Facultad de Ciencias Económicas y Empresariales

ICADE

MÁS ALLÁ DEL RATING: EL TIPO DE EXPERIENCIA COMO CLAVE DEL APEGO EMOCIONAL EN EL SECTOR HOTELERO

TRABAJO FIN DE GRADO

Autor: Cayetana Fernández de Valderrama Blanco

Director: Ana Isabel Jiménez-Zarco

MADRID | abril 2026

RESUMEN: Este trabajo analiza qué factores explican el apego emocional de los huéspedes con los hoteles a partir de 416.032 reseñas de TripAdvisor en Londres, Nueva York y Shanghái. Para ello, se construye el *Love Score*, un indicador que combina la valoración numérica del cliente, el tono emocional del texto y la presencia de señales de lealtad. Mediante un modelo de tópicos (LDA) se identifican ocho tipos de experiencia, agrupados en cuatro categorías: emocional, funcional, mixta y friccional. Los resultados muestran que el tipo de experiencia tiene mayor impacto en el apego emocional que la categoría del hotel. En particular, las experiencias relacionadas con el trato humano y el entorno generan más vínculo que los atributos puramente funcionales, mientras que la fricción administrativa lo reduce de forma significativa. Además, el análisis confirma que la relación entre satisfacción y apego no es lineal, y que el tipo de experiencia aporta capacidad explicativa adicional incluso cuando se controla por la valoración numérica. En conjunto, los resultados sugieren que la clave para generar apego emocional no está solo en ofrecer un buen servicio, sino en diseñar experiencias que conecten con el huésped.

ABSTRACT: This study examines the factors that drive emotional attachment between guests and hotels using a dataset of 416,032 TripAdvisor reviews from London, New York, and Shanghai. To operationalize this concept, a composite indicator called the *Love Score* is developed, combining customer ratings, textual sentiment, and explicit signals of loyalty. Using Latent Dirichlet Allocation (LDA), eight dominant experience types are identified and grouped into four categories: emotional, functional, mixed, and frictional. The results show that the type of experience has a stronger impact on emotional attachment than the hotel category itself. Human interactions and the surrounding environment generate higher levels of attachment, while administrative friction significantly reduces it. The findings also reveal that the relationship between satisfaction and emotional attachment is not linear, and that the type of experience provides additional explanatory power beyond numerical ratings. Overall, the study suggests that building emotional attachment in the hotel industry depends not only on delivering quality service, but on creating experiences that meaningfully connect with guests.

PALABRAS CLAVE: apego emocional, experiencia del cliente, sector hotelero, análisis de sentimiento, modelado de tópicos, reseñas online, Love Score

KEY WORDS: emotional attachment, customer experience, hospitality industry, sentiment analysis, topic modeling, online reviews, Love Score

1. Introducción 1
2. Marco Teórico 3
2.1 El concepto de Love Brand 3
2.2 Factores emocionales en la fidelización hotelera 4
2.3 Marketing experiencial y tipología de experiencias 6
2.4 Engagement digital y señales de lealtad en reseñas online 7
2.5 Síntesis y justificación del modelo empírico 8
3. Metodología 10
3.1 Descripción del dataset 10
3.2 Procesamiento y limpieza del dataset 11
3.3 Modelado de experiencias mediante análisis de tópicos 12
3.3.1 Procesamiento textual para el modelado de tópicos 13
3.3.2 Construcción del diccionario y del corpus 14
3.3.3 Muestra de entrenamiento 15
3.3.4 Selección del número de tópicos 15
3.3.5 Asignación de tópicos y control de ambigüedad temática 16
3.3.6 Clasificación conceptual de los tópicos 17
3.4 Construcción de las variables analíticas 17
3.4.1 Satisfacción funcional: rating_norm y rating_100 18
3.4.2 Tono emocional: sent_norm 18
3.4.3 Señales de lealtad: kw_loyal_strict 19
3.4.4 El Love Score: indicador compuesto de apego emocional 20
3.4.5 Variables complementarias 21
3.4.6 Modelo predictivo del apego emocional 22
4. Resultados 24
4.1 Análisis descriptivo del dataset 24
4.2 Identificación de patrones experienciales mediante LDA 27
4.3 Love Score y satisfacción funcional por tipo de experiencia 29
4.4 Divergencia entre satisfacción funcional y apego emocional 32
4.5 Asociación del tipo de experiencia con el apego emocional (regresión OLS) 35
4.6 Modelo predictivo del apego emocional 37
5. Discusión 40
5.1 El tipo de experiencia como determinante del apego emocional 40
5.2 La paradoja del gap: satisfacción sin apego y fricción con emoción 42
5.3 El Love Score como métrica de engagement: aportaciones y limitaciones ... 44
6. Conclusiones 47
6.1 Síntesis de hallazgos principales 47
6.2 Implicaciones estratégicas para la gestión hotelera 48

6.3 Limitaciones del estudio 50
6.4 Líneas de investigación futura 51
7. Bibliografía 53
8. Anexos 55

1. Introducción

Hay hoteles que se olvidan al poco tiempo. Otros, en cambio, logran que el cliente quiera volver. No necesariamente porque sean los más lujosos o los más baratos, sino porque algo de la experiencia dejó una huella en el cliente. Esa diferencia, entre un hotel que simplemente satisface necesidades y uno que genera vínculo, es el objeto de estudio de este trabajo.

En el sector turístico actual, los viajeros tienen más opciones que nunca y, en muchos destinos, la oferta hotelera es difícilmente diferenciable en términos de precio y características. En ese contexto, las marcas que consiguen destacar no lo hacen únicamente por atributos funcionales, sino porque la experiencia que ofrecen termina teniendo un significado personal para el huésped. Son las llamadas *Love Brands*.

Una *Love Brand* no es solo una marca preferida; además, es recomendada y perdonada por el cliente cuando esta comete errores. En la industria hotelera este concepto es especialmente importante porque la experiencia es el producto en sí mismo. Los huéspedes interactúan con el personal, habitan las estancias y viven momentos que pueden ser memorables, decepcionantes o simplemente correctos. Las propias vivencias pesan más en la relación con la marca que cualquier campaña de comunicación o marketing que pueda llevar a cabo el hotel.

Al mismo tiempo, las plataformas de reseñas online han transformado la manera en que los viajeros comunican y evalúan sus experiencias. Una reseña en TripAdvisor no es una encuesta de satisfacción, es un relato que cuenta el huésped de manera espontánea, en el que declara gratitud, frustración, lealtad o afecto. Señales que ningún cuestionario estructurado podría capturar. "El personal se acordaba de mi nombre", "Este hotel ha arruinado mi viaje", "volvería sin dudarlo". Estas expresiones revelan no solo lo que ocurrió, sino cómo le hizo sentir al huésped. Por esta razón, las reseñas son una fuente muy valiosa para analizar el apego emocional.

Este trabajo analiza, a partir de un corpus de 416.032 reseñas de TripAdvisor procedentes de hoteles en Londres, Nueva York y Shanghái, cómo se relacionan la satisfacción funcional, es decir, la valoración de aspectos tangibles como la habitación, la ubicación o

el servicio, y el apego emocional de los huéspedes en función del tipo de experiencia que describen. Para evaluar el apego emocional, se desarrolla un indicador compuesto, el *Love Score*, que integra tres dimensiones: satisfacción declarada (*rating*), tono emocional del texto (análisis de sentimiento) e intención explícita de retorno o recomendación (marcadores lingüísticos). Para identificar los tipos de experiencia dominantes en las reseñas, se aplica un modelo de análisis textual LDA que agrupa las reseñas según los patrones de contenido que comparten. El posicionamiento del hotel (de lujo, intermedio o más económico) se utiliza solo como referencia, ya que los datos revelan algo llamativo: la categoría del establecimiento apenas explica las diferencias en el apego emocional entre los huéspedes. Lo que sí las explica, y de forma mucho más clara, es el tipo de experiencia vivida durante la estancia.

Este estudio busca dar respuesta a tres preguntas. Primera: ¿qué tipos de experiencias hoteleras generan mayor vínculo emocional con el huésped? Segunda: ¿hasta qué punto un huésped puede valorar positivamente un hotel (habitación correcta, servicio eficiente) sin llegar a desarrollar un apego real hacia él? Tercera: ¿qué hay en ciertas experiencias que convierte una estancia satisfactoria en algo que se recuerda y se recomienda?

La estructura del trabajo sigue la lógica de las preguntas de investigación planteadas. En primer lugar, el marco teórico define el concepto de *Love Brand* y justifica por qué la experiencia hotelera y las reseñas online son un contexto adecuado para estudiar el apego emocional. La metodología describe el dataset, el proceso de limpieza, la construcción del *Love Score* y la estrategia analítica empleada. El núcleo de esa estrategia es identificar patrones de experiencia recurrentes en las reseñas, a los que en adelante llamaremos tópicos experienciales, es decir, tipos de vivencia que aparecen de forma sistemática en el corpus. Los resultados analizan cómo varía el apego emocional entre esos tipos de experiencia, qué distancia existe entre valorar bien un hotel y desarrollar un vínculo real con él, y qué capacidad tienen esos patrones para predecir la conexión emocional del huésped. La discusión interpreta esos hallazgos de forma integrada y los traduce en implicaciones teóricas y de gestión. Por último, el trabajo recoge sus limitaciones y plantea futuras líneas de investigación.

2. Marco Teórico

2.1 El concepto de *Love Brand*

Algunas marcas se eligen por lo que ofrecen, mientras que otras generan una relación más profunda con el consumidor. La diferencia no está solo en la calidad del producto o en el precio, sino en si la relación con la marca va más allá de la transacción y se ha convertido en algo con carga afectiva. Este fenómeno, conocido como *brand love* en la literatura académica, describe un vínculo duradero en el que la marca adquiere un significado personal para quien la consume.

Carroll y Ahuvia (2006) lo describen como un vínculo caracterizado por el apego emocional y una fuerte conexión afectiva. No basta con que la marca funcione bien o resulte satisfactoria: lo que define el *brand love* es una sensación de cercanía o identificación similar a la que se da en relaciones personales. Por eso los consumidores que lo experimentan no solo repiten compra, sino que recomiendan la marca, toleran errores puntuales y mantienen su lealtad incluso cuando existen alternativas más económicas o convenientes.

Una de las explicaciones sobre cómo se forma este vínculo es el concepto de personalidad de marca. Según Aaker (1997), los consumidores atribuyen rasgos humanos a las marcas, los cuales se organizan en cinco dimensiones: sinceridad, entusiasmo, competencia, sofisticación y robustez. Esta última hace referencia a marcas percibidas como fuertes o resistentes asociadas al contexto estadounidense, como Marlboro o Levi's. Cuando esa personalidad se alinea con los valores del consumidor es más probable que se genere una conexión emocional, y que esta sea más estable en el tiempo. Al elegir marcas con determinada personalidad, los consumidores también expresan quiénes son o quiénes quieren ser, convirtiendo el consumo en un acto de identidad. Sin embargo, la aplicabilidad del modelo no es completamente universal: la dimensión de robustez, está construida sobre referencias culturales muy específicas del mercado estadounidense, no se traslada con el mismo significado a otros contextos. Esto es relevante en un estudio que abarca mercados tan distintos como Londres, Nueva York y Shanghái.

Desde una perspectiva complementaria, Albert y Merunka (2013) amplían este enfoque al introducir dimensiones como la integración de la marca en la identidad del consumidor,

la idealización y el compromiso emocional. Según estos autores, el *brand love* se desarrolla cuando la marca ocupa un lugar relevante en la vida del consumidor, más allá de la compra: cuando representa aspiraciones, aporta inspiración o forma parte de cómo alguien se define a sí mismo. En esos casos la relación se construye sobre emociones, recuerdos y asociaciones simbólicas, no sobre atributos funcionales.

Es importante distinguir, sin embargo, entre personalidad y experiencia de marca. La personalidad es lo que el consumidor imagina o proyecta sobre una marca antes de tener contacto directo con ella, los rasgos que le atribuye desde fuera. La experiencia, en cambio, ocurre en el contacto real: es lo que se siente, se percibe y se recuerda después de haber estado ahí. Schmitt, Brakus y Zarantonello (2009) identifican cuatro dimensiones en esa experiencia: sensorial, afectiva, intelectual y conductual. En la industria hotelera esta distinción cobra especial peso: la identidad de una cadena puede construirse desde la comunicación y el marketing, pero la experiencia solo ocurre dentro del hotel, y es esa experiencia, no la imagen previa, la que determina si el huésped vuelve.

Los consumidores que desarrollan este tipo de vínculo no solo repiten compra: recomiendan la marca, la defienden y comparten su experiencia de forma espontánea. En el entorno digital esto se traduce en reseñas online, que son precisamente el tipo de comportamiento que este estudio analiza.

2.2 Factores emocionales en la fidelización hotelera

En hostelería, la fidelización es un fenómeno más complejo de lo que parece. Los huéspedes no vuelven a un hotel únicamente por factores funcionales como la comodidad o el precio. Estos elementos son necesarios, pero no suficientes para generar fidelidad. Lo que convierte a un huésped en cliente habitual suele ser algo menos tangible: la sensación de que ese lugar, de alguna manera, te conoce. La investigación así lo confirma: la dimensión emocional predice la intención de volver mejor que cualquier evaluación funcional (Mattila, 2006), y lo que hace que un cliente quiera volver no es solo el servicio en sí, sino sentirse bien tratado, reconocido y cómodo durante la estancia.

El mecanismo que explica este fenómeno tiene que ver con la memoria. Las experiencias emocionalmente intensas (tanto positivas como negativas) se almacenan de forma

duradera y el consumidor las recupera cuando toma decisiones futuras. No es la evaluación de la calidad lo que lleva a volver, sino el recuerdo de cómo esa estancia hizo sentir. Un gesto personalizado, una gestión empática de un problema, o una atmósfera que hizo sentirse como en casa pueden pesar más en esa decisión que el precio o las instalaciones.

De todos los factores que moldean esa experiencia, el componente humano es probablemente el más determinante. Las interacciones con el personal; desde el saludo al llegar hasta la resolución de un problema, definen el tono emocional de la estancia de una forma que ningún *amenity* puede reemplazar. Parasuraman, Zeithaml y Berry (1988) identifican la empatía y la garantía, entendidas como la atención individualizada y la cortesía del personal, como dimensiones clave de la calidad percibida en los servicios. Cuando esas interacciones se perciben como auténticas generan confianza, y la confianza es el primer paso hacia la fidelidad. Un trato frío o distante, en cambio, puede arruinar una estancia incluso en instalaciones de primera categoría.

Otro elemento relevante es la personalización. Los huéspedes valoran cuando un hotel recuerda sus preferencias, anticipa sus necesidades o adapta pequeños detalles a su situación. Esa sensación de ser reconocido añade valor más allá de la eficiencia del servicio y convierte la experiencia en algo difícil de replicar en otro establecimiento.

La experiencia hotelera también está condicionada por el entorno social. Verhoef et al. (2009) señalan que el comportamiento de otros clientes del hotel puede influir en la percepción de la estancia, construyendo o deteriorando involuntariamente el apego emocional. Es un factor que este estudio no controla directamente, pero que conviene reconocer como parte del contexto experiencial.

Aun así, los aspectos funcionales siguen siendo la base. La limpieza, la calidad de la habitación o el ruido actúan como requisitos mínimos: su ausencia genera insatisfacción, pero su presencia no genera afecto ni lealtad por sí sola (Zeithaml, Berry & Parasuraman, 1996). La diferencia entre un cliente satisfecho y uno fiel está, casi siempre, en lo que ocurre por encima de esa base.

2.3 Marketing experiencial y tipología de experiencias

Pine & Gilmore (1998) argumentan que los consumidores han dejado de comprar productos o servicios para optar por experiencias. En la industria hotelera esto es literal: el huésped no adquiere una cama, sino la vivencia completa que solo existe en su memoria una vez que ha terminado la estancia.

La literatura distingue entre experiencias de naturaleza funcional, que giran en torno a aspectos prácticos: que la habitación esté lista, que el check-in sea rápido, que la ubicación permita llegar a donde se necesita; y experiencias de naturaleza emocional, generan afecto, dejan recuerdos y construyen un vínculo que va más allá de si todo funcionó correctamente. Verhoef et al. (2009) subrayan que la experiencia del cliente no se limita al momento del consumo: abarca desde la búsqueda inicial hasta la valoración posterior, e integra dimensiones cognitivas, afectivas, sensoriales y relacionales. En el contexto hotelero, esto significa que la experiencia empieza antes de entrar al hotel y termina después de abandonarlo. Schmitt, Brakus y Zarantonello (2009) concretan esas dimensiones en cuatro: sensorial, afectiva, intelectual y conductual, pueden activarse en distinta medida según el tipo de experiencia vivida, lo que justifica tratarlas como categorías analíticamente diferenciadas y no como un continuo uniforme.

Un tercer tipo de experiencia relevante para este estudio es la experiencia friccional, caracterizada por la presencia de problemas, conflictos o frustraciones durante la estancia. Estas experiencias generan una carga emocional negativa que se refleja en valoraciones bajas, pero también, en reseñas emocionalmente intensas. La fricción no es simplemente ausencia de satisfacción funcional, es una experiencia que activa emociones como la frustración, la indignación o la decepción. En el análisis textual aparece asociada a lo que la literatura denomina aspectos críticos: dimensiones de importancia para el huésped que reciben valoración negativa y cuya mejora tiene el mayor impacto potencial sobre la satisfacción global.

Esta tipología (experiencias emocionales, funcionales, mixtas y friccionales) es el marco interpretativo del análisis empírico de este estudio. A través de modelado de tópicos (LDA) se identifican los patrones experienciales dominantes en las reseñas y se analiza

cómo cada tipo de experiencia se asocia con distintos tipos de niveles de apego emocional medido a través del *Love Score*.

La hipótesis inicial es que el tipo de experiencia vivida durante la estancia explica el vínculo emocional del huésped con el hotel mejor que la categoría del establecimiento. Es decir, lo que ocurre dentro importa más que el número de estrellas. Los datos de este estudio permiten poner esa hipótesis a prueba.

2.4 Engagement digital y señales de lealtad en reseñas online

La identificación de estos tipos de experiencia descritos en el apartado anterior requiere instrumentos de medición capaces de capturar su dimensión emocional de forma sistemática. Las reseñas online ofrecen esa capacidad, y de una forma que ningún otro método puede replicar.

Meyer y Schwager (2007) distinguen entre contactos directos con la marca, los que ocurren durante la compra o el uso del servicio, y contactos indirectos, como las recomendaciones entre conocidos o las reseñas online. Estos últimos ocurren sin que la marca los controle, y su impacto en la percepción del consumidor puede ser tan real como el de la experiencia directa. Una reseña en TripAdvisor no es una encuesta de satisfacción; es un relato espontáneo donde el huésped decide, sin que nadie se lo pida, qué merece ser contado y cómo.

Van Doorn et al. (2010) definen el *customer engagement behaviour* como las manifestaciones conductuales del cliente hacia una marca más allá de la transacción directa. Escribir una reseña online es precisamente una de estas manifestaciones: el huésped dedica tiempo y esfuerzo a describir su experiencia de forma voluntaria, lo que convierte el texto de la reseña en una expresión auténtica de su relación con el establecimiento. Un indicador relevante de ese compromiso es la extensión del texto, lo que la literatura denomina profundidad de la reseña, ya que reseñas más largas y detalladas reflejan un mayor nivel de implicación emocional y resultan más influyentes en las decisiones de otros consumidores (Mudambi & Schuff, 2010).

En el contenido textual de las reseñas, las expresiones de intención de retorno y de recomendación constituyen señales lingüísticas directas de lealtad conductual. Frases como "volveré sin duda", "mi hotel de referencia" o "lo recomiendo a todos" no solo expresan satisfacción: revelan una relación proyectada hacia el futuro con la marca. Oliver (1999) describe la lealtad como un proceso que evoluciona en etapas. Primero el consumidor reconoce racionalmente que una marca es mejor que las alternativas: lealtad cognitiva. Después desarrolla un vínculo afectivo hacia ella. Más tarde aparece la intención firme de volver: lealtad conativa. Finalmente esa intención se convierte en comportamiento repetido. Las expresiones de retorno y recomendación que aparecen en las reseñas, son manifestaciones de esa fase conativa: el huésped no solo está satisfecho, ha decidido que va a volver. Y esa decisión, expresada de forma espontánea en un texto público, es una de las señales más claras de vínculo emocional real con la marca.

La combinación de la satisfacción declarada, el tono emocional del texto y las señales lingüísticas de lealtad constituye la base del indicador compuesto desarrollado en este estudio. No basta con que el huésped esté satisfecho, ni con que su texto sea emocionalmente positivo; la presencia simultánea de ambas dimensiones junto con la intención de retorno es lo que permite aproximarse de forma operativa el concepto de *brand love* en el contexto de las reseñas digitales.

2.5 Síntesis y justificación del modelo empírico

El *brand love* en hostelería no surge de un único factor sino de una cadena: la personalidad de marca crea expectativas antes de la estancia, las experiencias durante la estancia las confirman o las rompen, y las reseñas online son el momento en que el huésped decide, con sus propias palabras, qué tipo de relación tiene con ese hotel. El vínculo se construye en esa secuencia, no en un solo instante.

Esta cadena tiene una implicación directa para el diseño del análisis: medir el apego emocional únicamente con la puntuación numérica no es suficiente. La puntuacion numérica otorgada por el huésped en la plataforma, captura si las expectativas se cumplieron, pero no la intensidad del vínculo que esa experiencia dejó. Dos huéspedes pueden dar exactamente cuatro estrellas a un hotel por razones completamente distintas, uno porque la habitación era correcta y el desayuno aceptable, otro porque el personal

gestionó con una amabilidad inesperada un problema que arruinó su primer día. El número no los distingue. Por eso el *Love Score* integra tres dimensiones complementarias: la satisfacción declarada a través de la puntuación, el tono emocional extraído del texto, y las señales lingüísticas de intención de retorno o recomendación. Cada una captura algo que las otras no.

El modelado de tópicos LDA completa el modelo desde otra dirección: en lugar de medir el apego emocional directamente, permite identificar qué tipos de experiencia describen los huéspedes en sus reseñas y analizar cómo varía ese apego entre ellos. La hipótesis de partida es que el tipo de experiencia (emocional, funcional, mixta o friccional) puede aportar una mayor capacidad explicativa del vínculo con la marca que la categoría del establecimiento.

3. Metodología

3.1 Descripción del dataset

El análisis empírico de este trabajo se basa en una base de datos con 416.032 reseñas de hoteles, extraídas de TripAdvisor y publicadas en 2024 a través del Science Data Bank (SciDB), repositorio de datos científicos de acceso abierto[1]. El dataset fue exportado a formato Excel (.xlsx) para su procesamiento local en el entorno de análisis.

El dataset cubre más de 100 hoteles distribuidos en tres grandes ciudades de referencia internacional: Shanghái, Londres y Nueva York. Estas ciudades fueron seleccionadas por los autores de la base de datos por su relevancia dentro del sector hotelero global y por la diversidad de segmentos y cadenas presentes en cada mercado.

Las reseñas de la base de datos fueron recopiladas mediante técnicas de *web scraping* aplicadas directamente sobre TripAdvisor, combinando información estructurada (fecha, *rating*, nombre del hotel, datos de usuario) con texto libre generado espontáneamente por los huéspedes. Esta combinación permite analizar tanto variables cuantitativas como lingüísticas en un mismo marco analítico. Las variables disponibles en el dataset original son las siguientes:

- *Unnamed 0.1 y Unnamed 0*: columna generada automáticamente durante la exportación del dataset, sin contenido analítico.
- *hotel_name*: nombre del hotel tal como aparece en TripAdvisor.
- *id_review:* identificador único de la reseña, utilizado para el control de duplicados.
- t*itle*: título de la reseña escrito por el usuario.
- *date:* fecha de publicación de la reseña.
- *location:* ciudad y país de origen del usuario.
- *user_name:* nombre del usuario que escribe la reseña.
- *user_link:* enlace al perfil público del usuario en TripAdvisor.
- *date_of_stay:* fecha de la estancia en el hotel.
- *rating*: puntuación numérica otorgada por el usuario en escala 1-5.
- *review*: texto completo de la reseña, variable central del análisis lingüístico.

[1] Disponible en: https://www.scidb.cn/detail?dataSetId=61761d4278cf4572aae2bccaf7acbe49 [Consultado: diciembre 2025]

- *rating_review:* valoración de utilidad asignada por otros usuarios a la reseña.
- *n_review_user:* número total de reseñas publicadas por ese usuario en la plataforma.
- *n_votes_review*: votos útiles recibidos por la reseña.

El volumen del dataset y su distribución entre tres ciudades de referencia internacional proporcionan una base sólida para identificar patrones en las experiencias descritas y analizar cómo se relacionan la valoración del servicio con el vínculo emocional que el huésped desarrolla hacia el hotel.

3.2 Procesamiento y limpieza del dataset

Antes de iniciar la fase analítica, fue necesario realizar un proceso de limpieza y estandarización del dataset original. Dado que los datos procedían de un proceso de *web scraping*, el archivo original tenía inconsistencias habituales en este tipo de extracción que podían comprometer la validez de los análisis posteriores.

En primer lugar, se eliminaron dos columnas vacías (*Unnamed*: 0 y *Unnamed*: 0.1) que correspondían a índices automáticos generados durante la exportación del archivo y no contenían información. A continuación, se estandarizaron los nombres de todas las variables, convirtiéndolos a minúsculas y sustituyendo espacios por guiones bajos, con el fin de garantizar consistencia.

Las variables temporales: *date* y *date_of_stay*, que inicialmente estaban almacenadas como cadenas de texto, se convirtieron a formato de fecha estándar (año-mes-día). Esta conversión es necesaria para asegurar su correcta interpretación en análisis temporales y para evitar errores en operaciones de ordenación o filtrado.

Adicionalmente, se excluyeron las reseñas con menos de 30 caracteres, al considerarse insuficientes para el análisis textual y la asignación temática fiable.

La verificación mediante *id_review* identificó 484 reseñas potencialmente duplicadas (0,12% del corpus). Dado su volumen reducido y su distribución entre hoteles y ciudades, se evaluó que su presencia no altera de forma significativa los indicadores agregados ni

los patrones identificados, por lo que el análisis se realiza sobre el corpus completo de 416.032 observaciones.

Algunas variables, en particular *location* y *date_of_stay*, presentaban valores faltantes. Dado que éstas no son esenciales para el estudio, se conservaron las observaciones con datos ausentes para mantener el mayor volumen y diversidad posible. En el caso de variables textuales como *review*, los valores nulos se imputaron como cadenas vacías, para evitar errores en las fases posteriores de procesamiento lingüístico.

Una variable crítica para el estudio es el *rating* que cada huésped otorga. En el conjunto original, algunas puntuaciones aparecían expresadas en múltiplos de diez (por ejemplo 40 o 50 en lugar de 4 o 5). Todos los valores fueron normalizados a escala original 1-5, dividiendo entre 10 cuando era necesario, garantizando la homogeneidad de la variable de satisfacción funcional en todo el dataset.

Tras este proceso de limpieza, el dataset quedó compuesto por 416.032 observaciones válidas, cada una correspondiente a una reseña individual del corpus. El número de observaciones se mantiene respecto al dataset original porque el proceso de limpieza no implicó eliminación de filas: los valores nulos en variables no esenciales se conservaron para maximizar el volumen analítico y los nulos en variables textuales se imputaron como cadenas vacías.

3.3 Modelado de experiencias mediante análisis de tópicos

El objetivo central del estudio es analizar cómo varía la relación entre satisfacción funcional y apego emocional, según el tipo de experiencia descrita por los huéspedes. Para capturar estas experiencias de forma sistemática a partir del texto, se empleó el modelo de análisis *Latent Dirichlet Allocation* (LDA), una técnica de aprendizaje no supervisado que permite identificar temas latentes en grandes colecciones de documentos (Blei, Ng & Jordan, 2003).

La unidad de análisis principal es la reseña individual, clasificada según el tipo de experiencia dominante que ésta describe. Este enfoque permite identificar tipos de

experiencia a partir de cómo los propios huéspedes las describen, en lugar de recurrir a categorías definidas de antemano.

3.3.1 Procesamiento textual para el modelado de tópicos

Antes de entrenar el modelo, el texto de las reseñas fue sometido a un proceso de preparación en varias etapas, implementado mediante la librería *spaCy* (modelo *en_core_web_sm*). El análisis textual se realiza sobre el conjunto de reseñas previamente filtrado, excluyendo aquellas con menos de 30 caracteres, al considerarse insuficientes para una asignación temática fiable.

1. Tokenización y lematización

Cada reseña fue procesada mediante el modelo de procesamiento lingüístico de *spaCy*, que realiza simultáneamente la tokenización (división del texto en unidades léxicas), y la lematización (reducción de cada término a su forma canónica). La lematización mejora la coherencia temática al agrupar variantes morfológicas bajo un mismo término: "staying", "stayed" y "stays" se tratan como una única unidad léxica "stay".

2. Filtrado de tokens no informativos

Se eliminaron *stopwords* (palabras muy frecuentes sin contenido semántico relevante, como “el”, “la” o “de”), signos de puntuación, espacios, números y *tokens* de menos de tres caracteres. Este filtrado reduce el ruido léxico y evita que términos muy frecuentes, pero poco informativos distorsionen los patrones temáticos del modelo.

3. Limpieza semántica específica del dominio

Adicionalmente se aplicó una limpieza semántica en dos niveles. Por un lado, se eliminaron términos de dominio muy genéricos en el contexto hotelero, como "hotel", "room", "staff", "stay", "night" o "bed", cuya presencia en casi todas las reseñas impediría la diferenciación temática.

Por otro lado, con el parámetro *REMOVE_GEO = True*, se eliminaron referencias geográficas específicas como "london", "new_york" o "shanghai", evitando que el modelo generase tópicos de naturaleza geográfica en lugar de experiencial. Esta decisión es relevante para el enfoque del estudio: el interés no es dónde ocurre la experiencia sino qué tipo de experiencia se describe.

4. Construcción de n-gramas

Para preservar expresiones compuestas semánticamente relevantes, se construyeron bigramas y trigramas mediante el modelo *Phrases* de la librería *Gensim*, con umbrales *min_count=20* y *threshold=10*. El primero establece el número mínimo de apariciones necesarias para que una combinación de palabras sea considerada, evitando la inclusión de expresiones poco frecuentes. El segundo controla el grado de asociación requerido entre términos para formar una unidad compuesta, asegurando que solo se mantengan aquellas combinaciones con suficiente coherencia estadística. Este paso permite que términos como "customer service" o "hot water" sean tratados como unidades léxicas únicas, mejorando la interpretabilidad de los tópicos resultantes.

Tras la construcción de bigramas y trigramas, se aplicó una limpieza adicional para eliminar fragmentos geográficos residuales y términos genéricos con poca capacidad discriminante.

3.3.2 Construcción del diccionario y del corpus

Una vez generados los *tokens* finales, se creó un diccionario de términos en el que cada palabra única del conjunto de reseñas recibe un identificador numérico único.

Sobre este diccionario se aplicaron filtros de frecuencia estándar: se eliminaron términos con aparición en menos de 50 documentos (*no_below=50*) y términos presentes en más del 20% de los documentos (*no_above=0,20*). El primer filtro elimina términos demasiado raros para ser estadísticamente relevantes; el segundo elimina términos tan frecuentes que no aportan capacidad discriminante entre tópicos. El diccionario resultante contenía 2.595 términos sobre un corpus de 30.000 documentos.

Cada reseña fue convertida a formato *bag-of-words*, representándola como una lista de pares (identificador de término, frecuencia de aparición). Esta representación es la entrada estándar para el modelo LDA.

3.3.3 Muestra de entrenamiento

El modelo LDA se entrenó sobre una muestra aleatoria de 30.000 observaciones del corpus total de 416.032, lo que representa aproximadamente el 7,2% del corpus. Esta decisión se tomó por motivos de eficiencia computacional: entrenar LDA sobre el corpus completo requeriría recursos enormes sin una ganancia muy significativa en la calidad de los tópicos, dado que el modelo converge correctamente con muestras de tamaño suficiente.

La muestra fue seleccionada mediante muestreo aleatorio simple con semilla fija (*random_state=42*) garantizando la reproducibilidad del análisis. La representatividad de la muestra respecto al corpus completo fue verificada mediante test de *Kolmogorov-Smirnov* bilateral sobre las tres variables principales del modelo (*rating*, sent_*norm* y *love_score*). En los tres casos el estadístico D fue inferior a 0,006 y el p-valor superior a 0,27, lo que indica que la muestra no presenta diferencias significativas respecto al corpus completo en sus distribuciones.

3.3.4 Selección del número de tópicos

Para determinar el número óptimo de tópicos se entrenaron modelos LDA para valores de k entre 6 y 12, evaluando cada uno mediante la métrica de coherencia semántica *c_v* (Röder, Both & Hinneburg, 2015). Esta métrica se basa en la idea de que, si las palabras de un tópico aparecen juntas con frecuencia en el corpus, ese tópico será más coherente y fácil de interpretar. Los resultados del grid de coherencia fueron los siguientes:

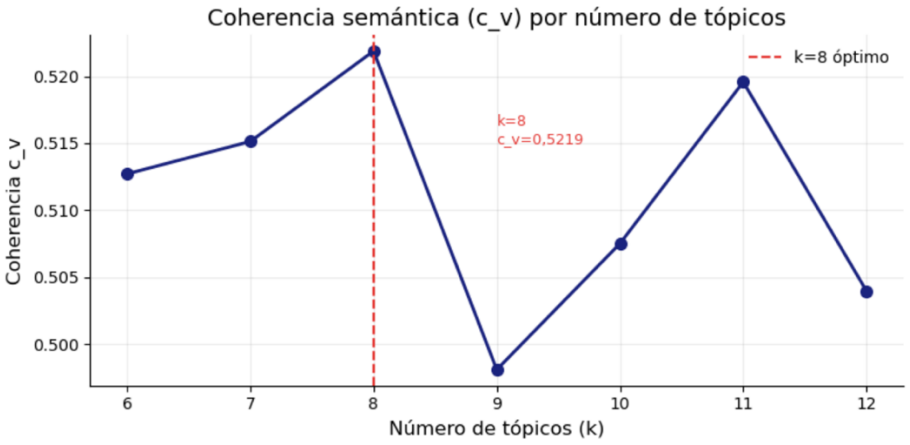


Figura 1. Coherencia semántica por número de tópicos

El modelo con k=8 obtuvo la mayor coherencia (c_v=0,5219). Aunque k=11 presenta un valor próximo (0,5196), la diferencia es marginal y el principio de parsimonia justifica preferir el modelo más simple. La elección de k=8 fue además validada mediante el análisis de solapamiento entre tópicos: el máximo solapamiento observado entre cualquier par de tópicos fue de 3 palabras sobre 12, inferior al 25%, lo que confirma una diferenciación temática adecuada.

3.3.5 Asignación de tópicos y control de ambigüedad temática

Cada reseña recibe del modelo LDA una distribución de probabilidad sobre los 8 tópicos. Para el análisis se asignó a cada reseña su tópico dominante, definido como aquel con mayor probabilidad. Sin embargo, dado que una misma reseña puede abordar varias dimensiones experienciales simultáneamente, se evaluó la claridad de asignación temática mediante dos criterios complementarios: la probabilidad del tópico dominante (*topic_prob*) y el margen respecto al segundo tópico más probable (*topic_margin*). Las reseñas cuya probabilidad dominante o cuyo margen se situaban por debajo del percentil 40 de sus respectivas distribuciones, fueron clasificadas como ambiguas (*topic_ambiguous=true*). Los umbrales obtenidos fueron p40=0,441 para la probabilidad y p40=0,154 para el margen. Al aplicar estos valores, el 46,68% de las reseñas se clasificaron como ambiguas, es decir, sin una asignación clara a un único tópico. La muestra analítica principal quedó finalmente compuesta por 15.996 reseñas con dominancia temática definida.

El percentil 40 se utiliza como umbral para equilibrar dos objetivos: asegurar que cada reseña esté claramente asociada a un tópico y, al mismo tiempo, mantener un tamaño de muestra suficiente en cada grupo. Umbrales más exigentes (percentil 50 o superior) reducirían el número de observaciones en algunos tópicos minoritarios, como T6 (Restauración, n=268) o T4 (Movilidad, n=285), hasta niveles poco adecuados para el análisis.

Además, el análisis realizado sobre las reseñas más ambiguas muestra que los patrones principales se mantienen, lo que refuerza la consistencia de los resultados

3.3.6 Clasificación conceptual de los tópicos

Como toda técnica no supervisada, el LDA identifica patrones estadísticos en los datos, pero no asigna significado a los tópicos resultantes. La interpretación conceptual requiere un paso adicional de análisis cualitativo. Cada tópico fue etiquetado manualmente a partir de dos fuentes de evidencia: las doce palabras con mayor peso probabilístico en la distribución del tópico y ejemplos de reseñas con alta probabilidad de pertenencia (*topic_prob* > 0,80 y *topic_margin* > 0,40).

A partir de este análisis, los ocho tópicos fueron clasificados en cuatro categorías conceptuales alineadas con el marco teórico de la sección 2.3:

Tabla 1. Tópicos LDA

Tópico	Nombre	Tipo	Palabras clave principales
T1	Habitación y espacio físico	funcional	small, shower, apartment, space, large, kitchen, size, floor
T2	Check-in, reservas y gestión administrativa	friccional	check, tell, ask, book, pay, charge, find, desk, arrive
T3	Entorno cercano y accesibilidad a pie	mixto	close, restaurant, area, quiet, friendly, near, minute_walk
T4	Movilidad urbana y conexiones	funcional	walk, street, elevator, subway, airport, floor, block
T5	Experiencia emocional y trato recibido	emocional	amazing, experience, love, thank, lovely, wonderful, perfect
T6	Restauración y bebidas	mixto	food, order, drink, bar, table, coffee, morning, serve
T7	Problemas de descanso y confort	friccional	sleep, noise, door, bad, smell, shower, hot, window
T8	Servicios del hotel y experiencia complementaria	mixto	restaurant, view, bund, gym, lounge, pool, business, bar

Para comprobar la coherencia de esta clasificación, se analizaron varios indicadores en cada tópico: el *rating* medio, el *Love Score* medio, la diferencia entre ambos y el porcentaje de reseñas con señales de lealtad.

Los tópicos considerados emocionales presentan valores más altos de *Love Score* y mayor proporción de lealtad, mientras que los friccionales muestran los niveles más bajos. Este patrón es coherente con la interpretación de cada grupo.

3.4 Construcción de las variables analíticas

Tras la limpieza del dataset y el procesamiento textual, se construyeron las variables analíticas para analizar los conceptos descritos en el marco teórico. El objetivo era disponer de medidas cuantitativas que capturasen simultáneamente la dimensión

funcional, la dimensión emocional y señales conductuales de lealtad presentes en cada reseña.

3.4.1 Satisfacción funcional: *rating_norm* y *rating_100*

La variable de satisfacción funcional se construyó a partir del *rating* original en escala 1-5. Para permitir su integración en el indicador compuesto, se normalizó a una escala continua 0-1 mediante la siguiente transformación lineal:

$$\textbf{rating_norm} = (\textbf{rating} - 1) / 4$$

Esta transformación asigna el valor 0 a la puntuación mínima (1 estrella) y el valor 1 a la máxima (5 estrellas), manteniendo la proporción entre las distintas puntuaciones.

Para facilitar la interpretación en los análisis descriptivos, se construyó una versión reescalada: *rating_100* = 100 × *rating_norm*, expresada en escala 0–100.

La media del corpus es de 80,41 puntos (desviación típica: 27,41), lo que refleja el sesgo positivo característico de las plataformas de reseñas; los huéspedes con experiencias neutras tienden a no escribir reseñas, lo que sobre representa las experiencias extremas, especialmente las positivas.

3.4.2 Tono emocional: *sent_norm*

Para capturar la dimensión emocional del texto se construyó un indicador de sentimiento léxico basado en un diccionario propio de 30 términos positivos y 27 términos negativos, seleccionados por su frecuencia y relevancia semántica en el sector hotelero. Para cada reseña se contabilizaron las ocurrencias de términos positivos (*pos_hits*) y negativos (*neg_hits*).

La puntuación neta de sentimiento se calculó como la diferencia entre ambos tipos de términos, normalizada por la raíz cuadrada de la longitud del texto. Esta normalización permite evitar que las reseñas más extensas obtengan puntuaciones extremas por acumulación de palabras. Sobre esta puntuación se aplicó una función (*tanh*) que acota el resultado a un rango entre -1 y +1, y reduce la influencia de valores extremos.

Finalmente, el resultado se reescaló al rango 0-1 mediante una transformación min-max para permitir su integración en el indicador compuesto.

El diccionario alcanzó cobertura en el 93,89% de las reseñas del corpus, con al menos un término positivo o negativo detectado por observación. La media de *sent_norm* en el corpus es de 0,608 (desviación típica: 0,120).

Este enfoque presenta una limitación que conviene reconocer. Al basarse en un diccionario construido ad hoc, y no en modelos preentrenados como *VADER* o *BERT*, su capacidad para interpretar correctamente contextos de negación o ironía es limitada.

En particular, puede identificar términos positivos en reseñas con valoración global negativa cuando el huésped menciona aspectos favorables antes de describir el problema principal. Por ejemplo, en una reseña como “la habitación era amplia y la ubicación buena, pero el servicio fue pésimo”, el modelo puede captar las palabras positivas iniciales sin reflejar adecuadamente la valoración final. Esta limitación se ha tenido en cuenta en el análisis de cobertura y debe considerarse al interpretar los resultados.

Esta limitación podría reducirse mediante reglas de negación, tratamiento de conectores de contraste o el uso de modelos preentrenados de análisis de sentimiento. Sin embargo, dado el objetivo exploratorio del trabajo y la necesidad de mantener un procedimiento interpretable y reproducible, se optó por un enfoque léxico más sencillo.

3.4.3 Señales de lealtad: *kw_loyal_strict*

Para capturar señales explícitas de lealtad se construyó la variable binaria *kw_loyal_strict*, que toma valor 1 cuando la reseña contiene al menos una de las 19 expresiones predefinidas asociadas a intención de retorno o recomendación, y 0 en caso contrario. Las expresiones más frecuentes en el corpus fueron "stay here again" (21.264 ocurrencias), "highly recommend" (21.041) y "come back" (12.528).

El criterio de selección consistió en priorizar expresiones claras de intención de retorno o recomendación, evitando términos ambiguos que pudieran aparecer en otros contextos. El 21,22% de las reseñas del corpus contiene al menos una de estas expresiones.

Esta variable captura intención declarada de retorno o recomendación, lo que constituye una aproximación conductual al apego emocional alineada con el concepto de lealtad conativa de Oliver (1999). No obstante, debe interpretarse como intención y no como lealtad efectiva observada, dado que el dataset no permite verificar si el huésped repitió su estancia.

3.4.4 El *Love Score*: indicador compuesto de apego emocional

Las tres dimensiones anteriores se integraron en un indicador compuesto denominado *Love score,* diseñado para aproximar de forma operativa el grado de apego emocional del huésped hacia el hotel. Se construyó mediante una combinación ponderada de los tres componentes:

$$\textit{Love Score} = 100 \cdot (0{,}6 \cdot \textit{rating_norm} + 0{,}3 \cdot \textit{sent_norm} + 0{,}1 \cdot \textit{kw_loyal_strict})$$

El *Love Score* toma valores entre 0 y 100, lo que facilita su interpretación y su comparación directa con *rating_100*. La media del *Love Score* en el corpus es de 68,61 puntos (desviación típica: 19,89).

La distribución de pesos refleja tres consideraciones:

En primer lugar, el 60% asignado al *rating* se justifica porque es la medida más directa y validada de la satisfacción del cliente, con mayor fiabilidad que las variables derivadas del texto (Parasuraman, Zeithaml y Berry, 1988). En segundo lugar, el 30% asignado al sentimiento refleja su capacidad para captar matices emocionales que no aparecen en la puntuación numérica, aunque se reconoce su menor precisión al basarse en un diccionario léxico construido ad hoc. Por último, el 10% correspondiente a *kw_loyal_strict* responde a su naturaleza binaria y a su menor variabilidad. Su función es reforzar el indicador

cuando existen señales explícitas de lealtad, sin distorsionar el resultado cuando estas no están presentes.

El *Love Score* no mide la lealtad comportamental observada, sino el apego emocional declarado. Dado que no se dispone de datos que permitan verificar la repetición efectiva de estancia, el indicador combina tres dimensiones complementarias que aproximan la vinculación actitudinal del huésped con el establecimiento.

La alta correlación entre *rating_100* y *Love Score* (r=0,962) es una consecuencia directa de su construcción, ya que el *rating* representa el 60% del indicador. Esto no debe interpretarse como redundancia. La diferencia entre ambas variables, medida a través de *gap_love_vs_rating = Love Score − rating_100*, constituye precisamente la dimensión analítica de mayor interés en este estudio.

Adicionalmente, se estimó un modelo de regresión OLS en el que *Love Score* se utiliza como variable dependiente, las *dummies* de tópico como variables independientes y *rating_100* como variable de control. El objetivo es estimar hasta qué punto cada tipo de experiencia influye en el apego emocional del huésped, una vez controlado el nivel de satisfacción funcional (*rating_100*).

3.4.5 Variables complementarias

Adicionalmente, se construyeron dos variables derivadas del *Love Score* para analizar las diferencias entre el *Love Score* y el *rating_100*, es decir, en qué medida el apego emocional difiere de la valoración funcional:

1. ***love_nonrating = 100 · (0,3 · sent_norm + 0,1 · kw_loyal_strict)***. Componente del *Love Score* independiente del *rating*, que captura exclusivamente las dimensiones emocional y conductual. Su media en el corpus es de 20,37 puntos.

2. ***gap_love_vs_rating = love_score − rating_100***. Mide la diferencia entre el apego emocional declarado y la satisfacción funcional. Valores positivos indican que el huésped expresa un mayor apego emocional del que sugiere su puntuación numérica,

mientras que valores negativos lo contrario. La media del gap en el corpus es de −11,80 puntos, lo que sugiere que el *rating* tiende a situarse por encima *del Love Score*, especialmente en reseñas con puntuación máxima, donde el sentimiento léxico no alcanza el mismo nivel.

Es importante tener en cuenta que el rating forma parte del propio cálculo del Love Score, por lo que ambas variables están directamente relacionadas. Esto implica que los resultados de los modelos que las utilizan conjuntamente no deben interpretarse como efectos causales independientes, sino como relaciones parciales. En este contexto, el objetivo del análisis no es identificar un efecto causal puro, sino analizar en qué medida el tipo de experiencia se asocia con cambios en el apego emocional una vez considerada la satisfacción funcional.

Con este diseño metodológico, se dispone de un marco analítico que permite examinar la relación entre satisfacción funcional y apego emocional a partir de las reseñas. A continuación, se presentan los resultados obtenidos.

3.4.6 Modelo predictivo del apego emocional

Con el objetivo de complementar el análisis explicativo, se construyó un modelo predictivo para clasificar las reseñas según su nivel de apego emocional. Conviene precisar que el modelo opera sobre características observadas en la reseña ya escrita: no anticipa el apego antes de que ocurra la experiencia, sino que identifica, a partir de variables textuales, si una reseña corresponde a un perfil de alta o baja vinculación emocional. Su aportación principal es demostrar que el tipo de experiencia descrita aporta capacidad clasificatoria adicional más allá de la valoración numérica.

La variable dependiente es *love_nonrating*, el componente del *Love Score* independiente del *rating*, que captura exclusivamente la dimensión emocional y conductual de la reseña.

Para definir las clases, se establecieron dos grupos extremos. Se considera alta vinculación emocional cuando *love_nonrating* supera el percentil 75 (≥23,34), y baja cuando se sitúa por debajo del percentil 25 (≤16,14). Las observaciones con valores

intermedios se excluyen del análisis para asegurar una separación clara entre ambas categorías.

La muestra resultante está compuesta por 15.029 reseñas, distribuidas de forma equilibrada entre ambas clases (50% cada una).

Este planteamiento permite evitar un problema importante: si se utilizara el *Love Score* completo como variable a predecir, el modelo estaría, en parte, reconstruyendo el propio indicador, ya que algunas de sus variables forman parte de su cálculo. Al centrarse en *love_nonrating*, se evalúa de forma más limpia la capacidad explicativa de las variables.

Como variables explicativas se incluyeron tres dimensiones complementarias: *rating_100*, como medida de satisfacción funcional; *review_length*, como aproximación al nivel de implicación del huésped en la reseña; y *topic_dom*, que recoge el tipo de experiencia dominante identificado mediante LDA.

Se estimaron tres modelos de complejidad creciente: Random Forest, Gradient Boosting y red neuronal (MLP), con el objetivo de comparar su capacidad clasificatoria. El entrenamiento se realizó sobre el 80% de la muestra y la evaluación sobre el 20% restante.

4. Resultados

4.1 Análisis descriptivo del dataset

El corpus analítico está formado por 416.032 reseñas de TripAdvisor procedentes de más de 100 hoteles ubicados en Londres, Nueva York y Shanghái. La distribución de la satisfacción funcional, medida a través del *rating* normalizado (*rating_100*), presenta un marcado sesgo positivo: la media se sitúa en 80,41 puntos (desviación típica: 27,41), con una concentración notable en los valores más altos. Este patrón refleja el denominado sesgo de autoselección, los huéspedes con experiencias neutras tienden a no escribir reseñas, y debe tenerse en cuenta en la interpretación de los resultados.

El indicador de tono emocional (*sent_norm*) presenta una distribución relativamente equilibrada, con una media de 0,608 y una desviación típica de 0,120. Dado que la variable está acotada entre 0 y 1, una media por encima de 0,5 indica que predominan las valoraciones positivas, aunque sin concentrarse en los valores extremos, lo que sugiere un tono moderadamente positivo en la mayoría de las reseñas.

La correlación entre *rating_100* y *sent_norm* es de 0,567. Este valor indica una relación positiva de intensidad moderada: en general, a mayores puntuaciones, más positivo es el tono del texto, pero la relación no es perfecta. Esto implica que ambas variables están relacionadas, pero capturan aspectos distintos de la experiencia del huésped.

El *Love Score* presenta una media de 68,61 puntos (desviación típica: 19,89). Su correlación con *rating_100* es de 0,962, un valor muy alto que se explica por la propia construcción del indicador, ya que el *rating* representa el 60% de su composición. Por tanto, esta correlación no debe interpretarse como redundancia, sino como una consecuencia esperada del diseño.

La variable analíticamente más relevante es la diferencia entre ambas medidas. El *gap_love_vs_rating* presenta una media de −11,80 puntos, lo que indica que, en promedio, el *Love Score* es inferior al *rating*.

Este resultado sugiere que una valoración funcional alta no implica necesariamente un nivel equivalente de apego emocional. En particular, en muchas reseñas con puntuación

máxima, el lenguaje utilizado es positivo, pero no lo suficientemente intenso como para reflejar un vínculo emocional fuerte. Por ello, el *rating* tiende a situarse por encima del *Love Score*, generando una divergencia sistemática entre ambas medidas.

Las señales explícitas de lealtad aparecen en el 21,22% de las reseñas. Entre las expresiones más frecuentes se encuentran "stay here again" (21.264 ocurrencias), "highly recommend" (21.041) y "come back" (12.528).

Su correlación con el *Love Score* es de 0,379, la más baja de los tres componentes. Esto refleja que la presencia de estas expresiones es menos frecuente y más puntual que la satisfacción o el tono positivo, que aparecen en la mayoría de las reseñas. Sin embargo, cuando el huésped expresa de forma explícita su intención de volver o recomendar, introduce una señal clara de vínculo que va más allá de una valoración positiva general.

La relación entre la satisfacción funcional y el apego emocional es claramente positiva: a medida que aumenta el *rating*, también lo hace el *Love Score*. Por ejemplo, el valor medio pasa de 13,64 puntos en reseñas con 1 estrella a 82,34 en reseñas con 5 estrellas.

Sin embargo, este aumento no es uniforme. El salto en el *Love Score* entre 4 y 5 estrellas (de 64,93 a 82,34, +17,4 puntos) es mayor que entre 1 y 2 estrellas (de 13,64 a 29,62, +16,0 puntos). Esto sugiere que, en los niveles más altos de satisfacción, el apego emocional crece de forma más intensa: no solo aumenta, sino que se amplifica.

El tono emocional sigue un patrón similar. El valor medio de *sent_norm* pasa de 0,432 en reseñas con 1 estrella a 0,657 en reseñas con 5 estrellas, lo que refuerza la idea de que las valoraciones más altas están asociadas a un lenguaje más positivo.

Esto sugiere que las experiencias muy positivas generan una respuesta emocional desproporcionadamente mayor.

Para complementar el análisis descriptivo, la siguiente figura muestra la distribución de las principales variables del estudio: *rating_100*, *sent_norm* y *Love Score*.

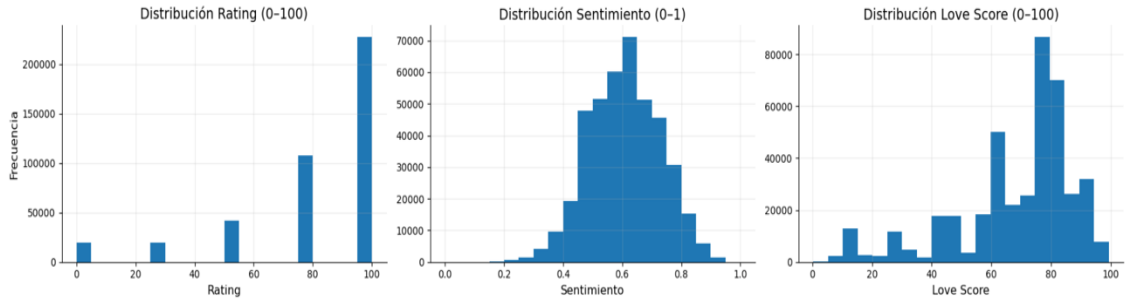


Figura 2. Distribución de *rating_100*, *sent_norm* y *Love Score*

En primer lugar, la distribución del *rating_100* presenta una clara concentración en los valores más altos, con un pico muy pronunciado en torno a 100. Este patrón confirma el sesgo positivo característico de las plataformas de reseñas, donde predominan las valoraciones muy altas.

En contraste, la distribución del *sent_norm* es más equilibrada y presenta una forma aproximadamente normal, centrada en valores intermedios (en torno a 0,6). Esto indica que, aunque las valoraciones numéricas tienden a concentrarse en niveles altos, el lenguaje utilizado en las reseñas refleja mayor variabilidad en el tono emocional.

Por último, el *Love Score* muestra una distribución intermedia entre ambas: mantiene un sesgo positivo, pero con mayor dispersión que el *rating*. Esto sugiere que el indicador captura matices adicionales más allá de la puntuación numérica, evitando la saturación observada en el *rating* y permitiendo diferenciar mejor entre niveles de apego emocional.

Tabla 2. Estadísticos descriptivos principales.

	Media	Desv. típica	Mínimo	Máximo
rating_100	80.415	27.408	0.000	100.000
sent_norm	0.608	0.120	0.000	1.000
kw_loyal_strict	0.212	0.409	0.000	1.000
love_score	68.613	19.886	0.285	99.371
gap_love_vs_rating	-11.801	9.902	-35.986	34.906

La tabla 2 resume los principales estadísticos descriptivos del análisis. Los resultados son consistentes con los patrones observados anteriormente: una distribución sesgada del *rating*, un tono emocional más equilibrado y una diferencia sistemática entre satisfacción funcional y apego emocional.

