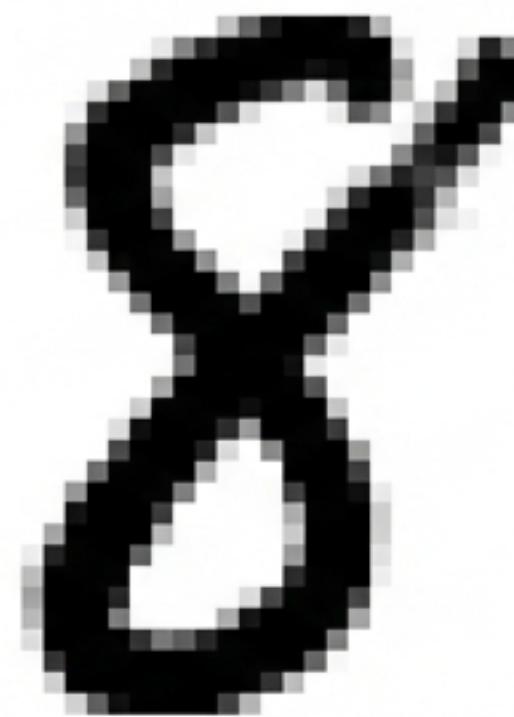


Del Dígito a la Moda: La Evolución de un Clásico de la IA

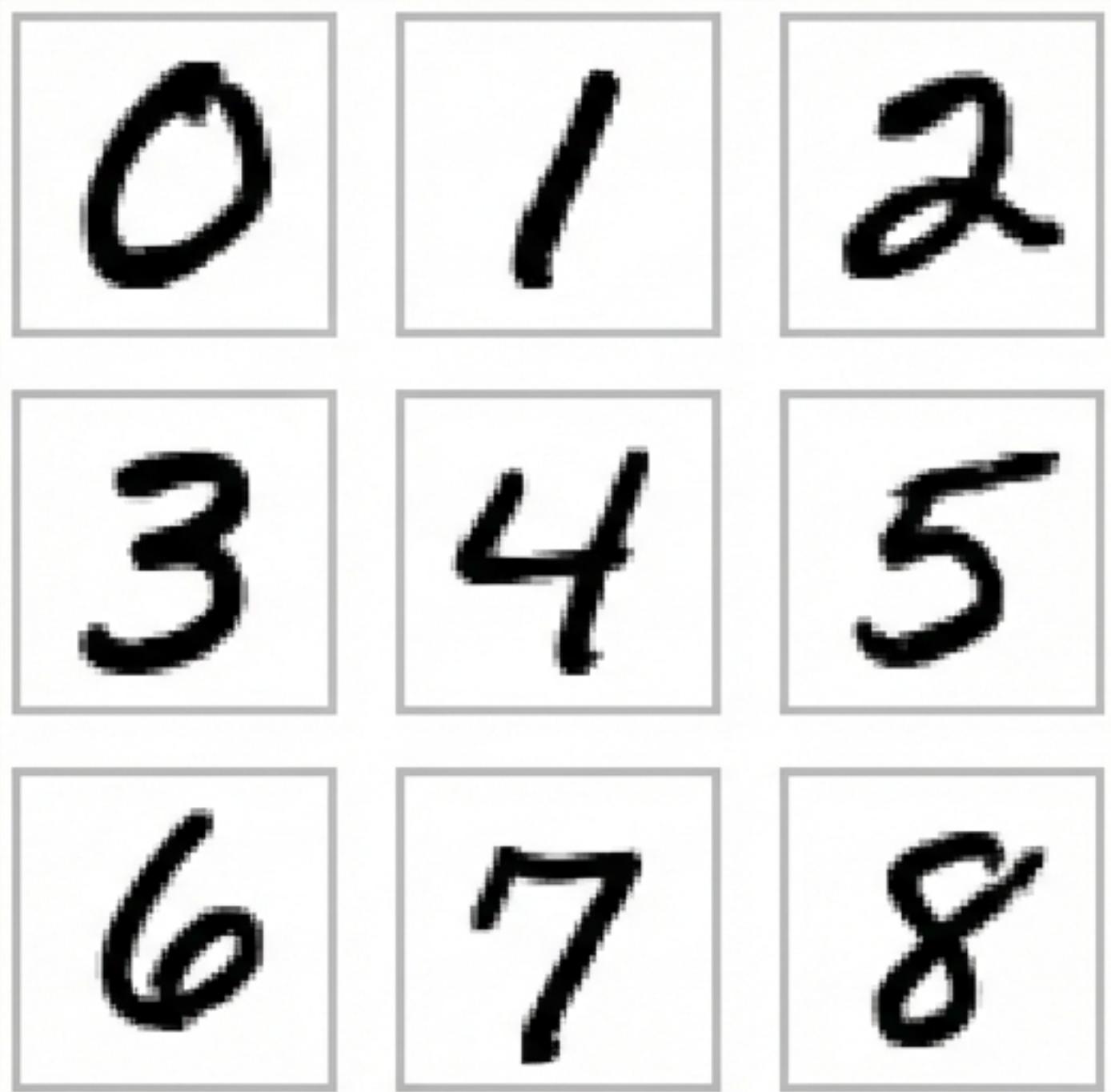
Un análisis práctico del dataset Fashion-MNIST con TensorFlow



