



GRADO EN BUSINESS ANALYTICS

TRABAJO FIN DE GRADO

Shein como Lovebrand: Un análisis de la disonancia ética,
la gamificación y la paradoja del consumidor digital

Autor: Patricia Sánchez Martín

Director: Ana Isabel Jiménez Zarco

Madrid

Abril de 2026

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título
**Shein como Lovebrand: Un análisis de la disonancia ética, la gamificación y la
paradoja del consumidor digital**

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2025/26 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido

tomada de otros documentos está debidamente referenciada.



Fdo.: Patricia Sánchez Martín

Fecha: 18/ 04/ 2026

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO



Fdo.: Ana Isabel Jiménez Zarco

Fecha: 18/ 04/ 2026



GRADO EN BUSINESS ANALYTICS

TRABAJO FIN DE GRADO

Shein como Lovebrand: Un análisis de la disonancia ética,
la gamificación y la paradoja del consumidor digital

Autor: Patricia Sánchez Martín

Director: Ana Isabel Jiménez Zarco

Madrid

Abril de 2026

Agradecimientos

A mi tutora, Ana Isabel Jiménez Zarco, por su guía y por la confianza depositada en este proyecto desde el primer día.

A mis padres, por su apoyo incondicional y por ser el motor que me ha permitido llegar hasta aquí. Este logro es, en gran parte, vuestro.

A mis amigos, por el apoyo constante y por hacer que el camino fuera mucho más ligero.

Gracias a todos por ayudarme a cerrar esta etapa.

Shein como Lovebrand: Un análisis de la disonancia ética, la gamificación y la paradoja del consumidor digital

Autor: Sánchez Martín, Patricia.

Director: Jiménez Zarco, Ana Isabel.

Entidad Colaboradora: ICADE – Universidad Pontificia Comillas.

RESUMEN DEL PROYECTO

Este Trabajo de Fin de Grado investiga la paradoja de marca que representa Shein en el mercado español, analizando cómo una corporación bajo constante escrutinio ético logra consolidarse como una *Lovemark* en el ecosistema digital. El objetivo principal es desglosar los pilares de esta lealtad mediante una metodología avanzada de Business Analytics y Procesamiento de Lenguaje Natural (PLN).

Para ello, se ha consolidado un corpus de 11.717 reseñas extraídas de dos plataformas con arquitecturas de incentivos opuestas: iGraal (incentivada mediante *cashback*) y Trustpilot (orgánica). Técnicamente, el estudio somete a examen el rendimiento de dos arquitecturas de deep learning (BETO y RoBERTuito), actuando como una auditoría técnica que revela las limitaciones actuales del PLN para procesar el ruido semántico, las dobles negaciones y el lenguaje informal del consumidor.

Tras verificar mediante dicha auditoría que la valoración declarada (*ground truth*) ofrece una mayor fiabilidad estratégica ante la complejidad del discurso, los hallazgos revelan una dualidad crítica en la reputación de la marca. Mientras que en iGraal existe un sesgo de positividad artificial del 94,9%, el análisis en Trustpilot confirma un estatus de *Lovemark* orgánica con un 71,5% de valoraciones positivas. Los resultados demuestran que la lealtad de los usuarios no se fundamenta en los valores éticos o medioambientales de la firma –factores que resultan marginales en el discurso del consumidor (< 0.5%)–, sino en la fiabilidad operativa y logística excepcional.

En conclusión, Shein ha logrado proyectar los pilares de Kevin Roberts (Misterio, Sensualidad e Intimidad) a través de la gamificación y la eficiencia, logrando que la gratificación emocional inmediata actúe como un anestésico ante la disonancia ética del modelo de consumo *ultra-fast fashion* y los objetivos de sostenibilidad (ODS-12).

Palabras clave: Shein, Análisis de Sentimiento, Lovemarks, BETO, RoBERTuito, Deep Learning, Comportamiento del Consumidor, E-commerce, Business Analytics.

Shein as a Lovebrand: An analysis of ethical dissonance, gamification, and the digital consumer paradox

Author: Sánchez Martín, Patricia.

Supervisor: Jiménez Zarco, Ana Isabel.

Collaborating Entity: ICADE – Universidad Pontificia Comillas.

ABSTRACT

This bachelor's Thesis investigates the brand paradox represented by Shein in the Spanish market, analyzing how a corporation under constant ethical scrutiny has managed to consolidate itself as a Lovemark within the digital ecosystem. The primary objective is to deconstruct the pillars of this brand loyalty through an advanced methodology based on Business Analytics and Natural Language Processing (NLP).

To this end, a corpus of 11,717 reviews was consolidated from two platforms with opposing incentive architectures: iGraal (incentivized via cashback) and Trustpilot (organic). From a technical perspective, the study examines the performance of two deep learning architectures (BETO and RoBERTuito), acting as a technical audit that reveals current NLP limitations in processing semantic noise, double negatives, and informal consumer language.

After verifying through this audit that the declared rating (ground truth) provides greater strategic reliability given the complexity of the discourse, the findings reveal a critical duality in the brand's reputation. While iGraal exhibits an artificial positivity bias of 94.9%, the analysis in Trustpilot confirms an organic Lovemark status with a 71.5% positive rating. The results demonstrate that user loyalty is not founded on the firm's ethical or environmental values –factors that remain marginal in consumer discourse (< 0.5%)– but rather on exceptional operational and logistical reliability.

In conclusion, Shein has successfully projected Kevin Roberts' pillars (Mystery, Sensuality, and Intimacy) through gamification and efficiency, allowing immediate emotional gratification to act as an anesthetic against the ethical dissonance of the ultra-fast fashion model and sustainability goals (SDG-12).

Keywords: Shein, Sentiment Analysis, Lovemarks, BETO, RoBERTuito, Deep Learning, Consumer Behavior, E-commerce, Business Analytics.

Índice de la memoria

Capítulo 1. Introducción	7
1.1 Contextualización del fenómeno Lovebrand	7
1.1.1 De la marca a la Lovemark: La teoría de Kevin Roberts	8
1.1.2 La adaptación del concepto al entorno digital	8
1.2 Justificación del caso de estudio: Shein	8
1.2.1 El gigante del ultra-fast fashion	8
1.2.2 Disonancia ética: violaciones a los ODS y normativas de la UE	9
1.2.3 La paradoja: web scraping y brecha actitud-comportamiento	9
1.3 Objetivos del proyecto	9
1.3.1 Objetivo general	9
1.3.2 Objetivos específicos	10
1.4 Estructura del trabajo	10
Capítulo 2. Marco teórico	11
2.1 El concepto de Lovebrand y su evolución	11
2.1.1 Los pilares de Kevin Roberts	11
2.1.2 La evolución digital: “App Love”	12
2.2 Comportamiento del consumidor digital y economía de la atención	13
2.2.1 De la atención a la adicción dopaminérgica	13
2.2.2 El sesgo de reciprocidad en reseñas	13
2.3 El modelo de consumo fast fashion y sus implicaciones éticas	14
2.3.1 Del Fast Fashion al Ultra-Fast Fashion	14
2.3.2 Violaciones éticas y medioambientales	14
2.3.3 La brecha actitud-comportamiento (cognitive dissonance)	15
2.4 Estrategias de marketing emocional en el entorno online	16
2.4.1 Gamificación y programas de lealtad	16
2.4.2 Del fenómeno del “haul” al Live Shopping	16
Capítulo 3. Metodología	17
3.1 Diseño de investigación y justificación del enfoque	17
3.1.1 La necesidad de modelos contextuales (BERT) frente a léxicos (VADER)	17
3.2 Recolección de datos: trabajo de campo digital	18

3.2.1 Herramienta de extracción: Instant Data Scraper	18
3.2.2 Fuentes de datos y caracterización del corpus	19
3.3 Preprocesamiento y limpieza del corpus	20
3.4 Pipeline de procesamiento para el análisis de sentimientos	20
3.4.1 Normalización de texto para transformers.....	21
3.4.2 Normalización léxica para análisis de frecuencias.....	21
3.5 Modelos seleccionados: BETO y RoBERTuito	21
3.5.1 Modelo A: BETO (Spanish BERT)	22
3.5.2 Modelo B: RoBERTuito (Arquitectura RoBERTa).....	22
3.6 Limitación metodológica: la ética invisible	23
3.7 Métricas de evaluación y análisis estratégico.....	24
Capítulo 4. Análisis de resultados.....	25
4.1 Introducción al análisis.....	25
4.2 Análisis descriptivo y sesgo de positividad.....	25
4.2.1 El ecosistema incentivado: el caso iGraal	26
4.2.2 El ecosistema orgánico: el caso Trustpilot	27
4.2.3 Conclusión comparativa: el impacto del incentivo	27
4.2.4 Coherencia de la inferencia: El modelo como espejo del ecosistema.....	28
4.3 Auditoría técnica: evaluación de modelos.....	31
4.3.1 Desbalanceo de clases y la insuficiencia de la exactitud.....	31
4.3.2 Elección del F1-Score Macro como indicador de fiabilidad	32
4.3.3 Análisis cualitativo de discrepancias entre usuario y modelos.....	33
4.3.4 Rendimiento comparado: por clase y longitud del texto	36
4.4 Impacto del incentivo económico en el discurso.....	40
4.4.1 Análisis conceptual: el desplazamiento de la marca.....	40
4.4.2 Análisis sintáctico: La polarización de la transacción	42
4.4.3 La perspectiva Lovebrand	43
4.5 Reputación orgánica y construcción de la lovebrand	45
4.5.1 El retorno al producto y al servicio.....	45
4.5.2 Radiografía de la experiencia: fricciones vs aciertos	46
4.5.3 El veredicto Lovebrand	48
4.6 La disonancia ética y la ceguera del consumidor	51
4.6.1 Análisis cualitativo del discurso.....	51

4.6.2	<i>Conclusión de la dimensión ética</i>	54
Capítulo 5. <i>Discusión y conclusiones</i>		55
5.1	La anatomía de una Lovemark digital	55
5.1.1	<i>La dualidad del ecosistema incentivado: El caso de iGraal</i>	55
5.1.2	<i>El veredicto de la reputación orgánica: El caso de Trustpilot</i>	56
5.1.3	<i>Construcción de los pilares: Misterio, Sensualidad, Intimidación</i>	57
5.2	Evaluación de los objetivos específicos	58
5.3	Aportaciones y conclusiones finales	59
5.4	Limitaciones y líneas futuras de investigación	59
5.4.1	<i>Limitaciones del estudio</i>	60
5.4.2	<i>Líneas de investigación futuras</i>	60
Capítulo 6. <i>Bibliografía</i>		61
ANEXO I: <i>Pipeline de Análisis de Datos</i>		67
ANEXO II: <i>Declaración de Uso de IA</i>		76

Índice de figuras

Figura 1: Eje Amor/Respeto de Kevin Roberts	12
Figura 2: Distribución de sentimientos en iGraal basada en puntuación de estrellas	26
Figura 3: Distribución de sentimientos en Trustpilot basada en puntuación de estrellas ..	27
Figura 4: Inferencia global en iGraal (BETO)	28
Figura 5: Inferencia global en iGraal (RoBERTuito).....	29
Figura 6: Inferencia Global en Trustpilot (BETO).....	30
Figura 7: Inferencia Global en Trustpilot (RoBERTuito)	30
Figura 8: Comparativa de matrices de confusión por modelo en iGraal	33
Figura 9: Comparativa de matrices de confusión por modelo en Trustpilot	33
Figura 10: Ground truth: 5 estrellas vs Predicción IA: Negativo	34
Figura 11: Ground truth: 5 estrellas vs Predicción IA: Negativo	35
Figura 12: Ground-truth: 5 estrellas vs Predicción IA: Negativo.....	35
Figura 13: Comparativa de sensibilidad (recall) por clase en iGraal	36
Figura 14: Comparativa de sensibilidad (recall) por clase en Trustpilot	37
Figura 15: Longitud del comentario frente al ground truth en iGraal	38
Figura 16: Longitud del comentario frente al ground truth en Trustpilot	38
Figura 17: Nubes de palabras más frecuentes por estrellas en iGraal	41
Figura 18: Expresiones más frecuentes (bigramas) por estrellas en iGraal	42
Figura 19: Proporción de promotores transaccionales (falsos positivos) en valoraciones altas de iGraal	44
Figura 20: Nubes de palabras más frecuentes por estrellas en Trutspilot	45
Figura 21: Bigramas por estrellas en Trustpilot	46
Figura 22: Frecuencia de puntos de dolor en reseñas negativas	47
Figura 23: Frecuencia de puntos fuertes en reseñas positivas	48
Figura 24: Análisis de emociones Lovebrand en promotores orgánicos de Trustpilot	49
Figura 25: Mapas semánticos de la lealtad emocional: análisis cualitativo por categorías	49
Figura 26: Frecuencias exactas de términos vinculados a la ética laboral en iGraal	52

Figura 27: Frecuencias exactas de términos vinculados a la ética laboral en Trustpilot ...	52
Figura 28: Frecuencias exactas de términos vinculados a la sostenibilidad y medioambiente en iGraal	53
Figura 29: Frecuencias exactas de términos vinculados a la sostenibilidad y medioambiente en Trustpilot	53

Índice de tablas

Tabla 1: Comparativa de las fuentes de datos 19

Capítulo 1. INTRODUCCIÓN

La industria de la moda contemporánea se encuentra en una encrucijada sin precedentes, definida por una tensión polarizante entre la aceleración digital del consumo y la urgencia global por la sostenibilidad. En el epicentro de esta tensión se sitúa Shein, un gigante del comercio electrónico que ha redefinido los límites de la producción textil y la relación marca-consumidor. En este Trabajo de Fin de Grado se propone diseccionar el fenómeno de Shein bajo la óptica de la teoría de las *Lovemarks* de Kevin Roberts, explorando cómo una marca rodeada de controversias éticas y medioambientales logra generar niveles de lealtad y adoración comparables a los de las marcas de culto tradicionales.

La relevancia de esta investigación radica en la paradoja observable en el comportamiento del consumidor digital: mientras la conciencia social sobre los Objetivos de Desarrollo Sostenible (ODS) y la ética empresarial alcanza máximos históricos en el discurso público, el éxito comercial de Shein (basado en el modelo de *ultra-fast fashion*) sugiere una disociación crítica entre los valores declarados y los hábitos de compra reales. A través del análisis de miles de reseñas en plataformas con arquitecturas de incentivos divergentes, este trabajo busca arrojar luz sobre los mecanismos psicológicos, como la gamificación y el incentivo económico, que sostienen esta “lealtad más allá de la razón”.

1.1 CONTEXTUALIZACIÓN DEL FENÓMENO LOVEBRAND

Para comprender la posición de Shein en el mercado actual, es necesario establecer primero el marco teórico que define las relaciones emocionales entre consumidores y corporaciones. La evolución del marketing en las últimas dos décadas ha transitado desde una economía de la atención (centrada en la visibilidad), hacia una economía de la atracción, donde la conexión emocional es la moneda de cambio fundamental.

1.1.1 DE LA MARCA A LA LOVEMARK: LA TEORÍA DE KEVIN ROBERTS

El concepto de *Lovemark* fue acuñado por Kevin Roberts, entonces CEO de Saatchi & Saatchi, en su trabajo pionero de 2004. Roberts postuló que las marcas tradicionales se habían “mercantilizado” (*commoditized*), centrándose excesivamente en el rendimiento, la eficiencia y la confianza racional, descuidando el componente emocional que impulsa la lealtad a largo plazo [1], [2].

Según la teoría de Roberts, para que una marca trascienda su categoría y se convierta en una *Lovemark*, debe poseer tres elementos intangibles críticos: Misterio, Sensualidad e Intimidación. Esta tríada genera lo que Roberts denomina “lealtad más allá de la razón”, donde el consumidor perdona fallos y defiende la marca activamente [3].

1.1.2 LA ADAPTACIÓN DEL CONCEPTO AL ENTORNO DIGITAL

En la era del comercio electrónico, la teoría de Roberts ha evolucionado. Investigadores han identificado que, en el entorno móvil, conceptos como “*App Love*” y “*App Respect*” actúan como medidores de lealtad. Sin embargo, surge una cuestión crítica al aplicar esta teoría a Shein: ¿Puede existir una *Lovemark* cuando el “Respeto” institucional –entendido como respeto por la ética, la ley y el medio ambiente– está comprometido? [4]

1.2 JUSTIFICACIÓN DEL CASO DE ESTUDIO: SHEIN

La elección de Shein como objeto de estudio no es arbitraria: representa uno de los casos más extremos y exitosos de la digitalización de la moda, y simultáneamente, el mayor desafío a los paradigmas de sostenibilidad global.

1.2.1 EL GIGANTE DEL ULTRA-FAST FASHION

Shein ha evolucionado de ser un minorista desconocido a convertirse en la encarnación del *ultra-fast fashion*. Con una facturación de alrededor de 38.000 millones de dólares en 2024 y un crecimiento interanual del 18%, Shein añade miles de nuevos estilos diarios a su plataforma, superando la capacidad de cualquier competidor tradicional como H&M o Zara.

Además, su dominio se refleja en la penetración de mercado, siendo la aplicación de compras la más descargada en 2024 [5], [6], [7].

1.2.2 DISONANCIA ÉTICA: VIOLACIONES A LOS ODS Y NORMATIVAS DE LA UE

El modelo de negocio de Shein entra en conflicto directo con los ODS, y ha provocado reacciones legislativas sin precedentes en países como Francia. Investigaciones recientes de Greenpeace (2024/2025) hallaron que el 32% de los productos testados contenían sustancias químicas peligrosas por encima de los límites legales de la UE, incluyendo niveles alarmantes de ftalatos y PFAS. Asimismo, informes de Public Eye (2024) confirmaron jornadas laborales de 75 horas semanales en sus fábricas proveedoras, violando las leyes laborales chinas y los propios códigos de conducta de la empresa [8], [9], [10].

La Unión Europea ha respondido con la “Estrategia para Textiles Sostenibles” y el “Reglamento de Ecodiseño” (ESPR), buscando prohibir la destrucción de textiles no vendidos y combatir el *greenwashing* [11], [12], [13].

1.2.3 LA PARADOJA: WEB SCRAPING Y BRECHA ACTITUD-COMPORTAMIENTO

Originalmente, el diseño de esta investigación contemplaba un *web scraping* masivo para analizar conversaciones sobre ética. Sin embargo, la escasez significativa de comentarios orgánicos sobre estos temas en plataformas masivas como iGraal y Trustpilot valida la teoría de la brecha actitud-comportamiento (*attitude-behavior gap*): los consumidores suprimen la información ética para disfrutar de la gratificación del consumo barato e inmediato, un fenómeno que este TFG explorará en profundidad [14], [15].

1.3 OBJETIVOS DEL PROYECTO

1.3.1 OBJETIVO GENERAL

El objetivo principal de este trabajo es analizar la construcción de marca de Shein y la lealtad de sus consumidores mediante el análisis de sentimientos en plataformas de reseñas, contrastando la percepción del usuario con la realidad ética y sostenible de la empresa.

