Propuesta de un marco metodológico para la extracción de patrones frecuentes y creación de un sistema de recomendación de productos para restauración

Erik Bola Sánchez[1,2], Sergi Ramírez[3], and Dante Conti[3]

[1] Universitat de Barcelona, Barcelona, España
ebolasan7@alumnes.ub.edu, erikbola04@gmail.com
[2] Universitat Politècnica de Catalunya, Barcelona, España
erik.bola@estudiantat.upc.edu
[3] Knowledge Engineering and Machine Learning group, Intelligent Data Science and Artificial Intelligence (IDEAI-UPC), Barcelona, España

Resumen Este trabajo propone un marco metodológico para la extracción de patrones frecuentes y episodios a partir de datos transaccionales de un punto de venta (TPV) de restauración, con el objetivo de dar soporte a la recomendación de productos en tiempo real durante el servicio y, con ello, a la optimización de la gestión de ventas. La propuesta cubre el ciclo completo: la definición precisa del problema en un escenario presencial *face-to-face*, la centralización de los datos dispersos en varias tablas del TPV, su transformación en una base de datos transaccional secuencial y la aplicación de algoritmos de minería de patrones secuenciales (GSP, PrefixSpan y SPADE) como primer modelo de referencia. Sobre datos reales de un restaurante (791.078 eventos transformados en 117.642 secuencias) se obtienen los primeros resultados y se diagnostican sus limitaciones: las reglas extraídas tienden a ser triviales, con antecedentes muy cortos y dominadas por los productos más frecuentes, con un *lift* medio bajo (1,68). Este diagnóstico, lejos de ser un resultado negativo, constituye la principal aportación del trabajo: caracteriza con evidencia por qué el modelado secuencial clásico es insuficiente en este dominio y justifica una hoja de ruta hacia representaciones densas de sesión (embeddings) con *reranking* orientado al margen de beneficio, que se plantea como línea futura.

Keywords: Gestión de ventas · Minería de patrones secuenciales · Reglas asociativas y patrones frecuentes · Restauración · Sistemas de recomendación.

1 Introducción

1.1 Contexto: la venta sugerida en restauración presencial

La gestión de ventas en restauración presencial descansa, en buena medida, sobre un proceso que rara vez se sistematiza: la toma de comanda. Durante ese proceso,

el personal de sala dispone de un margen de decisión considerable sobre qué productos sugerir a cada mesa —un entrante adicional, una bebida, un postre o un acompañamiento— y esas sugerencias tienen un impacto directo tanto en la experiencia del cliente como en la cuenta de resultados del establecimiento. Sin embargo, el personal suele contar con poca preparación específica para la venta sugerida, y la rotación característica del sector dificulta consolidar un conocimiento experto estable sobre qué recomendar en cada situación.

El resultado es que la venta sugerida queda, en la práctica, a criterio individual y poco reproducible de cada camarero. Dos mesas con el mismo consumo inicial pueden recibir recomendaciones muy distintas, y los productos de mayor valor o margen tienden a quedar infrarrepresentados frente a los más evidentes. En un sector con márgenes ajustados y fuerte estacionalidad, esta variabilidad representa una oportunidad de mejora cuantificable: una herramienta que proponga, en tiempo real y a partir de los productos ya pedidos en una mesa, un conjunto reducido de sugerencias relevantes constituye una palanca directa sobre la calidad de la comanda y sobre el ticket medio por comensal.

A diferencia de la recomendación *Business-to-Consumer* (B2C) *online* clásica —donde el usuario interactúa directamente con una aplicación o una web y deja un rastro persistente de navegación— en restauración presencial la interacción es *face-to-face* y está mediada por un profesional, un canal mucho menos explorado en la literatura que el de la venta digital [1]. La sugerencia no se muestra al consumidor en una pantalla, sino que se canaliza a través del camarero, que decide cómo y cuándo activarla durante la conversación con el cliente. Esta diferencia, aparentemente menor, condiciona todo el diseño del sistema: la unidad de análisis no es un usuario persistente con histórico, sino una *cuenta efímera* —una secuencia temporal corta de productos consumidos en una mesa durante un único servicio— con observabilidad limitada y una ventana de actuación muy breve.