Tabla 3. Matriz de correlaciones (Pearson)

	rating_100	sent_norm	kw_loyal_strict	love_score	gap_love_vs_rating
rating_100	1.000	0.567	0.158	0.962	-0.836
sent_norm	0.567	1.000	0.239	0.699	-0.165
kw_loyal_strict	0.158	0.239	1.000	0.379	0.325
love_score	0.962	0.699	0.379	1.000	-0.654
gap_love_vs_rating	-0.836	-0.165	0.325	-0.654	1.000

La matriz de correlaciones confirma los patrones descritos anteriormente. Destaca la alta correlación entre *rating_100* y *Love Score* (r = 0,962), coherente con la construcción del indicador.

La relación entre *rating_100* y *sent_norm* (r = 0,567) es positiva pero moderada, lo que refuerza que ambas variables capturan dimensiones distintas. Las señales de lealtad presentan correlaciones más bajas, lo que refleja su menor frecuencia, pero mayor especificidad.

El *gap_love_vs_rating* muestra una correlación negativa elevada con ambas variables, confirmando su papel como medida de diferencia entre satisfacción funcional y apego emocional.

4.2 Identificación de patrones experienciales mediante LDA

Para identificar los principales tipos de experiencia descritos por los huéspedes, se aplicó un modelo LDA sobre una muestra de reseñas del corpus.

Del modelo resultaron ocho tópicos interpretables, clasificados en cuatro categorías conceptuales a partir de sus palabras clave más representativas y de ejemplos de reseñas con alta probabilidad de pertenencia (véase Tabla 1).

Los ocho tópicos abarcan las principales dimensiones de la experiencia hotelera urbana. Dos de ellos concentran la mayor parte de las reseñas con asignación temática definida: T5, experiencia emocional (6.383 reseñas, 39,9%), y T3, entorno cercano (5.443, 34,0%).

Esto sugiere que las reseñas hoteleras tienden a organizarse en torno a dos grandes ejes: la interacción con el personal y el contexto urbano y accesibilidad a pie del establecimiento.

En el extremo opuesto, T6 Restauración (n=268) y T4 Movilidad (n=285) presentan tamaños mucho más pequeños, lo que indica que las reseñas centradas exclusivamente en gastronomía o conectividad urbana son menos frecuentes.

La agrupación en cuatro tipos (emocional, funcional, mixto y friccional) permite ordenar los tópicos según la naturaleza de la experiencia descrita. Los tópicos friccionales (T2 y T7) recogen reseñas centradas en problemas o incidencias. Los funcionales (T1 y T4) describen aspectos operativos del servicio. Los mixtos (T3, T6 y T8) combinan ambos elementos. Por último, el tópico emocional (T5) agrupa experiencias marcadas por una fuerte carga afectiva.

Para asegurar una asignación clara, el análisis principal se centra en las reseñas con dominancia definida de un único tópico. De las 30.000 reseñas modeladas, 15.994 (53,31%) cumplen este criterio y constituyen la muestra analítica principal. El análisis sobre el resto no altera los patrones observados, lo que refuerza la consistencia de los resultados.

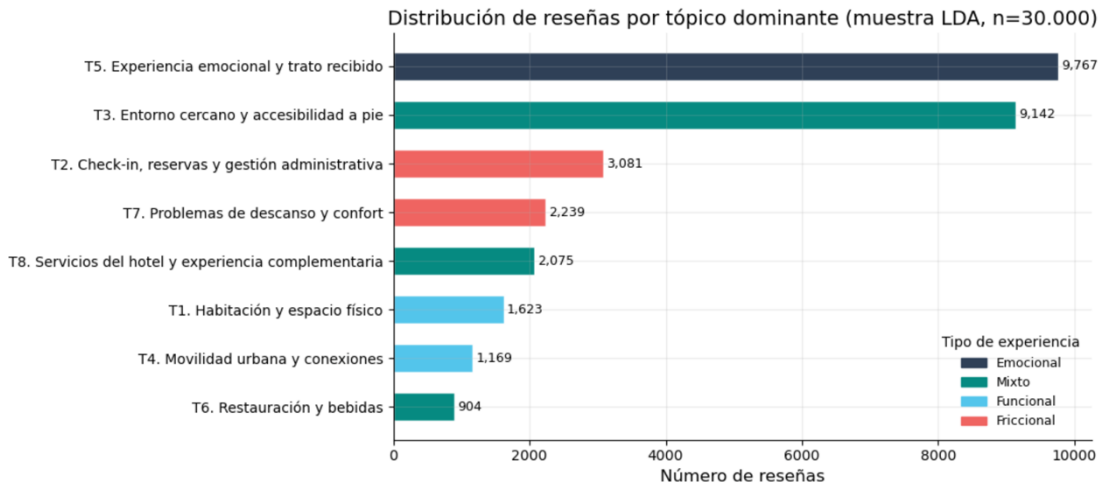


Figura 3. Distribución de reseñas por tópico dominante

Una vez identificados los ocho tópicos, el siguiente paso es analizar si estas categorías se reflejan en diferencias reales en satisfacción funcional y apego emocional. Es decir, no

basta con identificar tipos de experiencia, sino comprobar si esas diferencias se traducen en cómo el huésped valora el hotel y en el vínculo que llega a desarrollar con él.

4.3 *Love Score* y satisfacción funcional por tipo de experiencia

El modelado de tópicos permite clasificar las reseñas según el tipo de experiencia dominante descrita por los huéspedes. Sin embargo, esta clasificación por sí sola no permite evaluar si todos los tipos de experiencia generan el mismo nivel de satisfacción o vínculo emocional.

Por ello, el siguiente paso consiste en comparar los ocho tópicos en términos de satisfacción funcional, apego emocional, diferencia entre ambas dimensiones y presencia de señales explícitas de lealtad. Conviene recordar que los análisis por tópico se basan en la muestra analítica de 15.994 reseñas con dominancia temática clara, extraída de las 30.000 modeladas mediante LDA, y no sobre el corpus completo de 416.032 reseñas. Las conclusiones sobre patrones por tópico deben entenderse en ese contexto muestral.

Tabla 4. Métricas por tópico (dominancia clara)

	topic_dom	n	rating_mean	sent_mean	love_mean	gap_mean	loyal_pct	prob_mean	margin_mean
0	5	6382	97.38	0.666	80.99	-16.38	25.98	0.645	0.453
1	3	5441	87.78	0.648	74.61	-13.17	24.94	0.621	0.415
2	2	1518	30.98	0.474	33.69	2.71	8.70	0.613	0.424
3	7	893	36.28	0.454	36.25	-0.03	8.51	0.568	0.362
4	8	764	84.26	0.596	69.34	-14.92	9.03	0.551	0.326
5	1	443	75.23	0.560	63.84	-11.39	18.96	0.542	0.319
6	4	285	78.16	0.574	65.63	-12.53	15.09	0.531	0.307
7	6	268	63.99	0.540	55.88	-8.12	12.69	0.543	0.310

Los datos muestran diferencias entre los distintos tópicos. El rango del *Love Score* abarca desde 33,69 puntos en T2 (Check-in y gestión administrativa) hasta 80,99 en T5 (Experiencia emocional y trato recibido), una brecha de casi 47 puntos que supera ampliamente las diferencias observadas entre segmentos hoteleros (4,8 puntos entre Luxury y Conventional, con *Love Score* de 71,28 y 66,48 respectivamente). Esto confirma que el tipo de experiencia descrita explica mejor el apego emocional que el posicionamiento del establecimiento.

El tópico con mayor *Love Score* es T5 (Experiencia emocional y trato recibido), con 80,99 puntos, y concentra además la mayor proporción de señales explícitas de lealtad (25,99%). Las reseñas asociadas a este tópico incluyen referencias al trato personalizado, al reconocimiento por parte del personal o a momentos especialmente memorables durante la estancia.

T3 (Entorno cercano y accesibilidad) ocupa la segunda posición (74,61 puntos, *loyal_pct* = 24,94%), con un nivel de lealtad muy próximo al de T5, pero con un *rating* medio inferior (87,78 frente a 97,38). Esta diferencia anticipa que la valoración funcional y el apego emocional no siempre evolucionan al mismo ritmo, algo que se analiza en detalle en el apartado siguiente.

En los tópicos friccionales el patrón es claramente distinto. T2 (Check-in y gestión administrativa) y T7 (Problemas de descanso y confort) presentan *rating* medios bajos (30,98 y 36,28 respectivamente), *Love Scores* también reducidos (33,69 y 36,25) y proporciones de lealtad inferiores al 9%. Estos resultados reflejan que cuando fallan elementos básicos del servicio, como los procesos de llegada o la calidad del descanso, el impacto negativo afecta tanto a la satisfacción como al vínculo con el hotel, sin que otros aspectos positivos logren compensarlo.

Entre los tópicos funcionales aparece un resultado relevante. T1 (Habitación y espacio físico) presenta una proporción de lealtad (25,98%) superior a la de otros tópicos funcionales como T4 (Movilidad y conexiones, 15,09%) o T6 (Restauración y bebidas, 12,69%), pese a mostrar niveles de *Love Score* similares. Esto sugiere que la experiencia en la habitación, como núcleo de la estancia, tiene un peso mayor en la intención de volver que otros atributos funcionales más secundarios.

La validación sobre la submuestra de reseñas ambiguas (n = 14.004) muestra resultados consistentes. T5 sigue siendo el tópico con mayor *Love Score* (76,61) y mayor proporción de lealtad (23,84%), mientras que T2 continúa siendo uno de los tópicos con menor apego. En esta submuestra, T7 (Problemas de descanso y confort) presenta incluso niveles de *Love Score* más bajos, lo que refuerza su impacto negativo sobre la experiencia del huésped. En conjunto, los resultados entre tópicos se mantienen, lo que aporta robustez a los resultados.

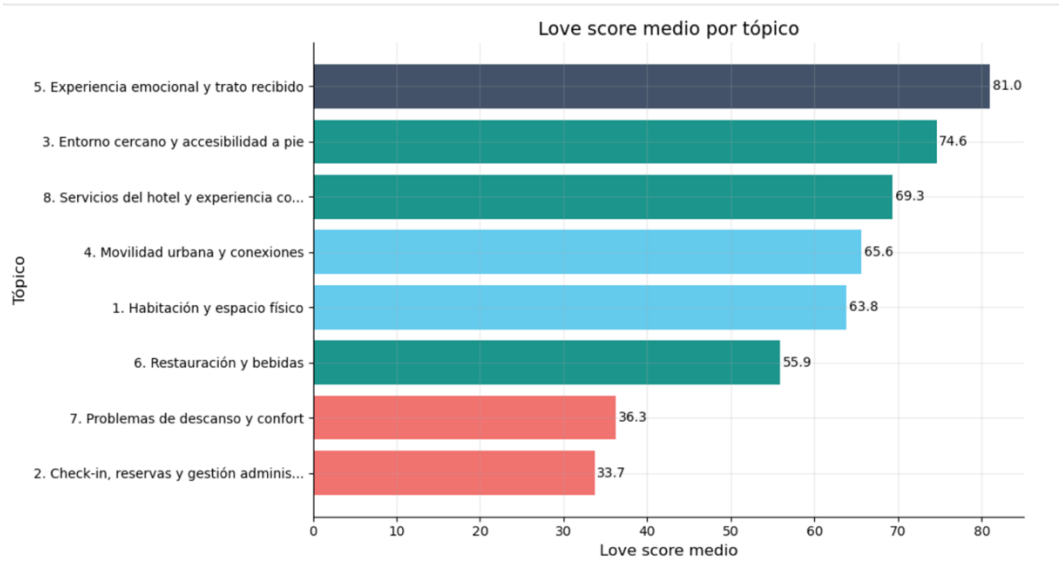


Figura 4. Barras horizontales de *Love Score* medio por tópico

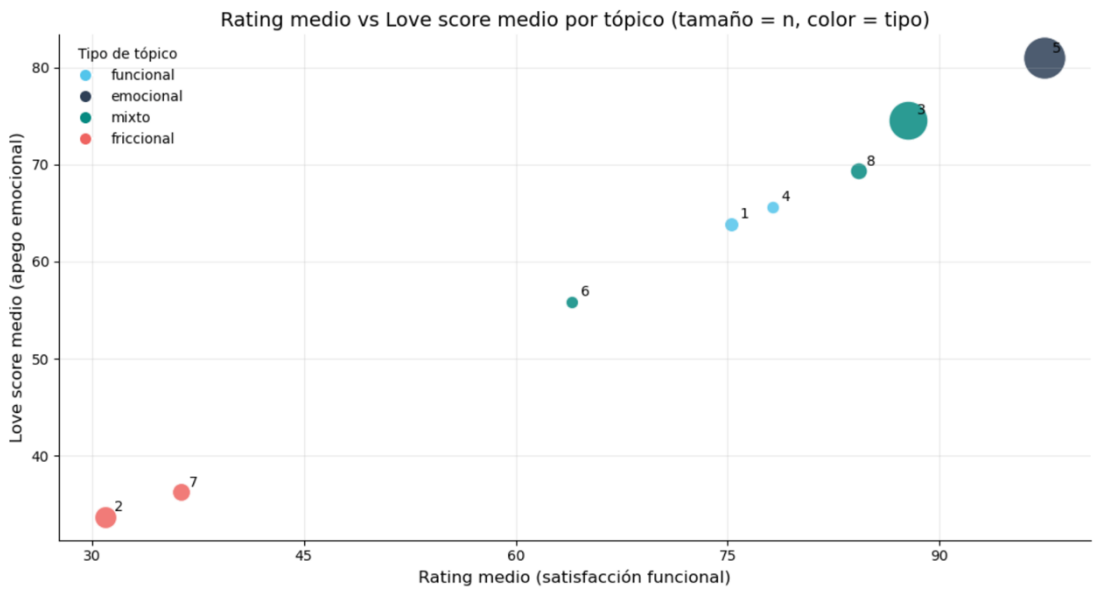


Figura 5. *Rating* medio vs *Love score* medio

La Figura 5 permite analizar de forma conjunta la relación entre satisfacción funcional y apego emocional para cada tipo de experiencia.

En términos generales, se observa una relación positiva: a mayor *rating*, mayor *Love Score*. Sin embargo, los puntos no se alinean perfectamente, lo que indica que ambas dimensiones no evolucionan al mismo ritmo.

Los tópicos emocionales y relacionales (especialmente T5 y T3) se sitúan en la parte superior del gráfico, con niveles elevados de apego. En particular, T5 destaca como el único tópico que combina valores muy altos en ambas dimensiones.

Por el contrario, los tópicos friccionales (T2 y T7) se concentran en la zona inferior, reflejando bajos niveles tanto de satisfacción como de vínculo.

Los tópicos funcionales y mixtos se agrupan en posiciones intermedias, pero con diferencias relevantes en el eje vertical. En especial, el caso de T3, con un nivel de lealtad cercano al de T5, pero con un *rating* inferior, muestra que una buena valoración no siempre se traduce en el mismo nivel de apego emocional.

Por ello, el siguiente apartado se centra en analizar de forma específica esta diferencia entre satisfacción y apego a través del *gap* entre *Love Score* y *rating*.

4.4 Divergencia entre satisfacción funcional y apego emocional

El análisis del *Love Score* por tópico revela no solo diferencias en el nivel de apego emocional, sino también patrones distintos en su relación con la satisfacción funcional.

Para profundizar en esta cuestión, se analiza la diferencia entre el *Love Score* y el *rating* medio en cada tópico. En términos generales, esta diferencia es negativa en la mayoría de los casos, lo que indica que la valoración funcional tiende a situarse por encima del apego emocional.

Sin embargo, esta distancia no es constante y varía de forma significativa entre tipos de experiencia, lo que permite identificar en qué contextos la satisfacción se traduce, o no, en un vínculo más fuerte con el hotel.

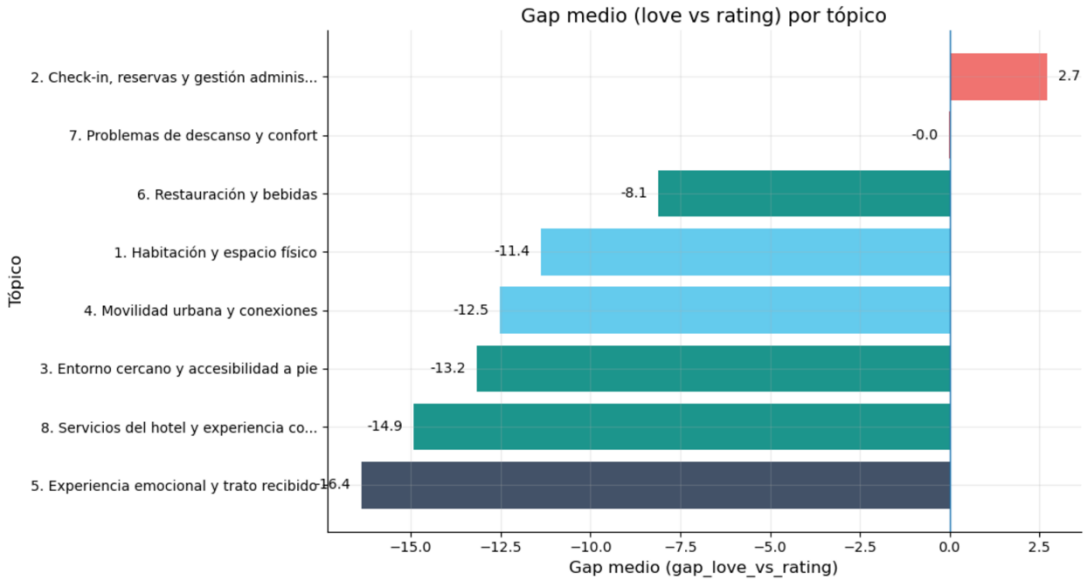


Figura 6. Barras horizontales del gap medio por tópico

El análisis permite identificar tres patrones diferenciados.

Primer patrón - Experiencias de alta satisfacción con gap negativo pronunciado.
T5 (Experiencia emocional y trato recibido) presenta el gap más negativo (−16,38), pese a tener el *rating* más alto (97,38). Este resultado se explica porque, cuando la puntuación está muy cerca del máximo, apenas puede aumentar más, mientras que el *Love Score* todavía tiene margen para variar. Como consecuencia, aunque el apego emocional es muy alto, no alcanza el mismo nivel que la puntuación, lo que hace que la diferencia entre ambas medidas se amplíe.

Segundo patrón - Experiencias friccionales con gap próximo a cero o positivo.
El caso más llamativo es T2 (Check-in y gestión administrativa), que presenta un gap positivo (+2,71), siendo el único tópico donde el *Love Score* supera al *rating*.