1.3.2 OBJETIVOS ESPECÍFICOS

- I. **Desarrollar un corpus robusto:** a partir de opiniones de consumidores españoles provenientes de fuentes heterogéneas (iGraal y Trustpilot) mediante técnicas de extracción automatizada (*web scraping*), garantizando la representatividad de los datos.
- II. **Evaluar el impacto de los incentivos económicos (*cashback*):** Evaluar cómo el modelo de recompensa de iGraal distorsiona el sentimiento del consumidor, generando una positividad artificial.
- III. **Determinar el sentimiento real actual:** Analizar reseñas recientes de Trustpilot para identificar puntos de fricción reales (logística, calidad...) sin el filtro del incentivo monetario inmediato, evaluando la eficacia de los modelos de lenguaje basados en la arquitectura Transformer (PLN) para capturar esta polaridad frente a la valoración real del usuario.
- IV. **Verificar la brecha ética:** Confirmar empíricamente la hipótesis de que la ética corporativa es un factor irrelevante en la evaluación poscompra del consumidor español, contrastando la ausencia de términos éticos en las reseñas con la magnitud de las controversias reales.

1.4 ESTRUCTURA DEL TRABAJO

El trabajo se divide en cinco capítulos. Tras esta introducción, el Capítulo 2 aborda el marco teórico. El Capítulo 3 detalla la metodología de *web scraping* y análisis de sentimientos. El Capítulo 4 presenta la discusión de resultados comparando iGraal y Trustpilot. Finalmente, el Capítulo 5 ofrece conclusiones y recomendaciones.

Capítulo 2. MARCO TEÓRICO

Este capítulo establece los fundamentos conceptuales necesarios para analizar la paradoja de Shein. Se examina la evolución de la relación marca-cliente desde la teoría clásica de las *Lovemarks* hasta la psicología del consumidor en la economía de la atención digital, cerrando con las implicaciones éticas del modelo de producción que sustenta este “amor” digital.

2.1 EL CONCEPTO DE LOVEBRAND Y SU EVOLUCIÓN

La teoría de las marcas ha sufrido una transformación radical en el siglo XXI, pasando de un enfoque transaccional, a uno relacional y emocional.

2.1.1 LOS PILARES DE KEVIN ROBERTS

Kevin Roberts revolucionó el marketing en 2004 al afirmar que: “*lo que mueve a los seres humanos es la emoción, no la razón*”. Para Roberts, las marcas que logran sobrevivir a la mercantilización son aquellas que se convierten en *Lovemarks*, definidas por tres componentes esenciales [16], [17]:

- **Misterio:** La capacidad de contar historias, evocar sueños, mitos e iconos, aprovechando el pasado, el presente y el futuro.
- **Sensualidad:** La estimulación de los cinco sentidos. En el contexto digital, esto se traduce en una experiencia visual saturada y en una interfaz de usuario táctil altamente reactiva.
- **Intimidad:** La creación de empatía, compromiso y pasión. Es la capacidad de la marca de hacerse indispensable en la vida diaria del consumidor.

Roberts propuso el Eje Amor/Respeto, en el cual se identifican cuatro tipos de productos diferentes:

1. **Bajo Respeto/Bajo Amor:** *Commodities* (productos genéricos).

2. **Alto Respeto/Bajo Amor:** Marcas funcionales (buena calidad-precio, pero sin vínculo emocional).
3. **Bajo Respeto/Alto Amor:** Modas pasajeras (*Fads*).
4. **Alto Respeto/Alto Amor:** *Lovemarks*.



Figura 1: Eje Amor/Respeto de Kevin Roberts

Fuente: Roberts (2004)

2.1.2 LA EVOLUCIÓN DIGITAL: “APP LOVE”

En la última década, la teoría se ha adaptado al ecosistema móvil. Investigaciones recientes introducen los conceptos de “*App Love*” y “*App Respect*”. En este entorno, la “intimidad” se logra mediante la presencia constante del icono de la app en el dispositivo personal del usuario y con la personalización algorítmica. Shien ha capitalizado esta evolución, convirtiendo su aplicación no solo en una tienda, sino en una red social de consumo donde la interacción diaria genera un sucedáneo de intimidad [4].

2.2 COMPORTAMIENTO DEL CONSUMIDOR DIGITAL Y ECONOMÍA DE LA ATENCIÓN

El éxito de Shein no puede explicarse solo por el precio; se basa en la explotación experta de los sesgos cognitivos en la llamada “economía de la atención”.

2.2.1 DE LA ATENCIÓN A LA ADICCIÓN DOPAMINÉRGICA

En el entorno digital, la atención es un recurso escaso. El modelo de negocio de las plataformas actuales se basa en maximizar el tiempo en pantalla. Shein utiliza un sistema de recompensa variable (similar a las máquinas tragaperras) a través de ofertas *flash*, juegos para obtener descuentos, cuentas atrás y puntos por log in diario [18].

Psicológicamente, esto activa el circuito de la dopamina. La compra de *fast fashion* desencadena una gratificación instantánea, en la que el placer reside más en “la caza de la ganga” que en el uso posterior de la prenda. Este ciclo crea una adicción al consumo que nubla el juicio racional sobre la calidad o la necesidad real del producto [19], [20].

2.2.2 EL SESGO DE RECIPROCIDAD EN RESEÑAS

Un componente clave del comportamiento digital es la validación social mediante reseñas. Sin embargo, plataformas como iGraal introducen incentivos económicos (como el *cashback*) que alteran este comportamiento. Estudios académicos demuestran que los incentivos financieros activan la norma de reciprocidad, en la que el consumidor siente una obligación psicológica inconsciente de “devolver el favor” a la plataforma que le paga, lo que se traduce en reseñas artificialmente altas que no reflejan la experiencia real del producto. Además, el incentivo monetario puede introducir un sesgo de positividad, donde el usuario suprime los aspectos negativos para justificar su ganancia económica [21].

2.3 EL MODELO DE CONSUMO FAST FASHION Y SUS IMPLICACIONES ÉTICAS

Para contrastar el “amor” del consumidor con la realidad corporativa, es necesario definir el modelo operativo de Shein y sus costes externos, y por qué los consumidores ignoran sus fallos éticos.

2.3.1 DEL FAST FASHION AL ULTRA-FAST FASHION

Mientras que el *fast fashion* tradicional (Inditex, H&M, ...) opera con ciclos de semanas, el modelo de “*real-time retail*” de Shein produce lotes pequeños en días, basándose en big data para predecir micro tendencias. Esto se traduce en la adición diaria de miles de nuevos estilos a la web, fomentando una obsolescencia programada psicológica en la que la ropa se percibe como desechable casi de inmediato. Este volumen masivo es lo que imposibilita un control de calidad y ético riguroso [22].

2.3.2 VIOLACIONES ÉTICAS Y MEDIOAMBIENTALES

La investigación periodística y de ONGs reciente ha documentado extensamente el impacto negativo de este modelo, chocando frontalmente con los ODS de la UE:

- I. **Impacto ambiental (ODS 12):** El informe de Greenpeace Alemania (noviembre 2024/2025) reveló que el 32% de las prendas de Shein analizadas contenían sustancias químicas peligrosas por encima de los límites REACH de la UE, incluyendo disruptores hormonales (ftalatos) y PFAS en niveles hasta 3.000 veces superiores a lo permitido [8], [23].
- II. **Explotación laboral (ODS 8):** La investigación de seguimiento de Public Eye en 2024 confirmó que, a pesar de las promesas de mejora, los trabajadores en la cadena de suministro de Shein en Guangzhou continuaban trabajando hasta 75 horas a la semana, con un solo día libre al mes, bajo un sistema de pago por pieza que incentiva la autoexplotación [9], [10], [24].

Además, un ejemplo reciente y paradigmático de la tensión entre Shein y la normativa europea es la situación en Francia:

III. **Reacción legislativa y social en el caso de Francia:** Coincidiendo con la apertura de tiendas pop-up y planes de expansión física en París, el gobierno francés ha impulsado una legislación pionera “*anti-fast fashion*”.

- **La “ley anti-Shein”:** La Asamblea Nacional francesa aprobó medidas que incluyen una prohibición de productos de *ultra-fast fashion* y la imposición de una penalización ecológica de hasta 10 euros por prenda vendida para 2030, destinada a compensar el impacto ambiental [25], [26].
- **Rechazo social:** La apertura de espacios físicos de Shein en Francia ha sido recibida con protestas y grafitis con la leyenda “*Shein kills*”, lo que evidencia un rechazo activo por parte de sectores concienciados de la sociedad.

Este contexto es crucial para el estudio: demuestra que la información sobre los abusos de Shein es pública y notoria en Europa. Si el análisis posterior de las reseñas de usuarios españoles no refleja esta preocupación, confirmará una desconexión total entre el debate político-social y la realidad del consumidor digital promedio [27].

2.3.3 LA BRECHA ACTITUD-COMPORTAMIENTO (COGNITIVE DISSONANCE)

A pesar de la disponibilidad de esta información, los consumidores siguen comprando. Esto se explica mediante la teoría de la disonancia cognitiva y la brecha actitud-comportamiento: el consumidor experimenta una incomodidad psicológica al conocer los abusos éticos, pero la resuelve no dejando de comprar, sino trivializando la información ética (“todas las marcas lo hacen”) o priorizando la necesidad económica y estética inmediata sobre sus valores morales a largo plazo [28].

2.4 ESTRATEGIAS DE MARKETING EMOCIONAL EN EL ENTORNO ONLINE

Shein mitiga la falta de “respeto” ético mediante una estrategia de marketing emocional agresiva que simula los atributos de una *Lovemark*.

2.4.1 GAMIFICACIÓN Y PROGRAMAS DE LEALTAD

Shein ha gamificado la experiencia de compra. Los usuarios ganan puntos no solo por comprar, sino también por acciones que requieren tiempo y compromiso emocional, como escribir reseñas, subir fotos de sus “*hauls*”, ver transmisiones en vivo y participar en minijuegos dentro de la app. Esto crea un ecosistema cerrado donde el usuario invierte “trabajo” en la marca, aumentando su coste de cambio y su vinculación emocional (efecto IKEA: “amamos lo que ayudamos a construir”) [18], [29], [30].

2.4.2 DEL FENÓMENO DEL “HAUL” AL LIVE SHOPPING

En el ecosistema de Shein, el marketing de influencia y el contenido generado por el usuario (UGC) no son solo herramientas, sino los principales motores de crecimiento. Los vídeos de “*Shein haul*” y las sesiones de Live Shopping en plataformas como Instagram y TikTok han dejado de ser mera publicidad para convertirse en rituales de pertenencia a una comunidad global.

Al mostrar volúmenes masivos de prendas baratas, el consumidor obtiene una gratificación inmediata y validación social. Este fenómeno refuerza la norma de que el sobreconsumo no solo es aceptable, sino deseable, logrando desplazar con éxito la conversación desde un plano ético hacia uno puramente estético y lúdico.

A través de este discurso de “democratización de la moda”, Shein logra construir una narrativa de empoderamiento que sustituye la crítica a su modelo de negocio por una experiencia de usuario gratificante. Esto consolida su posición como *Lovemark* digital: una relación intensa, adictiva y emocionalmente gratificante, construida sobre la negación de su realidad productiva [31], [32].

Capítulo 3. METODOLOGÍA

Este capítulo describe el diseño de la investigación, y justifica la elección de un enfoque cuantitativo-computacional basado en técnicas de minería de datos y de procesamiento del lenguaje natural (PLN) para analizar una cantidad considerable de datos no estructurados. Dada la complejidad del lenguaje en redes sociales y plataformas de opinión (caracterizado por el uso de jerga, ironía, emojis y errores gramaticales), se ha optado por modelos de deep learning basados en la arquitectura Transformer, específicamente adaptados al idioma español.

3.1. DISEÑO DE INVESTIGACIÓN Y JUSTIFICACIÓN DEL ENFOQUE

El estudio adopta un diseño no experimental, descriptivo y transversal. Se basa en el análisis de fuentes secundarias de datos masivos (*User Generated Content*, UGC) extraídos de plataformas de reseñas. No se manipulan variables, sino que se analizan las manifestaciones naturales de opinión de los usuarios en dichos entornos digitales preexistentes.

3.1.1 LA NECESIDAD DE MODELOS CONTEXTUALES (BERT) FRENTE A LÉXICOS (VADER)

Tradicionalmente, el análisis de sentimiento en estudios de mercado se ha realizado mediante enfoques basados en el léxico, siendo VADER (*Valence Aware Dictionary and sEntiment Reasoner*) una de las herramientas más populares. VADER asigna puntuaciones de polaridad a palabras individuales y aplica reglas heurísticas para manejar negaciones y modificadores [33].

Sin embargo, para el presente estudio centrado en el español y en un contexto de *Lovebrand* donde la ironía y el sarcasmo pueden ser prevalentes, VADER presenta limitaciones críticas:

1. **Dependencia del idioma:** VADER es nativo del inglés. Su aplicación en español requiere traducción automática previa o el uso de diccionarios adaptados, que pierden matices culturales y dialectales, lo que a menudo da lugar a clasificaciones erróneas.
2. **Incapacidad contextual:** VADER no comprende el contexto global de la oración. Ante frases como: “Me encanta esperar dos meses por mi pedido”, VADER detectaría “encanta” como positivo, fallando en identificar la ironía evidente. Los modelos BERT (*Bidirectional Encoder Representations from Transformers*), por el contrario, analizan la secuencia completa bidireccionalmente, permitiendo capturar estas dependencias semánticas complejas [34].
3. **Riqueza morfológica del español:** El español es un idioma con una morfología rica. Modelos como BETO o MarIA, entrenados específicamente con corpus masivos en español, manejan mejor las conjugaciones verbales y el género gramatical que los enfoques basados en reglas simples [35], [36].

Por tanto, se ha determinado que el uso de modelos BERT pre-entrenados en español es metodológicamente superior y necesario para cumplir los objetivos de precisión del estudio.

3.2 RECOLECCIÓN DE DATOS: TRABAJO DE CAMPO DIGITAL

La construcción del corpus se ha realizado mediante técnicas de *web scraping*, seleccionando dos fuentes que representan dos motivaciones de usuario distintas: iGraal (motivación extrínseca/económica) y Trustpilot (motivación intrínseca/expresiva).

3.2.1 HERRAMIENTA DE EXTRACCIÓN: INSTANT DATA SCRAPER

Para la extracción de datos, se utilizó *Instant Data Scraper*, una herramienta heurística de extracción de datos web basada en inteligencia artificial. La elección de esta herramienta se justifica por:

- **Capacidad de detección de patrones:** Utiliza IA para analizar la estructura HTML (DOM) y detectar automáticamente contenedores de datos repetitivos (reseñas), lo

cual es ideal para plataformas dinámicas que actualizan frecuentemente su código frontend [37].

- **Gestión de paginación y de desplazamiento (scroll) infinito:** Permite la navegación automatizada a través de miles de páginas de reseñas mediante la carga dinámica de contenido, superando las limitaciones de la recolección manual y asegurando la obtención de un volumen de datos estadísticamente significativo.
- **Validez académica:** El uso de extensiones de *scraping* para la recolección de datos con fines académicos ha sido validada en literatura reciente, siempre que se respeten los términos de servicio y la privacidad de los usuarios (anonimización).

3.2.2 FUENTES DE DATOS Y CARACTERIZACIÓN DEL CORPUS

Se ha consolidado un dataset final de 11.717 registros, dividido de la siguiente manera:

Fuente de datos	Volumen	Periodo	Naturaleza del incentivo	Hipótesis de sesgo
iGraal	5.264 reseñas	Dic2020 – Dic2025	Económico (<i>cashback</i>)	Sesgo de positividad: Los usuarios tienden a valorar positivamente la transacción para asegurar o validar su recompensa económica.
Trustpilot	6.453 reseñas	Oct2025 – Ene2026	Reputacional y orgánico	Polarización realista: Plataforma utilizada típicamente para quejas posventa o recomendaciones genuinas. Se espera una mayor presencia de críticas sobre temas como logística y calidad.

Tabla 1: Comparativa de las fuentes de datos

Este diseño comparativo permite aislar la variable “incentivo” y analizar cómo distorsiona la percepción de la marca Shein.

3.3 PREPROCESAMIENTO Y LIMPIEZA DEL CORPUS

Antes de la ingesta de datos por parte de los algoritmos, se aplicó un protocolo de limpieza manual y automatizado para asegurar la calidad del dataset:

1. **Filtrado de ruido:** Se eliminaron del corpus todas las reseñas que consistían únicamente en una puntuación de estrellas sin contenido textual, ya que no aportan valor semántico para el análisis de PLN.
2. **Anonimización:** Cumpliendo con los principios éticos de investigación digital, se eliminaron variables sensibles como nombres de usuario reales que no son relevantes para el análisis de sentimiento.
3. **Generación del “ground truth”:** Para evaluar la precisión de las IAs, se creó una variable objetivo, basada en la puntuación de estrellas original otorgada por el usuario. Se realizó una discretización de esta variable (1 – 5) a categórica:
 - **1 – 2 estrellas:** Etiquetado como “Negativo”.
 - **3 estrellas:** Etiquetado como “Neutral”.
 - **4 – 5 estrellas:** Etiquetado como “Positivo”.

3.4 PIPELINE DE PROCESAMIENTO PARA EL ANÁLISIS DE SENTIMIENTOS

El flujo de trabajo técnico se implementó utilizando el lenguaje de programación Python, apoyándose en la librería *pandas* para la ingesta y manipulación de datos, la arquitectura *transformer* de Hugging Face para la ejecución de los modelos de deep learning, y *matplotlib* y *seaborn* para la representación visual de los hallazgos. Asimismo, se integró el uso de expresiones regulares (*regex*) y diccionarios léxicos personalizados para la categorización temática del discurso (logística, economía, ética, emociones), permitiendo cuantificar dimensiones que los modelos de sentimiento general no aíslan por sí solos.

3.4.1 NORMALIZACIÓN DE TEXTO PARA TRANSFORMERS

A diferencia de los métodos clásicos, los modelos Transformer se benefician del contexto completo. Por tanto, el preprocesamiento de texto fue conservador:

- **No se aplicó *stemming* ni eliminación de *stop words*:** Eliminar preposiciones o conjunciones en modelos como BERT puede destruir el contexto sintáctico necesario para detectar ironía.
- **Gestión de emojis:** Se mantuvieron los caracteres emoji, ya que los modelos seleccionados (especialmente RoBERTuito) son capaces de interpretar su carga semántica.
- **Limpieza de espacios:** Se normalizaron espacios en blanco, tabulaciones y saltos de línea residuales del *scraping* mediante expresiones regulares.

3.4.2 NORMALIZACIÓN LÉXICA PARA ANÁLISIS DE FRECUENCIAS

Aunque los modelos Transformer ingirieron el texto bruto para no perder contexto, para los análisis cualitativos posteriores (extracción de bigramas y nubes de palabras) se aplicó un preprocesamiento intensivo adicional. Se eliminaron signos de puntuación, se unificó el texto a minúsculas y se aplicó un filtrado de *stop words* (palabras vacías como artículos, preposiciones y pronombres) utilizando diccionarios especializados de la librería NLTK (*Natural Language Toolkit*). Esto garantizó que el análisis de frecuencias destacara sustantivos y adjetivos con carga semántica real.

3.5 **MODELOS SELECCIONADOS: BETO Y ROBERTUITO**

Para el análisis de sentimiento, se descartó el uso de modelos multilingües genéricos (como mBERT) a favor de modelos monolingües optimizados para el español, basándose en la evidencia de que estos últimos capturan mejor la sutileza, la ironía y la gramática compleja del español [38].

Por tanto, el análisis técnico se fundamenta en el empleo de dos arquitecturas de referencia para el procesamiento del lenguaje en español. No se ha realizado un reentrenamiento (*fine-tuning*) desde cero, sino que se ha procedido a la inferencia directa mediante modelos ya optimizados para el análisis de sentimiento en español. Este enfoque permite establecer un marco de contraste técnico entre la interpretación automatizada y el *ground truth* declarado por el usuario.

3.5.1 MODELO A: BETO (SPANISH BERT)

Desarrollado por la Universidad de Chile, BETO es el modelo de referencia de la academia hispana. Se utiliza el modelo: `finiteautomata/beto-sentiment-analysis` [39].

- **Base:** BERT-Base-Spanish-WWM-Cased.
- **Justificación:** Es el modelo canónico entrenado con corpus formales masivos (Wikipedia, fuentes académicas y OPUS Project). En esta investigación, actúa como punto de referencia para evaluar cómo una IA entrenada con español estándar/académico se enfrenta a la jerga de internet.

Al ser una arquitectura basada en BERT y entrenada con la técnica de *Whole Word Masking* (WWM), BETO posee una ventaja analítica fundamental. Mientras que otros modelos se entrenan prediciendo fragmentos de palabras sueltas (*tokens*), WWM obliga al algoritmo a predecir palabras completas en función del contexto circundante. Esta exigencia durante su fase de preentrenamiento le otorga una comprensión superior de las relaciones de dependencia en oraciones complejas. Esto le permite identificar matices y ambigüedades en reseñas donde el sentimiento no es evidente de forma aislada, sino que depende de la estructura sintáctica total.

3.5.2 MODELO B: ROBERTUITO (ARQUITECTURA ROBERTA)

Para esta segunda aproximación, se optó por la arquitectura RoBERTa, que optimiza el procedimiento de entrenamiento de BERT al eliminar la predicción de la siguiente frase y emplear un *masking* dinámico. Aunque el modelo MarIA (desarrollado por la Biblioteca Nacional de España y el Barcelona Supercomputing Center) es el estándar de esta

arquitectura, para este TFG se ha seleccionado una variante especializada: RoBERTuito, implementada mediante la librería *pysentimiento* [40], [41].

- **Base:** RoBERTa-base-cased entrenada específicamente con 500 millones de tuits en español.
- **Justificación:** Al compartir la arquitectura robusta de MarIA, pero haber sido entrenado con textos de redes sociales, este modelo representa la mejor aproximación posible para entender el lenguaje informal, los errores ortográficos y la ironía esperables de las reseñas de los usuarios de *fast fashion*.

3.6 LIMITACIÓN METODOLÓGICA: LA ÉTICA INVISIBLE

Originalmente, se contempló un módulo de análisis para comentarios sobre ética y sostenibilidad. Sin embargo, el trabajo de campo preliminar reveló que estos temas son marginales en las plataformas de reseñas de consumo masivo. La gente puntúa “genial” basándose en la logística y el precio, ignorando la ética. Esta ausencia de datos no es un fallo de recolección, sino un resultado en sí mismo que será discutido cualitativamente: la invisibilidad de los ODS y el escándalo en un país vecino en la experiencia poscompra del usuario de *fast fashion*.

Por tanto, la dimensión ética se aborda mediante la inferencia negativa (la ausencia de mención) y se contrasta con la literatura teórica, en lugar de mediante un modelo de clasificación específico. Para medir empíricamente este silencio, la metodología se apoya en un enfoque léxico: se utilizan diccionarios temáticos y expresiones regulares con límites de palabra para aislar, cuantificar y medir la concentración exacta de estos términos en la reducida minoría de usuarios que los emplean.

3.7 MÉTRICAS DE EVALUACIÓN Y ANÁLISIS ESTRATÉGICO

Dado el fuerte desbalanceo de clases detectado en la exploración inicial (predominio masivo de reseñas positivas sobre todo en iGraal), la evaluación no podía basarse únicamente en la exactitud (*accuracy*):

1. **Matrices de confusión:** Se generaron mapas de calor para visualizar no solo los aciertos, sino la direccionalidad del error.
2. **Métrica F1-Score macro:** Se seleccionó esta métrica como el indicador principal de fiabilidad, ya que penaliza a los modelos que ignoran las clases minoritarias (comentarios negativos o neutrales) y ofrece una visión más honesta del rendimiento real.
3. **Análisis de discrepancias (auditoría humana):** Se aisló un subconjunto de datos en el que los modelos discrepaban para realizar una inspección cualitativa manual y detectar matices como sarcasmo o ambigüedad. Esta fase de auditoría resultó determinante para validar la fiabilidad de los modelos frente al *ground truth* y establecer el criterio de análisis para las conclusiones estratégicas finales.
4. **Detección de “ceguera ética”:** Dado que la ética es una dimensión latente y no una categoría etiquetada por el usuario, se implementó un algoritmo de búsqueda de palabras clave y bigramas para cuantificar la frecuencia de términos relacionados con derechos laborales y sostenibilidad vs términos relacionados con otros temas como precio, calidad y logística. Adicionalmente, para validar el contexto real de estas escasas menciones, se generaron representaciones semánticas (nubes de palabras) basadas en frecuencias exactas tras una limpieza profunda de *stop words* con NLTK. Este proceso permitió confirmar que los términos detectados correspondían efectivamente al contexto analizado, aportando una necesaria capa de validación cualitativa al análisis.

Capítulo 4. ANÁLISIS DE RESULTADOS

4.1 INTRODUCCIÓN AL ANÁLISIS

En este capítulo se exponen los hallazgos obtenidos del análisis computacional de 11.717 reseñas extraídas de iGraal y Trustpilot. Siguiendo la metodología descrita en el Capítulo 3, los datos han sido procesados mediante dos modelos de lenguaje basados en la arquitectura Transformer: BETO (basado en la arquitectura BERT) y RoBERTuito (basado en la arquitectura RoBERTa), para contrastar la solidez del discurso del usuario frente a su valoración numérica. La exposición de los resultados se estructura en tres bloques:

1. **Auditoría de modelos IA:** Se evalúa el rendimiento técnico de los modelos, evaluando si son capaces de detectar matices lingüísticos que escapan a la puntuación de estrellas tradicional.
2. **Validación del sesgo del incentivo:** Se contrasta la distribución de valoraciones entre una plataforma incentivada económicamente (iGraal) y una orgánica (Trustpilot), demostrando cuantitativamente la hipótesis del “sesgo de positividad”.
3. **Minería de opinión y ética:** Se diseccionan los temas de conversación para identificar los puntos de dolor reales del cliente y verificar la invisibilidad de las preocupaciones éticas en el discurso del consumidor *fast fashion*.

4.2 ANÁLISIS DESCRIPTIVO Y SESGO DE POSITIVIDAD

El primer nivel de análisis se centra en el *ground truth*, es decir, la valoración explícita que el usuario otorga mediante el sistema de estrellas (1 – 5), antes de aplicar cualquier inferencia algorítmica. Este análisis no solo caracteriza la naturaleza de la muestra y el impacto del canal de recolección, sino que actúa como una prueba inicial para validar la hipótesis del sesgo por incentivo económico.

4.2.1 EL ECOSISTEMA INCENTIVADO: EL CASO IGRAAL

La **Figura 2** muestra la distribución de sentimientos en la plataforma iGraal, basada en las 5.264 reseñas recolectadas. Como se puede observar, existe una hegemonía casi absoluta de las valoraciones positivas.

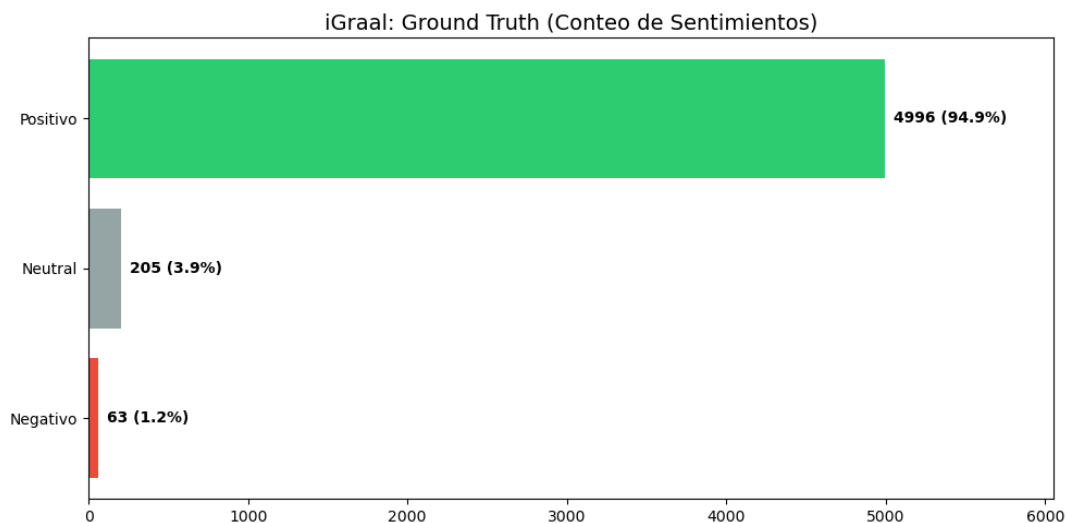


Figura 2: Distribución de sentimientos en iGraal basada en puntuación de estrellas

Fuente: Elaboración propia

Los datos revelan un comportamiento extremadamente sesgado: casi el 95% de las reseñas son etiquetadas como positivas (4 – 5 estrellas). Las valoraciones negativas y neutrales son estadísticamente irrelevantes en este entorno. Este fenómeno, a priori, confirma la hipótesis del sesgo de positividad: en un entorno donde el usuario interactúa con la marca motivado por una recompensa económica (*cashback*), la valoración tiende a convertirse en un trámite “de agradecimiento” o validación del servicio financiero, más que una crítica honesta del producto adquirido. El usuario “premia” la devolución del dinero con 5 estrellas, invisibilizando cualquier posible defecto en la experiencia de compra.

4.2.2 EL ECOSISTEMA ORGÁNICO: EL CASO TRUSTPILOT

Por el contrario, al analizar la distribución en Trustpilot en la **Figura 3**, donde la motivación del usuario es intrínseca (expresar una opinión sin recompensa directa), el panorama cambia significativamente.

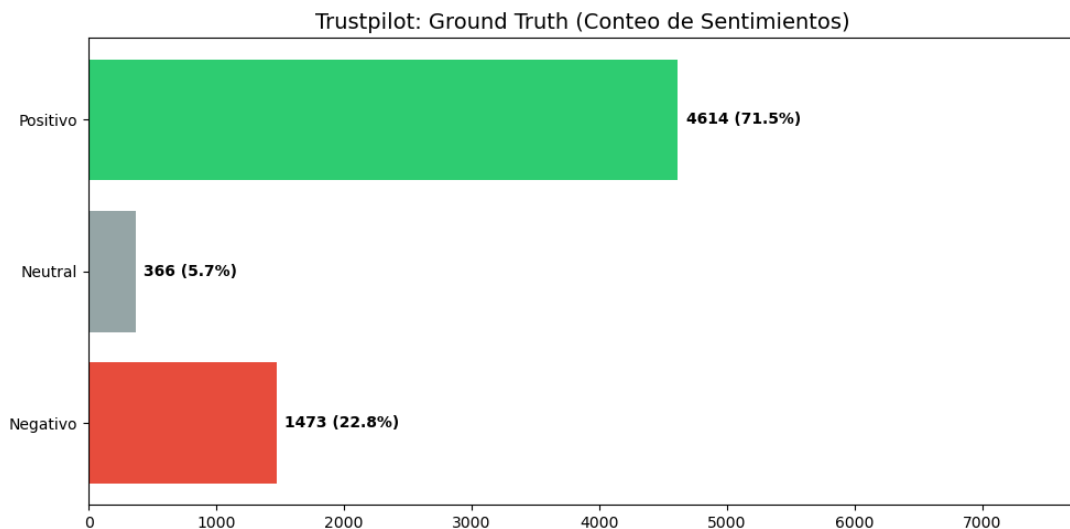


Figura 3: Distribución de sentimientos en Trustpilot basada en puntuación de estrellas

Fuente: Elaboración propia

En esta plataforma, observamos una polarización más realista. Aunque sigue existiendo un volumen considerable de reseñas positivas, emergen con fuerza las categorías “Neutral” y “Negativo”. La presencia significativa de valoraciones de 1 – 2 estrellas demuestra que los usuarios acuden a Trustpilot para visibilizar problemas que en iGraal quedaron silenciados.

4.2.3 CONCLUSIÓN COMPARATIVA: EL IMPACTO DEL INCENTIVO

La yuxtaposición entre ambas figuras anteriores ofrece una de las conclusiones clave de este estudio: el incentivo económico distorsiona la métrica de reputación digital.

Para una marca como Shein, confiar exclusivamente en métricas provenientes de plataformas de fidelización como iGraal generaría una falsa sensación de excelencia, ignorando puntos de dolor críticos que solo afloran cuando se elimina la variable del dinero.

4.2.4 COHERENCIA DE LA INFERENCIA: EL MODELO COMO ESPEJO DEL ECOSISTEMA

Antes de proceder a la auditoría técnica detallada de los algoritmos, es fundamental observar la distribución global de sentimientos proyectada por los modelos de inteligencia artificial. Esta comparativa visual permite validar si la inferencia automatizada es capaz de replicar la naturaleza de los de los datos brutos analizados anteriormente.

4.2.4.1 Ecosistema iGaal

En iGaal, ambos modelos muestran una consistencia casi total con el sesgo de positividad detectado en el *ground truth* de 94,9% (**Figura 2**):

- **BETO**: Identifica un 86,5% de reseñas positivas.
- **RoBERTuito**: Detecta un 86,7% de positividad.

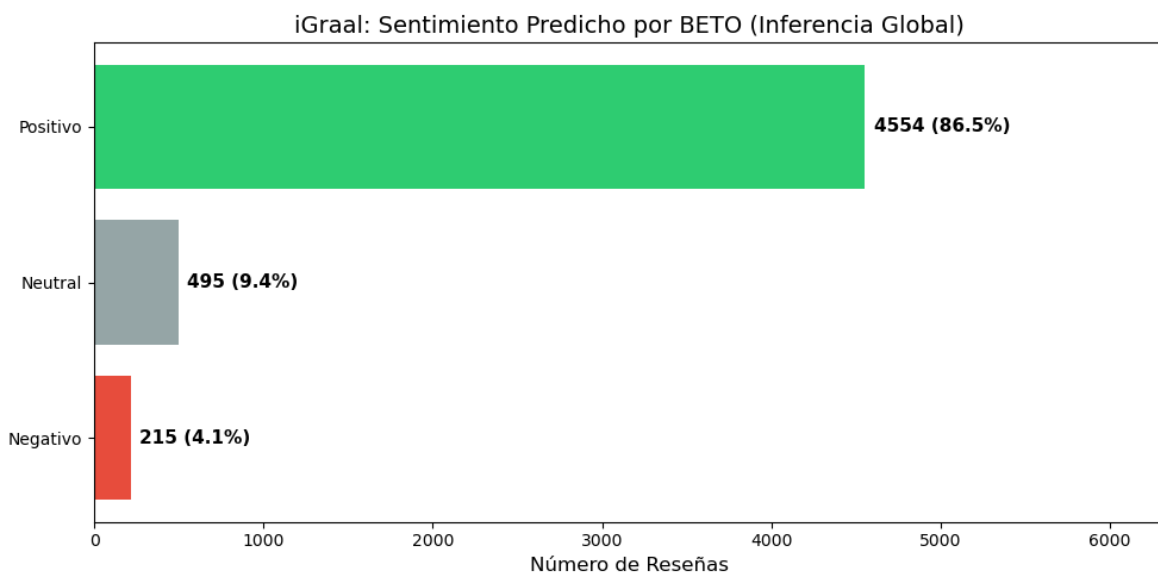


Figura 4: Inferencia global en iGaal (BETO)

Fuente: Elaboración propia

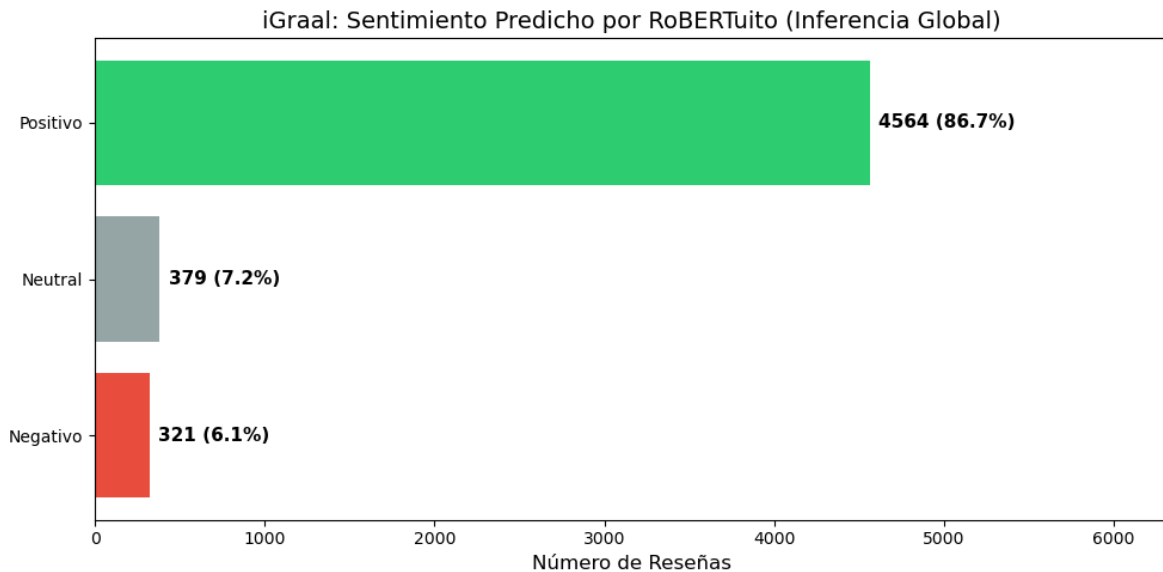


Figura 5: Inferencia global en iGaal (RoBERTuito)

Fuente: Elaboración propia

Estos resultados confirman que la narrativa textual en entornos incentivados está muy polarizada hacia la validación del *cashback*. La inteligencia artificial, al analizar únicamente el lenguaje, corrobora que el discurso del usuario en esta plataforma es, mayoritariamente, de agradecimiento.

4.2.4.2 Ecosistema Trustpilot

Al analizar la reputación de Shein en Trustpilot, los modelos reflejan una apertura hacia el discurso crítico propio de los entornos orgánicos, donde el *ground truth* positivo baja al 71,5% (Figura 3):

- **BETO:** La positividad detectada desciende al 56,7%.
- **RoBERTuito:** Muestra una visión aún más crítica, con un 52,8% de positivos y un 28,9% de negativos.