1.2 El problema: recomendación intra-servicio

El problema que se aborda es la *recomendación intra-servicio*: sugerir productos mientras la mesa sigue abierta, tomando como entrada el prefijo parcial del pedido (los artículos ya servidos, en su orden cronológico) y produciendo como salida un conjunto reducido (Top-k) de candidatos a sugerir a continuación. El criterio de éxito es que esos candidatos aparezcan efectivamente en el consumo futuro de la mesa, lo que permite evaluar el sistema de forma *offline* sobre datos históricos antes de cualquier despliegue.

Formular el problema en estos términos plantea retos específicos. Las secuencias son cortas y muy variables, el catálogo de productos es amplio, y el comportamiento de consumo depende fuertemente del contexto (turno, día de la semana, festivos). Además, la recomendación debe ser interpretable y accionable por una persona: un sistema de caja negra que sugiera productos sin una lógica comprensible difícilmente será adoptado por el personal de sala. Estas restricciones hacen del dominio un banco de pruebas exigente para las técnicas clásicas

de descubrimiento de patrones, que históricamente se han aplicado en escenarios *online* con sesiones más largas y estructuradas [2,3].

1.3 Contribuciones y estructura del documento

Este artículo no propone aún el sistema de recomendación definitivo, sino que sienta sus bases metodológicas y establece el punto de partida del proyecto. Su contribución es doble. En primer lugar, formula un **marco metodológico reproducible** que va desde la centralización de datos heterogéneos de TPV hasta la obtención de una base transaccional secuencial lista para minería de patrones, con una clara separación entre etapas independientes de la herramienta concreta de TPV. En segundo lugar, aplica sobre esa base un **primer modelo de referencia** basado en minería de patrones frecuentes y secuenciales, analiza sus resultados y diagnostica de forma cuantitativa sus límites, estableciendo así la justificación empírica de las líneas de trabajo posteriores —en particular, el salto hacia representaciones densas de sesión.

El resto del documento se organiza como sigue. La Sección 2 revisa la lógica de las reglas de asociación y los patrones secuenciales, recoge casos de estudio de minería de patrones aplicada a la recomendación y sitúa el problema en el sector. La Sección 3 describe el marco metodológico propuesto. La Sección 4 presenta la experimentación: la aplicación concreta de la metodología y los resultados obtenidos. La Sección 5 discute sus limitaciones. Finalmente, la Sección 6 recoge las conclusiones y las líneas futuras.

2 Estado del arte

2.1 Reglas de asociación y patrones frecuentes

El análisis de la cesta de la compra (*Market Basket Analysis*) y la minería de reglas de asociación constituyen el punto de partida natural para descubrir relaciones de coocurrencia entre productos. Una regla $X \Rightarrow Y$ se evalúa habitualmente con tres medidas: el *soporte*, que mide la frecuencia conjunta de X e Y; la *confianza*, que estima la probabilidad condicionada $P(Y \mid X)$; y el *lift*, que cuantifica cuánto se aleja esa asociación de la independencia estadística (lift $= 1$ indica independencia, y valores claramente superiores a 1 indican asociación positiva). Algoritmos como Apriori [4] y FP-Growth [5] permiten extraer de forma eficiente los conjuntos de ítems frecuentes a partir de los cuales se derivan las reglas, controlando la explosión combinatoria mediante umbrales mínimos de soporte.

2.2 Minería de patrones secuenciales y episodios

Cuando el orden de consumo es relevante —como ocurre en una comanda, donde los productos se piden secuencialmente— el problema se enriquece con la minería de patrones secuenciales (*Sequential Pattern Mining*, SPM) y de episodios

frecuentes. Algoritmos como GSP [6], PrefixSpan [7] y SPADE [8] buscan subsecuencias frecuentes respetando el orden temporal, lo que en teoría permite capturar reglas del tipo "si una mesa ha pedido A y luego B, tenderá a pedir C". La lógica subyacente es la misma que en las reglas de asociación, pero condicionada a la posición: el interés no es solo *qué* aparece junto, sino *en qué orden*. GSP aborda el problema mediante generación y prueba de candidatos por niveles; PrefixSpan evita la generación explícita de candidatos proyectando bases de datos por prefijos; y SPADE emplea una representación vertical (listas de identificadores) que reduce el número de pasadas sobre los datos. Estos métodos son atractivos por su interpretabilidad directa, una propiedad especialmente valiosa en un entorno donde la recomendación debe ser comprensible y accionable por una persona.

