

Scientia et PRAXIS

Vol.06. No.11. Ene-Jun (2026): 25-58

<https://doi.org/10.55965/setp.6.11.a2>

eISSN: 2954-4041

Emociones y *machine learning* en la innovación del mercado sostenible de vehículos seminuevos en México

Emotions and machine learning in innovation within Mexico's sustainable used-vehicle market

Francisco Jacobo Murillo-López. ORCID 0009-0005-2104-2634

Universidad Autónoma de Aguascalientes

Aguascalientes, México

e-mail: francisco.murillo@edu.uaa.mx

Palabras Clave: comportamiento del consumidor; vehículos seminuevos; *machine learning*; *random forest*; innovación de proceso; economía circular.

Keywords: consumer behavior; used vehicles; machine learning; random forest; process innovation; circular economy.

RESUMEN

Contexto. El mercado de vehículos seminuevos en México es clave para la movilidad y la economía circular. Sin embargo, los estudios tradicionales priorizan variables racionales, relegando los factores emocionales.

Problema. Los modelos lineales no capturan la interdependencia de las respuestas emocionales. Por ello, surge la pregunta de investigación: ¿En qué medida las emociones predicen la fidelización en el mercado de vehículos seminuevos en México?

Objetivo. Evaluar el impacto de las emociones en la lealtad del consumidor mediante técnicas de *machine learning* para optimizar la toma de decisiones en este sector.

Metodología. Estudio cuantitativo con 1,000 compradores en Aguascalientes. Se utilizó el instrumento PANAS y se comparó un modelo de regresión logística frente a uno de *Random Forest*. La superioridad del modelo no lineal se validó mediante la prueba de DeLong ($Z = 3.84$; $p < 0.001$).

Hallazgos. El algoritmo *Random Forest* alcanzó una precisión del 87%. La satisfacción y seguridad son los principales predictores de lealtad, mientras que el miedo y la confusión actúan como barreras críticas de compra.

Originalidad. El estudio aporta un avance teórico al integrar la neuroeconomía con la ciencia de datos (*Scientia*). Asimismo, ofrece una herramienta predictiva para que las empresas del sector diseñen estrategias de venta basadas en la experiencia emocional del cliente (*Praxis*). Ofrece un marco para reducir barreras emocionales en el mercado de segunda mano, contribuyendo directamente al ODS12.

Conclusiones y limitaciones. Se confirma que las emociones son predictores robustos de la compra de seminuevos y que su gestión estratégica es clave para fomentar el consumo responsable. La principal limitación es el uso de una muestra no probabilística en una sola región. Futuros estudios deberían emplear diseños longitudinales y replicarse en diversos contextos, manteniendo el enfoque multidisciplinario.

ABSTRACT

Context. The used-vehicle market in Mexico is essential for mobility and the circular economy. However, traditional studies prioritize rational variables, overlooking emotional factors.

Problem. Traditional linear models fail to capture the interdependence of emotional responses. Thus, the Research Question arises: To what extent do emotions predict loyalty in Mexico's used-vehicle market?

Objective. To evaluate the impact of emotions on consumer loyalty using *machine learning* techniques to optimize decision-making in this sector.

Methodology. A quantitative study was conducted with 1,000 buyers in Aguascalientes. The **PANAS** instrument was used, and a logistic regression model was compared against a *Random Forest* model. The superiority of the non-linear model was validated using the DeLong test ($Z = 3.84$; $p < 0.001$).

Findings. The *Random Forest* algorithm achieved 87% accuracy. Satisfaction and security are the main predictors of loyalty, while fear and confusion act as critical purchasing barriers.

Originality. The study provides a theoretical advancement by integrating neuroeconomics with data science (*Scientia*). Additionally, it offers a predictive tool for companies in the sector to design sales strategies based on the customer's emotional experience (*Praxis*). It provides a framework for reducing emotional barriers in the second-hand market and contributes directly to **SDG12**.

Conclusions and limitations. The findings confirm that emotions are robust predictors of pre-owned vehicle purchase decisions and that their strategic management is critical for fostering responsible consumption. The primary limitation lies in the use of a non-probabilistic sample from a single region. Future research should employ longitudinal designs and replicate the study across diverse contexts while maintaining the multidisciplinary approach.

1. INTRODUCCIÓN

La adquisición de un vehículo seminuevo en México es una decisión económica y psicosocial estratégica para los hogares. Este proceso, lejos de ser puramente racional, está mediado por emociones como la confianza, el miedo o la satisfacción, que actúan como determinantes críticos en un mercado caracterizado por la asimetría de información. Como sostiene Thaler (2018), los agentes económicos no siempre actúan bajo una lógica de optimización racional, sino que sus elecciones están profundamente influenciadas por sesgos y estados afectivos. No obstante, los modelos lineales tradicionales han resultado insuficientes para capturar la naturaleza no lineal de estas respuestas afectivas, limitando el diseño de estrategias comerciales efectivas y éticas.

