

Motor de Recomendación con Deep Learning sobre Dataset Simulado

Arnau Sastre

[linkedin.com/in/arnausastre](https://www.linkedin.com/in/arnausastre)

August 9, 2025

Abstract

Este proyecto demuestra cómo construir un sistema de recomendación basado en técnicas de Deep Learning, utilizando un enfoque de **Neural Collaborative Filtering (NCF)**. El sistema ha sido desarrollado sobre un dataset simulado de usuarios y productos, simulando un entorno realista de ecommerce. Se implementa tanto el modelo como el análisis interpretativo y visual de las recomendaciones, embeddings y agrupaciones de productos.

1 Introducción

Los motores de recomendación son fundamentales en plataformas digitales como ecommerce, suscripciones o contenido bajo demanda. Aumentan la conversión, mejoran la retención y personalizan la experiencia de usuario. Este trabajo propone un modelo de recomendación entrenado sobre un dataset sintético que replica interacciones entre usuarios y productos.

2 Dataset simulado

Se generaron interacciones positivas y negativas entre usuarios y productos:

- **1,000 usuarios**
- **500 productos**
- **10,000 interacciones positivas** (compras)
- **10,000 interacciones negativas** (no compró)

El conjunto final se estructura como pares (**usuario**, **producto**) con etiqueta binaria 0 o 1.

3 Modelo de recomendación

Se utiliza un modelo de tipo **Neural Collaborative Filtering (NCF)** con las siguientes características:

- **Embeddings** para representar usuarios y productos en un espacio latente
- **Capas densas** no lineales para capturar relaciones complejas
- **Función de pérdida binaria** y métrica AUC para evaluación

4 Evaluación

El modelo se evaluó utilizando métricas propias de sistemas de recomendación:

- **Precision@10**: proporción de productos recomendados que realmente fueron comprados
- **Recall@10**: cobertura sobre los productos relevantes para cada usuario

La evaluación se realizó **usuario por usuario**, comparando las recomendaciones con su historial simulado.

5 Ejemplo de resultados

Usuario 42 – Top recomendaciones:

Product ID	Score
287	0.7950
355	0.7721
...	...

Estas recomendaciones reflejan los productos que el modelo considera más relevantes para el usuario.

6 Análisis avanzado

Además del modelo, se integran análisis adicionales sobre los embeddings del sistema:

Visualización de Embeddings

- Reducción de dimensionalidad con **PCA**
- Visualización del espacio latente de productos en 2D

Clustering de productos

- Algoritmo **K-Means** aplicado a los embeddings de producto
- Agrupación automática de productos similares según su comportamiento

Este análisis ayuda a entender cómo el modelo "agrupa" productos relacionados aunque no estén etiquetados explícitamente.

7 Tecnologías utilizadas

- Python 3
- TensorFlow / Keras
- NumPy, Pandas, Scikit-learn
- Matplotlib, PCA, KMeans

8 Aplicaciones reales

Este sistema es fácilmente adaptable a datos reales en contextos como:

- Ecommerce: recomendaciones de productos personalizadas
- Apps de fidelización: sugerencias de contenido o premios
- Retail digital: ventas cruzadas y upselling inteligente
- Plataformas de suscripción: mejora de engagement

9 Conclusiones

Se ha desarrollado un motor de recomendación robusto basado en Deep Learning y entrenado sobre un dataset simulado. El sistema incluye tanto el modelo como herramientas de análisis y visualización que pueden trasladarse fácilmente a un entorno productivo con datos reales.

Este enfoque permite validar capacidades técnicas sin depender de datasets privados, y demuestra cómo aplicar técnicas modernas de machine learning y análisis interpretativo en proyectos de negocio.

Contacto

Si te interesa implementar un motor de recomendación en tu empresa o necesitas ayuda con modelos de personalización, machine learning o sistemas inteligentes, puedes escribirme por **LinkedIn** o **Malt**. También puedes consultar otros proyectos técnicos en mi **GitHub**.