2.3 Casos de estudio: SPM aplicada a la recomendación

La minería de patrones secuenciales se ha aplicado con éxito a la recomendación en diversos dominios, lo que permite contrastar su comportamiento con el que cabe esperar en restauración. En *e-commerce*, se ha empleado PrefixSpan para extraer patrones secuenciales de compra y construir un recomendador a partir de las reglas temporales descubiertas, aprovechando que las secuencias de navegación y compra presentan regularidades repetitivas [2]. En *online retail*, varios trabajos proponen sistemas híbridos que combinan el algoritmo GSP con datos de comportamiento y *clickstream*, integrando la señal secuencial con información transaccional para mejorar la cobertura de las recomendaciones [3]. En esa misma línea, otros autores incorporan esquemas de ponderación multicriterio —por ejemplo, *Simple Additive Weighting* (SAW)— sobre los patrones extraídos con GSP a partir de transacciones y *clickstream*, de modo que las reglas no solo se ordenan por su frecuencia sino también por criterios adicionales de relevancia [9]. Por último, en el dominio del turismo y la recomendación de puntos de interés (POI), se han combinado el modelado de tópicos y PrefixSpan para la recomendación secuencial de itinerarios, añadiendo contexto demográfico a la minería de patrones de viaje [10].

Estos casos comparten un patrón común: la SPM rinde bien cuando las secuencias son suficientemente largas y repetitivas y cuando el vocabulario de ítems es moderado, condiciones que favorecen la aparición de patrones de orden superior estadísticamente robustos. El dominio de la restauración presencial se aleja precisamente de ese escenario —secuencias cortas (del orden de cinco productos), vocabulario amplio (varios miles de artículos) y fuerte concentración del consumo en unos pocos productos—, lo que anticipa las limitaciones que se documentan empíricamente en la Sección 4 y motiva la comparación con técnicas más expresivas.

Es importante subrayar que, en la mayoría de estos trabajos, la señal secuencial se complementa con fuentes adicionales: *clickstream* de navegación, datos demográficos, ponderaciones multicriterio o información de catálogo. Es decir, la SPM rara vez sostiene en solitario la recomendación, sino que actúa como un componente dentro de un sistema híbrido. En nuestro caso de uso, en cambio, la

única evidencia disponible en el instante de la recomendación es el prefijo de la cuenta y un puñado de variables de contexto, sin histórico persistente del comensal ni rastro de navegación. Esta pobreza relativa de señal explícita refuerza la necesidad de un modelo capaz de generalizar a partir de la propia secuencia, y no solo de reproducir coocurrencias frecuentes. En paralelo, el estado del arte reciente converge hacia representaciones densas de la sesión —*embeddings*— donde el problema se formula como recuperación de vecinos en un espacio latente [11], e incluso se exploran modelos de lenguaje de gran tamaño (LLM) para enriquecer dichas representaciones secuenciales [12]; esta dirección se retoma como línea futura en la Sección 6.

3 Marco metodológico propuesto

El marco propuesto (Fig. 1) articula cinco etapas: definición del problema, centralización de datos, transformación a base transaccional secuencial, derivación de variables de contexto y aplicación de los modelos de patrones. Cada etapa es independiente de la herramienta concreta de TPV y reproducible sobre otros establecimientos.

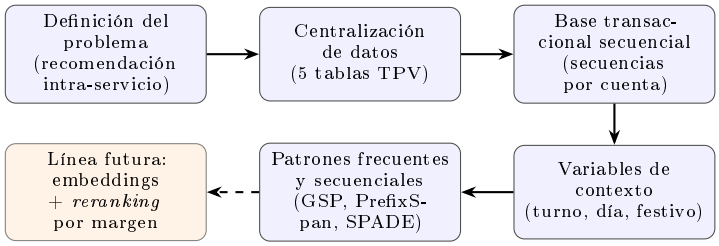


Fig. 1: Marco metodológico propuesto. En azul, las etapas desarrolladas en este trabajo; en naranja, la extensión planteada como línea futura.

3.1 Definición del problema

Se formaliza la entrada como el prefijo parcial de una cuenta —la lista ordenada de productos ya servidos en una mesa— y la salida como un conjunto reducido (Top-k) de productos candidatos a ser sugeridos a continuación. El criterio de éxito en esta fase es que los candidatos aparezcan efectivamente en el consumo futuro de la mesa; la incorporación explícita de criterios económicos se reserva para fases posteriores.

3.2 Centralización de los datos

La base analítica se construye a partir de cinco tablas del TPV: `Documentos_Cabecera`, `Documentos_Lineas`, `Articulos`, `Caja_Lineas` y `Clientes`. Cada ticket (`Documentos_Cabecera`) representa la comanda completa de una mesa en un servicio, desglosada por líneas en el orden en que se pidieron los productos (`Documentos_Lineas`). La centralización exige resolver las claves que relacionan cabecera y líneas y unificar la codificación de artículos.

3.3 Transformación a base transaccional secuencial

El proceso ETL extrae las tablas del sistema *tpv-net* a un formato manejable, empareja temporalmente cabecera y líneas por el identificador del ticket y representa cada cuenta como una secuencia cronológica de productos. Se filtran las secuencias con menos de dos productos y los servicios con duración anómala. El resultado es una base transaccional secuencial con **791.078 eventos** transformados en **117.642 secuencias**, con una longitud media de **5,3 productos** y un vocabulario de **2.346 artículos**. Esta base es la unidad de entrada común a todos los modelos posteriores.

3.4 Derivación de variables de contexto

A partir de la cabecera se derivan atributos que condicionan el comportamiento de consumo: turno del servicio, día de la semana, indicador de festivo y duración del ticket. Estas variables permiten distinguir patrones propios de comida, cena o días especiales y enriquecen tanto la minería de patrones como las extensiones futuras.

3.5 Modelos de referencia: minería de patrones

Sobre la base transaccional se aplican tres algoritmos de minería de patrones secuenciales ampliamente contrastados —GSP, PrefixSpan y SPADE— configurados con umbrales de soporte y confianza, y filtrando las reglas por *lift* para descartar asociaciones espurias. El objetivo de esta fase no es maximizar una métrica de recomendación, sino obtener un conjunto interpretable de reglas que sirva como referencia y como diagnóstico del dominio.

4 Experimentación

4.1 Aplicación de la metodología

La experimentación se realiza sobre la base transaccional secuencial descrita en la Sección 3, es decir, sobre las 117.642 secuencias resultantes del proceso ETL. Cada secuencia se trata como una transacción ordenada de productos, y los *tokens* de contexto (turno, día, festivo) se incorporan como atributos asociados a la cuenta. Sobre este conjunto se ejecutan los tres algoritmos de referencia:

- **GSP**, mediante generación y verificación de candidatos por niveles, con poda basada en el soporte mínimo.
- **PrefixSpan**, mediante proyección sucesiva de la base de datos por prefijos, evitando la generación explícita de candidatos.
- **SPADE**, mediante una representación vertical de las secuencias que reduce el número de pasadas sobre los datos.

En todos los casos se fijan umbrales mínimos de soporte y confianza para acotar el número de patrones, y las reglas resultantes se filtran por *lift* con el fin de descartar asociaciones que no superan de forma apreciable la independencia estadística. El producto de esta fase es, para cada algoritmo, un conjunto de reglas secuenciales caracterizadas por su antecedente, su consecuente y sus medidas de soporte, confianza y *lift*, que constituye la base del análisis de la siguiente sección.