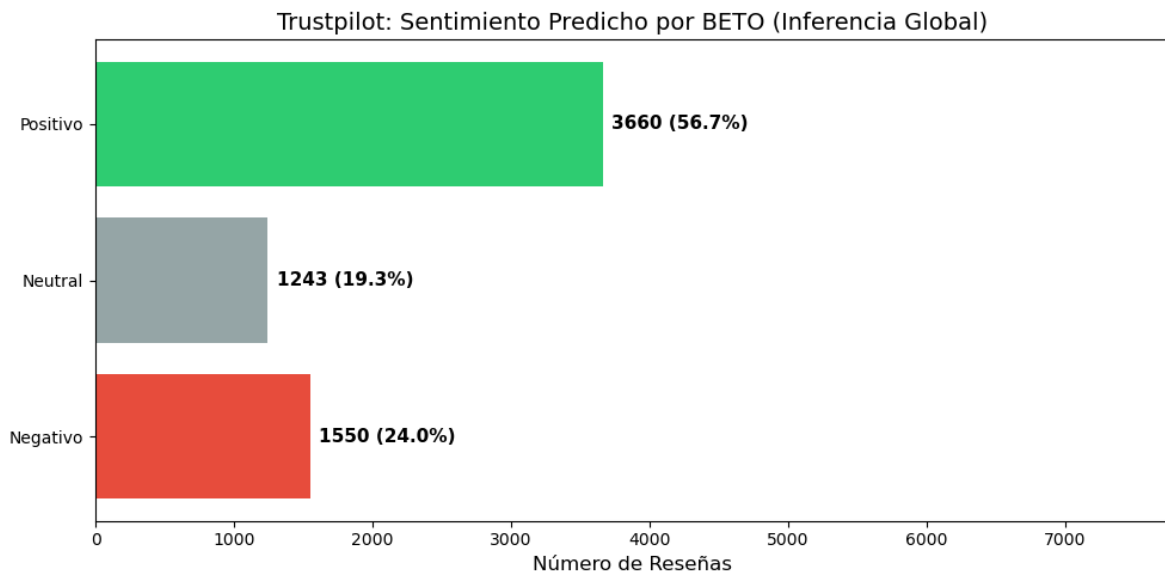


Figura 6: Inferencia Global en Trustpilot (BETO)

Fuente: Elaboración propia

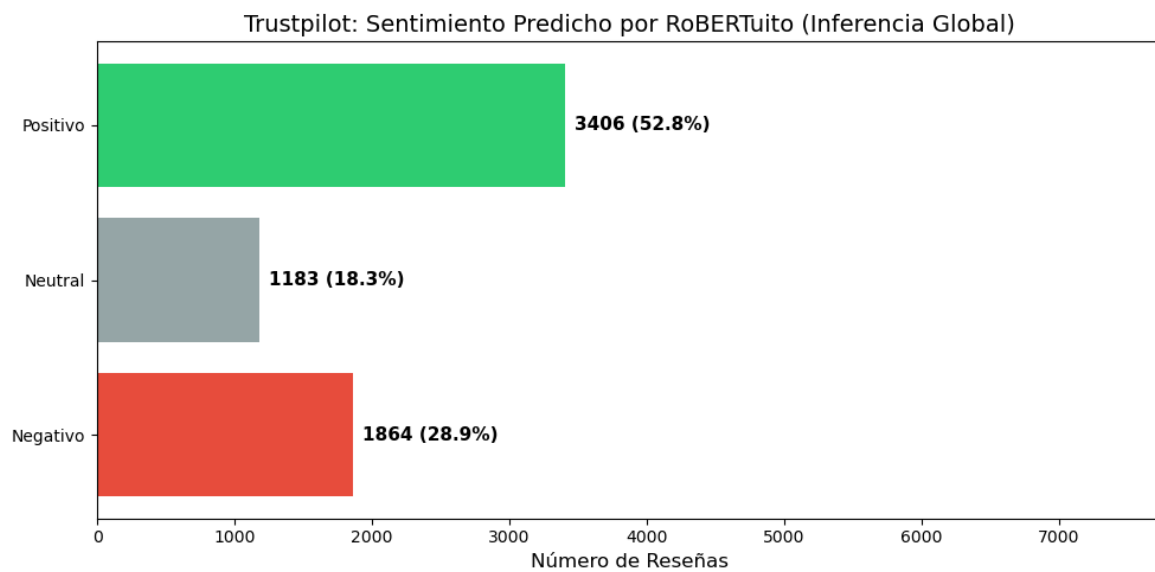


Figura 7: Inferencia Global en Trustpilot (RoBERTuito)

Fuente: Elaboración propia

Esta alineación entre la percepción humana y la interpretación de la máquina demuestra que los modelos capturan fielmente la atmósfera comunicativa de cada plataforma. No obstante, la existencia de discrepancias entre la inferencia algorítmica y la realidad del usuario (que

se explorarán en la auditoría técnica a continuación) justifica que el diagnóstico estratégico posterior se fundamente en la certidumbre del *ground truth* proporcionado por los consumidores.

4.3 AUDITORÍA TÉCNICA: EVALUACIÓN DE MODELOS

Tras establecer el *ground truth* basado en la puntuación de estrellas otorgada por los usuarios y contrastar la distribución proporcional de las categorías (“Positivo”, “Neutral” y “Negativo”) proyectada por la inteligencia artificial en cada plataforma, la siguiente fase de la investigación consistió en evaluar empíricamente la capacidad predictiva de los modelos Transformer seleccionados (BETO y RoBERTuito).

Auditar el rendimiento de un modelo de PLN en plataformas de opinión requiere una comprensión de las métricas de evaluación estadísticas, ya que un análisis superficial puede conducir a conclusiones metodológicas erróneas, especialmente en ecosistemas altamente sesgados como iGraal.

4.3.1 DESBALANCEO DE CLASES Y LA INSUFICIENCIA DE LA EXACTITUD

En la evaluación estándar de modelos de clasificación, la métrica más intuitiva y comúnmente reportada es la exactitud (*accuracy*), la cual mide el porcentaje total de predicciones correctas sobre el total del corpus:

$$Accuracy = \frac{Predicciones\ correctas}{Total\ de\ muestras}$$

Al observar los resultados iniciales de los modelos en la plataforma iGraal, ambos algoritmos presentaban un *accuracy* extraordinariamente alto, superando el 85%. No obstante, en el contexto de esta investigación, esta métrica presenta un sesgo estadístico. Como se demostró en el apartado descriptivo, el corpus de iGraal sufre de un gran desbalanceo, donde casi el 95% de las reseñas están etiquetadas como positivas por los usuarios. En este escenario, si un algoritmo careciera de inteligencia artificial y se limitara a clasificar sistemáticamente el 100% de los textos como reseñas positivas, obtendría automáticamente un 95% de *accuracy*.

Por lo tanto, esta métrica es incapaz de evaluar si el modelo realmente sabe identificar las críticas (la clase minoritaria), que es el objetivo central del análisis de reputación.

4.3.2 ELECCIÓN DEL F1-SCORE MACRO COMO INDICADOR DE FIABILIDAD

Para solventar esta limitación matemática y penalizar a los modelos que ignoran las clases minoritarias, la evaluación técnica de este estudio se fundamenta en la métrica F1-Score en su variante macro.

El F1-Score es la medida armónica entre dos conceptos fundamentales en la ciencia de datos:

- **Precisión (*precision*):** Mide la calidad de la predicción y penaliza los falsos positivos (FP) frente a los verdaderos positivos (VP).

$$Precision = \frac{VP}{VP + FP}$$

- **Sensibilidad (*recall*):** Mide la exhaustividad y penaliza los falsos negativos (FN) frente a los verdaderos positivos (VP).

$$Recall = \frac{VP}{VP + FN}$$

Al aplicar la variante macro, el sistema calcula el F1-Score de forma independiente para cada una de las tres clases y luego extrae la media aritmética simple. Esto obliga a la IA a ser igual de precisa detectando quejas que detectando halagos.

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

Al aplicar esta métrica, el rendimiento real de los modelos descendió a valores cercanos a 0,45 en iGraal y 0,58 en Trustpilot, desvelando la dificultad real del ejercicio, debido principalmente a la dificultad para detectar la clase “Neutral” correctamente.

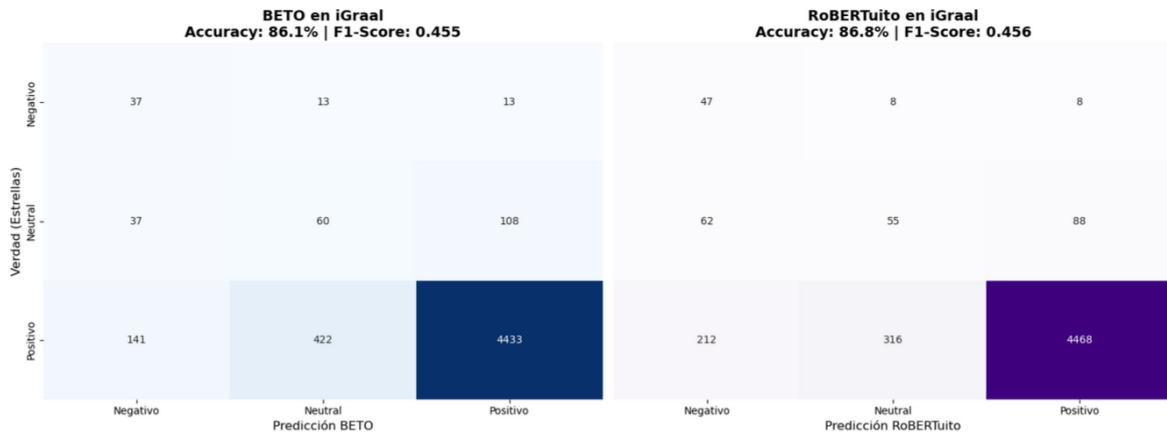


Figura 8: Comparativa de matrices de confusión por modelo en iGraal

Fuente: Elaboración propia

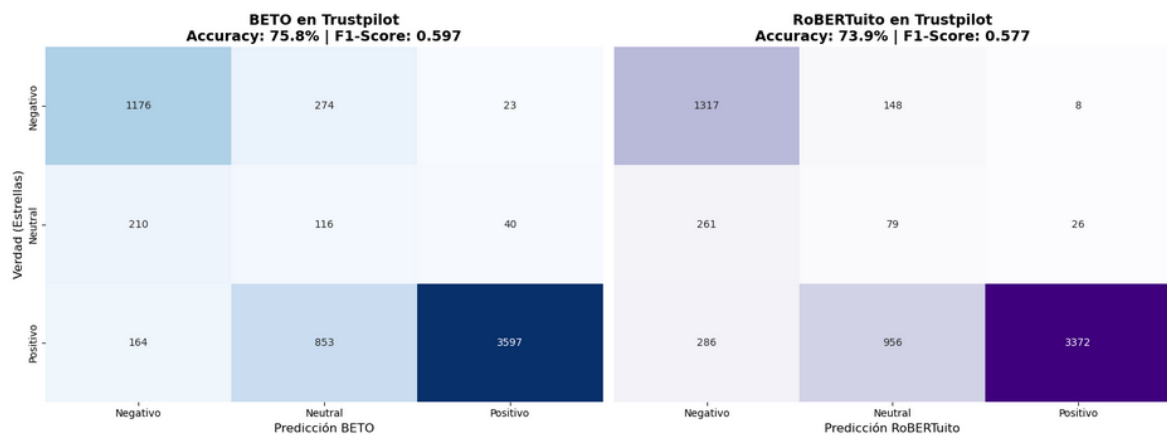


Figura 9: Comparativa de matrices de confusión por modelo en Trustpilot

Fuente: Elaboración propia

4.3.3 ANÁLISIS CUALITATIVO DE DISCREPANCIAS ENTRE USUARIO Y MODELOS

Al analizar las matrices de confusión (Figura 8 y Figura 9), se observaron discrepancias significativas entre la valoración numérica otorgada por el usuario y la inferencia realizada por los modelos. Resultó de especial interés analítico el subconjunto de registros en los que, a pesar de contar con un *ground truth* de 4 – 5 estrellas, tanto BETO como RoBERTuito coincidieron en una predicción de sentimiento negativo.

Para comprender la raíz de esta contradicción entre la puntuación y la carga semántica del texto, se procedió a realizar una auditoría manual sobre esos casos de aparente divergencia.

El análisis de 100 casos aislados en iGraal reveló que este fenómeno no responde a un error predictivo unidimensional, sino que evidencia dos factores críticos: las limitaciones intrínsecas del PLN y la disonancia cognitiva del consumidor en ecosistemas incentivados.

A) Limitaciones algorítmicas

En primer lugar, se constató que las arquitecturas Transformer aún presentan vulnerabilidades al procesar la riqueza morfológica y las expresiones idiomáticas del español. Un porcentaje considerable de las discrepancias corresponde a errores legítimos de la IA, que fracasa al interpretar dobles negaciones o ironías. Un ejemplo representativo del dataset que ilustra este sesgo algorítmico es el siguiente:



SHEIN cada vez ofrece más variedad de productos (ropa, cosmética, etc) a unos precios increíbles y la calidad no tiene nada que envidiarle a muchas tiendas de renombre que están en nuestro país desde hace décadas. El concepto de "producto chino" como producto de baja calidad ha desaparecido gracias a SHEIN.

Figura 10: Ground truth: 5 estrellas vs Predicción IA: Negativo

Fuente: iGraal

En este caso, el usuario está redactando un elogio absoluto. Sin embargo, la acumulación de términos con fuerte valencia negativa (“no”, nada”, “envidiarle”, “baja”, “desaparecido”) satura la capacidad de los modelos para capturar la polaridad global de la frase, provocando un falso negativo técnico.

B) Disonancia cognitiva

En contraposición, otra parte de las discrepancias audita a favor de los modelos de IA, revelando que el “fallo” no reside en el algoritmo, sino en el comportamiento transaccional del usuario motivado por el *cashback*. Se identificaron dos subtipos de comportamiento anómalo que los modelos detectaron correctamente como no positivos:

- a) **Queja silenciada:** Usuarios que puntúan con 5 estrellas para asegurar la validación de su recompensa, pero utilizan el texto libre para exponer puntos de fricción reales, sin comentar nada positivo de Shein explícitamente.



Desde que descubrí esta app siempre compro a Shein a través de ella. Como única pega es que Shein no valida el cashback de las compras y tengo que esperar 3 meses a que se valide automáticamente, pero no me importa esperar y mientras voy acumulando

Figura 11: Ground truth: 5 estrellas vs Predicción IA: Negativo

Fuente: iGraal

- b) **Generación de texto aleatorio para relleno:** Un hallazgo cualitativo relevante ha sido la detección de patrones de escritura automatizada o pseudoaleatoria (como por el uso de texto predictivo de los smartphones). Este fenómeno se observa especialmente en perfiles de usuarios recurrentes que, ante la obligación de alcanzar un mínimo de caracteres para percibir el incentivo económico de iGraal, optan por introducir cadenas de texto inconexas. Estos registros se han mantenido en el corpus de estudio por dos razones:

- Han permitido evaluar la robustez de los algoritmos ante ruido semántico puro, observándose que la IA tiende a clasificar estas estructuras caóticas como “Negativo” o “Neutral” al no detectar marcadores de satisfacción.
- Confirman que el modelo de recompensa de la plataforma iGraal altera no solo la polaridad de la reseña, sino la calidad lingüística de la misma, forzando la creación de contenido sin valor informativo para el comercio, pero funcional para el usuario.



Todo bien

Todo bien. La idea es arriesgar. La idea es arriesgar no merecen la pena que te toque el lunes a dado un ataque de ansiedad que me fuera de plazo y que te dijo que te toque el lunes

Figura 12: Ground-truth: 5 estrellas vs Predicción IA: Negativo

Fuente: iGraal

4.3.4 RENDIMIENTO COMPARADO: POR CLASE Y LONGITUD DEL TEXTO

Al evaluar las métricas cuantitativas globales (*accuracy* y F1-Score), los datos revelan que no existe una superioridad absoluta y transversal de un modelo frente al otro. Por el contrario, BETO y RoBERTuito presentan rendimientos muy parejos, lo que confirma la robustez de ambas arquitecturas Transformer para el análisis de UGC. La verdadera diferenciación entre ambos no radica en proclamar un ganador global, sino en observar cómo fluctúa su eficacia predictiva dependiendo de dos variables clave: la clase semántica a la que se enfrenten y la longitud del texto analizado.

Para ilustrar esta variación, que responde directamente a la naturaleza de su preentrenamiento (*Domain Adaptation*), se generó una comparativa del porcentaje de acierto de cada modelo segmentado por clase y por plataforma.

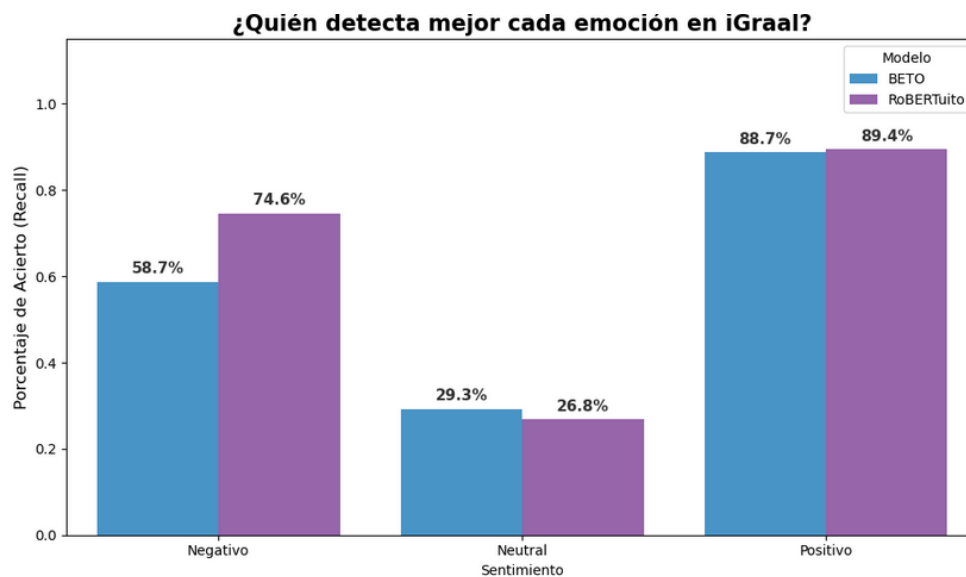


Figura 13: Comparativa de sensibilidad (*recall*) por clase en iGraal

Fuente: Elaboración propia

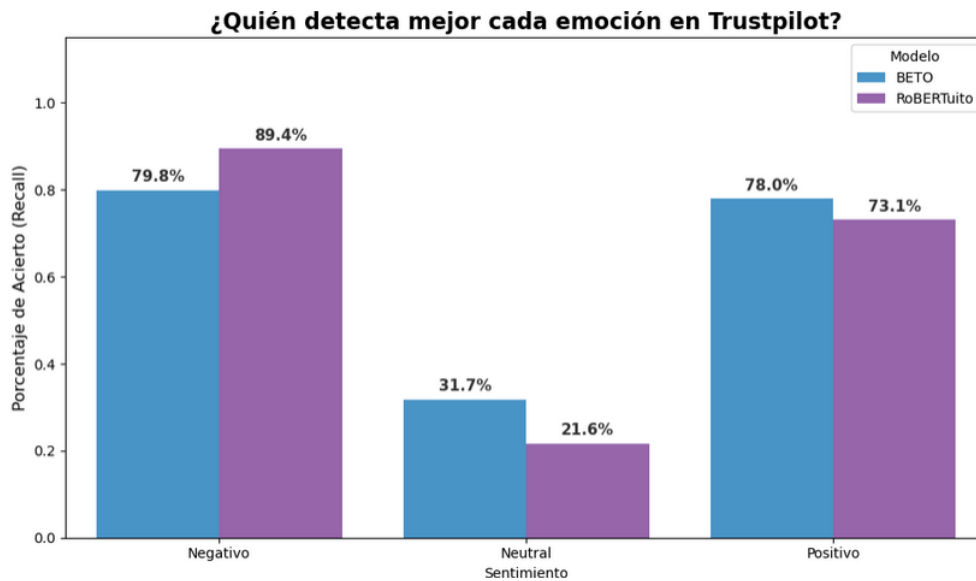


Figura 14: Comparativa de sensibilidad (recall) por clase en Trustpilot

Fuente: Elaboración propia

La especialización arquitectónica condiciona drásticamente los resultados en dos grandes vertientes que se desarrollan a continuación.

