Exportación de modelos ML

Freddy Rospigliosi Cohaila

¿Por qué exportar un modelo?

- Permite reutilizar modelos entrenados sin reentrenar.
- Facilita la integración en aplicaciones reales (web, móviles, APIs).
- Acelera la inferencia en producción.
- Favorece el versionado y la colaboración entre equipos.

¿Qué se exporta exactamente?

- Estructura del modelo (algoritmo y arquitectura).
- Pesos/coeficientes (los valores aprendidos).
- Preprocesamiento (opcional pero recomendado).
- Configuración.

Formatos comunes de exportación

- .pkl o .joblib: Python nativo (scikit-learn, XGBoost, etc.).
- h5 o SavedModel (TensorFlow/Keras)
- ONNX: formato interoperable entre frameworks
- PMML: para uso en entornos empresariales

TensorFlowJS

- Es una biblioteca (library) de JavaScript para entrenar y ejecutar modelos de ML en el navegador o en Node.js.
- Permite usar ML sin servidores, directamente en el frontend.
- Reutiliza modelos entrenados en Python o entrena modelos desde cero con JS.

Ventajas

- V No se necesita backend
- **V** Corre en tiempo real, en el navegador
- Interactivo y accesible (solo se necesita un navegador)

¿Dónde se usa TensorFlow.js?

- Reconocimiento facial o de gestos en apps web
- Clasificación de texto en tiempo real
- Aplicaciones educativas interactivas
- Juegos que usan inteligencia artificial
- Realidad aumentada (AR)

Código para cargar modelo en el navegador

```
<script src="https://cdn.jsdelivr.net/npm/@tensorflow/tfjs"></script>
<script>
  async function cargarModelo() {
   const model = await tf.loadLayersModel('modelo_web/model.json');
   const prediccion = model.predict(tf.tensor2d([[5.1, 3.5, 1.4, 0.2]], [1,
 4]));
   prediccion.print();
  cargarModelo();
• </script>
```

¿Qué es un tensor?

- Un tensor es solo una forma de organizar datos en diferentes dimensiones.
 - Puedes imaginarlo como una caja o contenedor de números.
- Escalar (0D): estatura = 1.65 Es un solo número → Tensor de 0 dimensiones
- Vector (1D): estaturas = [1.65, 1.70, 1.58] Lista → Tensor de 1 dimensión (vector)
- **Matriz (2D):** aulas = [
- [1.65, 1.70, 1.58], # Aula 1
- [1.75, 1.60, 1.68] # Aula 2
-]

Tensor 3D

La estatura de estudiantes en 2 aulas, en 2 turnos (mañana y tarde): Cubo → Tensor de 3 dimensiones

```
turnos = [
  [ # Turno mañana
    [1.65, 1.70, 1.58], # Aula 1
    [1.75, 1.60, 1.68] # Aula 2
  [ # Turno tarde
   [1.55, 1.62, 1.59], # Aula 1
    [1.80, 1.65, 1.66] # Aula 2
```

¿Por qué importa esto?

- En Machine Learning, los **datos se convierten en tensores** para que los modelos puedan entenderlos y operar sobre ellos. Por ejemplo:
 - Imágenes → tensor 3D (alto, ancho, canales de color)
 - Audio → tensor 1D o 2D
 - Tablas → tensor 2D

¿Qué es Teachable Machine?

- Teachable Machine es una herramienta gratuita creada por Google que permite a cualquier persona entrenar modelos de machine learning sin necesidad de programar.
- Ideal para aprender los conceptos básicos de IA y machine learning
- Funciona desde el navegador
- Puedes entrenar modelos con imágenes, sonidos o poses corporales

¿Cómo funciona?

- Teachable Machine sigue un proceso simple de tres pasos:
- Entrena:
 - Captura o sube tus datos (fotos, sonidos, poses)Organiza los datos en clases (por ejemplo: "Gato" y "Perro")Haz clic en "Entrenar modelo" y el sistema aprende a reconocer las diferencias.
- Prueba
 - Prueba en tiempo real usando tu cámara o micrófono.
 - El modelo intentará predecir a qué clase pertenece tu entrada.
- Exporta
- Puedes descargar el modelo entrenado para usarlo en:
 - Tu sitio web (con **TensorFlow.js**)
 - Aplicaciones (Android, Python)
 - Herramientas como P5.js, Scratch, etc.

¿Qué puedes entrenar con Teachable Machine?

Tipo de proyecto	Qué analiza	Ejemplo
Imagen	La cámara o fotos cargadas	Clasificar emojis con tu rostro
Audio	Sonidos captados por micrófono	Distinguir palmadas vs. Silbidos
Pose	Movimiento corporal con cámara	Controlar un juego con tus brazos