4.2 Resultados

Los tres algoritmos extraen reglas, pero la inspección de los resultados revela un patrón consistente y poco favorable para el modelado del orden de pedido. En términos agregados:

- SPADE extrajo 146 reglas, de las cuales el 94,5 % tenían un antecedente de longitud 1.
- PrefixSpan produjo 1.416 reglas filtradas, pero solo 9 presentaban un antecedente de longitud mayor o igual que 3.
- El *lift* medio de las reglas, 1,68, refleja asociaciones débiles y poco discriminativas.

La predominancia de antecedentes de longitud 1 es especialmente reveladora: indica que, en la práctica, los algoritmos apenas capturan dependencias de orden superior y se limitan en su mayoría a relacionar pares de productos. La Tabla 1 recoge las reglas más representativas de cada método y la Fig. 2 las sitúa en el plano confianza–*lift*.

Tabla 1: Reglas representativas extraídas por cada método. Conf. = confianza.

Método	**Antecedente → Consecuente**	**Conf.**	**Lift**
GSP	BRUNCH → CROISSANT	92,8 %	49,67
GSP	ZAMBURIÑAS → TEJA DE ALMENDRA	11,3 %	6,29
GSP	ALMEJAS MARINERA → PAN	27,6 %	5,32
PrefixSpan	MEJILLONES MARINERA → PAN	28,4 %	4,64
PrefixSpan	COPA CERVEZA → COPA CERVEZA	28,2 %	17,32
PrefixSpan	CAÑA | AGUA 1 L → CAÑA	35,9 %	10,22
SPADE	COPA V. BLANCO → COPA V. BLANCO	22,5 %	3,79
SPADE	ARROZ NEGRO → CAFÉ HIELO	24,8 %	2,85
SPADE	AGUA 1 L | PATATAS BRAVAS → PAN TOMATE	34,1 %	2,30

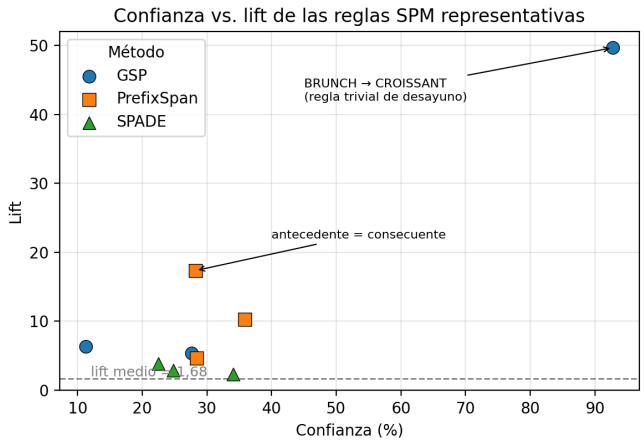


Fig. 2: Confianza frente a *lift* de las reglas representativas. Solo las reglas triviales de desayuno alcanzan *lift* alto; el resto se agrupa en valores bajos, próximos al *lift* medio.

El análisis cualitativo de las reglas confirma el diagnóstico cuantitativo. Por un lado, las reglas con *lift* elevado corresponden casi siempre a productos de desayuno (p. ej. BRUNCH → CROISSANT, con un *lift* de 49,67), un nicho de consumo muy estereotipado pero de escaso interés para el objetivo de negocio, que se concentra en comidas y cenas. Por otro lado, varias reglas repiten el mismo producto como antecedente y consecuente (COPA CERVEZA → COPA CERVEZA, COPA V. BLANCO → COPA V. BLANCO), lo que indica que se está capturando la mera repetición de artículos muy vendidos —el cliente repite bebida— más que un patrón de consumo informativo que anticipe un producto distinto. El resto de reglas, con *lift* próximo al valor medio, sugieren mayoritariamente productos de alta frecuencia general (pan, pan con tomate), que se pedirían con o sin recomendación. En conjunto, ninguna de las reglas extraídas permite inferir de forma fiable qué artículo diferencial pedirá una mesa en función de su sesión completa, que es precisamente la pregunta que el sistema de recomendación pretende responder.