El "Hola, Mundo" de la Visión por Computadora

Durante casi dos décadas, el dataset MNIST fue el punto de partida universal para cualquiera que se adentrara en el aprendizaje automático.

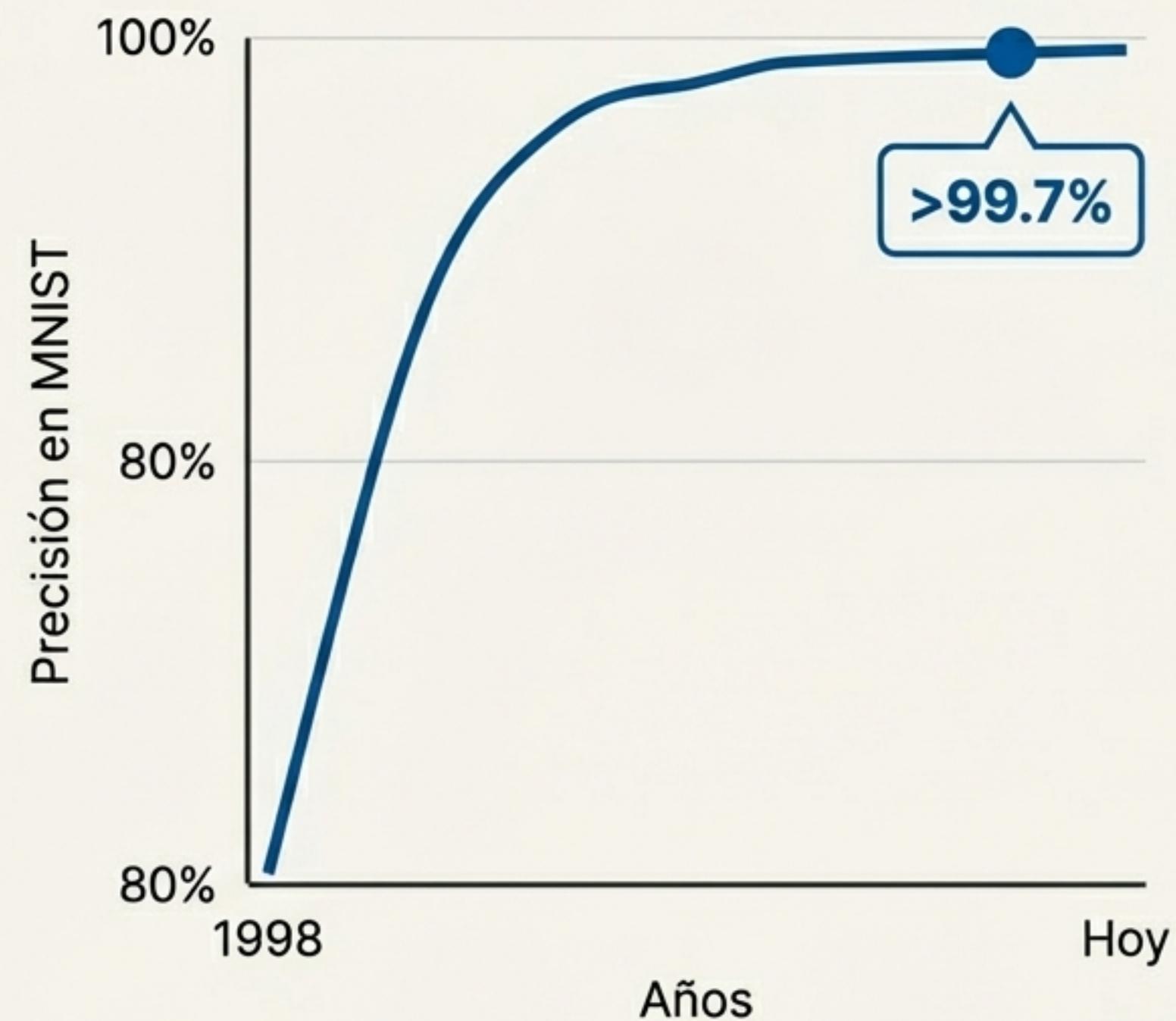
- Contiene 70,000 imágenes de dígitos manuscritos (0-9).
- Introducido en 1998 por LeCun et al., se convirtió en la prueba de fuego para nuevos algoritmos de clasificación.
- Su simplicidad y tamaño manejable permitieron la creación rápida de prototipos y la democratización de la investigación en el campo.