A pesar de su baja valoración media (30,98), las reseñas muestran una carga emocional fuerte. Esto puede deberse, por un lado, a descripciones de incidencias con un lenguaje muy expresivo y, por otro, a clientes que, pese a una mala experiencia puntual, mantienen cierta lealtad hacia el hotel.

T7 (Problemas de descanso y confort), en cambio, presenta un gap prácticamente nulo (−0,03), lo que indica que, en este tipo de experiencias, la baja satisfacción va acompañada de un nivel de apego igualmente bajo.

Tercer patrón - Experiencias funcionales y mixtas con gap intermedio.

Los tópicos T1, T3, T4, T6 y T8 presentan gaps negativos de nivel intermedio, es decir, el *Love Score* se sitúa por debajo del *rating*, pero sin diferencias tan extremas como en T5 ni tan próximas a cero como en los tópicos friccionales. En estos casos, la relación entre satisfacción funcional y apego emocional es más estable: cuando la valoración del hotel mejora, también mejora el vínculo con el huésped, y ambas medidas evolucionan de forma bastante alineada.

Esto sugiere que, en este grupo de tópicos, una buena experiencia funcional sí se acompaña de una respuesta emocional positiva, pero sin aumentar especialmente el apego. Dicho de otro modo, son experiencias que satisfacen y pueden gustar, pero no producen el mismo efecto emocional que las experiencias claramente relacionales o especiales.

Dentro de este grupo, T3 (Entorno cercano y accesibilidad) merece una mención especial. Aunque también presenta un gap negativo, este es relativamente elevado en términos absolutos, lo que indica que su componente emocional es más fuerte que en otros tópicos funcionales o mixtos. En la práctica, esto sugiere que el entorno del hotel y su accesibilidad no solo se valoran por utilidad, sino también por la experiencia que hacen posible durante la estancia.

Un elemento adicional que ayuda a interpretar este patrón es la proporción de señales de lealtad. Los tópicos con mayor apego emocional relativo, como T5 y T3, también presentan las tasas más altas de intención de retorno (25,99% y 24,94%), mientras que los tópicos friccionales se sitúan por debajo del 9%. Esta coincidencia entre *Love Score*, *gap* y señales de lealtad refuerza la idea de que el indicador capta de forma razonable el vínculo emocional del huésped con el hotel.

4.5 Asociación del tipo de experiencia con el apego emocional (regresión OLS)

Hasta este punto, el análisis ha mostrado dos ideas principales: por un lado, que los distintos tópicos presentan niveles muy diferentes de apego emocional; y por otro, que también varían en la distancia entre satisfacción funcional y vínculo. Sin embargo, estos resultados son descriptivos. Para estimar en qué medida el tipo de experiencia se asocia con diferencias en el apego emocional, una vez considerada la satisfacción funcional, se utiliza una regresión OLS. Este modelo permite comparar los distintos tipos de experiencia manteniendo constante el nivel de satisfacción, y así observar si algunos de ellos se asocian con mayores niveles de apego emocional.

Es importante tener en cuenta que el rating forma parte del propio cálculo del *Love Score*, por lo que ambas variables están directamente relacionadas. Por ello, los resultados deben interpretarse como asociaciones parciales, y no como efectos causales independientes.

En este contexto, el modelo responde a la siguiente pregunta: ¿qué tipo de experiencia se asocia con un mayor nivel de apego emocional cuando se compara a igualdad de satisfacción funcional?

El modelo se estima sobre 15.994 observaciones de la muestra analítica, utilizando T2 (*Check-in* y gestión administrativa) como categoría de referencia, al ser el tópico con menor *Love Score*.

Tabla 5. Resultados de la regresión OLS

Tópico	Coef.	p-valor	Sig.
T3 — Entorno cercano y accesibilidad a pie	+2,896	<0,001	***
T5 — Experiencia emocional y trato recibido	+2,859	<0,001	***
T1 — Habitación y espacio físico	+0,530	0,091	n.s.
T4 — Movilidad urbana y conexiones	+0,358	0,334	n.s.
T6 — Restauración y bebidas	+0,089	0,811	n.s.
T8 — Servicios complementarios	-0,018	0,947	n.s.
T7 — Problemas de descanso y confort	-0,986	<0,001	***
rating_100 (control)	+0,669	<0,001	***

Ref: T2 — Check-in y gestión administrativa | N=15.994 | R^2 ajustado=0,929 | F=26.283 (p<0,001)
*** p<0,001 ** p<0,01 * p<0,05 n.s. no significativo

El modelo presenta un R^2 ajustado de 0,9293, lo que indica que explica el 92,93% de la variación del *Love Score*. Es decir, la mayor parte de las diferencias en apego emocional

entre reseñas puede explicarse por las variables incluidas en el modelo, principalmente el *rating* y el tipo de experiencia descrita.

No obstante, este elevado nivel de ajuste debe interpretarse con cautela. Dado que el rating forma parte de la construcción del *Love Score*, existe una relación estructural entre ambas variables que contribuye a aumentar la varianza explicada del modelo.

El estadístico F es de 26.283 ($p<0,001$), lo que confirma que el modelo es globalmente significativo. En otras palabras, las variables incluidas, en conjunto, tienen capacidad explicativa sobre el *Love Score* y el modelo no es resultado del azar.

Como era esperable, el *rating_100* es el predictor principal (coeficiente = 0,669, $p<0,001$). Esto significa que, en promedio, por cada punto adicional en la valoración del hotel, el *Love Score* aumenta en 0,669 puntos, manteniendo constante el tipo de experiencia descrita. En otras palabras, una parte importante del apego emocional está directamente asociada al nivel de satisfacción funcional: cuanto mejor se valora el hotel, mayor tiende a ser el vínculo con el huésped.

El resultado más relevante es que solo tres tópicos tienen un efecto estadísticamente significativo sobre el *Love Score* una vez que se tiene en cuenta la satisfacción funcional. Este resultado es interesante por dos motivos:

En primer lugar, T3 (Entorno cercano y accesibilidad a pie) presenta el coeficiente más alto (+2,896), ligeramente por encima de T5 (Experiencia emocional y trato recibido, +2,859). Esto implica que, a igualdad de rating, las reseñas centradas en el entorno generan algo más de apego emocional que las centradas en la experiencia directa con el hotel.

Este resultado no cambia lo observado en el análisis descriptivo, donde T5 presenta el mayor nivel de apego en términos absolutos, pero ayuda a entenderlo mejor. Parte de ese resultado se debe a su *rating* especialmente alto (97,38). Cuando se tiene en cuenta este efecto, T3 pasa a ser el tópico con mayor capacidad de generar apego emocional adicional. Una posible explicación es que el entorno urbano (la ubicación, la accesibilidad o la

proximidad a puntos de interés) influye en cómo el huésped vive el viaje en su conjunto, generando una valoración emocional que va más allá del propio hotel.

En segundo lugar, T7 (Problemas de descanso y confort) presenta un efecto negativo significativo (−0,986, p<0,001). Esto indica que, a igualdad de *rating*, este tipo de experiencias reduce el *Love Score* en comparación con T2 (categoría de referencia). En otras palabras, los problemas relacionados con el descanso (ruido, temperatura o estado de la habitación) tienen un impacto especialmente negativo en el vínculo emocional, incluso mayor que otras incidencias administrativas.

El resto de los tópicos (T1, T4, T6 y T8) no muestran efectos significativos una vez controlado el *rating*. Esto sugiere que dimensiones como la habitación, la movilidad, la restauración o los servicios complementarios influyen en la valoración funcional, pero no generan un efecto adicional sobre el apego emocional. Son aspectos que, en general, cumplen o fallan, pero no marcan la diferencia en términos de vínculo.

En conjunto, la regresión refuerza una idea clave del trabajo: el apego emocional no depende solo del nivel de satisfacción. El tipo de experiencia también influye. Algunas dimensiones, como el entorno o la interacción con el personal, tienen una capacidad real de generar vínculo más allá de la valoración numérica, mientras que otras se limitan a explicar la satisfacción sin aportar un componente emocional adicional.

4.6 Modelo predictivo del apego emocional

Tabla 6. Resultados comparativos de los modelos predictivos

Modelo	AUC-ROC	F1 macro	Accuracy
Random Forest	0.8471	0.7752	0.7768
Gradient Boosting	0.8533	0.7802	0.7821
Red Neuronal (MLP)	0.8502	0.7794	0.7804

Los tres modelos presentan resultados muy similares, con valores de AUC entre 0,847 y 0,8533, lo que indica una buena capacidad para diferenciar entre reseñas con alto y bajo apego emocional. El modelo de Gradient Boosting obtiene el mejor resultado (AUC =

0,853), aunque la diferencia frente al resto es mínima. Esto sugiere que la capacidad predictiva no depende tanto del algoritmo utilizado como de la calidad de las variables incluidas en el modelo.

Para entender mejor qué factores explican esta capacidad clasificatoria, se analizó la importancia de variables en el modelo Random Forest:

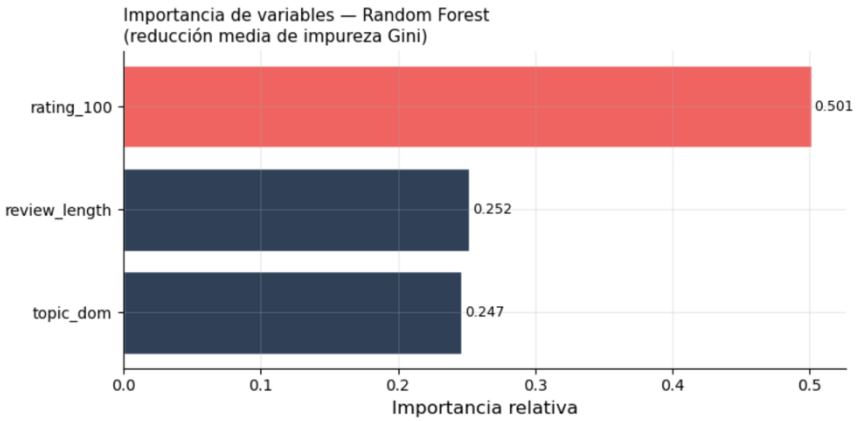


Figura 7. Importancia de variables - Random Forest

El resultado muestra tres contribuciones claramente diferenciadas. El *rating_100* concentra el 50,1% de la importancia total, lo que confirma que la satisfacción funcional es el principal factor asociado al apego emocional en la clasificación de las reseñas.

El resto de la capacidad clasificatoria se reparte entre dos variables: la longitud de la reseña (*review_length*, 25,2%) y el tópico dominante (*topic_dom*, 24,7%). Ambas aportan un peso muy similar, lo que indica que no solo importa cuánto se valora el hotel, sino también cómo se describe la experiencia.

En particular, el peso del tópico dominante resulta relevante. Su contribución confirma que el tipo de experiencia descrita aporta información adicional sobre el apego emocional que no se recoge en la puntuación numérica. Es decir, incluso considerando el nivel de satisfacción, el contenido de la reseña permite diferenciar mejor entre distintos niveles de vinculación emocional.

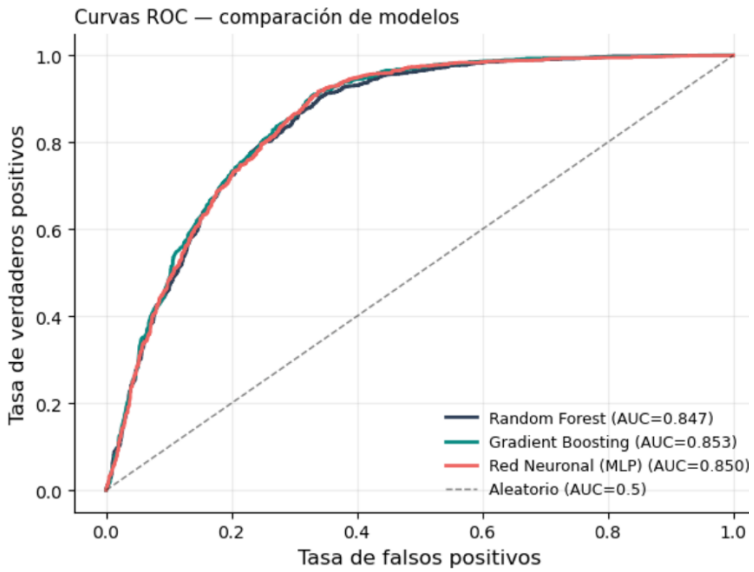


Figura 8. Curvas ROC comparativas de los tres modelos

Las curvas ROC muestran que los tres modelos presentan un comportamiento muy similar a lo largo de todos los umbrales de clasificación. En todos los casos, las curvas se sitúan claramente por encima de la línea aleatoria, lo que confirma una buena capacidad de discriminación. Además, la proximidad entre las curvas indica que el rendimiento es estable y no depende de forma significativa del algoritmo utilizado.

En conjunto, estos resultados refuerzan la evidencia obtenida previamente. La satisfacción funcional es el principal factor asociado al apego emocional, pero no lo explica por completo. El tipo de experiencia descrita añade una dimensión adicional que permite entender mejor por qué algunas estancias generan vínculo y otras no, incluso cuando la valoración es similar.

5. Discusión

5.1 El tipo de experiencia como determinante del apego emocional

Una de las preguntas de fondo de este trabajo es sencilla, pero relevante desde el punto de vista de gestión: ¿por qué algunos hoteles consiguen que el cliente quiera volver y recomendarlos, mientras que otros se quedan en una experiencia correcta sin generar ningún vínculo?

Los resultados muestran una idea clara. La diferencia no está tanto en la categoría del hotel como en el tipo de experiencia que vive el huésped durante la estancia.

La magnitud del efecto es significativa. El *Love Score* medio varía en 47 puntos entre el tópico con mayor apego emocional, T5 (experiencia emocional y trato recibido), y el de menor apego, T2 (check-in y gestión administrativa). En comparación, la diferencia entre segmentos hoteleros dentro del mismo corpus es mucho menor, apenas 4,8 puntos entre hoteles de lujo y convencionales.

Esto muestra que una categoría hotelera más alta no implica automáticamente una mayor capacidad para generar apego emocional. La categoría fija unas expectativas iniciales, pero no determina por sí sola la relación que el huésped acaba desarrollando con el hotel. Ese vínculo depende de la experiencia concreta vivida durante la estancia y del grado en que esta cumple o supera lo que el cliente esperaba.

Esto puede entenderse con un ejemplo sencillo. Un huésped que se aloja en un hotel convencional puede tener una experiencia completamente alineada con sus expectativas: una habitación funcional, un servicio correcto y, sobre todo, un trato cercano o un detalle inesperado. Esa combinación puede generar un nivel de apego emocional similar al de un huésped en un hotel de lujo.

En cambio, un hotel de mayor categoría, aunque de base ofrezca mejores instalaciones o estándares de servicio más altos, no garantiza por sí mismo la creación de ese vínculo. Si la experiencia resulta impersonal, el resultado puede ser una estancia satisfactoria, pero sustituible.

Este resultado encaja con la idea de Pine y Gilmore (1998): en hostelería, la experiencia no acompaña al servicio, es el servicio. Y no todas las experiencias tienen la misma capacidad de generar vínculo.

Las reseñas del tópico T5 así lo reflejan. En estos casos, el huésped no describe tanto el hotel como lo que le ha pasado dentro de él: interacciones con el personal, momentos concretos o detalles que han marcado la estancia. Ese tipo de contenido no es simplemente una evaluación sino un recuerdo. Y ese recuerdo es el que explica por qué algunos clientes vuelven o recomiendan.

La regresión OLS ayuda a afinar esta interpretación. Una vez que se tiene en cuenta el nivel de satisfacción, solo tres tipos de experiencia tienen un efecto propio sobre el apego emocional. T5 (experiencia emocional y trato recibido) y T3 (entorno cercano y accesibilidad) se asocian con niveles de apego superiores a los que cabría esperar solo por su *rating,* mientras que T7 (problemas de descanso y confort) lo reduce.

El resto de las dimensiones, como la habitación, la restauración o los servicios complementarios, no muestran un efecto adicional. Son necesarias para que la experiencia funcione, pero por sí solas no generan apego.

El modelo predictivo refuerza esta misma idea desde otro ángulo. El tipo de experiencia, reflejado en el tópico dominante, explica una parte importante del componente emocional de las reseñas, incluso teniendo en cuenta la valoración numérica. Es decir, no solo está asociado al apego emocional, también ayuda a anticiparlo.

Un resultado especialmente interesante es el papel del entorno. Una vez considerada la satisfacción funcional, T3 muestra una relación ligeramente más fuerte con el apego emocional que T5. Esto sugiere que la ubicación, la accesibilidad o la cercanía a puntos de interés influyen directamente en cómo el huésped vive el viaje en su conjunto. En estos casos, el hotel no se valora solo por lo que ofrece, sino por cómo encaja en la experiencia global del cliente. Una estancia puede no ser excepcional en términos de servicio, pero sí resultar especialmente positiva porque facilita el viaje o lo hace más cómodo.

En conjunto, los resultados apuntan a una idea clara. El apego emocional no depende únicamente de que el hotel funcione bien, sino de que la experiencia tenga significado para el huésped.

Desde una perspectiva de gestión, esto tiene implicaciones directas. La categoría del hotel marca el punto de partida, pero es la experiencia la que determina el resultado. Un hotel bien ubicado, con un trato cercano y capaz de generar momentos que el cliente recuerde, puede tener más capacidad de generar vínculo que un hotel de mayor categoría donde todo funciona correctamente, pero la experiencia resulta impersonal.

5.2 La paradoja del gap: satisfacción sin apego y fricción con emoción

Si hay un resultado que merece pararse a interpretar, es el siguiente: el tópico con mayor valoración funcional (*rating*) es también el que presenta mayor distancia entre satisfacción y apego emocional. Y, en el extremo opuesto, el tópico con peor valoración es el único en el que el apego llega a superar a la satisfacción.
Más que una contradicción, estos dos patrones ayudan a entender mejor cómo se construye el vínculo entre el huésped y el hotel.

La primera paradoja aparece en T5 (experiencia emocional y trato recibido). Es el tópico con el *rating* más alto de todo el análisis, cerca del máximo, y sin embargo presenta el gap más negativo. Es decir, el nivel de apego emocional es alto, pero queda por debajo de la valoración funcional.

Esto tiene una explicación bastante intuitiva. Cuando la valoración está cerca del máximo, apenas puede seguir subiendo. En cambio, el apego emocional depende de algo distinto. No basta con que todo funcione bien, hace falta que el huésped lo exprese en el texto: que mencione momentos concretos, que utilice un lenguaje más emocional o que deje señales claras de lealtad. Y eso no ocurre en todas las reseñas, incluso cuando la experiencia ha sido excelente.