Conviene traducir estos números a comportamiento operativo. Si el 94,5 % de las reglas de SPADE tienen un antecedente de un solo producto y apenas 9 de las 1.416 reglas de PrefixSpan alcanzan un antecedente de longitud 3, en la práctica el sistema condicionaría su recomendación, como mucho, en uno o dos productos previos, ignorando el resto de la sesión. Esto tiene dos consecuencias inmediatas. La primera es una *cobertura* reducida: muchas secuencias activas no coinciden con el antecedente exacto de ninguna regla y, por tanto, no reciben recomendación alguna, justo cuando el camarero más la necesitaría. La segunda es que, dado el bajo *lift* medio, el conjunto de reglas que sí se disparan se comporta de forma muy próxima a un recomendador de popularidad global restringido a los pocos productos que coocurren con frecuencia. El modelo de referencia,

en suma, no aprovecha la estructura secuencial que en principio justificaba su elección.

5 Discusión: limitaciones del modelo de referencia

Los resultados anteriores permiten un diagnóstico preciso. En este dominio, la minería de patrones secuenciales se comporta más como una *detección de coocurrencias frecuentes* que como un modelado útil del orden de pedido. Cuatro factores lo explican:

- **Rigidez de las reglas.** Solo se dispara una recomendación cuando la secuencia activa coincide con el antecedente exacto de una regla, lo que limita drásticamente su cobertura ante combinaciones no vistas.
- **Dominancia de productos frecuentes.** Las reglas con soporte suficiente tienden a recomendar los artículos más vendidos, sin aportar señal diferencial sobre la cuenta concreta.
- **Dispersión (*sparsity*).** El amplio vocabulario (2.346 artículos) frente a secuencias cortas (5,3 productos de media) produce una matriz muy dispersa, donde las combinaciones nuevas no encuentran reglas aplicables.
- **Baja capacidad discriminativa.** Un *lift* medio de 1,68 indica que la mayoría de asociaciones apenas superan la independencia estadística.

Estas limitaciones son coherentes con el comportamiento documentado en los casos de estudio de la Sección 2.3: la SPM rinde cuando las secuencias son largas y repetitivas, condición que el dominio de la restauración presencial no satisface.

Desde la óptica de la gestión de ventas —objetivo último de este trabajo— el diagnóstico tiene una lectura directa. Un sistema de reglas con baja cobertura y dominado por productos frecuentes tendería a recomendar lo que la mesa pediría de todos modos, desaprovechando el verdadero potencial de la venta sugerida: orientar al comensal hacia productos diferenciales o de mayor valor que, sin la sugerencia, no habría considerado. En otras palabras, un *baseline* de este tipo es poco útil para mover el ticket medio, que es la métrica que el negocio quiere optimizar. Para que la herramienta aporte valor real al camarero, debe (i) ofrecer recomendaciones incluso ante combinaciones de productos no vistas previamente, (ii) discriminar entre cuentas con un mismo producto inicial pero contextos distintos, y (iii) admitir la incorporación de criterios económicos en la decisión final. Ninguna de estas tres propiedades es alcanzable con el modelo de reglas, lo que delimita con precisión los requisitos del enfoque posterior.

Conviene subrayar que este hallazgo no invalida el marco metodológico propuesto: las etapas de centralización, transformación y enriquecimiento contextual son plenamente reutilizables, y la propia base transaccional secuencial es el insumo que habilita las técnicas más expresivas que se plantean a continuación. El valor de esta fase es, precisamente, aportar la evidencia que justifica el cambio de paradigma con datos del propio establecimiento, en lugar de asumirlo a priori.

6 Conclusiones y líneas futuras

Se ha presentado un marco metodológico reproducible para transformar los datos transaccionales de un TPV de restauración en una base secuencial apta para la recomendación intra-servicio, y se ha aplicado sobre ella un primer modelo de referencia basado en minería de patrones frecuentes y secuenciales. Los primeros resultados, con reglas mayoritariamente triviales (un 94,5 % de antecedentes de longitud 1 en SPADE) y un *lift* medio bajo (1,68), demuestran de forma cuantitativa que el modelado secuencial clásico es insuficiente en este dominio. Lejos de ser un resultado negativo, este diagnóstico es la principal aportación del trabajo: caracteriza con precisión las limitaciones del *baseline* y fundamenta empíricamente, con datos del propio restaurante, la necesidad de un cambio de enfoque.