El Problema de la Perfección: Cuando un Benchmark Queda Obsoleto

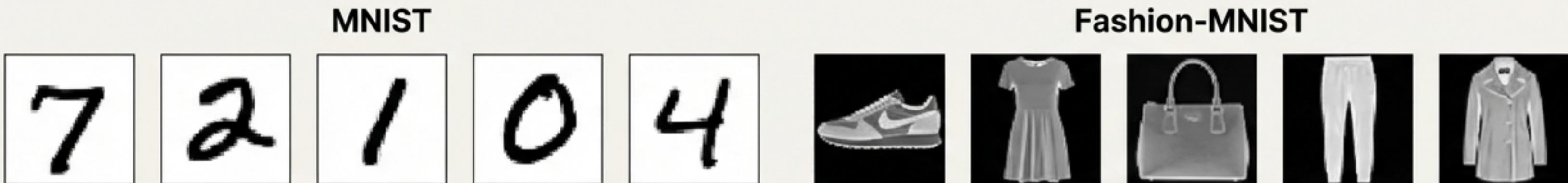
Con el avance de las redes neuronales profundas, los modelos comenzaron a alcanzar precisiones superiores al 99.7% en MNIST.

- Este nivel de rendimiento, casi perfecto, dificultaba la diferenciación entre las arquitecturas de modelos más nuevas y potentes.
- El dataset estaba ‘resuelto’, lo que limitaba su utilidad para la investigación de vanguardia.
- La comunidad de IA necesitaba un nuevo desafío que fuera igual de accesible pero más representativo de los problemas de visión del mundo real.



La Respuesta: Fashion-MNIST, un Reemplazo Directo

Creado por Zalando Research en 2017, Fashion-MNIST fue diseñado como un "drop-in replacement" (reemplazo directo) del MNIST original.

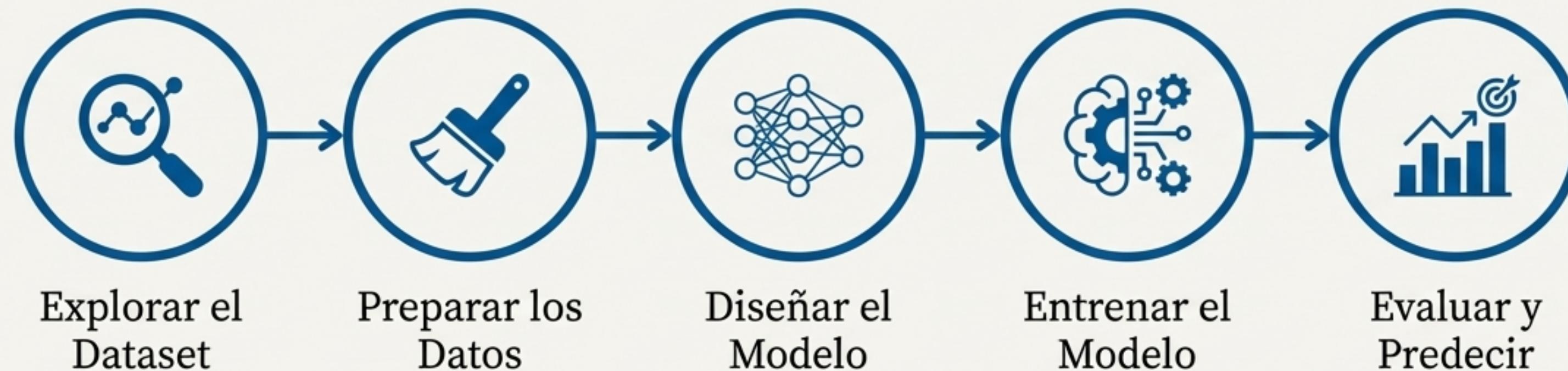


Característica	MNIST	Fashion-MNIST
Contenido	Dígitos manuscritos	Artículos de moda
Nº de Clases	10	10
Imágenes Entrenamiento	60,000	60,000
Imágenes Prueba	10,000	10,000
Formato Imagen	28x28 Escala de grises	28x28 Escala de grises

