectives: Understand what LLMs are, and how they differ from each other, their current state, frontiers and challenges: Pre-training, Hallucination, Personal Information Protection.

Topics Covered:

- 5.1 Foundational components of most modern LLMs: self-attention mechanisms, embeddings, and positional encodings.
- 5.2 Pre-training, fine-tuning, and reinforcement learning from human feedback (RLHF).
- 5.3 Prompt Engineering and In-Context Learning.

Module 6: Open-Source Models and On-Premise Inference

Learning Objectives: Navigate the open-source AI ecosystem to discover, evaluate, and select appropriate models for specific tasks. Articulate the strategic advantages and challenges of on-premise deployment, including privacy, cost, and customization.

Topics Covered:

- 6.1 An introduction to the key platforms (e.g., Hugging Face).
- 6.2 Foundational open-weight models (e.g., Llama, Mistral).
- 6.3 A review of hardware considerations (CPU vs. GPU, VRAM) and setting up a robust environment using tools like.

Module 7: Customizing Models with Parameter-Efficient Fine-Tuning (PEFT)

Learning Objectives: Learn the computational and financial limitations of full fine-tuning and articulate the value proposition of PEFT. Differentiate between key PEFT methods, with a primary focus on Low-Rank Adaptation (LoRA). Core problem that PEFT solves: customizing large models efficiently. Understand how to merge a trained LoRA adapter with the base model to create a standalone, customized model ready for inference.

Topics Covered:

- 7.1 PEFT.
- 7.2 LoRA.
- 7.3 Quantized LoRA (QLoRA).

Module 8: Integrating AI with LLM APIs

Learning Objectives: Authenticate and execute API calls to major commercial LLM providers (i.e.: OpenAI, Google, Anthropic). Integrate API-based function calling (tool use) to grant models access to external tools and data sources.

Topics Covered:

- 8.1 Python client libraries for the respective APIs (e.g., openai, google-generativeai, anthropic).
- 8.2 Deep dive on API REST paradigm.
- 8.3 AI Workflow orchestration with low-code/no-code platforms.

Module 9: Prompt Engineering Essentials

Learning Objectives: Deconstruct prompts into their core components (e.g., role, instruction, context, format) to analyze their effectiveness. Apply foundational prompting patterns, including Zero-Shot, Few-Shot, and Chain-of-Thought (CoT) prompting. Structure prompts to control for specific output

formats, tones, and personas. Construct high-quality prompts and evaluate prompt performance. Distinguish Zero-Shot from Few-Shot prompting.

Topics Covered:

- 9.1 Optimization principles on communicating with LLMs.
- 9.2 OpenAI API "Cookbook".
- 9.3 Google's "Prompt design" guides.
- 9.4 "Prompt Engineering Guide" by DAIR.AI.

Module 10: Advanced Prompting Engineering & Agent Design

Learning Objectives: Go beyond basic prompts using advanced techniques like Chain-of-Thought and ReAct to make models think step-by-step. Understand Function Calling to give your agents access to any API, from web search to custom databases.

Topics Covered:

- 10.1 Agentic Design Patterns & Principles.
- 10.2 Agent Tracing, Debugging, and Evaluation.
- 10.3 Chain-of-Thought.
- 10.4 LangChain library
- 10.5 Function Calling.

Module 11: Ethical & Trustworthy AI

Learning Objectives: Explore AI's risks, biases, and need for transparency. Examine ethical considerations, societal impacts, privacy, and autonomous decision-making. Investigate data, algorithmic, and human biases, discussing mitigation strategies for fairness and ethically deploy AI.

Topics Covered:

- 11.1 AI Ethics frameworks.
- 11.2 Fairness: How to practically avoid biases.
- 11.3 Human-in-the-loop: Boosting model's overall accuracy.
- 11.4 Explainable Artificial Intelligence (XAI) on NeMo.

Module 12: Retrieval-Augmented Generation (RAG) - Part 1

Learning Objectives: Build vector databases and use similarity search.

Topics Covered:

- 12.1 Embeddings: Numerical representations of real-world objects ML and AI systems use to understand complex knowledge domains as humans.
- 12.2 FAISS: Facebook AI Similarity Search.
- 12.3 Chroma: AI-native vector database designed for building AI applications.

Module 13: RAG - Part 2 (Application Building)

Learning Objectives: Develop full RAG applications using LangChain.

Topics Covered:

- 13.1 Vector store chaining,
- 13.2 Memory: An incentive or an obstacle.
- 13.3 Query response formats.
- 13.4 RAG in deployment.

Module 14: Error Evaluation, Handling and Explainability.

Learning Objectives: Evaluate, explain, and debug model behavior. Topics Covered:

- 14.1 Model's Explainability using SHAP (eSHapley Additive exPlanations).
- 14.2 Local Interpretable Model-Agnostic Explanations (LIME).
- 14.3 Hallucination detection, and prevention.
- 14.4 Human evaluation with feedback loops.

Module 15-16: App Prototyping with Streamlit (Phyton friendly)

Learning Objectives: Build deployable UIs for LLM workflows. Topics Covered: Gradio, Streamlit, basic UI integration.

Suggested Resources:

- 15.1 Streamlit and Gradio for rapid UI prototyping.
- 15.2 Hugging Face Spaces for LLM's deployment.

Module 17: Key Metris on MLOps

Learning Objectives: Differentiate between and select the right metrics across model performance, operational efficiency, and business impact. Track and analyze key operational metrics including inference latency, throughput, and resource utilization (CPU/GPU) to manage costs and ensure a quality user experience.

Topics Covered:

- 17.1 The Three Pillars of MLOps Metrics: Model Quality, Operational Health, and Business Impact.
- 17.2 Weights & Biases and MLOps platforms.
- 17.3 NVIDIA Inference Microservices (NIM) and NVIDIA NeMo Microservices.

Module 18-19: Capstone Project Development

Learning Objectives: Finalize a functional LLM-based project integrating multiple skills.

Topics Covered:

- 18.1 Planning your project: Smart documentation from start.
- 18.2 Sandbox, Staging and Production Environments.
- 18.3 Iterative testing/tuning for overall LLM's improvement.
- 18.4 Cloud-based deployment.

Module 20: Capstone Showcase & Certification Prep

Learning Objectives: Present and defend AI projects; prepare for NVIDIA NCA-GENL exam.

Topics Covered:

- 20.1 Project presentation.
- 20.2 Peer-sourced feedback.
- 20.3 Exam domains.
- 20.4 Sample test items.