A la vista del diagnóstico, la continuación natural del trabajo consiste en sustituir las reglas rígidas por una representación densa de la sesión. La dirección prevista es modelar cada cuenta parcial como un vector en un espacio latente (*embeddings* de sesión), recuperar las cuentas históricas más similares mediante similitud coseno y construir los candidatos a partir de ese vecindario, superando así los problemas de rigidez y dispersión que se han documentado. Sobre esos candidatos se plantea una fase de *reranking* que incorpore una señal explícita de negocio —margen de beneficio y precio unitario— para alinear la recomendación con la rentabilidad del establecimiento, y no solo con la probabilidad de acierto. Otras extensiones incluyen la validación *online* con actualización de la cuenta en tiempo real, la incorporación de nuevas variables de contexto (número de comensales, meteorología) y un objetivo complementario de regresión para predecir el volumen de trabajo de cada servicio. Estas líneas, centradas en los *embeddings* y en el filtrado por margen, constituyen el núcleo del desarrollo posterior del proyecto.

Agradecimientos

Esta investigación ha sido posible gracias al apoyo y soporte no financiero por parte de la *Grupo Capi Sericios de Hosteleria S.L.*, que gentilmente ha apoyado la idea del trabajo y ha facilitado los datos para la fase de experimentación, así como ideas de negocio y entendimiento de los resultados (impacto) en un entorno más orientado al negocio que a los aspectos académicos o técnicos.

Referencias

1. Capturing Latent Situational Influence in Offline Retail (2026). Embeddings + ItemCF, *offline retail*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6533792
2. Sequential Pattern Mining for E-commerce Recommender System (2020). PrefixSpan, *e-commerce*. https://scholar.ui.ac.id/en/publications/sequential-pattern-mining-for-e-commerce-recommender-system/
3. A Hybrid Recommender System Based on Customer Behavior and Transaction Data Using Generalized Sequential Pattern Algorithm (2024). GSP + *clickstream*, *online retail*. https://zenodo.org/records/7306895
4. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules. In: Proc. 20th International Conference on Very Large Data Bases (VLDB), pp. 487–499 (1994)
5. Han, J., Pei, J., Yin, Y.: Mining frequent patterns without candidate generation. In: Proc. ACM SIGMOD International Conference on Management of Data, pp. 1–12 (2000)
6. Srikant, R., Agrawal, R.: Mining sequential patterns: Generalizations and performance improvements. In: Proc. 5th International Conference on Extending Database Technology (EDBT), pp. 1–17 (1996)
7. Pei, J., Han, J., Mortazavi-Asl, B., Pinto, H., Chen, Q., Dayal, U., Hsu, M.-C.: PrefixSpan: Mining sequential patterns efficiently by prefix-projected pattern growth. In: Proc. 17th International Conference on Data Engineering (ICDE), pp. 215–224 (2001)
8. Zaki, M.J.: SPADE: An efficient algorithm for mining frequent sequences. Machine Learning **42**(1/2), 31–60 (2001)
9. A Hybrid Recommender System Based on Customer Behavior and Transaction Data Using Generalized Sequential Pattern Algorithm (2022). SAW + GSP sobre transacciones y *clickstream*, *e-commerce*/*online retail*. https://beei.org/index.php/EEI/article/download/4021/3038
10. A Hybrid Recommender System Using Topic Modeling and PrefixSpan in Social Media (2023). *Topic modeling* + PrefixSpan + contexto demográfico, turismo/POI. https://link.springer.com/article/10.1007/s40747-022-00958-5
11. Embeddings Methods for Next-Item and Last-Basket Session-Based Recommenders (2023). Item2vec + Doc2Vec, *e-commerce*/*food*. https://www.ijml.org/vol12/1092-TC4003.pdf
12. Improving Sequential Recommendations with LLMs (2024). *LLM embeddings* + *retrieval* + modelos secuenciales. https://arxiv.org/html/2402.01339v2