“Fashion-MNIST pretende ser un reemplazo directo del dataset MNIST original para comparar algoritmos de aprendizaje automático.” – Xiao et al., 2017

Construyendo un Estilista de IA: Un Caso Práctico

A continuación, seguiremos un flujo de trabajo completo de Machine Learning usando TensorFlow para entrenar una red neuronal que clasifique imágenes de Fashion-MNIST. Este proceso se divide en cinco pasos clave.



Paso 1: Explorando el Universo de la Moda

Puntos Clave del Dataset

- **Imágenes de Entrenamiento:** 60,000
- **Imágenes de Prueba:** 10,000
- **Dimensiones:** 28x28 píxeles
- **Canales de Color:** 1 (Escala de grises)
- **Valores de Píxel:** 0 a 255

Las 10 Clases

0: T-shirt/top
1: Trouser
2: Pullover
3: Dress
4: Coat

5: Sandal
6: Shirt
7: Sneaker
8: Bag
9: Ankle boot



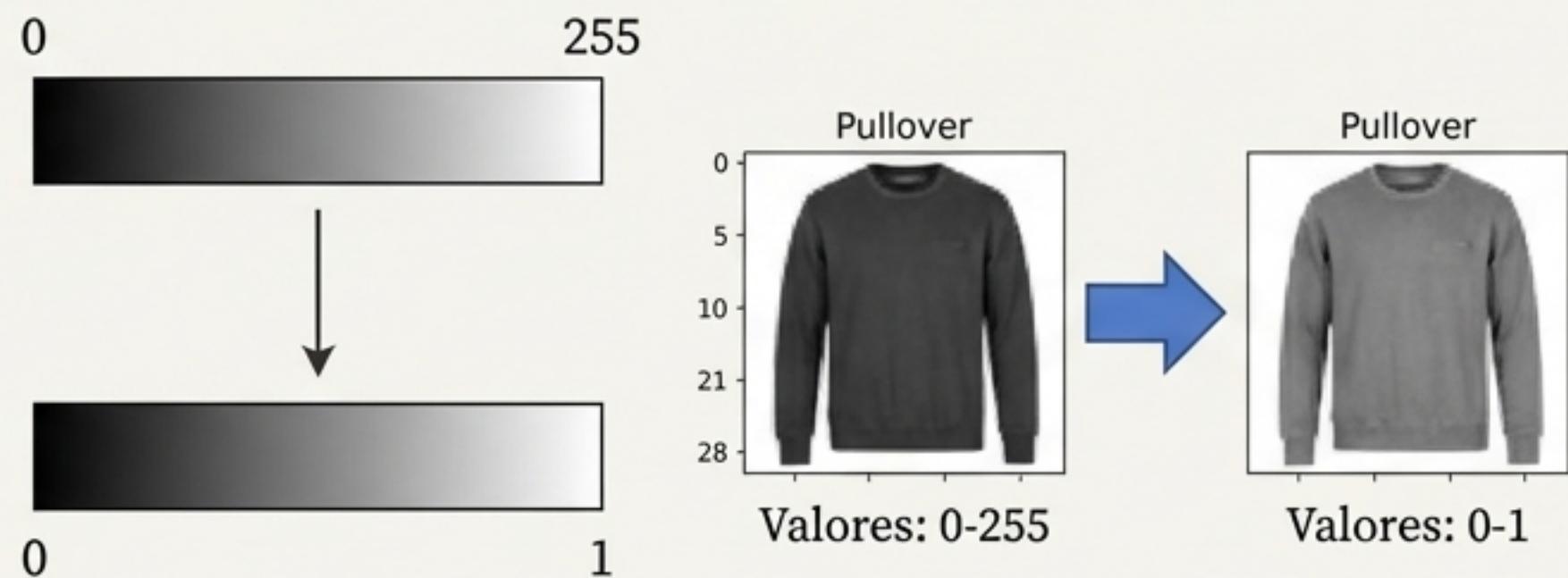
Paso 2: Preparando el Lienzo Digital

Normalización de Datos

Los valores de los píxeles de las imágenes van de 0 (negro) a 255 (blanco). Antes de alimentar la red neuronal, escalamos estos valores a un rango de 0 a 1.