En otras palabras, es más fácil dar una puntuación máxima que expresar un vínculo emocional fuerte. Esta diferencia encaja con la distinción clásica entre satisfacción y

apego: la satisfacción responde a si el servicio ha cumplido, mientras que el apego implica una conexión que no se genera automáticamente.

La segunda paradoja aparece en T2 (check-in y gestión administrativa). Es el tópico peor valorado, pero el único en el que el apego emocional llega a superar la satisfacción. Aquí el patrón es distinto. Por un lado, las incidencias administrativas generan reseñas muy intensas en términos emocionales. Los clientes describen problemas, frustraciones o conflictos con un lenguaje cargado, lo que eleva el tono emocional del texto, aunque la valoración sea muy baja. Esta es una limitación conocida de este tipo de análisis: capta intensidad, pero no siempre distingue bien entre emociones positivas y negativas.

Por otro lado, dentro de este mismo grupo aparecen clientes que, pese a una mala experiencia puntual, mantienen cierto vínculo con el hotel. Son huéspedes que ya han estado antes o que siguen considerando volver, lo que introduce una señal de apego que no encaja del todo con la baja puntuación. Cuando el cliente ya ha tenido buenas experiencias antes, una incidencia puntual no suele ser suficiente para romper el vínculo.

El contraste con T7 (problemas de descanso y confort) es especialmente claro. En este caso, el gap es prácticamente nulo. La baja satisfacción se traduce directamente en bajo apego emocional.

La diferencia es que este tipo de problemas no suele generar una reacción especialmente intensa. El cliente explica lo que ha pasado y, simplemente, decide no volver. No hay enfado explícito ni una reacción exagerada, pero sí una sensación clara de que no repetiría. Y ese tipo de experiencia, aunque sea menos llamativa, suele ser más difícil de corregir.

En conjunto, estos resultados muestran que la relación entre satisfacción y apego no es lineal. En el extremo positivo, la satisfacción alcanza su techo antes que el apego. En el extremo negativo, la reacción del cliente no siempre es la misma: depende del tipo de problema y de la relación previa que tenga con el hotel.

En este contexto, el gap entre ambas medidas es especialmente útil. No solo indica cuánto apego hay, sino cómo se construye ese apego en relación con la satisfacción.

Además, los datos confirman que las experiencias más intensas, tanto positivas como negativas, son las que más influyen en el vínculo con la marca. Las experiencias emocionales muy positivas generan recuerdo y recomendación. Las experiencias negativas intensas generan reacción. En cambio, las experiencias correctas pero neutras apenas dejan huella.

5.3 El *Love Score* como métrica de *engagement*: aportaciones y limitaciones

La construcción del *Love Score* plantea una cuestión: ¿aporta información nueva o simplemente reformula lo que ya mide el *rating*? La respuesta es que aporta valor, aunque con matices.

El principal aporte del indicador es captar una dimensión que el *rating* no recoge bien. La puntuación numérica refleja si la experiencia ha cumplido con lo esperado, pero no permite distinguir entre una valoración correcta y una experiencia que genera vínculo.

Dos reseñas pueden tener la misma puntuación y, sin embargo, reflejar situaciones muy distintas. En una, el cliente describe una estancia sin incidentes. En otra, cuenta una experiencia que le ha marcado, con interacción personal o momentos concretos. El *rating* las sitúa al mismo nivel; el *Love Score* introduce esa diferencia.

La alta correlación entre *rating_100* y *Love Score* es coherente con la construcción del indicador. Sin embargo, la diferencia entre ambos, capturada a través del *gap*, es donde aparece el valor analítico. Esa discrepancia refleja que la valoración funcional (*rating*) y el apego emocional no evolucionan de la misma forma.

Desde la perspectiva del *engagement*, el indicador combina tres elementos relevantes: la evaluación del servicio, el tono emocional del discurso y la presencia de señales explícitas de retorno o recomendación. Esta combinación permite entender mejor la relación del huésped con el hotel más allá de la puntuación.

Además, los resultados siguen un patrón claro. Los tópicos asociados a experiencias emocionales presentan niveles más altos de *Love Score* y mayor presencia de señales de lealtad, mientras que los tópicos friccionales muestran lo contrario. Esta consistencia refuerza la utilidad del indicador para analizar el vínculo emocional.

Ahora bien, el indicador también tiene limitaciones que es importante reconocer.

La primera está en el componente de sentimiento. Al basarse en un diccionario de términos, el análisis detecta la intensidad del lenguaje, pero no siempre interpreta correctamente el contexto. Expresiones con estructura negativa o frases con contraste pueden generar puntuaciones que no reflejan del todo la valoración real del cliente.

Por ejemplo, en una reseña como *"the room was nice, but everything else was a disaster"*, el término "*nice*" aporta carga positiva, aunque la valoración global sea claramente negativa. Es una limitación inherente a este tipo de enfoque.

La segunda limitación es más conceptual. El *Love Score* mide lo que el huésped expresa en la reseña, no su comportamiento posterior. Las señales de lealtad reflejan intención, pero no acción. Un cliente puede declarar que volvería y no hacerlo, o viceversa.

La tercera limitación se debe al diseño del propio indicador. Las ponderaciones asignadas a cada componente responden a una lógica razonada, pero no han sido ajustadas a partir de datos empíricos. Es decir, no se han estimado mediante un modelo estadístico que determine qué peso debería tener cada variable en función de su capacidad real para explicar el apego emocional.

Por ello, aunque la elección de pesos es coherente, otras combinaciones podrían dar resultados ligeramente distintos.

En este contexto, la aportación más sólida del análisis no está tanto en el valor absoluto del *Love Score* como en su comparación con el *rating*. El *gap* permite identificar cuándo una experiencia genera más o menos vínculo del que sugiere la valoración funcional, y resulta especialmente útil para detectar patrones entre tipos de experiencia.

En conjunto, el *Love Score* cumple bien su función dentro del estudio. No es una medida definitiva de apego emocional, pero sí una herramienta útil para complementar el *rating* y analizar cómo se construye el vínculo del huésped a partir de las experiencias descritas.

6. Conclusiones

6.1 Síntesis de hallazgos principales

Este estudio parte de una pregunta sencilla: ¿qué hace que un huésped desarrolle apego emocional hacia un hotel?

Los resultados permiten responder con claridad. El factor más importante no es la categoría del hotel, sino el tipo de experiencia que vive el cliente durante su estancia.

La diferencia es significativa. La brecha de 47 puntos en el *Love Score* entre el tópico de experiencia emocional (T5) y el de gestión administrativa (T2) es muy superior a la que existe entre hoteles de distintas categorías. En términos prácticos, un huésped puede desarrollar más apego emocional en un hotel de menor categoría si la experiencia que vive genera conexión. La categoría importa, pero menos de lo que suele asumirse.

El segundo hallazgo es que la relación entre satisfacción funcional y apego emocional no es lineal. Una valoración alta no garantiza vínculo, y una mala experiencia no siempre lo destruye de la misma forma.

En los niveles altos de satisfacción, el *rating* tiende a saturarse, mientras que el apego emocional depende de la intensidad del recuerdo generado. En los niveles bajos, los efectos dependen del tipo de problema. La fricción administrativa genera respuestas emocionales intensas, mientras que los problemas de descanso tienden a provocar una desconexión más silenciosa, es decir, no se suelen comunicar tanto. El *gap* entre *Love Score* y *rating* es clave para entender estas diferencias.

El tercer hallazgo tiene una implicación especialmente relevante. El entorno cercano y la accesibilidad (T3) se asocian con un nivel de apego emocional incluso superior al de la experiencia relacional directa cuando se tiene en cuenta la satisfacción funcional. En términos absolutos, T5 presenta el mayor *Love Score*. Sin embargo, cuando se compara a igualdad de nivel de satisfacción, T3 muestra una relación ligeramente más fuerte con el apego emocional. Esto indica que la ubicación es una parte central de la experiencia.

Un hotel que permite al huésped moverse con facilidad, explorar el entorno y sentirse integrado en la ciudad genera un vínculo que trasciende el propio servicio.

Por último, el *Love Score* demuestra que combinar la valoración, el tono emocional y las señales de lealtad permite captar aspectos de la experiencia que el *rating* no refleja. No es una medida perfecta, pero sí una herramienta útil para analizar el apego emocional en ausencia de datos de comportamiento real.

En conjunto, los resultados apuntan a una idea clara. En un sector donde la calidad funcional es cada vez más homogénea, la diferencia no está solo en hacerlo bien, sino en generar experiencias que el huésped recuerde y quiera repetir.

6.2 Implicaciones estratégicas para la gestión hotelera

Los resultados de este estudio no se quedan en lo descriptivo. Detrás de cada tópico hay decisiones concretas que los hoteles toman a diario. Este apartado traduce esos resultados en implicaciones prácticas para la gestión hotelera.

Implicación 1 - La dimensión humana es el principal generador de apego

El tópico T5 (experiencia emocional y trato recibido) concentra los niveles más altos de *Love Score* y de lealtad. No está asociado a instalaciones, sino a interacciones: atención personalizada, empatía o capacidad de resolver problemas.

La implicación es clara. El mayor impacto en el vínculo emocional no viene de la inversión en infraestructura, sino de cómo se trata al huésped. Un gesto sencillo, recordar una preferencia, anticiparse a una necesidad o gestionar bien un problema, tiene más efecto que muchas mejoras materiales.

A nivel de gestión, esto implica habilidades relacionales en la formación del personal. La eficiencia operativa es necesaria, pero no suficiente, y la automatización puede mejorar procesos, pero no sustituye la capacidad de generar conexión con el huésped.

Implicación 2 - La fricción administrativa destruye valor desproporcionadamente

El tópico T2 (check-in y gestión administrativa) presenta los peores resultados en apego y lealtad. Los problemas en esta fase no solo generan insatisfacción, sino que deterioran directamente la relación con el cliente. Es uno de los principales puntos de ruptura del vínculo.

La prioridad es eliminar este tipo de fricciones. Sistemas de reserva fiables, procesos ágiles y capacidad de resolución inmediata son críticos.

Además, cuando un problema se gestiona bien, puede reforzar la relación con el cliente. La recuperación del servicio no es solo una solución defensiva, sino una oportunidad real de generar vínculo.

Implicación 3 - La ubicación también genera apego emocional

El tópico T3 (entorno cercano y accesibilidad) es uno de los que más contribuye al apego emocional, incluso más allá del nivel de satisfacción.

Esto indica que la ubicación no es solo un factor logístico. Forma parte de la experiencia. Un hotel bien integrado en su entorno, que facilita al huésped moverse, explorar y sentirse parte del lugar, genera un vínculo que va más allá del propio servicio.

Para hoteles ya existentes, esto abre una vía clara. No se trata solo de dónde está el hotel, sino de cómo se activa ese entorno: recomendaciones locales, experiencias cercanas o conexión con el barrio.

Implicación 4 - Los problemas de descanso erosionan el vínculo

El tópico T7 (descanso y confort) tiene un efecto negativo claro sobre el apego emocional. A diferencia de otros problemas, no suele generar quejas intensas. El cliente simplemente tiene una mala experiencia y decide no volver.

Esto lo hace especialmente peligroso ya que no genera señales visibles, pero sí pérdida de clientes.

La implicación es priorizar el mantenimiento y las condiciones básicas de la estancia: ruido, temperatura, calidad del descanso. Son factores poco visibles, pero determinantes.

Implicación 5 - Medir solo la satisfacción es insuficiente

Las métricas habituales, como *rating* o NPS, capturan la valoración del servicio, pero no el vínculo que genera. Dos hoteles con la misma puntuación pueden tener niveles muy distintos de apego emocional.

Por eso, incorporar métricas que capten esta dimensión permite entender mejor qué experiencias generan retorno y recomendación.

El *Love Score*, con sus limitaciones, apunta en esta dirección. No sustituye al *rating*, pero lo complementa y permite detectar diferencias que de otro modo no serían visibles.

En conjunto, los resultados sugieren una idea clara. La ventaja competitiva no está solo en ofrecer un buen servicio, sino en generar experiencias que conecten con el huésped. En un contexto en el que la calidad funcional es cada vez más homogénea, la capacidad de crear vínculos emocionales se convierte en un factor diferencial clave.

6.3 Limitaciones del estudio

Toda investigación empírica tiene límites, y reconocerlos es tan importante como interpretar correctamente los resultados. Este estudio presenta cuatro limitaciones que conviene tener en cuenta.

La primera es el sesgo de autoselección propio de las reseñas voluntarias. TripAdvisor no recoge una muestra aleatoria de huéspedes, sino de aquellos que deciden escribir una reseña. La evidencia muestra que estos usuarios suelen tener experiencias más extremas, tanto positivas como negativas. Esto explica el sesgo positivo del dataset (*rating* medio de 80,41 sobre 100) y sugiere que las experiencias más neutras están mucho menos representadas. Por tanto, los resultados describen bien a los usuarios que escriben reseñas, pero no necesariamente al conjunto total de clientes.

La segunda es la representatividad geográfica. El dataset cubre tres ciudades: Londres, Nueva York y Shanghái que, pese a su relevancia internacional, representan mercados

hoteleros urbanos de alta densidad competitiva. Los patrones de apego emocional pueden diferir sustancialmente en mercados de turismo vacacional, en ciudades de menor tamaño o en contextos culturales distintos. La generalización de los hallazgos a otros mercados debe hacerse con cautela.

La tercera es la ausencia de datos longitudinales. El *Love Score* captura apego emocional declarado en el momento de escribir la reseña, pero no permite verificar si ese apego se tradujo en comportamiento real de retorno. La diferencia entre intención declarada y conducta real es uno de los problemas de la investigación en comportamiento del consumidor, y este estudio no puede resolverlo con los datos disponibles. Una extensión natural del trabajo requeriría datos de repetición de estancia vinculados a las reseñas.

La cuarta es el corte temporal del dataset. Las reseñas tienen una distribución temporal que no ha sido analizada en este estudio, y el mercado hotelero es sensible a disrupciones externas, crisis económicas, cambios en los patrones de viaje, eventos globales, que pueden alterar la relación entre tipos de experiencia y apego emocional. Los patrones identificados describen el corpus tal como está, sin controlar por efectos temporales que podrían matizar los resultados.

6.4 Líneas de investigación futura

Los hallazgos de este estudio abren varias direcciones de investigación que, de desarrollarse, permitirían avanzar de forma significativa en la comprensión del apego emocional en la industria hotelera.

La más inmediata es la mejora del componente de análisis de sentimiento. Sustituir el diccionario actual por modelos de lenguaje más avanzados, como RoBERTa o DeBERTa, permitiría captar mejor el contexto, la negación o la ironía. Esto solucionaría limitaciones detectadas en el análisis, como el caso de T2, donde la intensidad del lenguaje puede distorsionar la interpretación emocional. Esta mejora no requiere cambiar el *Love Score*, sino simplemente refinar uno de sus componentes.

La segunda línea es la validación longitudinal del indicador. Vincular el *Love Score* de una reseña con el comportamiento posterior del huésped, ¿volvió al hotel? ¿en qué plazo? ¿cuántas veces? permitiría validar si el indicador predice la repetición de estancia. Esta validación requeriría acceso a datos de reserva vinculados a perfiles de usuario. Para ello, sería necesario disponer de datos de reservas asociados a los usuarios, algo que las plataformas del sector tienen, pero que rara vez está disponible para investigación.

La tercera es la extensión a otros sectores de servicios. Los patrones identificados en este estudio, como el peso del tipo de experiencia o el impacto de la fricción, podrían aplicarse a otros servicios donde el cliente describe su experiencia, como restauración o aerolíneas. Aplicar el mismo marco analítico a otros corpus de reseñas permitiría comparar patrones entre sectores y consolidar una teoría más general del apego emocional basada en tipos de experiencia.

La cuarta línea es el análisis temporal del apego emocional. ¿Cómo evoluciona la relación entre satisfacción funcional y apego emocional a lo largo del tiempo en un mismo hotel? ¿Hay hoteles que construyen apego de forma consistente año tras año, y otros que lo erosionan gradualmente? El dataset de este estudio contiene información temporal que no ha sido explotada, y su análisis podría revelar como se construyen las *love brands* a lo largo del ciclo de vida de un establecimiento.

Finalmente, una línea especialmente relevante para el contexto actual es el análisis del impacto de la automatización sobre el apego emocional. El sector está incorporando cada vez más tecnología en la interacción con el cliente. El sector hotelero está incorporando de forma acelerada tecnologías de *check-in* automático y asistentes virtuales. Los datos de este estudio sugieren que la dimensión humana es el principal generador de apego emocional, lo que plantea una pregunta: ¿puede la inteligencia artificial replicar el tipo de conexión que genera T5, o la automatización de las interacciones destruirá el factor que más apego genera? Analizar este equilibrio será fundamental en los próximos años.

7. Bibliografía

Aaker, J. L. (1997). Dimensions of Brand Personality. Journal of Marketing Research, 34(3), 347–356. https://doi.org/10.2307/3151897

Albert, N., & Merunka, D. (2013). The role of brand love in consumer-brand relationships. Journal of Consumer Marketing, 30(3), 258–266. https://doi.org/10.1108/07363761311328928

Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993–1022.

Carroll, B. A., & Ahuvia, A. C. (2006). Some antecedents and outcomes of brand love. Marketing Letters, 17(2), 79–89. https://doi.org/10.1007/s11002-006-4219-2

Mattila, A. S. (2006). How Affective Commitment Boosts Guest Loyalty (and Promotes Frequent-guest Programs). Cornell Hotel and Restaurant Administration Quarterly, 47(2), 174–181. https://doi.org/10.1177/0010880405283943

Meyer, C., & Schwager, A. (2007). Understanding customer experience. *Harvard Business Review*, 85(2), 116–126.

Mudambi, S., & Schuff, D. (2010). What Makes a Helpful Online Review? a Study of Customer Reviews on Amazon.com. MIS Quarterly, 34(1), 185–200. https://doi.org/10.2307/20721420

Oliver, R. L. (1999). Whence Consumer Loyalty? Journal of Marketing, 63(4), 33–44.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40.

Pine, B. J., & Gilmore, J. H. (1998). Welcome to the experience economy. *Harvard Business Review*, 76(4), 97–105.