4.3.4.1 Superioridad de RoBERTuito en la queja directa

RoBERTuito, optimizado con lenguaje de redes sociales, destaca en los extremos de la polarización. Al intentar detectar las opiniones etiquetadas como “Negativo” muestra un rendimiento superior. En el entorno altamente sesgado de iGraal, logra un *recall* del 74,6%, (frente al 58,7% de BETO), y en Trustpilot escala hasta un 89,4%. Su entrenamiento con millones de interacciones informales le confiere una sensibilidad para detectar la queja cuando esta se manifiesta a través de jerga, insultos, ironía, neologismos o faltas de ortografía, barreras lingüísticas donde un modelo más académico como BETO comete más errores.

4.3.4.2 El valor de la estructura en BETO

Por otro lado, al evaluar las zonas grises (clase “Neutral”), la tendencia se invierte a favor de BETO, superando con creces a RoBERTuito en Trustpilot (31,7% vs 21,6%). Esto se

puede explicar porque las reseñas neutrales suelen ser textos argumentativos con sintaxis compleja y un balance fino entre aspectos positivos y negativos.

Por otro lado, para comprender la complejidad del reto lingüístico, se introdujo una variable adicional: el esfuerzo cognitivo (representado por la longitud de las reseñas).

Esfuerzo Cognitivo: Longitud del Comentario vs Sentimiento en iGraal

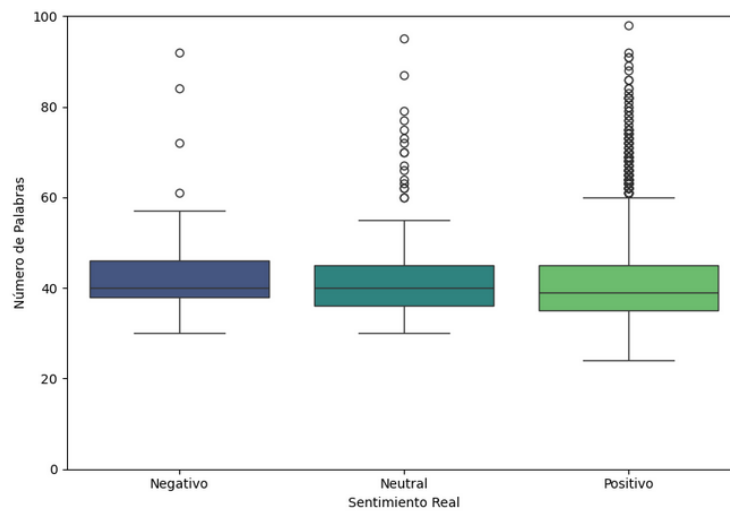


Figura 15: Longitud del comentario frente al ground truth en iGraal

Fuente: Elaboración propia

Esfuerzo Cognitivo: Longitud del Comentario vs Sentimiento en Trustpilot

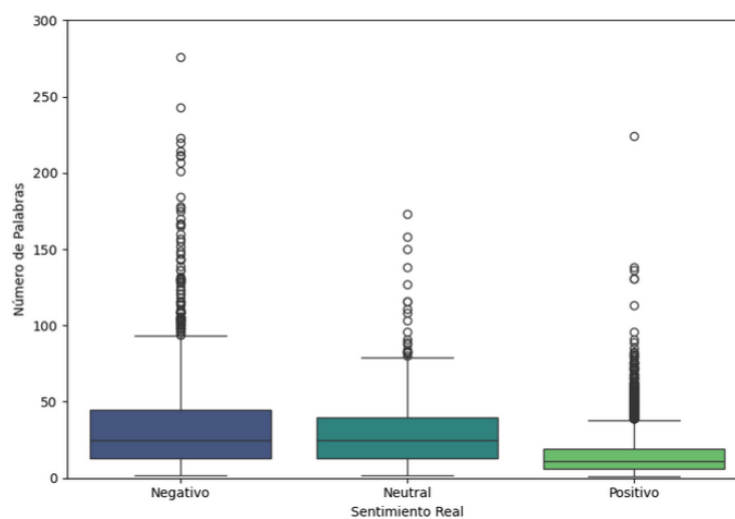


Figura 16: Longitud del comentario frente al ground truth en Trustpilot

Fuente: Elaboración propia

Este análisis descriptivo de los datos revela que el comportamiento del consumidor no es uniforme. En iGraal (**Figura 15**), domina la homogeneidad (en torno a 40 palabras), donde el incentivo económico motiva reseñas breves para la transacción rápida. Por el contrario, en Trustpilot (**Figura 16**), el comportamiento es orgánico: el esfuerzo cognitivo se dispara en las reseñas negativas, donde se encuentran los relatos más extensos, detallados y complejos.

Al cruzar este esfuerzo cognitivo con la capacidad de detección (*recall*) de cada modelo (**Figura 13** y **Figura 14**), emerge una paradoja reveladora que define la complementariedad de los modelos en este trabajo. Teóricamente, se esperaría que BETO, una arquitectura BERT con *Whole Word Masking* entrenada para procesar documentos formales, presentase una mayor resiliencia conforme aumenta la extensión del texto. Sin embargo, los resultados demuestran que las reseñas negativas de Trustpilot, aunque largas, no son “documentos formales”, sino relatos coloquiales para expresar frustración (**Figura 16**). En este escenario, es RoBERTuito (una variante de RoBERTa optimizada para lenguaje de redes sociales) quien maximiza su rendimiento y domina la clase “Negativo” de Trustpilot, incluso en textos extensos. Esta aparente contradicción se explica por su entrenamiento basado en micro mensajes coloquiales, que le permite ser mucho más sensible a los marcadores emocionales informales, independientemente de la longitud total.

4.3.4.3 Conclusión de uso de los modelos

No existe un modelo superior de forma absoluta, sino que sus arquitecturas son especialistas contextuales. Ambos modelos, por tanto, se complementan entre sí.

- **BETO (contexto formal):** Ofrece un análisis sintáctico más estable y conservador, ideal para reseñas largas y argumentadas donde la gramática es clave para no perder el hilo de la polaridad.
- **RoBERTuito (contexto coloquial):** Actúa como un radar más ágil para capturar el “ruido emocional”: la frustración extrema, la queja y la crítica orgánica, en textos informales, coloquiales y directos, incluso cuando se extienden en longitud.

Como balance de esta auditoría, se concluye que el PLN es una herramienta de apoyo masivo excepcional, pero en ecosistemas de alta polaridad como Shein, la valoración numérica del usuario (*ground truth*) sigue siendo la métrica de referencia más fiable para el diagnóstico estratégico de negocio. Por consiguiente, se establece dicha valoración como el eje de referencia para el desarrollo de los análisis cualitativos presentados en los bloques sucesivos de este capítulo.

4.4 IMPACTO DEL INCENTIVO ECONÓMICO EN EL DISCURSO

Una vez auditada la capacidad de los modelos predictivos, el análisis cualitativo se centró en la dimensión semántica. El objetivo de esta fase fue doble: evidenciar cómo el ecosistema de la plataforma altera el vocabulario del usuario y desvelar cuáles son los verdaderos puntos de fricción o palancas de satisfacción.

Como se planteó en la hipótesis inicial, el sistema de recompensas condiciona el discurso. Para validar empíricamente hasta qué punto el incentivo económico modifica las prioridades comunicativas del usuario en iGraal, se aplicó un análisis jerárquico que abarcó desde los conceptos aislados (nubes de palabras) hasta las estructuras semánticas mínimas (bigramas).

4.4.1 ANÁLISIS CONCEPTUAL: EL DESPLAZAMIENTO DE LA MARCA

Como primer paso, se generaron nubes de palabras (*word clouds*) segmentadas por la puntuación exacta de estrellas. Esta visualización pondera el tamaño de cada término en función de su frecuencia absoluta de aparición, permitiendo un diagnóstico visual inmediato de las temáticas dominantes.

4.4.2 ANÁLISIS SINTÁCTICO: LA POLARIZACIÓN DE LA TRANSACCIÓN

Para confirmar a nivel sintáctico lo detectado en los conceptos sueltos, se generó un análisis de bigramas (pares de palabras adyacentes). Al observar cómo se unen las palabras, se pudo decodificar el foco real de la queja o el elogio, revelando una interesante asimetría en el comportamiento del usuario.

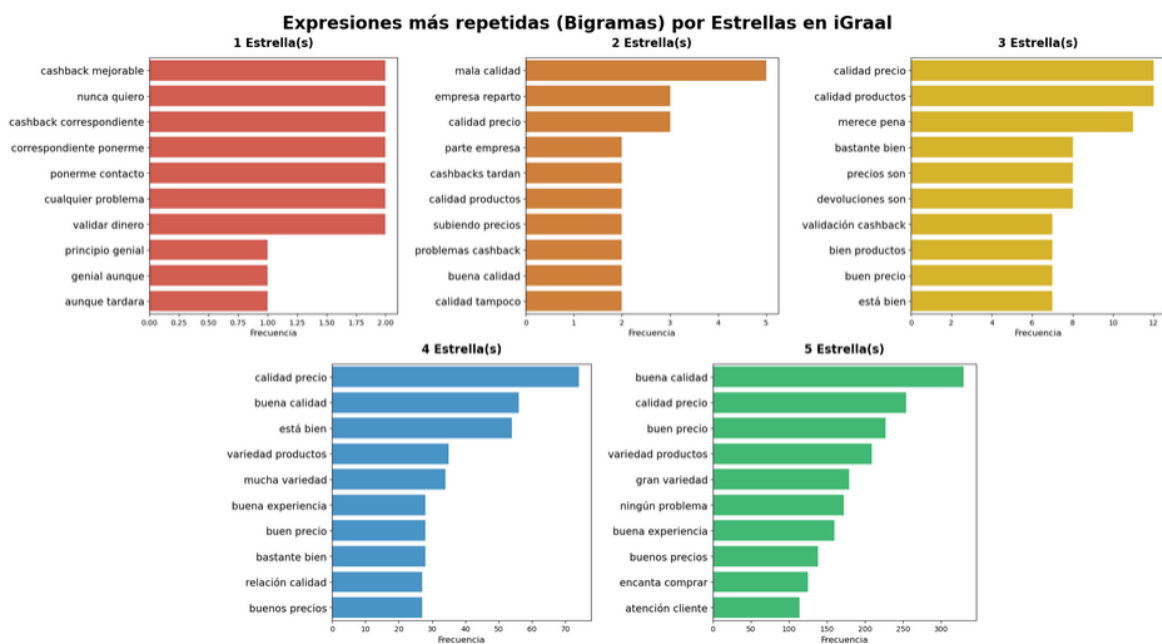


Figura 18: Expresiones más frecuentes (bigramas) por estrellas en iGraal

Fuente: Elaboración propia

Como se evidencia en la **Figura 18**, el incentivo económico no anula la opinión sobre la ropa, pero sí genera un sesgo asimétrico dependiendo del grado de satisfacción:

- **Insatisfacción (1 – 2 estrellas):** En las valoraciones más bajas, la crítica al producto está presente, como demuestra el bigrama “mala calidad” en las 2 estrellas. Sin embargo, esta insatisfacción con la prenda se mezcla y a menudo se ve superada por la frustración operativa. Expresiones sobre retrasos en la validación, problemas con el saldo o denegaciones del *cashback* demuestran que el usuario, en la mayoría de los casos, no está decepcionado con la ropa sino con el proceso de cobro.

- **Neutralidad (3 estrellas):** Actúa como zona de transición, donde el usuario comienza a estructurar opiniones duales, sopesando una buena “calidad precio” frente a los tiempos de espera de la plataforma.
- **Satisfacción (4 – 5 estrellas):** Es en la zona de máxima satisfacción donde la marca Shein recupera el protagonismo. A diferencia del sesgo negativo, cuando la experiencia general es buena el léxico principal vuelve a centrarse en el producto (con expresiones dominantes como “buena calidad” o “gran variedad”). A simple vista, este análisis sintáctico sugeriría que el usuario está valorando orgánicamente su experiencia con la moda de Shein.

4.4.3 LA PERSPECTIVA LOVEBRAND

En el marco de esta investigación, cruzar los datos sintácticos (bigramas) con el análisis conceptual previo (nubes de palabras) resulta clave para evaluar el verdadero estatus de Shein como *Lovebrand*.

Por un lado, el análisis de bigramas parece apuntar a la consolidación de una *Lovebrand*, mostrando a usuarios teóricamente muy satisfechos que utilizan expresiones como “buena calidad” o “buenos precios”. Sin embargo, las nubes de palabras advertían de que los términos aislados más repetidos seguían siendo financieros.

Para resolver esta dualidad y dimensionar el impacto real del incentivo, se procedió a cuantificar cuántas de estas valoraciones positivas estaban contaminadas por el vocabulario transaccional.

Proporción de Promotores Transaccionales (Falsos Positivos) en reseñas de 4 y 5 Estrellas (iGraal)

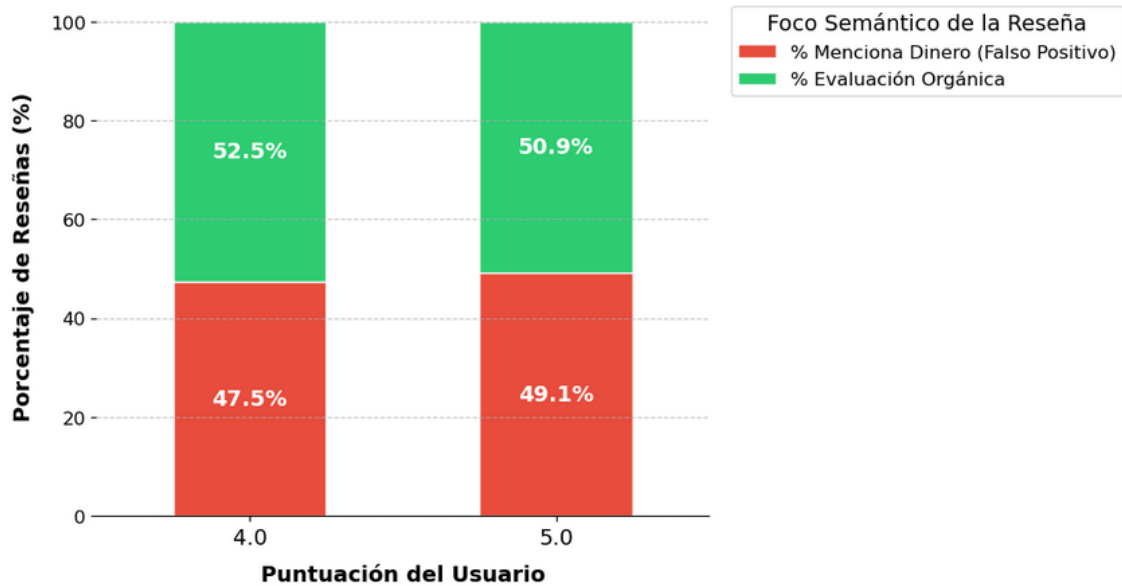


Figura 19: Proporción de promotores transaccionales (falsos positivos) en valoraciones altas de iGraal

Fuente: Elaboración propia

Como revela la **Figura 19**, casi la mitad de las reseñas más altas (47,5% en las de 4 estrellas y 49,1% en las de 5 estrellas) mencionan explícitamente términos vinculados a la recompensa económica (“cashback”, “dinero”, “saldo”, “descuento”, ...).

Este hallazgo empírico demuestra que una parte considerable de estos comentarios son, en realidad, falsos positivos en términos de lealtad de marca; no nos encontramos ante verdaderos defensores de marca (*brand advocates*) impulsados por una conexión emocional y genuina con Shein, sino con “promotores transaccionales”: usuarios cuya motivación principal es validar su cashback.

En definitiva, el ecosistema de iGraal infla las métricas de vanidad de la marca, creando la ilusión de una *Lovebrand* que se sostiene, en casi un 50% de sus valoraciones de excelencia, por un interés puramente económico y no por el valor real del producto.

4.5 REPUTACIÓN ORGÁNICA Y CONSTRUCCIÓN DE LA LOVEBRAND

Una vez demostrado empíricamente cómo el incentivo económico de iGraal condiciona la opinión del usuario, se plantea el análisis del corpus de Trustpilot como el “grupo de control” orgánico de esta investigación.

En este entorno, libre de recompensas financieras, la intención de la reseña vuelve a su origen natural: advertir a otros consumidores de un problema o recomendar genuinamente la marca. Para desgranar este discurso y garantizar una comparativa rigurosa, se ha replicado la misma metodología analítica de categorización semántica aplicada en el ecosistema incentivado.

4.5.1 EL RETORNO AL PRODUCTO Y AL SERVICIO

Como primer paso, se procedió a replicar la extracción jerárquica de nubes de palabras y bigramas para observar cuál es la base del diálogo orgánico.