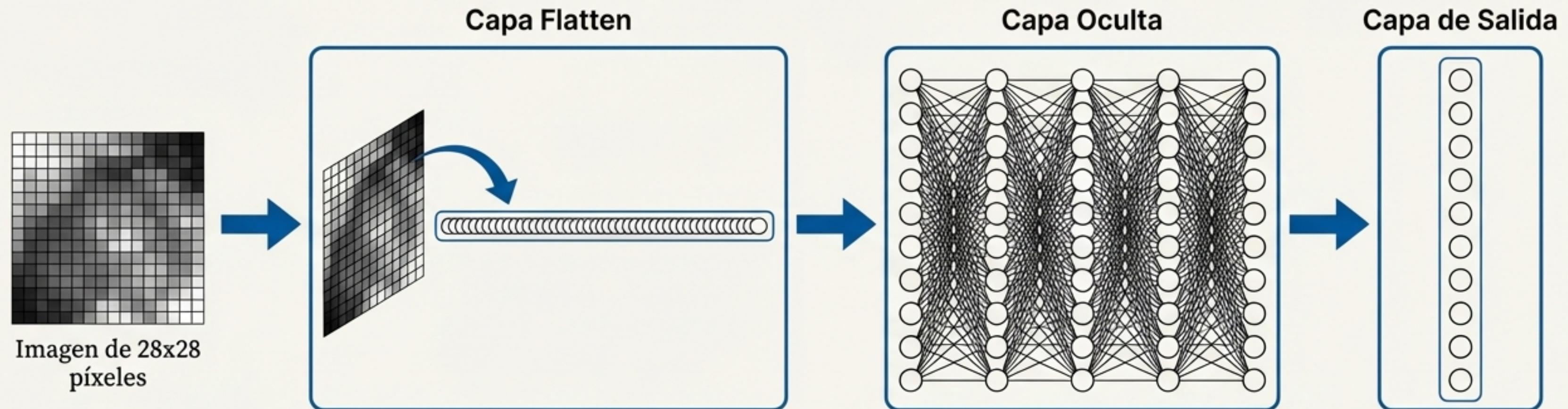
- **Operación**: `valor_pixel / 255.0`

La normalización asegura que todos los datos de entrada tengan una escala similar. Esto ayuda a que el modelo converja más rápido y de manera más estable durante el entrenamiento, ya que evita que algunos pesos de la red crezcan de forma desproporcionada. Es crucial aplicar la misma transformación tanto al conjunto de entrenamiento como al de prueba.



Paso 3: Diseñando la Arquitectura de la Red Neuronal

Construimos un modelo secuencial simple usando `tf.keras`, que consiste en una secuencia de capas encadenadas.



`tf.keras.layers.Flatten(input_shape=(28, 28))`

Función: Transforma la matriz 2D (28x28) en un vector 1D (784 píxeles). No aprende parámetros, solo reformatea los datos.

Salida: Vector de 784 neuronas

`tf.keras.layers.Dense(128, activation='relu')`

Función: Capa densamente conectada con 128 neuronas. La activación ReLU introduce no linealidad, permitiendo aprender relaciones complejas.

Salida: Vector de 128 neuronas

`tf.keras.layers.Dense(10)`

Función: Capa final con 10 neuronas, una por cada clase. Devuelve un array de 'logits' o puntuaciones brutas.