Schmitt, B. H., Brakus, J. J., & Zarantonello, L. (2009). Brand experience: What is it? How is it measured? Does it affect loyalty? *Journal of Marketing*, 73(3), 52–68. https://doi.org/10.1509/jmkg.73.3.052

Van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P. C. (2010). Customer engagement behavior: Theoretical foundations and research directions. Journal of Service Research, 13(3), 253–266. https://doi.org/10.1177/1094670510375599

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer Experience Creation: Determinants, Dynamics and Management Strategies. Journal of Retailing, 85(1), 31–41. https://doi.org/10.1016/j.jretai.2008.11.001

Xuhui, Z. (2024). TripAdvisor Hotel Review Dataset. ScienceDB. https://doi.org/10.57760/sciencedb.j00133.00327

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The Behavioral Consequences of Service Quality. Journal of Marketing, 60(2), 31–46.

8. Anexos

Anexo A. Nubes de palabras por tópico



Anexo B. Diccionario de sentimiento (términos positivos y negativos)

	Términos positivos (30)	Términos negativos (27)
0	amazing	awful
1	central	bad
2	clean	bedbugs
3	comfortable	bugs
4	convenient	cramped
5	cozy	dirty
6	delicious	disappointed
7	enjoyed	disappointing
8	excellent	disappointment
9	fantastic	filthy
10	favorite	horrible
11	favourite	loud
12	fresh	mold
13	friendly	mould
14	great	noisy
15	helpful	poor
16	highly	rude
17	love	small
18	loved	smelly
19	outstanding	stained
20	perfect	terrible
21	professional	tiny
22	recommend	uncomfortable
23	recommended	unfriendly
24	safe	unhelpful
25	satisfied	unprofessional
26	spacious	worst
27	superb	
28	walkable	
29	wonderful	

Anexo C. Lista completa de expresiones *kw_loyal_strict*

	Expresiones de lealtad (19)
0	come back
1	definitely come back
2	definitely recommend
3	highly recommend
4	i recommend this hotel
5	my favorite hotel
6	my favourite hotel
7	our favorite hotel
8	our favourite hotel
9	stay again
10	stay here again
11	we recommend this hotel
12	we will be back
13	we will stay again
14	we'll be back
15	will stay again
16	would definitely stay again
17	would recommend
18	would stay again

