

Sistema de recomendación explicable y sostenible basado en texto para escenarios cold-start

Pablo Pérez-Núñez¹, Verónica Bolón-Canedo², Oscar Luaces¹ y Jorge Díez¹

¹ Universidad de Oviedo, Escuela Politécnica de Ingeniería, Gijón
{pabloperrez, oluaces, jdiez}@uniovi.es

² Universidade da Coruña, Campus de Elviña, A Coruña
veronica.bolon@udc.es

Resumen

Los sistemas de recomendación (SR) se han consolidado como una tecnología clave para personalizar la experiencia del usuario y optimizar la búsqueda de productos o servicios en plataformas digitales. No obstante, su desarrollo se enfrenta a múltiples retos. Uno de los más relevantes es la capacidad de ofrecer recomendaciones personalizadas en ausencia de información previa sobre el usuario, lo que se conoce como *cold-start*. A este desafío se suma la creciente necesidad de que los sistemas sean transparentes y ofrezcan explicaciones comprensibles sobre el origen de sus recomendaciones, en respuesta a la desconfianza generada por la inteligencia artificial en los últimos años. Por último, el impacto energético y ambiental de los modelos de inteligencia artificial está adquiriendo una importancia creciente, tanto en el ámbito científico como en el social. Con estos desafíos en mente presentamos ETeR-X, un sistema de recomendación explicable, multilingüe y más eficiente que otras alternativas de rendimiento comparable.

1 Introducción

Los SR son herramientas de gran valor tanto para las empresas, que ven aumentadas sus ventas, como para los clientes, que disfrutan de una experiencia de consumo más satisfactoria. Sin embargo, la mayoría de los sistemas modernos se basan en modelos de aprendizaje profundo que operan como “cajas negras”. Esta naturaleza opaca no solo los vuelve ineficientes debido a su gran número de parámetros, sino que también oculta el origen de sus decisiones, mermando la confianza del usuario. Esta problemática ha impulsado la necesidad de una IA confiable, que debe ser explicable y sostenible. En respuesta, han surgido la Inteligencia Artificial Explicable (XAI) y la IA Verde (Green AI). Nuestro trabajo, ETeR-X, se sitúa en la intersección de estos campos, abordando tres retos clave: (1) Ofrecer recomendaciones sin información previa del usuario (cold-start), (2) generar recomendaciones transparentes y justificables (explicabilidad) y (3) minimizar el impacto computacional y energético (eficiencia).

2 Método propuesto

ETeR-X es un sistema de recomendación basado en texto que aprende la relación entre los ítems y las palabras relevantes extraídas de las reseñas textuales de clientes anteriores. El sistema se entrena para predecir a qué ítem (p.ej., un restaurante) pertenece una reseña determinada. Para representar el texto, primero eliminamos todas aquellas palabras que no sean sustantivos o adjetivos, que consideramos los términos más descriptivos. A continuación, seleccionamos el 10% más frecuente de estos términos para construir el vocabulario. Cada reseña se codifica utilizando Bag-of-Words (BoW) con TF-IDF y se normaliza con L1. La arquitectura del modelo es una red neuronal simple y por tanto computacionalmente ligera.

El vector BoW normalizado alimenta una única capa (con matriz de pesos W) seguida de una activación Sigmoide para permitir que el modelo pueda recomendar más de un ítem. La simplicidad del modelo es su principal fortaleza para la explicabilidad. A diferencia de los métodos post-hoc que intentan explicar una caja negra, la explicabilidad de ETeR-X es inherente. La matriz de pesos W aprendida durante el entrenamiento contiene la información necesaria. Cada fila p de la matriz representa un ítem, y cada columna i representa un término del vocabulario (sustantivo o adjetivo). Así, el valor $W_{p,i}$ indica la relevancia del término i (p.ej., “pulpo”) para el producto p (p.ej., el restaurante “Pulpeira de Melide”). Cuando un usuario introduce una consulta (p.ej., “Busco una pulpería gallega típica”), el sistema recomienda el producto “Pulpeira de Melide” porque los términos “pulpo”, “gallego” y “típico” tienen pesos altos en la fila correspondiente a ese restaurante en la matriz W . La explicación consiste en mostrar al usuario cómo los términos relevantes de su propia consulta coinciden con los términos que el modelo ha aprendido que definen ese producto.

3 Experimentación

Se evaluó ETeR-X en doce conjuntos de datos de dominios distintos (productos, POIs y restaurantes) y en tres idiomas (inglés, español y francés). Comparamos nuestro modelo con varios baselines, incluyendo sistemas que no operan en cold-start y grandes modelos de caja negra capaces de operar con textos como USEM y BERT. En el escenario de cold-start (evaluando solo reseñas de usuarios no vistos durante el entrenamiento), ETeR-X demuestra un rendimiento (medido con NDCG@10 y Recall@10) altamente competitivo, igualando en la mayoría de los casos a los complejos modelos BERT y USEM. Los resultados confirman que la sencillez de ETeR-X no implica una penalización significativa en la precisión de la recomendación.

El verdadero potencial de ETeR-X está en la eficiencia. Nuestra propuesta tiene drásticamente menos parámetros que los modelos BERT y USEM. Su número de parámetros escala con el número de ítems del dataset, mientras que los LLM tienen un tamaño fijo y masivo ante cualquier número de ítems. Durante el entrenamiento, se han medido las emisiones de carbono. Si bien los grandes modelos de texto pre-entrenados convergen en menos épocas, sus costes energéticos ocultos son enormes. Se estima que solo el pre-entrenamiento del modelo BERT generó 652.265,8 gramos de CO_2 (≈ 652 kg). En cambio, ETeR-X emitió, en su peor escenario de entrenamiento, solo 9,8 gramos de CO_2 . A esto se suma el coste de entrenar desde cero el modelo de PoS que ETeR-X necesita para el español, que estimamos en 10,013 gramos de CO_2 . Incluso sumando los costes del pre-procesamiento y del entrenamiento de nuestro modelo, ETeR-X es órdenes de magnitud más eficiente y sostenible que las alternativas de propósito general basadas en grandes modelos.

4 Conclusiones

Este trabajo presenta ETeR-X, un sistema de recomendación que ofrece un equilibrio excepcional entre rendimiento, explicabilidad y eficiencia. Resuelve eficazmente el problema del cold-start utilizando únicamente una consulta textual. Demostramos que ETeR-X alcanza una precisión de recomendación comparable a la de modelos de caja negra mucho más complejos, pero con un coste computacional y una huella de carbono drásticamente inferiores. Su arquitectura simple proporciona explicaciones inherentes y transparentes, fomentando la confianza del usuario. Esto posiciona a ETeR-X como una alternativa viable y sostenible para sistemas de recomendación en el mundo real.