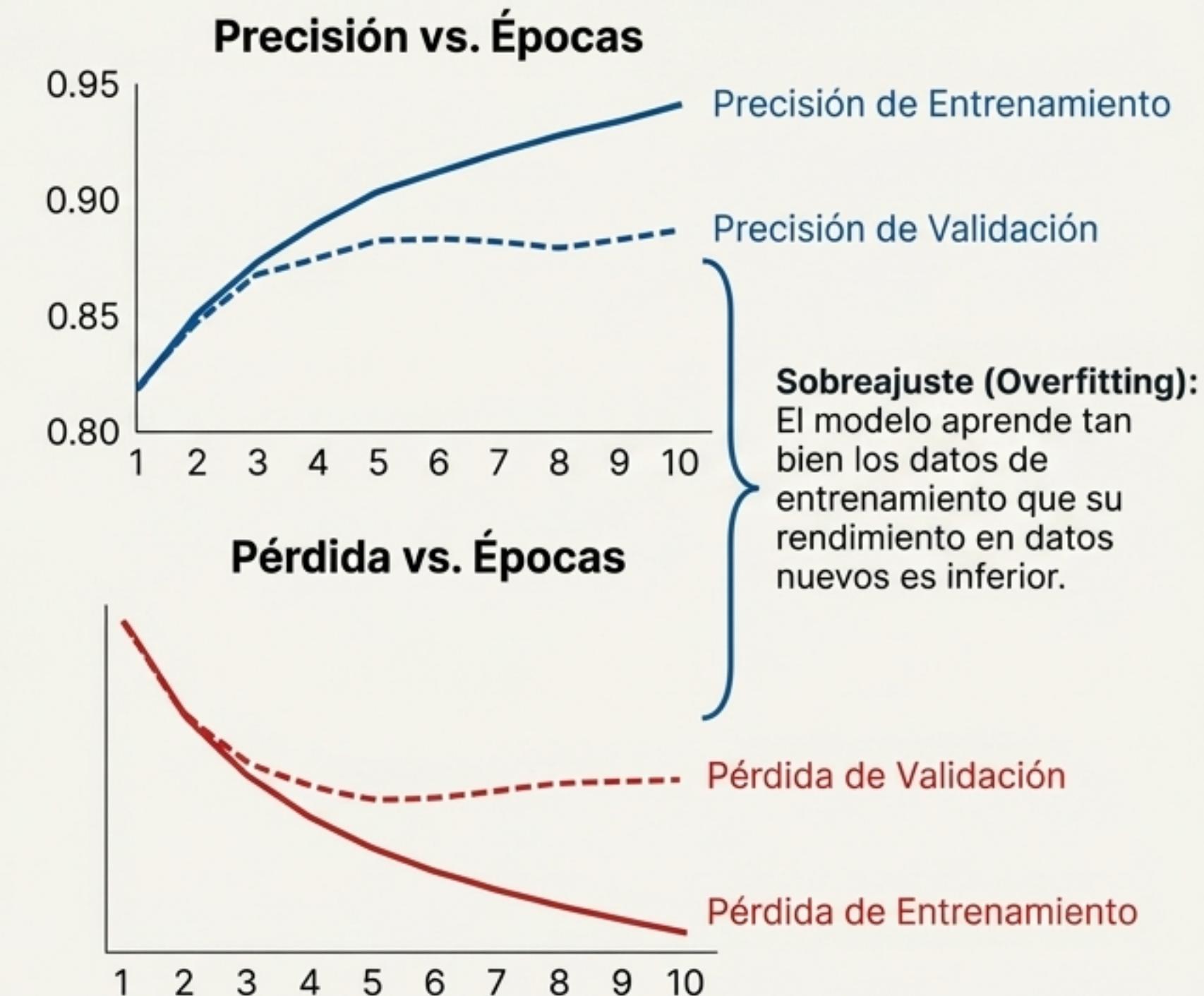
Salida: Vector de 10 logits

Paso 4: El Proceso de Aprendizaje y el Sobreajuste

Configuración del Entrenamiento

- **Optimizador:** 'adam', un algoritmo eficiente para ajustar los pesos del modelo.
- **Función de Pérdida:** **SparseCategoricalCrossentropy**, que mide la imprecisión del modelo.
- **Métricas:** ['accuracy'], para monitorear la fracción de imágenes clasificadas correctamente.
- **Épocas (Epochs):** 10, es decir, 10 pasadas completas sobre el conjunto de entrenamiento.

A medida que el modelo entrena, su precisión en los datos de entrenamiento aumenta y la pérdida disminuye. Sin embargo, la brecha entre el rendimiento en el entrenamiento y en la validación/prueba es clave.



Paso 5: La Hora de la Verdad, Evaluación del Modelo

Una vez entrenado, evaluamos el rendimiento final del modelo en el conjunto de prueba, que contiene 10,000 imágenes que nunca ha visto antes.

91.07%

Precisión en Entrenamiento



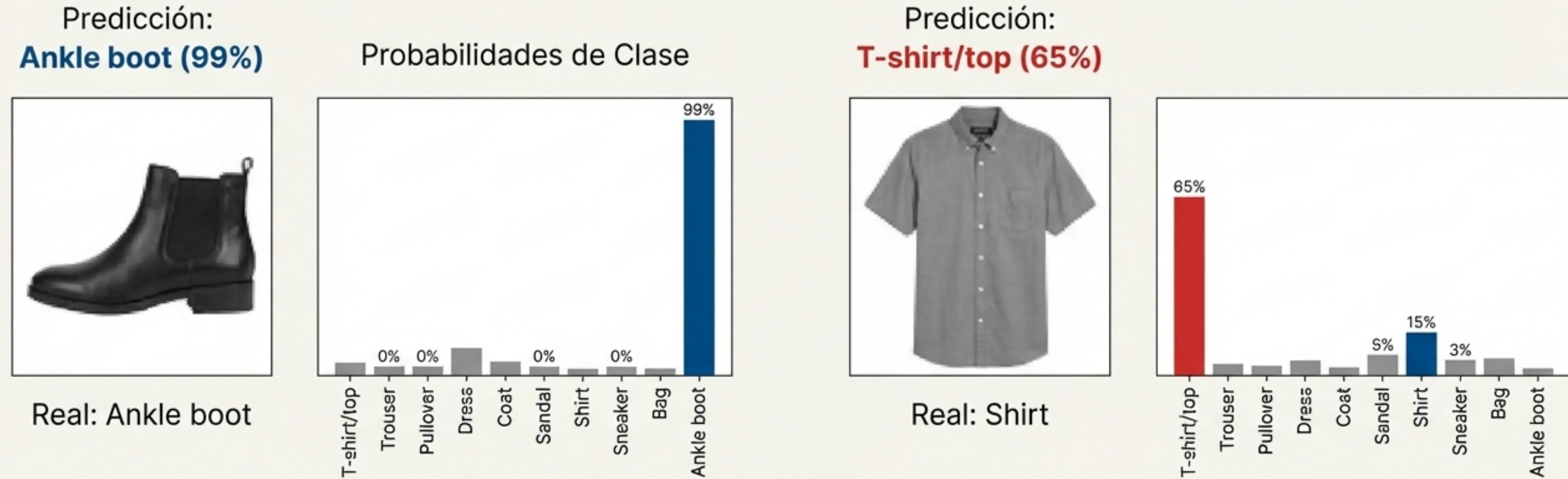
88.63%

Precisión en Prueba

- La precisión en el conjunto de prueba es ligeramente inferior a la del conjunto de entrenamiento.
- Esta diferencia confirma el **sobreajuste**: el modelo 'memorizó' detalles específicos de los datos de entrenamiento que no se generalizan perfectamente a datos nuevos. Esto es un comportamiento normal y un desafío central en machine learning.

Viendo al Modelo en Acción: Verificando Predicciones

Con el modelo entrenado, podemos usarlo para predecir la clase de cualquier imagen. Adjuntamos una capa Softmax para convertir los logits en probabilidades.



El modelo puede equivocarse, especialmente en clases visualmente similares como "Shirt" y "T-shirt/top".

Más Allá de la Pasarela: El Impacto de un Nuevo Estándar

Fashion-MNIST no es solo un dataset sobre ropa; es una herramienta fundamental para la comunidad de Machine Learning.

- **Benchmark Robusto:** Se ha utilizado en miles de artículos de investigación para validar nuevos algoritmos, desde Redes Neuronales Convolucionales (CNN) hasta modelos más avanzados.
- **Adopción Generalizada:** Utilizado por instituciones de investigación líderes como Google, la Universidad de Cambridge e IBM Research.
- **Un Desafío Persistente:** A diferencia de MNIST, sigue siendo un problema desafiante. El rendimiento de última generación (SOTA) alcanza el **96.91%**, demostrando que hay margen para la mejora y la innovación.



UNIVERSITY OF
CAMBRIDGE

IBM Research



Hoy, la moda. Mañana, el futuro.



Dominar un dataset como Fashion-MNIST es más que un ejercicio académico.

Es desarrollar una habilidad fundamental para resolver problemas de visión por computadora mucho más complejos y con un impacto real.

Las mismas técnicas y principios se aplican en campos que van desde el diagnóstico por imágenes médicas y la conducción autónoma hasta el análisis de datos satelitales.