Anexo D. Código Python del análisis

Se incluye a continuación una selección del código principal utilizado en el análisis. El anexo recoge los bloques que sustentan directamente los resultados presentados en el trabajo, omitiendo fragmentos exploratorios o de validación interna que no alteran las conclusiones.

```
# ============================================================================
# ANEXO D — CÓDIGO PYTHON DEL ANÁLISIS
# Entorno: Python
# Librerías principales: pandas, numpy, spaCy, gensim, scikit-learn,
# statsmodels y matplotlib
# Módulos auxiliares: re, warnings
# Recurso lingüístico utilizado: en_core_web_sm (spaCy)
# ============================================================================


# ============================================================================
# FASE 1 — PREPARACIÓN DEL DATASET Y CONSTRUCCIÓN DE VARIABLES ANALÍTICAS
# ============================================================================

import pandas as pd
import numpy as np
import re

# ----------------------------------------------------------------------------
# Bloque 1.1 — Carga del dataset
# ----------------------------------------------------------------------------

# Dataset original: 416.032 reseñas de TripAdvisor procedentes de hoteles
# en Londres, Nueva York y Shanghái, publicado en Science Data Bank (SciDB).

df = pd.read_excel("hoteles_reviews.xlsx")


# ----------------------------------------------------------------------------
# Bloque 1.2 — Limpieza estructural
# ----------------------------------------------------------------------------

# Se eliminan columnas vacías generadas automáticamente por Excel,
# columnas duplicadas y se normalizan los nombres de todas las variables
# a minúsculas con guiones bajos para garantizar consistencia.

unnamed_cols = [c for c in df.columns if str(c).startswith("Unnamed")]
if unnamed_cols:
    df = df.drop(columns=unnamed_cols, errors="ignore")

df = df.loc[:, ~df.columns.duplicated()]

df.columns = (
    df.columns.astype(str)
    .str.strip()
    .str.lower()
    .str.replace(r"\s+", "_", regex=True)
    .str.replace(r"__+", "_", regex=True)
)

# ----------------------------------------------------------------------------
# Bloque 1.3 — Formato de fechas y normalización del rating
# ----------------------------------------------------------------------------

# Conversión de variables temporales a formato datetime.
```

```
for col in ["date", "date_of_stay"]:
    if col in df.columns:
        df[col] = pd.to_datetime(df[col], format="mixed", errors="coerce")

# Normalización del rating a escala 1–5.
# Algunos valores aparecían expresados en múltiplos de 10 (ej. 40 en lugar
# de 4), probablemente por un error en el proceso de web scraping.
# Se detectan y corrigen dividiendo entre 10 cuando el máximo supera 5.

if "rating" in df.columns:
    df["rating"] = pd.to_numeric(df["rating"], errors="coerce")
    rating_max = df["rating"].max(skipna=True)
    if pd.notna(rating_max) and rating_max > 5:
        df["rating"] = df["rating"] / 10


# ----------------------------------------------------------------------------
# Bloque 1.4 — Segmentación hotelera
# ----------------------------------------------------------------------------

# Clasificación de hoteles en tres segmentos de posicionamiento:
# Luxury / Premium / Conventional, a partir del nombre del hotel.
# Esta variable se utiliza como referencia contextual en el análisis
comparativo, no como eje principal del estudio.

luxury_brands = [
    "ritz-carlton", "ritz", "st. regis", "st regis", "edition",
    "jw marriott", "park hyatt", "grand hyatt",
    "intercontinental", "kimpton",
    "sofitel", "gran melia", "me by melia"
]

premium_brands = [
    "marriott", "courtyard", "residence inn", "fairfield", "ac hotel",
    "sheraton", "westin", "le meridien",
    "hyatt regency", "hyatt place",
    "pullman", "novotel",
    "melia", "innside",
    "voco", "hotel indigo",
    "z hotel", "citizenm", "premier inn", "staycity"
]

conventional_brands = [
    "ibis", "ibis styles", "moxy",
    "holiday inn express",
    "la quinta", "ramada", "best western",
    "point a"
]

def classify_segment(hotel_name):
    n = "" if pd.isna(hotel_name) else str(hotel_name).lower()
    # Prioridad: Luxury > Premium > Conventional
    if any(b in n for b in luxury_brands):
        return "Luxury"
    if any(b in n for b in premium_brands):
        return "Premium"
    if any(b in n for b in conventional_brands):
        return "Conventional"
    return np.nan

df["segment"] = df["hotel_name"].apply(classify_segment)

# Submuestra con segmentación disponible (para análisis comparativo)
df_seg = df[df["segment"].notna()].copy()
```

```
# ------------------------------------------------------------------------------
# Bloque 1.5 — Análisis de sentimiento (sent_norm)
# ------------------------------------------------------------------------------

# Indicador de tono emocional basado en un diccionario léxico propio
# de 30 términos positivos y 27 negativos, seleccionados por frecuencia
# y relevancia semántica en el sector hotelero.
#
# La puntuación neta se calcula como:
#   sentiment = tanh((pos_hits - neg_hits) / sqrt(len(tokens)))
#
# El uso de tanh evita que reseñas más largas obtengan puntuaciones
# extremas por acumulación de términos. El resultado se reescala
# al intervalo [0, 1] mediante normalización min-max.
#
# Limitación reconocida: el enfoque léxico no capta bien la negación
# ni la ironía. Se documenta en el apartado de limitaciones del TFG.

positive_words = {
    "excellent", "great", "amazing", "wonderful", "fantastic", "perfect",
    "superb", "outstanding", "comfortable", "clean", "spacious", "cozy",
    "friendly", "helpful", "professional", "convenient", "central",
    "walkable", "safe", "delicious", "fresh", "enjoyed", "satisfied",
    "recommend", "recommended", "highly", "love", "loved", "favorite",
"favourite"
}

negative_words = {
    "bad", "terrible", "awful", "horrible", "poor", "worst",
    "dirty", "filthy", "smelly", "stained", "mould", "mold", "bugs", "bedbugs",
    "noisy", "loud", "uncomfortable", "small", "tiny", "cramped",
    "rude", "unfriendly", "unhelpful", "unprofessional",
    "disappointing", "disappointed", "disappointment"
}

pos_set, neg_set = set(positive_words), set(negative_words)

def sentiment_score(text):
    tokens = re.findall(r"\w+", str(text).lower())
    if len(tokens) == 0:
        return 0.0
    pos = sum(1 for t in tokens if t in pos_set)
    neg = sum(1 for t in tokens if t in neg_set)
    score_raw = (pos - neg) / np.sqrt(len(tokens))
    return float(np.tanh(score_raw))

df["sentiment"] = df["review"].apply(sentiment_score)

# Reescalado min-max al intervalo [0, 1]

smin = df["sentiment"].min(skipna=True)
smax = df["sentiment"].max(skipna=True)

df["sent_norm"] = (
    0.5 if (pd.isna(smin) or pd.isna(smax) or smax == smin)
    else (df["sentiment"] - smin) / (smax - smin)
)
df["sent_norm"] = df["sent_norm"].clip(0, 1)


# ------------------------------------------------------------------------------
# Bloque 1.6 — Señales de lealtad (kw_loyal_strict)
# ------------------------------------------------------------------------------

# Variable binaria que detecta expresiones explícitas de intención de retorno
# o recomendación. Solo se incluyen frases inequívocas; términos sueltos
# como "return" se excluyen deliberadamente para evitar falsos positivos.
#
```

```
# Fundamentación: captura la fase conativa de la lealtad según Oliver (1999),
# donde el huésped ha decidido volver o recomendar el hotel.

loyal_phrases_strict = [
    "stay again", "stay here again", "would stay again",
    "would definitely stay again", "will stay again",
    "we will stay again", "come back", "we will be back",
    "we'll be back", "definitely come back",
    "highly recommend", "would recommend",
    "definitely recommend", "i recommend this hotel",
    "we recommend this hotel",
    "my favourite hotel", "my favorite hotel",
    "our favourite hotel", "our favorite hotel"
]

pattern = re.compile(
    "|".join(re.escape(p) for p in loyal_phrases_strict),
    flags=re.IGNORECASE
)

df["kw_loyal_strict"] = df["review"].apply(
    lambda t: 1 if pattern.search(str(t)) else 0
).astype(int)


# ------------------------------------------------------------------------------
# Bloque 1.7 — Construcción del Love Score
# ------------------------------------------------------------------------------

# Indicador compuesto de apego emocional declarado.
#
# Fórmula:
#   Love Score = 100 * (0,6 * rating_norm + 0,3 * sent_norm + 0,1 *
kw_loyal_strict)


df["rating_norm"] = ((df["rating"] - 1) / 4).clip(0, 1)
df["rating_100"]  = 100 * df["rating_norm"]

df["love_score"] = 100 * (
    0.6 * df["rating_norm"] +
    0.3 * df["sent_norm"] +
    0.1 * df["kw_loyal_strict"]
)

# Componente emocional puro (sin rating): usado en el modelo predictivo
# para evitar reconstruir el propio indicador como variable dependiente

df["love_nonrating"] = 100 * (
    0.3 * df["sent_norm"] +
    0.1 * df["kw_loyal_strict"]
)

# Diferencia entre apego emocional y satisfacción funcional

df["gap_love_vs_rating"] = df["love_score"] - df["rating_100"]


# ==============================================================================
# FASE 2 — ANÁLISIS TEMÁTICO MEDIANTE LDA
# ==============================================================================

import spacy


# ------------------------------------------------------------------------------
# Bloque 2.1 — Preprocesamiento textual con spaCy
# ------------------------------------------------------------------------------
```

```
# Carga del modelo lingüístico en inglés.
# Se desactivan parser, NER y clasificador porque solo se necesita lematización:
reduce cada término a su forma canónica

nlp = spacy.load("en_core_web_sm", disable=["parser", "ner", "textcat"])

def preprocess_doc(doc):
    tokens_clean = []
    for token in doc:
        # Filtrar stopwords, puntuación, espacios y números
        if token.is_stop or token.is_punct or token.is_space or token.like_num:
            continue
        lemma = token.lemma_.strip().lower()
        # Descartar tokens muy cortos (< 3 caracteres)
        if len(lemma) < 3:
            continue
        tokens_clean.append(lemma)
    return tokens_clean

# Preparación del corpus: imputación de nulos y exclusión de reseñas con menos
de 30 caracteres, que no aportan contenido temático relevante

df_work = df.copy()
df_work["review"] = df_work["review"].fillna("").astype(str)
df_work = df_work[df_work["review"].str.len() >= 30].copy()

texts = df_work["review"].tolist()
tokens_list = []

# Procesamiento en lotes de 1.000 para eficiencia computacional
for doc in nlp.pipe(texts, batch_size=1000, n_process=1):
    tokens_list.append(preprocess_doc(doc))

df_work["tokens"] = tokens_list


# ------------------------------------------------------------------------------
# Bloque 2.2 — Limpieza semántica y construcción de n-gramas
# ------------------------------------------------------------------------------

from gensim.models.phrases import Phrases, Phraser

# Limpieza semántica posterior a spaCy en dos niveles:
#
# 1) Stopwords de dominio hotelero
#
# 2) Referencias geográficas. El interés del estudio es qué tipo de experiencia
#   se describe, no dónde ocurre.

REMOVE_GEO = True

DOMAIN_STOPWORDS = {
    "hotel", "room", "stay", "night", "bed", "bathroom", "staff",
    "place", "property", "guest", "time", "day"
}

GEO_STOPWORDS = {
    "london", "nyc", "new_york", "manhattan", "times_square", "central_park",
    "shanghai", "paris", "rome", "madrid", "barcelona", "berlin",
    "amsterdam", "tokyo", "singapore", "dubai", "boston", "chicago",
"los_angeles"
}

def clean_tokens_postspacy(tokens):
    return [
        t for t in (tokens or [])
        if t and len(t) >= 3
```

```
        and t not in DOMAIN_STOPWORDS
        and not (REMOVE_GEO and t in GEO_STOPWORDS)
    ]

df_work["tokens_clean"] = df_work["tokens"].apply(clean_tokens_postspacy)

# Construcción de bigramas y trigramas con Gensim Phrases.
# Permite tratar expresiones compuestas como unidades léxicas únicas
# ("customer_service", "hot_water"), mejorando la interpretabilidad del LDA.
# min_count=20: la expresión debe aparecer en al menos 20 documentos.

MIN_COUNT, THRESHOLD = 20, 10

bigram = Phrases(
    df_work["tokens_clean"].tolist(),
    min_count=MIN_COUNT,
    threshold=THRESHOLD
)

trigram = Phrases(
    Phraser(bigram)[df_work["tokens_clean"].tolist()],
    min_count=MIN_COUNT,
    threshold=THRESHOLD
)

bigram_phraser  = Phraser(bigram)
trigram_phraser = Phraser(trigram)

df_work["tokens_final"] = [
    trigram_phraser[bigram_phraser[t]]
    for t in df_work["tokens_clean"].tolist()
]


# ------------------------------------------------------------------------------
# Bloque 2.3 — Muestra de entrenamiento y construcción del corpus
# ------------------------------------------------------------------------------

import gensim
from gensim import corpora
from scipy.stats import ks_2samp

# Muestra aleatoria de 30.000 reseñas para entrenar el LDA
# La semilla fija garantiza la reproducibilidad del análisis.

df_sample = df_work.sample(
    n=min(30000, df_work.shape[0]),
    random_state=42
).copy().reset_index(drop=True)


# La representatividad de la muestra validada mediante test de Kolmogorov-
Smirnov

# Se compara la distribución de rating, sent_norm y love_score

ks_results = []

for var in ["rating", "sent_norm", "love_score"]:
    x_full = df_work[var].dropna()
    x_sample = df_sample[var].dropna()
    ks_stat, p_value = ks_2samp(x_full, x_sample)
    ks_results.append({
        "variable": var,
        "ks_stat": ks_stat,
        "p_value": p_value
    })
```

```
ks_df = pd.DataFrame(ks_results)
print(ks_df)

# Construcción del diccionario y corpus BoW (Bag of Words).
# Filtros de frecuencia aplicados:
#   no_below=50:   elimina términos presentes en menos de 50 documentos
#   no_above=0.20: elimina términos presentes en más del 20% de documentos
# El diccionario resultante contiene 2.595 términos.

dictionary = corpora.Dictionary(df_sample["tokens_final"].tolist())
dictionary.filter_extremes(no_below=50, no_above=0.20)

df_sample["bow"] = df_sample["tokens_final"].apply(dictionary.doc2bow)
corpus = df_sample["bow"].tolist()


# ------------------------------------------------------------------------------
# Bloque 2.4 — Selección del número de tópicos
# ------------------------------------------------------------------------------

from gensim.models.coherencemodel import CoherenceModel

# Selección del número óptimo de tópicos k mediante coherencia semántica c_v.
# La métrica c_v evalúa si las palabras de un tópico aparecen juntas
# con frecuencia en el corpus: mayor coherencia = tópico más interpretable.
# Se evalúan valores de k entre 6 y 12.

topic_grid = [6, 7, 8, 9, 10, 11, 12]
coherence_results = []

for k in topic_grid:
    lda_tmp = gensim.models.LdaModel(
        corpus=corpus,
        id2word=dictionary,
        num_topics=k,
        random_state=42,
        chunksize=2000,
        passes=10,
        alpha="auto"
    )
    cm = CoherenceModel(
        model=lda_tmp,
        texts=df_sample["tokens_final"].tolist(),
        dictionary=dictionary,
        coherence="c_v"
    )
    coherence_results.append({
        "topics": k,
        "coherence_cv": cm.get_coherence()
    })

coh_df =
pd.DataFrame(coherence_results).sort_values("topics").reset_index(drop=True)

n_topics_opt = int(coh_df.loc[coh_df["coherence_cv"].idxmax(), "topics"])


# ------------------------------------------------------------------------------
# Bloque 2.5 — Entrenamiento del modelo LDA definitivo
# ------------------------------------------------------------------------------

# Parámetros:
#   chunksize=2000: documentos procesados por iteración
#   passes=10:      pasadas completas sobre el corpus
#   alpha="auto":   aprende la distribución de tópicos por documento

lda_model = gensim.models.LdaModel(
    corpus=corpus,
```

```
    id2word=dictionary,
    num_topics=n_topics_opt,
    random_state=42,
    chunksize=2000,
    passes=10,
    alpha="auto"
)

# Visualización de las 12 palabras con mayor peso por tópico

TOPN = 12
for t in range(n_topics_opt):
    topic_terms = lda_model.show_topic(t, topn=TOPN)
    words = [w for w, _ in topic_terms]
    print(f"Tópico {t + 1}: {', '.join(words)}")


# ------------------------------------------------------------------------------
# Bloque 2.6 — Asignación de tópicos y control de ambigüedad
# ------------------------------------------------------------------------------

def get_top2_topics(bow, model):
    topic_probs = sorted(
        model.get_document_topics(bow, minimum_probability=0.0),
        key=lambda x: x[1],
        reverse=True
    )
    (t1, p1), (t2, p2) = topic_probs[0], topic_probs[1]
    return (t1 + 1, float(p1)), (t2 + 1, float(p2))

top2 = [get_top2_topics(bow, lda_model) for bow in df_sample["bow"].tolist()]

df_sample["topic_dom"]    = [t1 for ((t1, _), _) in top2]
df_sample["topic_prob"]   = [p1 for ((_, p1), _) in top2]
df_sample["topic_2"]      = [t2 for (_, (t2, _)) in top2]
df_sample["topic_margin"] = df_sample["topic_prob"] - [p2 for (_, (_, p2)) in
top2]

# Control de ambigüedad temática.
# Una reseña se considera ambigua cuando su probabilidad dominante o su
# margen respecto al segundo tópico quedan por debajo del percentil 40.

p_prob   = float(df_sample["topic_prob"].quantile(0.40))
p_margin = float(df_sample["topic_margin"].quantile(0.40))

df_sample["topic_ambiguous"] = (
    (df_sample["topic_prob"] < p_prob) |
    (df_sample["topic_margin"] < p_margin)
)

# Nombres definitivos de los 8 tópicos identificados

topic_names = {
    1: "Habitación y espacio físico",
    2: "Check-in, reservas y gestión administrativa",
    3: "Entorno cercano y accesibilidad a pie",
    4: "Movilidad urbana y conexiones",
    5: "Experiencia emocional y trato recibido",
    6: "Restauración y bebidas",
    7: "Problemas de descanso y confort",
    8: "Servicios del hotel y experiencia complementaria"
}

df_sample["topic_name"] = df_sample["topic_dom"].map(topic_names)


# ==============================================================================
# FASE 3 — LOVE SCORE POR TÓPICO Y REGRESIÓN OLS
```

```
# ============================================================================

# ----------------------------------------------------------------------------
# Bloque 3.1 — Métricas por tópico (muestra analítica principal)
# ----------------------------------------------------------------------------

# Muestra analítica principal: solo reseñas con dominancia temática clara
# (topic_ambiguous=False

df_final = df_sample[~df_sample["topic_ambiguous"]].copy()

# Clasificación conceptual de los 8 tópicos en 4 categorías

topic_type_final = {
    1: "funcional",    # Habitación y espacio físico
    2: "friccional",   # Check-in y gestión administrativa
    3: "mixto",        # Entorno cercano y accesibilidad
    4: "funcional",    # Movilidad urbana
    5: "emocional",    # Experiencia emocional y trato recibido
    6: "mixto",        # Restauración y bebidas
    7: "friccional",   # Problemas de descanso y confort
    8: "mixto"         # Servicios complementarios
}

df_final["topic_type"] = df_final["topic_dom"].map(topic_type_final)

# Tabla de métricas por tópico: rating medio, Love Score medio,
# gap (love - rating) y porcentaje de señales de lealtad

topic_table = (
    df_final.groupby("topic_dom")
    .agg(
        n=("topic_dom", "size"),
        rating_mean=("rating_100", "mean"),
        love_mean=("love_score", "mean"),
        gap_mean=("gap_love_vs_rating", "mean"),
        loyal_pct=("kw_loyal_strict", "mean")
    )
    .reset_index()
)

topic_table["topic_name"] = topic_table["topic_dom"].map(topic_names)
topic_table["topic_type"] = topic_table["topic_dom"].map(topic_type_final)
topic_table["loyal_pct"]  = topic_table["loyal_pct"] * 100


# ----------------------------------------------------------------------------
# Bloque 3.2 — Regresión OLS: asociación condicionada del tópico
# ----------------------------------------------------------------------------

import statsmodels.formula.api as smf
import warnings
warnings.filterwarnings("ignore")


# Variable dependiente: love_score
# Variable de control:  rating_100
# Variables de interés: dummies de tópico (referencia = T2)


df_ols = df_final.copy()
df_ols["topic_cat"] = pd.Categorical(
    df_ols["topic_dom"].astype(int),
    categories=[2, 1, 3, 4, 5, 6, 7, 8]   # T2 primero = categoría de referencia
)

formula  = "love_score ~ C(topic_cat) + rating_100"
resultado = smf.ols(
```

```
    formula=formula,
    data=df_ols.dropna(subset=["love_score", "rating_100", "topic_cat"])
).fit()

print(
    f"N={int(resultado.nobs):,} | "
    f"R² ajustado={resultado.rsquared_adj:.4f} | "
    f"F={resultado.fvalue:.2f} (p={resultado.f_pvalue:.4f})"
)
print(resultado.summary())


# ============================================================================
# FASE 4 — MODELO PREDICTIVO DEL APEGO EMOCIONAL
# ============================================================================

# ----------------------------------------------------------------------------
# Bloque 4.1 — Preparación de datos para clasificación
# ----------------------------------------------------------------------------

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler

# Variable dependiente: love_nonrating (componente emocional puro).
# Se usa en lugar del Love Score completo para evitar reconstruir el propio
# indicador: el rating (60% del Love Score) se usa como predictor, por lo
# que no puede formar parte también del target.
#
# Clasificación binaria en extremos:
#   Clase 1 (alta vinculación): love_nonrating >= percentil 75
#   Clase 0 (baja vinculación): love_nonrating <= percentil 25
# Las observaciones intermedias se excluyen para maximizar la separación.

p25 = df_sample["love_nonrating"].quantile(0.25)
p75 = df_sample["love_nonrating"].quantile(0.75)

df_model = df_sample[
    (df_sample["love_nonrating"] <= p25) |
    (df_sample["love_nonrating"] >= p75)
].copy()

df_model["love_high"]     = (df_model["love_nonrating"] >= p75).astype(int)
df_model["review_length"] = df_model["review"].astype(str).str.len()

# Predictores independientes del target:
#   rating_100:     satisfacción funcional declarada
#   review_length:  longitud de la reseña como proxy de implicación emocional
#   topic_dom:      tópico dominante identificado por LDA

features = ["rating_100", "review_length", "topic_dom"]

df_model = df_model.dropna(subset=features + ["love_high"])

X, y = df_model[features], df_model["love_high"]

# División 80/20 estratificada: garantiza la misma proporción de clases
# en entrenamiento y test

X_train, X_test, y_train, y_test = train_test_split(
    X, y,
    test_size=0.20,
    random_state=42,
    stratify=y
)


# ----------------------------------------------------------------------------
# Bloque 4.2 — Entrenamiento y comparación de modelos
```

```
# ------------------------------------------------------------------------------

from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.metrics import roc_auc_score, classification_report

# Estandarización: necesaria para la red neuronal MLP

scaler      = StandardScaler()
X_train_sc  = scaler.fit_transform(X_train)
X_test_sc   = scaler.transform(X_test)

# Random Forest: 200 árboles, profundidad máxima 10.
# Robusto e interpretable mediante importancia de variables (Gini).

rf = RandomForestClassifier(
    n_estimators=200,
    max_depth=10,
    random_state=42,
    n_jobs=-1
)

# Gradient Boosting: 200 estimadores, learning rate conservador (0,05).
gb = GradientBoostingClassifier(
    n_estimators=200,
    learning_rate=0.05,
    max_depth=4,
    random_state=42
)

# Red neuronal MLP: 2 capas ocultas (128 y 64 neuronas).
# Captura relaciones no lineales entre predictores.

nn = MLPClassifier(
    hidden_layer_sizes=(128, 64),
    activation="relu",
    solver="adam",
    max_iter=300,
    random_state=42
)

modelos = {
    "Random Forest":      (rf, X_train,    X_test),
    "Gradient Boosting":  (gb, X_train,    X_test),
    "Red Neuronal (MLP)": (nn, X_train_sc, X_test_sc),
}

resultados = {}

for nombre, (modelo, X_tr, X_te) in modelos.items():
    modelo.fit(X_tr, y_train)
    y_proba = modelo.predict_proba(X_te)[:, 1]
    auc     = roc_auc_score(y_test, y_proba)
    f1      = classification_report(
        y_test,
        modelo.predict(X_te),
        output_dict=True,
        zero_division=0
    )["macro avg"]["f1-score"]

    resultados[nombre] = {
        "modelo":  modelo,
        "AUC":     auc,
        "F1":      f1,
        "y_proba": y_proba
    }

    print(f"{nombre}: AUC={auc:.4f} | F1={f1:.4f}")
```

```
mejor = max(resultados, key=lambda k: resultados[k]["AUC"])
print(f"\nMejor modelo por AUC: {mejor} ({resultados[mejor]['AUC']:.4f})")


# ------------------------------------------------------------------------------
# Bloque 4.3 — Importancia de variables (Random Forest)
# ------------------------------------------------------------------------------

# Importancia basada en reducción media de impureza Gini.

rf_model    = resultados["Random Forest"]["modelo"]
importances = pd.Series(
    rf_model.feature_importances_,
    index=features
).sort_values(ascending=False)

print("Importancia de variables (Random Forest):")
for feat, imp in importances.items():
    print(f"  {feat}: {imp:.4f} ({imp * 100:.1f}%)")


# ==============================================================================
# FASE 5 — TABLAS Y VISUALIZACIONES
# ==============================================================================

import matplotlib.pyplot as plt
from sklearn.metrics import roc_curve

# Paleta de colores por tipo de experiencia
TYPE_COLORS = {
    "emocional":  "#2E4057",
    "mixto":      "#048A81",
    "funcional":  "#54C6EB",
    "friccional": "#EF6461",
}

def _get_color(topic_type):
    return TYPE_COLORS.get(topic_type, "#999999")

# DataFrame auxiliar para figuras
viz_df = topic_table.copy()
viz_df["label"]  = viz_df.apply(
    lambda r: f"T{int(r['topic_dom'])}. {r['topic_name']}", axis=1
)
viz_df["bubble"] = viz_df["n"] / 15


# Coherencia semántica por número de tópicos

fig, ax = plt.subplots(figsize=(8, 4))
ax.plot(
    coh_df["topics"], coh_df["coherence_cv"],
    marker="o", color="#1a237e", linewidth=2, markersize=7
)
ax.axvline(x=8, color="#e53935", linestyle="--", linewidth=1.5, label="k=8
óptimo")
ax.annotate(
    "k=8\nc_v=0,5219",
    xy=(8, coh_df.loc[coh_df["topics"] == 8, "coherence_cv"].values[0]),
    xytext=(9, 0.515), fontsize=9, color="#e53935"
)
ax.set_title("Coherencia semántica (c_v) por número de tópicos")
ax.set_xlabel("Número de tópicos (k)")
ax.set_ylabel("Coherencia c_v")
ax.set_xticks(coh_df["topics"])
ax.legend(frameon=False)
ax.grid(alpha=0.25)
```

```
ax.spines["top"].set_visible(False)
ax.spines["right"].set_visible(False)
plt.tight_layout()
plt.show()


# Estadísticos descriptivos y matriz de correlaciones
# Generadas a partir de df (corpus completo, 416.032 reseñas)

vars_tabla = [
    "rating_100", "sent_norm", "kw_loyal_strict",
    "love_score", "gap_love_vs_rating"
]

desc = df[vars_tabla].describe().T[["mean", "std", "min", "max"]]
desc.columns = ["Media", "Desv. típica", "Mínimo", "Máximo"]
display(desc.round(3))

display(df[vars_tabla].corr().round(3))


# Métricas por tópico
# Generada a partir de df_final (muestra analítica, 15.994 reseñas)

display(
    topic_table[
        ["topic_dom", "topic_name", "topic_type", "n",
         "rating_mean", "love_mean", "gap_mean", "loyal_pct"]
    ].round(2)
)


# Distribución de reseñas por tópico dominante

topic_counts = (
    df_sample.groupby(["topic_dom", "topic_name"])
    .size().reset_index(name="n")
    .sort_values("n", ascending=True)
)

type_map = {
    1: "funcional", 2: "friccional", 3: "mixto", 4: "funcional",
    5: "emocional", 6: "mixto",      7: "friccional", 8: "mixto"
}

topic_counts["tipo"]  = topic_counts["topic_dom"].map(type_map)
topic_counts["color"] = topic_counts["tipo"].map(TYPE_COLORS)
topic_counts["label"] = topic_counts.apply(
    lambda r: f"T{int(r['topic_dom'])}. {r['topic_name']}", axis=1
)

fig, ax = plt.subplots(figsize=(11, 5))
bars = ax.barh(
    topic_counts["label"], topic_counts["n"],
    color=topic_counts["color"], edgecolor="white", height=0.6
)
for bar, val in zip(bars, topic_counts["n"]):
    ax.text(
        bar.get_width() + 50, bar.get_y() + bar.get_height() / 2,
        f"{val:,}", va="center", fontsize=9
    )
ax.set_title("Distribución de reseñas por tópico dominante (muestra LDA,
n=30.000)")
ax.set_xlabel("Número de reseñas")
ax.grid(axis="x", alpha=0.25)
ax.spines["top"].set_visible(False)
ax.spines["right"].set_visible(False)
plt.tight_layout()
```

```
plt.show()


# Love Score medio por tópico

df2 = viz_df.sort_values("love_mean", ascending=True).copy()

fig, ax = plt.subplots(figsize=(11, 5))
ax.barh(
    df2["label"], df2["love_mean"],
    color=[_get_color(t) for t in df2["topic_type"]], alpha=0.9
)
for i, v in enumerate(df2["love_mean"].values):
    ax.text(v + 0.3, i, f"{v:.1f}", va="center", fontsize=10)
ax.set_title("Love Score medio por tópico")
ax.set_xlabel("Love Score medio (0-100)")
ax.grid(axis="x", alpha=0.25)
ax.spines["top"].set_visible(False)
ax.spines["right"].set_visible(False)
plt.tight_layout()
plt.show()


# Rating medio vs Love Score medio por tópico (scatter)

fig, ax = plt.subplots(figsize=(9, 6))
for _, r in viz_df.iterrows():
    ax.scatter(
        r["rating_mean"], r["love_mean"],
        s=r["bubble"], alpha=0.85,
        color=_get_color(r["topic_type"]),
        edgecolor="white", linewidth=0.8
    )
    ax.annotate(
        str(int(r["topic_dom"])),
        (r["rating_mean"], r["love_mean"]),
        textcoords="offset points", xytext=(6, 4), fontsize=10
    )
ax.set_title("Rating medio vs Love Score medio por tópico")
ax.set_xlabel("Rating medio (satisfacción funcional)")
ax.set_ylabel("Love Score medio (apego emocional)")
plt.tight_layout()
plt.show()


# Gap medio (Love Score - rating) por tópico

df3 = viz_df.sort_values("gap_mean", ascending=True).copy()

fig, ax = plt.subplots(figsize=(11, 5))
ax.barh(
    df3["label"], df3["gap_mean"],
    color=[_get_color(t) for t in df3["topic_type"]], alpha=0.9
)
ax.axvline(0, linewidth=1)
for i, v in enumerate(df3["gap_mean"].values):
    x = v + (0.3 if v >= 0 else -0.3)
    ax.text(x, i, f"{v:.1f}", va="center",
            ha="left" if v >= 0 else "right", fontsize=10)
ax.set_title("Gap medio (Love Score vs Rating) por tópico")
ax.set_xlabel("Gap medio (puntos)")
ax.grid(axis="x", alpha=0.25)
ax.spines["top"].set_visible(False)
ax.spines["right"].set_visible(False)
plt.tight_layout()
plt.show()
```

```
# Regresión OLS: resultados completos

print(resultado.summary())


# Importancia de variables (Random Forest)

importances_plot = pd.Series(
    rf_model.feature_importances_, index=features
).sort_values(ascending=True)

colors = [
    "#EF6461" if imp == importances_plot.max() else "#2E4057"
    for imp in importances_plot
]

fig, ax = plt.subplots(figsize=(8, 4))
bars = ax.barh(
    importances_plot.index, importances_plot.values,
    color=colors, edgecolor="white"
)
for bar, val in zip(bars, importances_plot.values):
    ax.text(
        bar.get_width() + 0.002,
        bar.get_y() + bar.get_height() / 2,
        f"{val:.3f}", va="center", fontsize=9
    )
ax.set_title(
    "Importancia de variables — Random Forest\n"
    "(reducción media de impureza Gini)",
    fontsize=11, loc="left"
)
ax.set_xlabel("Importancia relativa")
ax.grid(axis="x", alpha=0.25)
ax.spines["top"].set_visible(False)
ax.spines["right"].set_visible(False)
plt.tight_layout()
plt.show()


# Curvas ROC comparativas de los tres modelos

colores_roc = {
    "Random Forest":      "#2E4057",
    "Gradient Boosting":  "#048A81",
    "Red Neuronal (MLP)": "#EF6461",
}

fig, ax = plt.subplots(figsize=(7, 5))
for nombre, res in resultados.items():
    fpr, tpr, _ = roc_curve(y_test, res["y_proba"])
    ax.plot(fpr, tpr,
            label=f"{nombre} (AUC={res['AUC']:.3f})",
            color=colores_roc[nombre], linewidth=2)
ax.plot([0, 1], [0, 1], "k--", linewidth=1, alpha=0.5, label="Aleatorio
(AUC=0.5)")
ax.set_title("Curvas ROC — comparación de modelos", fontsize=11, loc="left")
ax.set_xlabel("Tasa de falsos positivos")
ax.set_ylabel("Tasa de verdaderos positivos")
ax.legend(frameon=False, fontsize=9)
ax.grid(alpha=0.25)
ax.spines["top"].set_visible(False)
ax.spines["right"].set_visible(False)
plt.tight_layout()
plt.show()
```

Anexo E. Declaración de Uso de Herramientas de Inteligencia Artificial Generativa

Por la presente, yo, Cayetana Fernández de Valderrama Blanco de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "**MÁS ALLÁ DEL RATING: EL TIPO DE EXPERIENCIA COMO CLAVE DEL APEGO EMOCIONAL EN EL SECTOR HOTELERO**", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. Brainstorming de ideas de investigación: Utilizado para idear y esbozar posibles áreas de investigación.
2. Metodólogo: Para descubrir métodos aplicables a problemas específicos de investigación.
3. Interpretador de código: Para realizar análisis de datos preliminares.
4. Constructor de plantillas: Para diseñar formatos específicos para secciones del trabajo.
5. Corrector de estilo literario y de lenguaje: Para mejorar la calidad lingüística y estilística del texto.
6. Sintetizador y divulgador de libros complicados: Para resumir y comprender literatura compleja.
7. Generador de datos sintéticos de prueba: Para la creación de conjuntos de datos ficticios.
8. Generador de problemas de ejemplo: Para ilustrar conceptos y técnicas.
9. Revisor: Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: abril de 2026

Firma: 