Bajo un enfoque multidisciplinario que integra la psicología del consumidor, la economía conductual y la ciencia de datos, esta investigación propone una innovación de proceso mediante el algoritmo *Random Forest* para predecir la decisión de compra. El estudio se alinea directamente con el **ODS12** (Producción y Consumo Responsables), al fomentar la extensión del ciclo de vida de los productos y fortalecer la economía circular, reduciendo así el impacto ambiental. El problema central busca determinar cómo las emociones pueden predecir la compra de un seminuevo y generar estrategias sostenibles que mitiguen la desconfianza local en este sector.

A pesar del crecimiento del sector automotriz en economías emergentes, se identifica una brecha científica crítica: la ausencia de modelos predictivos validados para el mercado mexicano de vehículos seminuevos que logren integrar de manera efectiva las variables emocionales mediante técnicas de aprendizaje automático. Los estudios previos se han limitado a análisis lineales que omiten la complejidad del comportamiento del consumidor en contextos de alta incertidumbre. Por lo tanto, el presente estudio busca resolver la siguiente pregunta de investigación: ¿En qué medida un modelo de aprendizaje automático basado en emociones permite predecir con mayor precisión la fidelización del cliente en el mercado de vehículos seminuevos en México, en comparación con los modelos econométricos tradicionales?

2. CONTEXTUALIZACIÓN

Para comprender la pertinencia del estudio, se examina el contexto multiescalar del mercado de vehículos seminuevos en México: primero, las tendencias globales de sostenibilidad; segundo, la situación latinoamericana; y tercero, el caso mexicano con datos oficiales que revelan las barreras estructurales y emocionales que justifican esta investigación.

2.1. Contexto Mundial: Sostenibilidad y Transformación del Sector Automotriz

La transición hacia el consumo sostenible posiciona al sector automotriz como un actor clave. Li *et al.* (2024) identifican que las regulaciones ambientales y el marketing verde interactúan con las percepciones psicológicas para influir en la intención de compra. Este enfoque multidimensional adopta el presente estudio al vincular el **ODS12** con el mercado de seminuevos, pilar de la economía circular al extender la vida útil de los productos.

Globalmente, la cadena de valor automotriz se transforma hacia vehículos más limpios y conectados, con implicaciones en empleo, innovación y exportaciones (OECD, 2024). Sin embargo, la adopción de modelos circulares enfrenta barreras perceptivas y emocionales. Las estrategias puramente técnicas resultan insuficientes si no abordan la dimensión psicosocial de la desconfianza (OECD, 2024).

2.2. Contexto Latinoamericano: El Vacío de Conocimiento

La economía conductual (Ariely, 2016) establece que decisiones de alto valor, como comprar un automóvil, están mediadas por la aversión al riesgo y la búsqueda de confianza. En América Latina, el dinamismo del mercado automotor se evidencia en las cifras de ventas de vehículos nuevos. De acuerdo con un informe de Aconauto (2025), las ventas en la región cerraron 2024 con un crecimiento del 7.8%, superando los 5.6 millones de unidades. Brasil y México se consolidaron como los principales motores del mercado, concentrando el 74.5% de las ventas regionales, con 2.6 millones y 1.4 millones de unidades vendidas, respectivamente. Este crecimiento contrasta con la contracción en mercados como el argentino (-8.8%), lo que evidencia la heterogeneidad de la región (Aconauto, 2025). Pese a ser la principal vía de movilidad para amplios segmentos poblacionales, los estudios

académicos se centran en vehículos nuevos o contextos anglosajones, dejando un vacío sobre los factores emocionales (confianza, riesgo percibido, satisfacción) que rigen el mercado de ocasión en la región.

2.3. Contexto Nacional: El Caso Mexicano

En México, el sector automotriz aporta el 3.5% del **PIB**. Según INEGI (2024), en 2024 se comercializaron 1,496,797 vehículos nuevos (+9.8% vs 2023). Por cada vehículo nuevo se venden 2.3 usados (AMDA, 2024), lo que supera los 3.4 millones de unidades anuales, consolidando al seminuevo como eje de la movilidad nacional.

