

NanoGPT con PyTorch: Arquitectura Transformer paso a paso

Taller • 8 horas • Zoom

Construye un modelo de lenguaje tipo NanoGPT completamente desde cero: sin modelos pre-entrenados, sin librerías de alto nivel — solo PyTorch. Para entender la arquitectura usaremos una analogía empresarial: el Transformer como una organización donde cada bloque es un departamento especializado que procesa, organiza y comunica información. Ensamblarás esa "empresa" bloque por bloque, entrenarás una versión reducida en vivo para formalizar el flujo completo, y servirás la inferencia en una interfaz web con Gradio. Aprenderás la teoría esencial, implementarás el modelo paso a paso y entenderás cómo se conectan sus principales elementos para generar texto.

Contenido

1. La arquitectura Transformer — el organigrama

Panorama de la arquitectura clásica: codificador, decodificador y cómo se conectan — la estructura general de la organización y cómo fluye la información entre sus áreas. Formulación probabilística del modelo de lenguaje: qué significa realmente predecir el siguiente token.

2. Embeddings y codificación posicional — el expediente

Capas de embeddings desde cero: cómo cada token recibe su "expediente" con la información que lo describe. Codificación posicional determinista (funciones sinusoidales) vs. aprendible (matriz optimizada durante el entrenamiento) — cómo el modelo sabe en qué lugar de la fila está cada quien.

3. Mecanismos de atención — las juntas de trabajo

El corazón del Transformer, implementado línea por línea: autoatención (Self-Attention), atención cruzada (Cross-Attention) y atención multicabeza (Multi-Head Attention) — la junta donde cada palabra decide a quién escuchar, y por qué conviene tener varias juntas en paralelo.

4. Ensamblado del modelo — la empresa en operación

Anidamos los departamentos construidos en un modelo generativo tipo NanoGPT. Revisión del flujo completo de información: del texto crudo a la predicción del siguiente token.

5. Entrenamiento autosupervisado — la capacitación

Entrenamiento como predictor de texto sobre fragmentos literarios. Entrenarás en vivo una versión reducida del modelo para dominar el pipeline completo, y trabajaremos con la versión completa ya entrenada (~12 hrs de cómputo) para los resultados finales. Evaluación del modelo como predictor de texto.

6. Inferencia y despliegue — la empresa atendiendo clientes

Generación de texto con tu modelo y despliegue de la inferencia en una interfaz web con Gradio.

7. Escalamiento y siguientes pasos + Q&A

Cómo llevar esta arquitectura a proyectos más ambiciosos: qué cambia al escalar datos, parámetros y cómputo. Sesión abierta de preguntas y respuestas.

Al terminar tendrás:

- Un Transformer implementado desde cero en PyTorch
- El pipeline completo dominado: arquitectura → entrenamiento → evaluación → despliegue
- Un modelo generativo funcionando en una interfaz web
- Constancia de participación verificable en línea

¿Para quién es este taller?

- **Desarrolladores e ingenieros de software** que usan LLMs vía API y quieren entender qué hay dentro de la caja: cómo funciona realmente el modelo que están consumiendo.
- **Estudiantes y practicantes de ML/Data Science** que conocen redes neuronales pero nunca han implementado un Transformer completo — y quieren pasar de leer el paper a escribir el código.
- **Profesionales técnicos** que buscan una base sólida antes de adentrarse en fine-tuning, RAG o arquitecturas más avanzadas. Aquí se construye el fundamento sobre el que descansa todo lo demás.

Este taller NO es para ti si:

- Buscas aprender a usar ChatGPT o herramientas de IA sin programar.
- Nunca has escrito código en Python — necesitas nociones básicas de PyTorch y redes neuronales.
- Quieres entrenar modelos de producción a gran escala de datos y de número de parámetros. Aquí construyes sólo el fundamento.

Clases

4 sesiones de 2 horas (8 hrs en total)

Lunes 3, miércoles 5, viernes 7 y lunes 10 de agosto

20:00–22:00, hora de la Ciudad de México

Las clases son impartidas vía Zoom y tendrás acceso a las grabaciones por 1 año después de terminado el curso.

Instructor

M. en C. Israel Prado — [ver biografía](#)

Testimonios

Hemos capacitado a más de 2,000 personas en IA — [Lee lo que dicen nuestros alumnos](#)

Medios de pago

Tarjeta de crédito o débito, Apple Pay, Google Pay y PayPal, en cursos.actumlogos.com

¿Problemas con tu pago?

Escríbenos por [WhatsApp](#) (te atiende Claudia Montaña)

Constancia de participación con folio único, verificable en línea

Al terminar recibirás una constancia de participación firmada por el instructor y el director de Actumlogos — agente capacitador registrado ante la Secretaría del Trabajo y Previsión Social. Cada constancia incluye un folio único con el que cualquier persona puede verificar su autenticidad en línea, y podrás compartirla directamente en LinkedIn.

Registro STPS: ZAGE-810930-FW2-0005

Ejemplo:

