

Propuesta de Microcredencial Universitaria

Título: Inteligencia Artificial y Análisis de Datos No-Code para Ciencias Sociales

Créditos: 3 ECTS (30 horas lectivas)

Centro: Universidad del País Vasco / Euskal Herriko Unibertsitatea (EHU)

Dirección Académica: Asier Amezaga Etxebarria (UPV/EHU) y Jordi Morales i Gras (EUDAN)

1. Justificación y Pertinencia

La Inteligencia Artificial (IA) ha trascendido su nicho originario para consolidarse como una herramienta transversal indispensable en prácticamente todas las facetas de la sociedad contemporánea. En el ámbito de las Ciencias Sociales, la capacidad de analizar grandes volúmenes de datos y utilizar modelos predictivos es crítica. Sin embargo, la barrera de entrada de la programación (Python/R) a menudo excluye a profesionales brillantes del análisis avanzado.

Esta microcredencial nace para **democratizar el acceso a la IA**, entendiendo por esta no solo los modelos generativos, sino también el Machine Learning (ML) y el Deep Learning (DL). Su filosofía se basa en el movimiento **No-Code/Low-Code**: permitir que profesionales de la sociología, la ciencia política y la comunicación diseñen y ejecuten flujos de trabajo analíticos utilizando entornos visuales (Orange Data Mining). Además, se integra la competencia de **trabajar con código junto a IAs**, capacitando a profesionales para orquestar procesos y automatizaciones mediante la colaboración con modelos generativos.

El curso aborda no solo la **capacitación técnica** instrumental, sino también la **soberanía tecnológica**: enseñando a desplegar modelos locales (*on-premise*) para garantizar la privacidad de los datos de investigación y fomentando una visión crítica sobre la gobernanza y la ética algorítmica.

2. Objetivos de Aprendizaje

Al finalizar la microcredencial, el alumnado será capaz de:

1. **Entender y Aplicar ML:** Comprender la lógica de los algoritmos de Machine Learning y entrenar modelos predictivos y de clustering mediante flujos de trabajo visuales.

2. **Analizar Texto (NLP):** Extraer insights de fuentes documentales masivas sin programación.
3. **Dominar la IA Generativa:** Utilizar estratégicamente modelos comerciales (ChatGPT, Gemini) y desplegar modelos Open Source (Llama, Mistral) en local para entornos seguros.
4. **Auditar y Gobernar:** Evaluar las implicaciones éticas, los sesgos y la infraestructura material que sostiene la IA.

3. Destinatarios

Profesionales y graduados en Ciencias Sociales y de la Comunicación: Sociología, Ciencia Política, Periodismo, Comunicación Audiovisual, Antropología, Relaciones Internacionales y otros campos afines que deseen incorporar la IA en sus metodologías de investigación o práctica profesional sin requerir conocimientos previos de programación.

4. Estructura del Programa (30 Horas / 3 ECTS)

Módulo 1: Fundamentos de IA y Machine Learning Visual (8 Horas)

- **Docencia:** Amaia Perez de Arriluzea (EUDAN) y Jordi Morales (EUDAN).
- **Metodología:** Taller práctico con Orange Data Mining (Software Open Source).

Este módulo rompe la barrera del código. Los estudiantes aprenderán a "programar visualmente" conectando widgets para cargar datos, procesarlos y extraer patrones.

- **Introducción al entorno No-Code:** Instalación y lógica de flujos de trabajo en Orange.
- **Aprendizaje Supervisado (Predicción):** Entrenamiento de modelos de clasificación y regresión sin escribir una línea de código.
- **Aprendizaje No Supervisado (Exploración):** Técnicas de Clustering para segmentación de grupos sociales y reducción de dimensionalidad.
- **Validación de Modelos:** Interpretación crítica de matrices de confusión y métricas de acierto (evitar el "engaño estadístico").

Módulo 2: Minería de Textos y NLP para Ciencias Sociales (8 Horas)

- **Docencia:** Amaia Perez de Arriluzea (EUDAN) y Jordi Morales (EUDAN).
- **Metodología:** Flujos de trabajo visuales en Orange para datos no estructurados.

El texto es una materia prima para las ciencias sociales. Este módulo enseña a transformarlo en dato cuantificable.

- **Preprocesamiento visual:** Limpieza de corpus, tokenización y filtrado.
- **Embeddings y Modelos Neuronales:** Introducción visual a cómo las máquinas "entienden" el significado transformando palabras en vectores matemáticos (embeddings), superando el análisis por palabras clave.
- **Análisis de Redes Semánticas:** Visualización y análisis de las relaciones entre conceptos para descubrir estructuras de significado en el texto.

Módulo 3: La Revolución de la IA Generativa: Nube y Soberanía Local (10 Horas)

- **Docencia:** Amaia Perez de Arriluzea (EUDAN) y Jordi Morales (EUDAN).
- **Metodología:** Híbrida (SaaS comercial + Despliegue Local).

Un recorrido desde el uso avanzado de chatbots hasta la infraestructura propia y la automatización.

- **Estrategias de IA Generativa:** Exploración del espectro de técnicas: desde el *Prompt Engineering* avanzado y la introducción al *Fine Tuning*, hasta el uso de RAG (*Retrieval-Augmented Generation*) para trabajar con documentos propios.
- **Soberanía del Dato (Local LLMs):** Taller de instalación y uso de **Ollama**. Cómo ejecutar una IA potente en el propio ordenador para analizar datos sensibles sin enviarlos a la nube.
- **Modelos Open Source vs. Open Weights:** Entender el ecosistema abierto (Hugging Face, Llama, Mistral) y las diferencias con los modelos comerciales.

Módulo 4: Gobernanza, Ética e Infraestructuras Tecnológicas (4 Horas)

- **Docencia:** Asier Amezaga (UPV/EHU) y Jordi Morales (EUDAN).

- **Metodología:** Seminario crítico y análisis de casos.

La IA no es magia, es industria e infraestructura.

- **Materialidad de la IA:** Centros de datos, consumo energético y geopolítica del hardware (GPUs).
- **Gobernanza y Regulación:** Impacto de la *EU AI Act* en la investigación social y la administración pública.
- **Ética y Sesgos:** Auditoría de datasets y modelos. ¿Cómo auditar una "caja negra"?
- **Modelos Abiertos:** La importancia del Open Source para la reproducibilidad científica y la transparencia pública.

5. Metodología Docente

- **Formato:** Virtual síncrono.
- **Enfoque Práctico:** Las sesiones teóricas se reducen al mínimo indispensable. El 80% del tiempo se dedica a resolver casos prácticos guiados utilizando software gratuito y de código abierto (Orange, Ollama).
- **Materiales:** Se proporcionarán datasets reales del ámbito social y flujos de trabajo (*workflows*) pre-diseñados.

6. Sistema de Evaluación

Para la obtención de la microcredencial no se realizará un examen tradicional, sino una **evaluación continua basada en retos**:

- Entrega de un flujo de trabajo en Orange (análisis de un dataset social).
- Pequeño ejercicio de *Prompt Engineering* o despliegue local documentado.
- Asistencia y participación activa en el módulo de gobernanza.

7. Calendario y Cronograma Propuesto

- **Fechas:** Del 21 de abril al 21 de mayo de 2026.
- **Horario:** Martes y jueves de 16:00 a 19:00h.
- **Duración:** 5 semanas (10 sesiones síncronas de 3 horas).