Trustpilot: Palabras más frecuentes por Estrellas



Figura 20: Nubes de palabras más frecuentes por estrellas en Trustpilot

Fuente: Elaboración propia

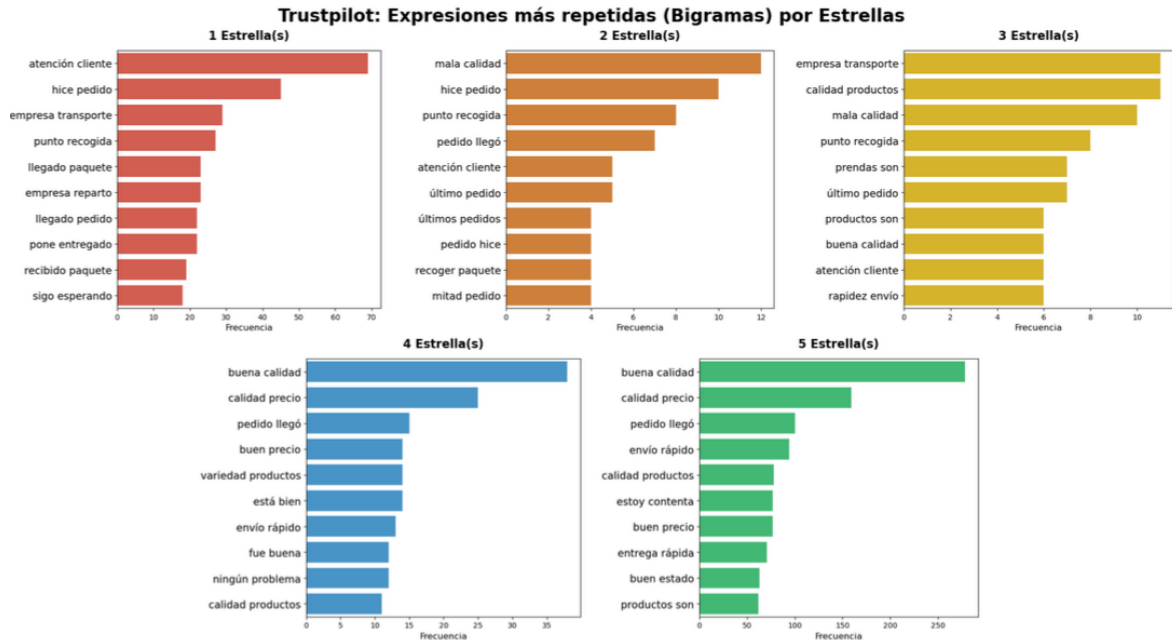


Figura 21: Bigramas por estrellas en Trustpilot

Fuente: Elaboración propia

Al observar la **Figura 20** y la **Figura 21**, el contraste con el ecosistema iGraal es absoluto. El dinero y las palabras relacionadas con el *cashback* desaparecen por completo del vocabulario. En su lugar, el léxico recobra su utilidad práctica para el sector del e-commerce:

- En el extremo de la **insatisfacción (1 – 2 estrellas)**, los conceptos y bigramas revelan a un consumidor frustrado por problemas operativos: “atención cliente”, “llegado pedido”, “sigo esperando”.
- Por el contrario, en la zona de **satisfacción (4 – 5 estrellas)**, el producto recupera su protagonismo: “buena calidad”, “calidad precio” o “entrega rápida”. Este es el lenguaje natural de un consumidor evaluando una marca de moda, no la eficiencia de una transacción bancaria.

4.5.2 RADIOGRAFÍA DE LA EXPERIENCIA: FRICCIONES VS ACIERTOS

Para profundizar en el análisis, se ha aplicado un algoritmo de categorización sobre el corpus verificado por el *ground truth* de Trustpilot, mediante diccionarios de términos, permitiendo

cuantificar los *pain points* (puntos de dolor) entre los detractores que puntuaron con 1 – 2 estrellas, y los *praise points* (puntos fuertes o de alabanza) entre los promotores que puntuaron con 4 – 5 estrellas.

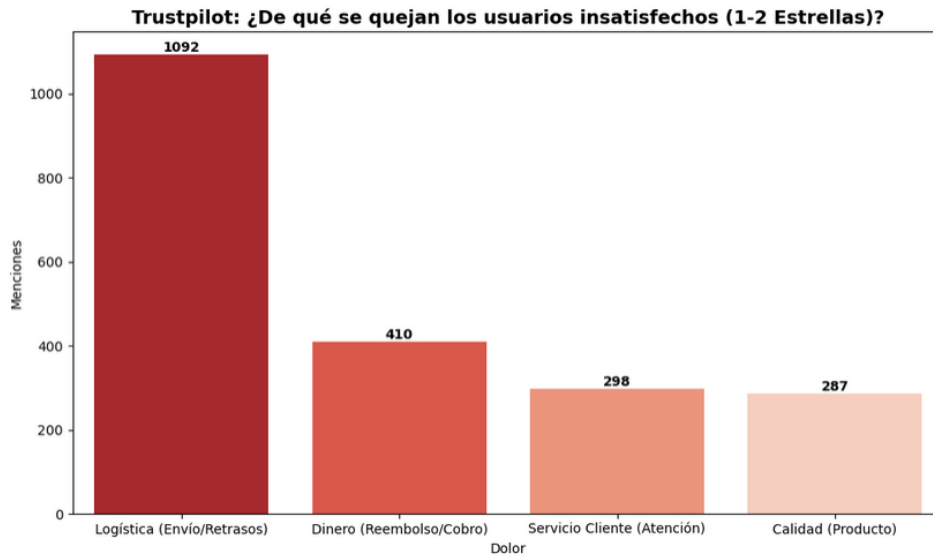


Figura 22: Frecuencia de puntos de dolor en reseñas negativas

Fuente: Elaboración propia

Los resultados de la insatisfacción mostrados en la **Figura 22** revelan que el gran “talón de Aquiles” de Shein es la logística, que con 1.092 menciones se posiciona como la queja principal. Le siguen los problemas relacionados con el dinero y reembolsos (410 menciones) y el servicio al cliente (298 menciones). Resulta revelador que las quejas sobre la calidad del producto ocupen el último lugar (287 menciones), lo que indica que el consumidor perdona una deficiencia textil, pero no la pérdida de su paquete o dinero. Este dato sugiere que el consumidor de Shein posee una tolerancia mayor a la baja calidad o fallos en el textil (posiblemente debido al bajo precio percibido), pero se muestra mucho más inflexible ante los fallos en la cadena de suministro o, en segundo lugar, en la gestión de su dinero.

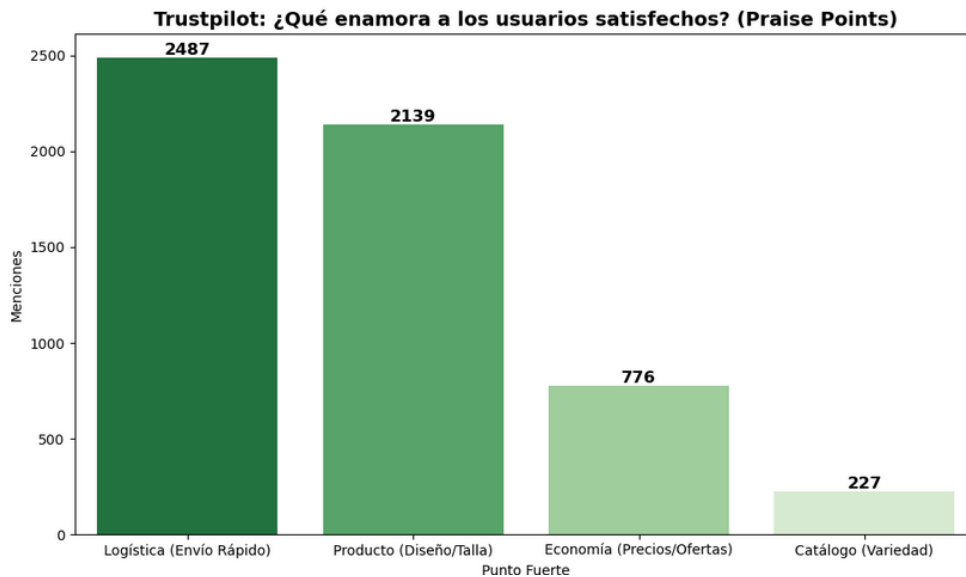


Figura 23: Frecuencia de puntos fuertes en reseñas positivas

Fuente: Elaboración propia

La centralidad logística se repite de forma simétrica en los aciertos. La logística en este caso, entendida como envíos rápidos, es el factor más alabado con 2.487 menciones, superando incluso al producto (2.139 menciones). Estos datos demuestran que la reputación de Shein no se construye únicamente sobre la ropa, sino sobre la eficiencia del servicio: el usuario premia la rapidez por encima de la economía (776 menciones) o la variedad del catálogo (227 menciones).

4.5.3 EL VEREDICTO LOVEBRAND

Tras identificar los puntos de dolor y halago, la investigación exigía determinar si Shein ha logrado generar vínculos emocionales genuinos en ausencia de incentivos económicos. Para ello, se desarrolló un modelo de categorización emocional basado en un diccionario léxico expandido que clasifica los testimonios positivos (4 – 5 estrellas) en cuatro dimensiones críticas de la teoría de Roberts: “Amor/Pasión”, “Felicidad/Alegría”, “Confianza/Seguridad” y “Anticipación/Sorpresa”.

Primero, se cuantificó el volumen de reseñas que activaban cada emoción (**Figura 24**) y, posteriormente, se generó un mosaico semántico (**Figura 25**) para visualizar de qué hablan exactamente los usuarios cuando manifiestan dichas emociones.

Las figuras anteriores ofrecen una perspectiva matizada sobre el estatus de Shein como *Lovebrand*. A diferencia de lo que sugería el análisis de iGraal, el entorno orgánico revela que la conexión emocional se basa, ante todo, en la Felicidad / Alegría (661 menciones) y la Confianza / Seguridad (639 menciones).

4.5.3.1 Análisis cualitativo de la lealtad

Al observar los mapas mentales de la **Figura 25**, se pueden extraer las siguientes conclusiones sobre la naturaleza de este “amor” y “respeto” por la marca:

- **El Respeto logístico (Confianza / Seguridad):** El análisis cualitativo de la nube verde confirma que el eje del “respeto” de Shein no es ético ni institucional, sino operativo. Los términos dominantes son “siempre”, “rápido”, “pedido”, “entrega”. El consumidor confía en la marca porque la logística funciona; la seguridad se deriva de la certidumbre de recibir el paquete en el tiempo esperado.
- **La gratificación estética (Amor / Pasión):** En la dimensión de pasión (nube rosa), los términos rotan hacia el producto: “encanta”, “precioso”, “ropa”. Aquí se observa que el vínculo afectivo nace de la satisfacción con el diseño y la estética de la prenda recibida, cumpliendo con el pilar de “sensualidad” de Roberts.
- **La adicción dopaminérgica (Felicidad / Alegría):** La nube naranja destaca por adjetivos como “genial”, “contenta” y “perfecto”. Esta categoría refleja el éxito del modelo de *ultra-fast fashion* como una experiencia recreativa: la felicidad no proviene solo de poseer la ropa, sino de la culminación exitosa de la transacción.
- **La ilusión del “haul” (Anticipación / Sorpresa):** La nube azul revela una dimensión psicológica clave: la gratificación por la espera. Palabras como “expectativas”, “llegó” y “gracias” demuestran que parte de la lealtad emocional reside en el deseo previo a la recepción del pedido. Shein logra capitalizar la anticipación del consumidor, convirtiendo el tiempo de envío en un elemento de “misterio” y deseo.

En conclusión, Shein ha logrado trascender la etiqueta de “tienda de bajo coste” para posicionarse como una marca de fiabilidad basada en el cumplimiento de expectativas. Sin

embargo, no se debe perder de vista que esta relación de confianza es extremadamente frágil: el usuario premia la eficiencia operativa con sentimientos de seguridad, pero no muestra tolerancia ante los fallos en la entrega. Así, la logística se confirma como el eje vertebrador sobre el cual bascula tanto la percepción de la *Lovebrand* como su mayor riesgo reputacional.

4.6 LA DISONANCIA ÉTICA Y LA CEGUERA DEL CONSUMIDOR

Después de haber analizado las dimensiones de satisfacción operativa y lealtad de marca, la investigación abordó un último factor crítico en la reputación corporativa actual: el impacto social y ambiental de Shein. Se planteó como hipótesis que, a pesar de las constantes controversias legislativas y mediáticas sobre el modelo de producción del *fast fashion*, el consumidor mantiene una “ceguera ética” casi total en el momento de evaluar públicamente su experiencia de compra.

Para verificar este fenómeno y medir su alcance real, se diseñó un algoritmo de extracción léxica basado en expresiones regulares, aislando menciones específicas sobre ética laboral (“explotación”, “moral”, “derechos laborales”, ...) y medioambiente (“sostenibilidad”, “contaminación”, “residuos”, “huella de carbono”, ...).

Con el objetivo de identificar si existía un debate generalizado o un interés minoritario, se contrastó el volumen total de palabras empleadas frente al número de usuarios únicos que las redactaron, midiendo así la concentración del discurso.

4.6.1 ANÁLISIS CUALITATIVO DEL DISCURSO

Para profundizar en la naturaleza de esta “ceguera ética” y comprender el contexto real en el que se producen estas escasas menciones, se llevó a cabo una extracción cualitativa mediante nubes de palabras basadas en frecuencias exactas. Esta técnica permite visualizar qué términos específicos lideran el escaso debate social en ambas plataformas.

iGraal: Conceptos sobre "Ética Laboral y Derechos"

Total Menciones: 4 | Comentarios Únicos: 3
(de un total de 5264 valoraciones)



Sobreexplotados
Discriminan
Moral

Figura 26: Frecuencias exactas de términos vinculados a la ética laboral en iGraal

Fuente: Elaboración propia

Trustpilot: Conceptos sobre "Ética Laboral y Derechos"

Total Menciones: 1 | Comentarios Únicos: 1
(de un total de 6453 valoraciones)



Explotación laboral

Figura 27: Frecuencias exactas de términos vinculados a la ética laboral en Trustpilot

Fuente: Elaboración propia

Al observar la **Figura 26** y la **Figura 27**, correspondientes a la dimensión de ética laboral, la principal conclusión es la ausencia casi total de discurso. En iGraal, el gráfico refleja un vacío evidente donde apenas logran formarse tres conceptos aislados (“moral”, “discriminan”, “sobreexplotados”). En Trustpilot, el escenario es aún más desolador: aflora un único término (“explotación laboral”). El consumidor, de forma sistemática, omite las condiciones de producción y derechos de sus trabajadores al evaluar su nivel de satisfacción con la marca.

Sin embargo, el hallazgo más revelador de esta investigación se produce al analizar la dimensión medioambiental.

iGraal: Conceptos sobre "Sostenibilidad y Medioambiente"

Total Menciones: 13 | Comentarios Únicos: 11
(de un total de 5264 valoraciones)

Medio ambiente
Plástico
Ecológico
Sostenible
Huella ecológica

Figura 28: Frecuencias exactas de términos vinculados a la sostenibilidad y medioambiente en iGraal

Fuente: Elaboración propia

Trustpilot: Conceptos sobre "Sostenibilidad y Medioambiente"

Total Menciones: 27 | Comentarios Únicos: 20
(de un total de 6453 valoraciones)

Plástico
Basura

Figura 29: Frecuencias exactas de términos vinculados a la sostenibilidad y medioambiente en Trustpilot

Fuente: Elaboración propia

Cuantitativamente, la preocupación ambiental parecía acumular un mayor volumen de menciones. No obstante, aunque en iGraal sí se ha detectado una fracción residual de comentarios genuinos que critican directamente la “huella ecológica” de la firma, la representación semántica destapa un espejismo en los datos. El concepto dominante en ambas plataformas (“plástico”) no denota un activismo ecológico por parte del usuario, sino que está profundamente ligado a la experiencia material de la compra.

La revisión cualitativa del corpus demuestra que el consumidor emplea el léxico ambiental como un adjetivo calificativo o despectivo para evaluar el producto, no el impacto corporativo. De este modo, el término “plástico” se utiliza frecuentemente para describir las bolsas en las que vienen doblados los artículos, mientras que “basura” se emplea para catalogar un artículo que no cumple las expectativas. El usuario vacía así estas palabras de su significado ecológico, utilizándolas como meros indicadores de insatisfacción material.

4.6.2 CONCLUSIÓN DE LA DIMENSIÓN ÉTICA

La triangulación de los datos cuantitativos (concentración del discurso en una minoría absoluta) y el análisis cualitativo (la resignificación del léxico ambiental para quejarse de la calidad textil) confirman empíricamente la hipótesis planteada en este estudio.

Shein opera bajo un paradigma de disonancia ética total por parte de su público. El consumidor, seducido por precios muy competitivos e impulsado por un sistema logístico rápido, seguro y confiable, desvincula por completo su relación de marca (*Lovebrand*) de las responsabilidades morales, laborales y ambientales de la corporación. La ética no es un factor que penalice la reputación pública de Shein porque, sencillamente, es una variable que el usuario ha normalizado y excluido de su proceso de evaluación.

Capítulo 5. DISCUSIÓN Y CONCLUSIONES

Este capítulo constituye la síntesis final de la investigación, donde se procede a la interpretación crítica de los hallazgos obtenidos tras el procesamiento de 11.717 reseñas. El propósito de este apartado no es meramente reiterar las métricas expuestas en el capítulo anterior, sino trascender la estadística descriptiva para ofrecer una respuesta integral a la paradoja planteada al inicio de este estudio: la coexistencia de una lealtad de marca ciega (*Lovebrand*) frente a un modelo productivo éticamente comprometido.

A continuación, se estructura la discusión en torno a la validación de los objetivos estratégicos del proyecto, contrastando la literatura teórica sobre el comportamiento del consumidor digital con la realidad empírica detectada en los ecosistemas de iGraal y Trustpilot. Se abordará el impacto distorsionador del incentivo económico en la reputación corporativa, la eficacia de las arquitecturas de inteligencia artificial para el análisis de sentimiento y, finalmente, se profundizará en el fenómeno de la “ceguera ética” como pilar fundamental de la experiencia de consumo en Shein, en el marco del *ultra-fast fashion*.

5.1 LA ANATOMÍA DE UNA LOVEMARK DIGITAL

A lo largo de esta investigación, se ha podido contrastar cómo el canal de comunicación y su arquitectura de incentivos actúan como filtros determinantes de la realidad reputacional de la corporación. La comparativa entre las dos plataformas analizadas revela una marca con una gran capacidad de fidelización, capaz de generar vínculos emocionales tanto en entornos incentivados como orgánicos.

5.1.1 LA DUALIDAD DEL ECOSISTEMA INCENTIVADO: EL CASO DE IGRAAL

El análisis de iGraal revela una hegemonía de valoraciones positivas del 94,9% (**Figura 2**). Si bien este dato inicial está influenciado por el sesgo de reciprocidad derivado del *cashback*, tras una mirada en detalle a la semántica del usuario, conviene matizar esta conclusión.

Como se observa en la auditoría realizada sobre la proporción de promotores (**Figura 19**), existe una división muy balanceada del discurso: mientras que un 47,5% – 49,1% de las reseñas se mencionan términos financieros y pueden considerarse “promotores transaccionales”, existe un sólido 50,9% – 52,5% de reseñas con puntuación de 5 y 4 estrellas, respectivamente, que realizan una evaluación orgánica. Este hallazgo es fundamental: demuestra que, incluso en un entorno donde el usuario tiene un interés económico directo, más de la mitad de los compradores satisfechos eligen utilizar su espacio de opinión para alabar genuinamente a Shein como marca de moda. Por tanto, iGraal no solo funciona como un operador financiero, sino que sirve de altavoz para una base de amantes de la marca verdaderos que conviven con el interés del descuento.

5.1.2 EL VEREDICTO DE LA REPUTACIÓN ORGÁNICA: EL CASO DE TRUSTPILOT

El análisis del corpus de Trustpilot actúa como la validación definitiva del estatus de Shein. En este entorno, libre de incentivos monetarios inmediatos, la marca mantiene una tasa de positividad del 71,5% (**Figura 3**). Esta cifra, obtenida en un escenario típicamente crítico, es un indicador robusto de que Shein es una marca que trasciende lo meramente funcional.