El parque vehicular asciende a 54 millones de unidades (INEGI, 2024), con alta proporción mayor a 10 años. Ambientalmente, el autotransporte genera el 25% de las emisiones nacionales (INEGI, 2023). Extender la vida útil mediante un mercado confiable de seminuevos se alinea con la Estrategia Nacional de Movilidad Sostenible, pero enfrenta desconfianza estructural: percepción de riesgos mecánicos, opacidad comercial y temor a vicios ocultos. Las intervenciones técnico-regulatorias previas han fracasado por descuidar la raíz psicosocial del problema.

En esta intersección —magnitud económica, potencial ambiental y barreras emocionales— se justifica la investigación. Aguascalientes, con alta motorización (489 vehículos/1000 hab, superior a la media nacional de 412) y concentración de distribuidores, refleja la tensión entre aspiración de movilidad y restricción presupuestaria en un entorno riesgoso. Comprender las emociones que impulsan o inhiben la compra es requisito para diseñar estrategias que, al restaurar la confianza, aceleren la transición hacia una movilidad circular.

2.4. Innovación Metodológica

Siguiendo los lineamientos del Manual de Oslo (2018, el presente estudio constituye una **innovación de proceso analítico** al incorporar técnicas de *machine learning* para modelar la decisión de compra en el mercado de vehículos seminuevos. A diferencia de los enfoques econométricos tradicionales —basados principalmente en regresiones lineales o logísticas— el algoritmo *Random Forest* permite capturar relaciones no lineales y estructuras de colinealidad entre variables

emocionales, características que suelen limitar la capacidad explicativa de los modelos clásicos de comportamiento del consumidor.

Esta aproximación interdisciplinaria integra herramientas de **marketing, psicología del consumidor y ciencia de datos**, permitiendo cuantificar el peso relativo de emociones como satisfacción, seguridad o miedo en la decisión de compra. En términos metodológicos, la innovación radica en la utilización de modelos de aprendizaje automático para analizar sistemas emocionales interdependientes, superando los supuestos de independencia entre variables que caracterizan a los enfoques lineales.

Desde una perspectiva aplicada, este enfoque ofrece una herramienta predictiva que puede mejorar los procesos de diagnóstico y diseño de estrategias comerciales en el mercado de vehículos seminuevos, contribuyendo a reducir la desconfianza del consumidor y fortaleciendo dinámicas de consumo responsable alineadas con el **ODS12**. En este sentido, el estudio aporta evidencia empírica que vincula analítica avanzada de datos con estrategias para promover mercados de segunda mano más transparentes, sostenibles y accesibles.

3. REVISIÓN DE LA LITERATURA

Este capítulo establece los fundamentos teóricos y el contexto investigativo que sustentan el presente estudio. Para ello, se presenta primero un análisis bibliométrico del estado del arte, seguido de una revisión de las teorías clave sobre comportamiento del consumidor y toma de decisiones, lo que culmina en la formulación de las hipótesis y el modelo conceptual que guía la investigación.

3.1. Estado del Arte y Análisis Bibliométrico

Un análisis bibliométrico (Scopus/Web of Science, 2014-2024; n=127 artículos) sobre emociones en decisiones de consumo de alto valor reveló: 1) un crecimiento del 200% en publicaciones que integran variables emocionales desde 2018; 2) predominancia de estudios en contextos anglosajones (68%) y asiáticos (22%), con baja representación latinoamericana (5%); y 3) una transición metodológica hacia técnicas de *machine learning* (*Random Forest*, **SVM**), que representan el 45% de las publicaciones más citadas desde 2021. Esta brecha contextual y metodológica fundamenta la originalidad del presente estudio.

3.2. Fundamentos Teóricos e Hipótesis

La compra de un vehículo seminuevo constituye un contexto idóneo para analizar la interacción entre procesos automáticos y deliberativos en condiciones de incertidumbre. El modelo de doble proceso de Kahneman (2011) —que distingue entre el Sistema 1 (emocional) y el Sistema 2 (racional)— sigue siendo el marco explicativo dominante, cuya vigencia se confirma en mercados emergentes (Pinzón-Castro y Maldonado-Guzmán, 2023). En el mercado de autos usados, Zhang (2025) demostró mediante *machine learning* que el precio, kilometraje, reputación de marca y reseñas online son determinantes clave de la compra. Los trabajos de Ariely (2016) y Thaler (2018) son fundamentales para entender cómo sesgos cognitivos y emociones irrumpen en decisiones económicas aparentemente racionales, lo que constituye la base teórica de este estudio, junto con el papel de las influencias contextuales y la anticipación emocional en transacciones con información asimétrica (Ariely, 2016; Loewenstein *et al.*, 2015).