A diferencia de otras firmas de bajo coste que solo inspiran valoraciones sobre el ahorro, el análisis de emociones “*Lovebrand*” en Trustpilot (**Figura 24**) revela que Shein ha logrado conquistar los dos ejes fundamentales de la teoría de Kevin Roberts: el amor y el respeto.

- **El Eje del Respeto (Confianza y Seguridad):** Con 636 menciones, la “Confianza / Seguridad” es el segundo pilar emocional más potente detectado. En el marco de las *Lovemarks*, el respeto se gana mediante el cumplimiento de la promesa de la marca. Los datos demuestran que Shein ha logrado este “Respeto institucional” no a través de sus integridad ética o medioambiental (dimensiones que, como se ha visto, son ignoradas), sino mediante la fiabilidad operativa y logística excepcional.
- **El Eje del Amor (Felicidad y Pasión):** La emoción dominante es la “Felicidad / Alegría” (661 menciones), complementada por un volumen significativo de “Amor / Pasión” (383 menciones). Esto confirma que Shein no es percibida como una simple

opción racional por precio, sino como una marca que genera una gratificación emocional intensa. Los usuarios utilizan un léxico apasionado que demuestra una “lealtad más allá de la razón”.

5.1.3 CONSTRUCCIÓN DE LOS PILARES: MISTERIO, SENSUALIDAD, INTIMIDAD

Tras el análisis semántico y emocional, es posible afirmar que Shein constituye su identidad digital reforzando la triada clásica de Roberts de una forma innovadora y adaptada al ecosistema del e-commerce digital:

- **Misterio:** Este pilar se alimenta de la gamificación. El flujo constante de miles de novedades diarias, junto a las ofertas *flash*, convierte el catálogo en un entorno de descubrimiento constante. Un hallazgo cualitativo importante en esta investigación ha sido la dimensión de “Anticipación / Sorpresa”, que revela que el consumidor disfruta del deseo previo y de la “expectativa” de recibir el paquete. Shein ha convertido así el tiempo de envío en una narrativa de misterio.
- **Sensualidad:** En el entorno digital, el pilar de la sensualidad se traduce en una interfaz visual saturada y altamente reactiva. La aplicación de Shein está diseñada para maximizar la estimulación sensorial a través de la pantalla. El proceso de “scroll infinito” se convierte en una experiencia gratificante y adictiva para el consumidor, que llega a priorizar la estimulación visual del producto sobre cualquier otra consideración ética o técnica.
- **Intimidad:** Lograda mediante la personalización algorítmica y la presencia móvil. Al incentivar para que el usuario invierta “trabajo” en la marca (como escribir reseñas, subir fotos reales de sus compras o participar en el programa de puntos), Shein crea un ecosistema cerrado de alta intimidad. Esto genera un sentido de pertenencia a una comunidad global que blindo a la marca frente a las críticas externas, convirtiendo al consumidor en un defensor que protege su relación con la marca.

En resumen, Shein se consolida como una “*Lovemark* de la era digital”, ya que ha logrado un equilibrio entre un alto **Respeto** (basado en la infalibilidad logística) y un alto **Amor**

(basado en la gratificación dopaminérgica). Esta dualidad permite a la marca operar bajo un paradigma de invulnerabilidad reputacional, donde la eficiencia del servicio actúa como escudo ante cualquier cuestionamiento moral.

5.2 EVALUACIÓN DE LOS OBJETIVOS ESPECÍFICOS

Tras la discusión de los hallazgos semánticos y emocionales, es conveniente evaluar el grado de consecución de los objetivos específicos planteados al inicio de esta investigación.

- I. **Desarrollo de un corpus robusto mediante técnicas de *web scraping*:** Se ha logrado consolidar un dataset final de 11.717 registros, superando ampliamente la masa crítica necesaria para garantizar la representatividad estadística para el estudio. La combinación de fuentes heterogéneas (iGraal y Trustpilot) ha sido fundamental para aislar variables psicológicas y económicas. El protocolo de limpieza garantizó la integridad de la ingesta por parte de los modelos de deep learning.
- II. **Evaluación del impacto de los incentivos económicos:** Este objetivo se ha cumplido con un alto grado de precisión. Se ha demostrado cuantitativamente que el modelo de recompensa de iGraal genera una positividad artificial. Sin embargo, la mayor aportación a este punto ha sido la capacidad de auditar este sentimiento mediante el uso de diccionarios léxicos, verificándose que cerca de la mitad de las valoraciones de excelencia están contaminadas por intereses transaccionales, confirmando que el incentivo financiero desplaza la conversación de marca hacia la validación del beneficio económico.
- III. **Determinación del sentimiento real y puntos de fricción y de alabanza:** A través del análisis del grupo de control en Trustpilot, se ha identificado la verdadera polaridad emocional del consumidor de Shein. Si bien se empleó la arquitectura Transformer como herramienta de auditoría técnica para contrastar la capacidad de la IA en la detección de matices, el diagnóstico final de los puntos de fricción y los de alabanza se ha basado en el *ground truth* de los datos. El estudio revela que la logística es el factor determinante de la experiencia: con 1.092 menciones negativas frente a 2.487 positivas, se confirma como el eje vertebrador de la satisfacción. Se

ha logrado desmitificar la idea de que Shein es amada solo por el precio, demostrando que la fiabilidad operativa juega el papel principal en sostener su reputación orgánica.

- IV. **Verificación de la brecha ética:** Se ha confirmado empíricamente que la ética corporativa es un factor absolutamente marginal en la evaluación poscompra del consumidor español. El uso de nubes de palabras permitió detectar que el usuario ha normalizado las controversias externas.

5.3 APORTACIONES Y CONCLUSIONES FINALES

Como conclusión general de este Trabajo de Fin de Grado, se sostiene que Shein representa un cambio de paradigma en la construcción de marcas de culto. Este estudio ha demostrado que es posible generar niveles de lealtad y seguridad en un entorno de disonancia ética sistémica. La marca ha logrado que el placer de la gratificación inmediata y la fiabilidad de su cadena de suministro actúen como escudos ante las violaciones de los ODS y las advertencias legislativas internacionales.

Desde una perspectiva técnica esta investigación pone a prueba el uso de modelos Transformer monolingües, revelando que, aunque arquitecturas como RoBERTuito muestran una mayor sensibilidad al lenguaje informal, la complejidad del discurso del consumidor (ironía, contradicciones y sesgos incentivados) todavía presenta desafíos significativos para una inferencia totalmente autónoma. Esto subraya la importancia de combinar la potencia del Big Data con la supervisión del analista para garantizar la precisión en el Business Analytics.

5.4 LIMITACIONES Y LÍNEAS FUTURAS DE INVESTIGACIÓN

A pesar del rigor metodológico aplicado y del volumen del corpus analizado, esta investigación presenta ciertas limitaciones que deben ser consideradas para la interpretación de los resultados y que, a su vez, puedan abrir vías para futuros estudios en el ámbito del Business Analytics.

5.4.1 LIMITACIONES DEL ESTUDIO

- **Naturaleza transversal del análisis:** Este trabajo ofrece una fotografía fija del sentimiento del consumidor en un periodo determinado (2020–2026). Al no ser un estudio longitudinal, no se puede observar con precisión cómo evoluciona la “ceguera ética” a medida que surgen nuevos escándalos o legislaciones.
- **Sesgo geográfico y cultural:** El estudio se ha centrado exclusivamente en el consumidor español. Los resultados podrían variar significativamente en mercados con una conciencia ambiental más arraigada o con legislaciones más restrictivas, como es el caso de Francia.
- **Limitación del formato UGC:** El análisis se ha restringido a reseñas textuales en plataformas de terceros. Se ha dejado fuera el impacto de los formatos visuales y efímeros (como vídeos de TikTok o de Instagram), que son el motor principal del fenómeno *haul* y de la validación social de la Generación Z.

5.4.2 LÍNEAS DE INVESTIGACIÓN FUTURAS

- **Análisis comparado internacional:** Sería de un gran valor académico replicar esta metodología de PLN comparada para analizar si la implementación de penalizaciones éticas y ecológicas en otros países europeos logra modificar efectivamente el léxico y las prioridades del consumidor, o si la ceguera ética es un fenómeno global intrínseco al modelo de consumo digital de Shein.
- **Integración del análisis multimodal:** Futuras investigaciones podrían integrar modelos de visión artificial para analizar los sentimientos expresados en los vídeos de “*Shein haul*”, comparando la narrativa visual (emociones positivas al mostrar la ropa) con la narrativa de las quejas logísticas.
- **Estudio de la persistencia de la lealtad:** Dada la fragilidad detectada en la relación de “Confianza” basada en la logística, una línea futura de interés sería investigar el umbral de tolerancia del consumidor: ¿cuántos fallos en la entrega o de qué magnitud son necesarios para que el “Respeto” y el “Amor” por la marca se desvanezcan?

Capítulo 6. BIBLIOGRAFÍA

- [1] A. Shuv-Ami, «A new scale of brand lovemarks».
- [2] Kevin Roberts, «Future Beyond Brands», SaatchiKevin.com. Accedido: 8 de enero de 2026. [En línea]. Disponible en: <https://saatchikevin.com/lovemarks/future-beyond-brands/>
- [3] Kevin Roberts, «Loyalty Beyond Reason», SaatchiKevin.com. Accedido: 8 de enero de 2026. [En línea]. Disponible en: <https://saatchikevin.com/speech/loyalty-beyond-reason-2008/>
- [4] A. K. Abu Farha, K. El Hedhli, I. Alnawas, H. Zourrig, y I. Becheur, «Drivers and outcomes of a shopper-retailer's app relationship», *Journal of Retailing and Consumer Services*, vol. 81, nov. 2024, doi: 10.1016/j.jretconser.2024.104002.
- [5] «Shein Revenue and Usage Stats», Backlinko. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://backlinko.com/shein-stats>
- [6] D. Curry, «Shein Revenue and Usage Statistics (2026)», Business of Apps. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://www.businessofapps.com/data/shein-statistics/>
- [7] «State of Mobile 2025».
- [8] «SHAME ON YOU, SHEIN HAZARDOUS CHEMICALS STILL FOUND IN FAST FASHION PRODUCTS».
- [9] «Follow-up investigation in China refutes Shein's promise of improvement», Public Eye. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://www.publiceye.ch/en/media-corner/press-releases/detail/follow-up-investigation-in-china-refutes-sheins-promise-of-improvement>

- [10] O. Classen y D. Hachfeld, «Interviews with factory employees refute Shein's promises to make improvements», Public Eye. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://www.publiceye.ch/en/topics/fashion/interviews-with-factory-employees-refute-sheins-promises-to-make-improvements>
- [11] L. Lüttin, «[Textiles] ESPR: How the Ecodesign for Sustainable Products Regulation will impact apparel and footwear brands», Carbonfact. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://www.carbonfact.com/blog/policy/espr-textile>
- [12] «Sustainable and Circular Textiles Strategy», European Commission. Accedido: 9 de enero de 2026. [En línea]. Disponible en: https://environment.ec.europa.eu/strategy/textiles-strategy_en
- [13] «European strategy to corner ultra-fast fashion», Ecomondo. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://www.ecomondo.com/en/news-detail/european-strategy-to-corner-ultra-fast-fashion?newsId=563571>
- [14] A. Charron, «Mind the Gap: Understanding Why Gen Z's Sustainability Values Do Not Align with Fashion Purchasing Behavior. Exploring the Do Not Align with Fashion Purchasing Behavior. Exploring the Factors that Contribute to Values-Behavior Gap Factors that Contribute to Values-Behavior Gap», Bryant University, 2025.
- [15] C. Tetik y O. Cherradi, «Attitude-Behavior Gap in sustainable clothing consumption», Jönköping University, 2020.
- [16] C. Martínez, «De Kevin Roberts a las Lovemarks: marcas que levantan pasiones», Puro Marketing. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://www.puromarketing.com/44/13142/kevin-roberts-lovemarks-marcas-levantan-pasiones>
- [17] «Love Brand», BrandTrust. Accedido: 12 de enero de 2026. [En línea]. Disponible en: <https://www.brand-trust.de/en/glossary/love-brand>

- [18] «SHEIN's Secret Weapon: How Ecommerce Gamification Turned Shoppers into Content Creators (And How You Can Too!)», Faisco. Accedido: 9 de enero de 2026. [En línea]. Disponible en: <https://faisco.com/marketing/2862.html>
- [19] M. Wijngaarden, «Psychology of Fast Fashion: Why You Can('t) Stop Buying It», ProjectCece. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.projectcece.com/blog/656/psychology-of-fast-fashion/>
- [20] «The Psychology of Fast Fashion: Why Conversation About Fast Fashion Evokes Such Strong Emotions In Us », The Sustainable Fashion Forum. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.thesustainablefashionforum.com/pages/the-psychology-of-fast-fashion-exploring-the-complex-emotions-fast-fashion-evokes-in-consumers>
- [21] Q. Guo, «From a Penny to Self-Disgust: How Cashback Rewards for Favorable Comments and Relationship Norms Affect Consumers' Post-Purchase Behavioral Intentions», *Journal of Theoretical and Applied Electronic Commerce Research* , vol. 19, n.º 2, pp. 1079-1094, jun. 2024, doi: 10.3390/jtaer19020056.
- [22] H. Worland, «Inside Shein: The disruption and opportunities behind real-time retail», *NEA*, may 2024, Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.nea.com/blog/shein-real-time-retail-success>
- [23] «Taking the shine off SHEIN: Hazardous chemicals in SHEIN products break EU regulations, new report finds», Greenpeace International. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.greenpeace.org/international/press-release/56979/taking-the-shine-off-shein-hazardous-chemicals-in-shein-products-break-eu-regulations-new-report-finds/>
- [24] L. Deppen, «Shein factory employees work 75-hour weeks for low wages: report», *Manufacturing Dive*. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.manufacturingdive.com/news/shein-public-eye-factory-workers-hours-wages/716430/>

- [25] T. Alonso, «Escándalo, multitudes e intervención del Gobierno: la rocambolesca llegada de Shein a París», *Modaes*. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.modaes.com/back-stage/escandalo-multitudes-e-intervencion-del-gobierno-la-rocambolesca-llegada-de-shein-a-paris>
- [26] «Francia aprueba ley para frenar el fast fashion.», *The Latin Issue*. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.thelatinissue.com/post/francia-aprueba-ley-para-frenar-el-fast-fashion#:~:text=El%20Parlamento%20franc%C3%A9s%20ha%20aprobado,la%20producci%C3%B3n%20de%20transporte%20y%20composici%C3%B3n.>
- [27] A. Pfeiffer, «‘A paroxysm of disposability’: France’s distaste for Shein’s ultra-fast fashion», *The Guardian*. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://www.theguardian.com/business/2025/oct/02/a-paroxysm-of-disposability-frances-distaste-for-sheins-ultra-fast-fashion>
- [28] W. Westwood y J. T. Auskalis, «Ethical Elegance: Navigating Consumer Morality and Production Practices in Fashion Through Analyzing Shein, Zara, & Vivienne Westwood». [En línea]. Disponible en: https://research.library.fordham.edu/environ_2015
- [29] Y. Li, «The Analysis of the Marketing Strategy of Shein and Accompanying Word of Mouth on Social Media», 2022.
- [30] «The power of loyalty schemes in the fashion industry», *VTEX*. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://vtex.com/en-us/blog/power-of-loyalty-schemes-fashion/>
- [31] R. Breen, «The Evil Fashion Giant: Shein and the Perpetuation of Overconsumption», *The University of Kansas*. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://opentext.ku.edu/environmentalgeopolitics/chapter/evil-fashion-giant-shein/>

- [32] K. Dillard, «Beneath the SHEIN Sheen: The Horrors of Ultra-Fast Fashion and Modern Garment Work», The Footnote. Accedido: 15 de enero de 2026. [En línea]. Disponible en: <https://the-footnote.org/2025/03/11/beneath-the-shein-sheen-the-horrors-of-ultra-fast-fashion-and-modern-garment-work/>
- [33] R. Hernández-Pérez, P. Lara-Martínez, B. Obregón-Quintana, L. S. Liebovitch, y L. Guzmán-Vargas, «Correlations and Fractality in Sentence-Level Sentiment Analysis Based on VADER for Literary Texts», *Information*, vol. 15, n.º 11, p. 698, nov. 2024, doi: 10.3390/info15110698.
- [34] R. A. Potamias, G. Siolas, y A.-G. Stafylopatis, «A transformer-based approach to irony and sarcasm detection», *Neural Comput. Appl.*, vol. 32, n.º 23, pp. 17309-17320, dic. 2020, doi: 10.1007/s00521-020-05102-3.
- [35] «Cuáles son los modelos de NLP (Procesamiento del Lenguaje Natural) aplicados al castellano», Hack a Boss. Accedido: 20 de enero de 2026. [En línea]. Disponible en: <https://www.hackaboss.com/blog/cuales-son-los-modelos-de-nlp-procesamiento-del-lenguaje-natural-aplicados-al-castellano>
- [36] C. Song, S. Chen, X. Cai, y H. Chen, «Sentiment Analysis of Spanish Political Party Tweets Using Pre-trained Language Models», nov. 2024.
- [37] S. Guan, «Instant Data Scraper Review and the Best Alternative for 2025», Thunderbit. Accedido: 20 de enero de 2026. [En línea]. Disponible en: <https://thunderbit.com/blog/instant-data-scraper-review-and-alternative>
- [38] R. Pan, J. A. García-Díaz, F. Garcia-Sanchez, y R. Valencia-García, «Evaluation of transformer models for financial targeted sentiment analysis in Spanish», *PeerJ Comput. Sci.*, vol. 9, 2023, doi: 10.7717/peerj-cs.1377.
- [39] J. Cañete, G. Chaperon, R. Fuentes, J.-H. Ho, H. Kang, y J. Pérez, «Spanish Pre-trained BERT Model and Evaluation Data», ago. 2023.

- [40] A. Gutiérrez-Fandiño *et al.*, «MarIA: Spanish Language Models», abr. 2022, doi: 10.26342/2022-68-3.
- [41] J. M. Pérez *et al.*, «pysentimiento: A Python Toolkit for Opinion Mining and Social NLP tasks», jul. 2024.

ANEXO I: PIPELINE DE ANÁLISIS DE DATOS

Este anexo detalla el procedimiento técnico seguido durante la investigación. El contenido se estructura siguiendo el ciclo de vida del dato: desde la normalización inicial del corpus, pasando por la implementación de los modelos Transformer, hasta la generación de métricas de rendimiento y visualizaciones semánticas. La presentación del código fuente garantiza la reproducibilidad del estudio y certifica el rigor metodológico aplicado en cada etapa del análisis.

I. Preprocesamiento de datos y generación del *ground truth*

Para garantizar la fiabilidad de la comparación entre la percepción humana y la inferencia algorítmica, se desarrolló un script de normalización y etiquetado. Este proceso permite transformar las valoraciones numéricas (estrellas) en categorías semánticas y limpiar el ruido textual derivado del proceso de *scraping*.

A continuación, se presenta el pipeline unificado aplicado tanto al corpus de iGraal como al de Trustpilot:

```
import pandas as pd
import re

# 1. CONFIGURACIÓN DE RUTAS
# Alternar entre 'iGraal_limpio.xlsx' y 'trustpilot_limpio.xlsx'
archivo_entrada = 'trustpilot_limpio.xlsx'
print(f"Iniciando preprocesamiento de: {archivo_entrada}")

# Carga de datos
df = pd.read_excel(archivo_entrada)

# 2. GENERACIÓN DEL GROUND TRUTH (Etiquetado lógico)
# Clasificamos las estrellas en categorías para la posterior matriz de confusión
def categorizar_sentimiento(estrellas):
    if estrellas <= 2:
        return 'Negativo' # Críticas y detractores
    elif estrellas == 3:
        return 'Neutral' # Indiferentes
    else:
        return 'Positivo' # Promotores (4 y 5 estrellas)

# Aplicación de la lógica (ajustando el nombre de la columna según fuente)
```

```
col_estrellas = 'Estrellas' if 'trustpilot' in archivo_entrada.lower() else
'Estrellas (sobre 5)'
df['ground_truth_sentiment'] = df[col_estrellas].apply(categorizar_sentimiento)

# 3. NORMALIZACIÓN TEXTUAL (Limpieza de ruido)
# Eliminación de espacios redundantes, saltos de línea y normalización de
codificación
def limpiar_texto(texto):
    if not isinstance(texto, str):
        return ""
    # Uso de expresiones regulares para normalizar espacios en blanco
    texto_limpio = re.sub(r'\s+', ' ', texto).strip()
    return texto_limpio

# Identificación de la columna de texto según la fuente
col_texto = 'Comentario'
df['Comentario_limpio'] = df[col_texto].apply(limpiar_texto)

# 4. EXPORTACIÓN DEL DATASET PREPARADO PARA NLP
archivo_salida = archivo_entrada.replace('.xlsx', '_preparado_nlp.csv')
columnas_finales = [col for col in df.columns if col in
                    ['Fecha', 'Fecha_Publicacion', 'Año', col_estrellas,
                    'ground_truth_sentiment', 'Comentario_limpio']]

df[columnas_finales].to_csv(archivo_salida, index=False)

print(f"Dataset normalizado guardado en: {archivo_salida}")
print("\nDistribución de clases para validación:")
print(df['ground_truth_sentiment'].value_counts())
```

II. Implementación técnica y ejecución de los modelos Transformer

Como se ha detallado en la metodología del trabajo, el procesamiento del lenguaje natural (PLN) se ha llevado a cabo mediante un pipeline unificado de ejecución. Dado que la lógica de inferencia, la arquitectura de los modelos y las funciones de normalización son idénticas para ambos conjuntos de datos, se presenta a continuación el código fuente consolidado.

Este script permite procesar tanto el corpus de iGraal como el de Trustpilot, simplemente alternando el archivo de entrada, garantizando así la reproducibilidad y la consistencia técnica de la investigación.

```
import pandas as pd
from transformers import pipeline
from pysentimiento import create_analyzer
from tqdm import tqdm

# Configuración de visualización de pandas
```

```
pd.set_option('display.max_colwidth', 100)

# 1. CARGA DEL CORPUS (Alternar según la fuente analizada)
# iGraal: 'iGraal_limpio_preparado_nlp.csv' | Trustpilot:
# 'trustpilot_limpio_preparado_nlp.csv'
archivo_entrada = 'trustpilot_limpio_preparado_nlp.csv'
df = pd.read_csv(archivo_entrada)
print(f"Iniciando procesamiento de {len(df)} registros...")

# -----
# 2. INICIALIZACIÓN DE MODELOS (BETO y RoBERTuito)
# -----

# MODELO A: BETO (Basado en arquitectura BERT)
# Referencia académica entrenada con fuentes formales.
print("\n--- Cargando Modelo A: BETO (Arquitectura BERT) ---")
beto_analyzer = pipeline("sentiment-analysis", model="finiteautomata/beto-
sentiment-analysis")

# MODELO B: RoBERTuito (Arquitectura RoBERTa - Familia MarIA)
# Optimizado para lenguaje informal, ironía y redes sociales.
print("--- Cargando Modelo B: RoBERTuito (Arquitectura RoBERTa) ---")
robertuito_analyzer = create_analyzer(task="sentiment", lang="es")

# -----
# 3. FUNCIONES DE NORMALIZACIÓN DE ETIQUETAS
# -----
def normalizar_beto(resultado):
    label = resultado['label'] # BETO emplea etiquetas POS/NEG/NEU
    if label == 'POS': return 'Positivo'
    if label == 'NEG': return 'Negativo'
    return 'Neutral'

def normalizar_robertuito(resultado):
    label = resultado.output # Pysentimiento extrae la predicción directa
    (output)
    if label == 'POS': return 'Positivo'
    if label == 'NEG': return 'Negativo'
    return 'Neutral'

# -----
# 4. EJECUCIÓN DEL PIPELINE DE INFERENCIA
# -----
tqdm.pandas() # Activación de barras de progreso para el seguimiento del proceso

print("\n>>> Ejecutando inferencia con BETO...")
# Se trunca el texto a 512 caracteres por limitación nativa de la arquitectura
BERT
df['BETO_Raw'] = df['Comentario_limpio'].progress_apply(
    lambda x: beto_analyzer(str(x)[:512], truncation=True, max_length=512)[0]
)
df['BETO_Sentiment'] = df['BETO_Raw'].apply(normalizar_beto)
```

```
print("\n>>> Ejecutando inferencia con RoBERTuito...")
df['RoBERTuito_Raw'] = df['Comentario_limpio'].progress_apply(
    lambda x: robertuito_analyzer.predict(str(x))
)
df['RoBERTuito_Sentiment'] = df['RoBERTuito_Raw'].apply(normalizar_robertuito)

# -----
# 5. EXPORTACIÓN DE RESULTADOS FINALES
# -----
archivo_final = archivo_entrada.replace('limpio_preparado_nlp',
' analisis_completo_PLN')
df.to_csv(archivo_final, index=False)

print(f"\nProceso finalizado con éxito. Resultados exportados a:
{archivo_final}")
```

III. Generación de visualizaciones y métricas de rendimiento

A. Análisis descriptivo y distribución del *ground truth*

Este script automatiza la representación visual de la distribución de sentimientos original (basada en estrellas) para ambas fuentes del dataset. Esta fase fue fundamental para identificar el desbalanceo de clases, lo cual justifica posteriormente el uso de métricas como el F1-Score en lugar de la exactitud simple.

```
import pandas as pd
import matplotlib.pyplot as plt

# 1. CARGA DE DATOS (Ajustar nombre de archivo según fuente)
archivo = 'trustpilot_limpio_preparado_nlp.csv'
df = pd.read_csv(archivo)

# 2. CONFIGURACIÓN ESTRATÉGICA
# Definimos el orden lógico y los colores corporativos para el estudio
conteo = df['ground_truth_sentiment'].value_counts()
orden = ['Positivo', 'Neutral', 'Negativo']
datos = [conteo.get(k, 0) for k in orden]
colores = ['#2ecc71', '#95a5a6', '#e74c3c']

# 3. GENERACIÓN DEL GRÁFICO DE DISTRIBUCIÓN
plt.figure(figsize=(10, 5))
barras = plt.barh(orden, datos, color=colores)

# Etiquetado dinámico de barras con valores absolutos y porcentajes
total = sum(datos)
for i, barra in enumerate(barras):
    ancho = barra.get_width()
    if total > 0:
        plt.text(ancho + (total * 0.01), barra.get_y() + barra.get_height() / 2,
```

```
f'{ancho} ({{ancho/total:.1%}})',  
va='center', fontweight='bold')  
  
plt.title(f'Distribución del Ground Truth: {archivo.split("_")[0].capitalize()}',  
fontsize=14)  
plt.xlim(0, total * 1.2) # Margen de seguridad para etiquetas  
plt.gca().invert_yaxis() # Positivo en la parte superior  
plt.tight_layout()  
  
# Exportación para la memoria  
plt.savefig(f'distribucion_sentimientos_{archivo.split("_")[0]}.png', dpi=300)  
plt.show()
```

B. Evaluación técnica y validación de los modelos: métricas de rendimiento

En este apartado se procede a la auditoría técnica para determinar la fiabilidad de los modelos. Se implementa el cálculo de matrices de confusión para visualizar la direccionalidad del error y la métrica de *recall* para asegurar la detección efectiva de las quejas (clase negativa).

```
import pandas as pd  
import matplotlib.pyplot as plt  
import seaborn as sns  
from sklearn.metrics import confusion_matrix, accuracy_score, f1_score  
  
# 1. CARGAR DATOS  
# Cambiar aquí: 'iGraal_analisis_completo_PLN.csv' o  
'trustpilot_analisis_completo_PLN.csv'  
archivo = 'Trustpilot_analisis_completo_PLN.csv'  
df = pd.read_csv(archivo)  
  
# Identificar la fuente para el título automático  
fuente = "iGraal" if "igraal" in archivo.lower() else "Trustpilot"  
  
# Definimos el orden lógico: De malo a bueno  
labels = ['Negativo', 'Neutral', 'Positivo']  
  
# 2. CALCULAR MÉTRICAS  
acc_beto = accuracy_score(df['ground_truth_sentiment'], df['BETO_Sentiment'])  
acc_robertuito = accuracy_score(df['ground_truth_sentiment'],  
df['RoBERTuito_Sentiment'])  
  
cm_beto = confusion_matrix(df['ground_truth_sentiment'], df['BETO_Sentiment'],  
labels=labels)  
cm_robertuito = confusion_matrix(df['ground_truth_sentiment'],  
df['RoBERTuito_Sentiment'], labels=labels)  
  
f1_beto = f1_score(df['ground_truth_sentiment'], df['BETO_Sentiment'],  
average='macro')
```

```
f1_robertuito = f1_score(df['ground_truth_sentiment'],
df['RoBERTuito_Sentiment'], average='macro')

# 3. GRÁFICA 1: MATRICES DE CONFUSIÓN COMPARATIVAS
fig, axes = plt.subplots(1, 2, figsize=(16, 6))

# Mapa de Calor BETO (Azules)
sns.heatmap(cm_beto, annot=True, fmt='d', cmap='Blues', xticklabels=labels,
yticklabels=labels, ax=axes[0], cbar=False)
axes[0].set_title(f'BETO en {fuente}\nAccuracy: {acc_beto:.1%} | F1-Score:
{f1_beto:.3f}', fontsize=14, fontweight='bold')
axes[0].set_ylabel('Verdad (Estrellas)', fontsize=12)
axes[0].set_xlabel('Predicción BETO', fontsize=12)

# Mapa de Calor RoBERTuito (Púrpuras)
sns.heatmap(cm_robertuito, annot=True, fmt='d', cmap='Purples',
xticklabels=labels, yticklabels=labels, ax=axes[1], cbar=False)
axes[1].set_title(f'RoBERTuito en {fuente}\nAccuracy: {acc_robertuito:.1%} | F1-
Score: {f1_robertuito:.3f}', fontsize=14, fontweight='bold')
axes[1].set_xlabel('Predicción RoBERTuito', fontsize=12)
axes[1].set_yticks([])

plt.tight_layout()
plt.show()

# 4. GRÁFICA 2: BAR CHART DE ACIERTO POR CATEGORÍA (RECALL)
report = []
for sent in labels:
    subset = df[df['ground_truth_sentiment'] == sent]
    total = len(subset)
    if total > 0:
        acc_b = len(subset[subset['BETO_Sentiment'] == sent]) / total
        acc_r = len(subset[subset['RoBERTuito_Sentiment'] == sent]) / total
    else:
        acc_b, acc_r = 0, 0

    report.append({'Sentimiento': sent, 'Modelo': 'BETO', 'Acierto': acc_b})
    report.append({'Sentimiento': sent, 'Modelo': 'RoBERTuito', 'Acierto':
acc_r})

df_bar = pd.DataFrame(report)

plt.figure(figsize=(10, 6))
# Paleta específica: Azul (#3498db) vs Morado (#9b59b6)
sns.barplot(data=df_bar, x='Sentimiento', y='Acierto', hue='Modelo',
palette={'BETO': '#3498db', 'RoBERTuito': '#9b59b6'})

plt.title(f'¿Quién detecta mejor cada emoción en {fuente}?', fontsize=16,
fontweight='bold')
plt.ylabel('Porcentaje de Acierto (Recall)', fontsize=12)
plt.ylim(0, 1.15)

# Inserción de porcentajes encima de las barras
```

```
for p in plt.gca().patches:
    height = p.get_height()
    if height > 0:
        plt.gca().text(p.get_x() + p.get_width()/2., height + 0.02,
                       f'{height:.1%}', ha="center", fontsize=11, fontweight='bold',
                       color='#333333')

plt.tight_layout()
plt.show()
```

IV. Metodología de análisis semántico avanzado (nubes de palabras)

En este apartado se describe el pipeline de procesamiento semántico diseñado para la exploración cualitativa del corpus objeto de la investigación. El flujo de trabajo propuesto permite trascender el análisis de frecuencias convencional mediante la integración de diccionarios léxicos específicos.

A. Descripción del pipeline de generación semántica

La metodología empleada para la creación de las nubes de palabras se estructura en cuatro fases críticas que garantizan la relevancia de los hallazgos:

1. **Segmentación estratégica:** Se procede a la filtración del dataset original para aislar los segmentos de interés (en este caso, los promotores con valoraciones de 4 – 5 estrellas), permitiendo un análisis granular de la satisfacción.
2. **Aplicación de diccionarios emocionales:** Se utiliza un sistema de clasificación basado en diccionarios de términos clave que vinculan el léxico del usuario con las cuatro dimensiones de la *Lovebrand* (“Amor”, “Felicidad”, “Confianza” y “Anticipación”).
3. **Normalización NLTK y Limpieza:** Se aplica un filtrado de *stop words* y una normalización mediante expresiones regulares para eliminar el ruido textual y términos genéricos que carecen de valor analítico.

4. **Representación Visual en Mosaico:** Se implementa una arquitectura de subplots que facilita la comparación directa entre dimensiones, utilizando escalas de color diferentes para mejorar la interpretabilidad de los datos.

Cabe destacar que el código para el mosaico presentado a continuación constituye un exponente del flujo de trabajo aplicado en esta investigación. El resto de las nubes de palabras distribuidas a lo largo de la memoria han sido generadas siguiendo esta misma lógica de procesamiento, asegurando la consistencia metodológica del estudio.

B. Código fuente del pipeline de visualización semántica

```
import pandas as pd
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import re
import nltk
from nltk.corpus import stopwords

# 1. PREPARACIÓN LÉXICA Y FILTRADO DE STOP WORDS
nltk.download('stopwords')
stop_words_es = set(stopwords.words('spanish'))
# Ampliación del diccionario de exclusión para optimizar el foco analítico
stop_words_es.update(['shein', 'si', 'bien', 'todo', 'pero', 'más', 'mas',
'cada', 'hace', 'hacer'])

# 2. CARGA DE DATOS Y SEGMENTACIÓN DE PROMOTORES (4-5 Estrellas)
df_tp = pd.read_excel('trustpilot_limpio.xlsx')
df_pos = df_tp[df_tp['Estrellas'] >= 4].copy()
df_pos['texto_low'] = df_pos['Comentario'].astype(str).str.lower()

# 3. DEFINICIÓN DE DICCIONARIOS EMOCIONALES (Ejes de la Lovemark)
emotions_keywords = {
    'Amor / Pasión': ['amo', 'amor', 'encanta', 'enamora', 'favorita', 'fan',
'obsesión', 'adoro', 'fascinada', 'encantada', 'maravilla', 'precioso', 'joya',
' enamorada', 'love'],
    'Felicidad / Alegría': ['feliz', 'content', 'alegría', 'maravilla', 'genial',
' ilusión', 'maravilloso', 'agradable', 'contentísima', 'maravillada', 'disfruto',
' super', 'súper'],
    'Confianza / Seguridad': ['confianza', 'confí', 'seguro', 'fiable',
' tranquilidad', 'garantía', 'sin duda', 'siempre', 'recomiendo', 'recomendable',
' fiel', 'seguridad', 'garantizado', 'nunca falla', 'siempre llega'],
    'Anticipación / Sorpresa': ['esperando', 'ganas', 'sorpresa', 'sorprend',
' supera', 'expectativas', 'deseando', 'descubrimiento', 'alucinada', 'flipando',
' no esperaba', 'superó']
}

# 4. CONFIGURACIÓN ESTRATÉGICA DE MAPAS DE COLOR
```

```
colormaps = {
    'Amor / Pasión': 'RdPu',
    'Felicidad / Alegría': 'autumn',
    'Confianza / Seguridad': 'Greens',
    'Anticipación / Sorpresa': 'winter'
}

# 5. GENERACIÓN DEL MOSAICO DE NUBES DE PALABRAS
fig, axes = plt.subplots(2, 2, figsize=(16, 10))
axes = axes.flatten()

for i, (emotion, keywords) in enumerate(emotions_keywords.items()):
    # Aplicación de máscara mediante expresiones regulares (regex)
    regex = '|'.join(keywords)
    mask = df_pos['texto_low'].str.contains(regex, case=False, na=False)
    text_emotion = "
.join(df_pos[mask]['Comentario'].astype(str).tolist()).lower()

    # Proceso de limpieza y tokenización
    text_cleaned = " ".join([w for w in re.findall(r'\b\w{3,}\b', text_emotion)
if w not in stop_words_es])

    if text_cleaned:
        # Instanciación de la nube de palabras con parámetros de diseño
optimizados
        wc = WordCloud(
            width=600, height=300,
            background_color='white',
            colormap=colormaps[emotion],
            max_words=50,
            prefer_horizontal=0.8
        ).generate(text_cleaned)

        # Representación en el mosaico (conversión a imagen para estabilidad
visual)
        axes[i].imshow(wc.to_image(), interpolation='bilinear')
        axes[i].set_title(f'Contexto de "{emotion}"\n({mask.sum()} menciones)',
fontsize=14, fontweight='bold', pad=10)
    else:
        axes[i].text(0.5, 0.5, 'Datos insuficientes para la categoría',
ha='center')

        axes[i].axis('off')

plt.suptitle('Trustpilot: Mapas Mentales de la Lovebrand Shein (4-5 Estrellas)',
fontsize=20, fontweight='bold', y=1.02)
plt.tight_layout()
plt.show()
```

ANEXO II: DECLARACIÓN DE USO DE IA

* Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, **Patricia Sánchez Martín**, estudiante de **Business Analytics** de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado “**Shein como Lovebrand: Un análisis de la disonancia ética, la gamificación y la paradoja del consumidor digital**”, 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. **Crítico:** Para encontrar contra-argumentos a una tesis específica que pretendo defender.
3. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
4. **Metodólogo:** Para descubrir métodos aplicables a problemas específicos de investigación.
5. **Interpretador de código:** Para realizar análisis de datos preliminares.
6. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
7. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
8. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
9. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
10. **Traductor:** Para traducir textos de un lenguaje a otro.

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: 18/04/2026

Firma: 