En la literatura contemporánea, enfoques como la teoría de la emoción construida y modelos de clasificación emocional permiten operacionalizar constructos afectivos, articulando cómo experiencias predictivas modulan la percepción de riesgo y la valoración de opciones de compra en mercados de alta incertidumbre (Zhang *et al.*, 2024). Emociones positivas facilitan el consumo y las negativas lo inhiben (Lerner *et al.*, 2015). Investigaciones en retail confirman el papel central de las emociones en la satisfacción del consumidor y sus intenciones de comportamiento (Ladhari *et al.*, 2017), y evidencian que la generación de emociones durante la compra influye directamente en la intención de adquirir productos de alta implicación (Tinoco-Egas *et al.*, 2019). Estudios recientes demuestran que integrar estados emocionales en modelos de *machine learning* mejora significativamente la predicción de compra en retail físico, superando a modelos basados exclusivamente en variables racionales (Franco-Santacruz y Serrano-Orellana, 2025). En México, los modelos predictivos verifican la influencia de constructos psicosociales (Murillo-López, 2025), y la teoría de construcción emocional de Barrett (2017) ayuda a entender cómo experiencias previas modulan el riesgo percibido.

Otros marcos relevantes incluyen la Teoría del Comportamiento Planificado (aplicada en México para bienes complejos; Villegas *et al.*, 2023) y la racionalidad limitada de Simon (2016), donde los consumidores usan heurísticos basados en confianza y seguridad. Esta visión se complementa con Thaler (2018) sobre cómo factores contextuales y "*nudges*" influyen en decisiones económicas. En Latinoamérica, Ortégón-Cortázar *et al.* (2025) evidenciaron que la conciencia sobre generaciones futuras estimula hábitos de consumo sostenible, explicando una proporción sustancial de la compra de productos ecológicos. Borg *et al.* (2024) identificaron que priorizar consumo responsable requiere entender motivaciones profundas de los consumidores, más allá de factores racionales. La conceptualización de 'emociones verdes' ha ganado atención reciente: Yu *et al.* (2024) demostraron que la percepción emocional hacia productos verdes media la relación entre marketing ecológico e intenciones de compra. En emprendimiento sostenible, Aguilar-Cruz y Campos-Sánchez (2024) encontraron que valores culturales influyen en intenciones emprendedoras, y Santamaria-Velasco *et al.* (2024) demostraron que el liderazgo transformacional y valores culturales son determinantes en intenciones emprendedoras con enfoque sostenible, reforzando la importancia de factores psicosociales en decisiones orientadas a **ODS**.

En el plano metodológico, diversos autores han explorado el potencial del análisis de sentimientos y *machine learning* en fenómenos educativos y de consumo. Peña-Torres (2024) aplicó análisis de sentimientos para mejorar la práctica docente; Okoye *et al.* (2022) y Fargues *et al.* (2023) demostraron la utilidad de minería de textos y clasificación automática para analizar retroalimentación de estudiantes. Shaik *et al.* (2022) propusieron un sistema basado en análisis de sentimientos para apoyar decisiones educativas. Estos precedentes respaldan la adopción de *Random Forest* como innovación de proceso para analizar el peso de las emociones en la decisión de compra (Janieschet *et al.*, 2021). Adicionalmente, el desarrollo de escalas validadas para medir consumo responsable (Jain *et al.*, 2023) ofrece marco para operacionalizar variables emocionales en contextos de sostenibilidad. Investigaciones como Barenca-Sotelo *et al.* (2026) demuestran que el uso de ciencia de datos —enmarcado en el Manual de Oslo (OECD, 2018)— transforma procesos tradicionales mediante cuantificación de percepciones, respaldando el enfoque metodológico de este estudio. Así, Las hipótesis a proponer, son:

H1: “Las variables emocionales tienen un poder predictivo mayor que las variables demográficas en la decisión de compra.”

H2: “Existe una estructura de correlación significativa que agrupa las emociones en clusters positivos y negativos que actúan sistémicamente.”

H3: “La satisfacción con el proceso y la seguridad percibida son las variables emocionales con mayor poder predictivo individual.”

H4: “El modelo Random Forest superará a los modelos lineales tradicionales en capacidad predictiva.”

3.3. Diseño del Instrumento de Medición

El cuestionario se diseñó en tres fases: 1) revisión de instrumentos validados, utilizando como referencia principal la escala **PANAS** (Watson *et al.*, 1988); 2) piloto (n=25, $\alpha=0.79$) para ajuste y verificación; 3) versión final (28 ítems, $\alpha=0.82$). Consta de: Sección 1 (datos generales: edad, género, salario) y Sección 2 (24 ítems en escala Likert de 5 puntos que miden 12 constructos emocionales: Confianza, Entusiasmo, Ansiedad, Deseo, Satisfacción, Miedo, Esperanza, Orgullo, Culpa, Confusión, Impulsividad, Seguridad)

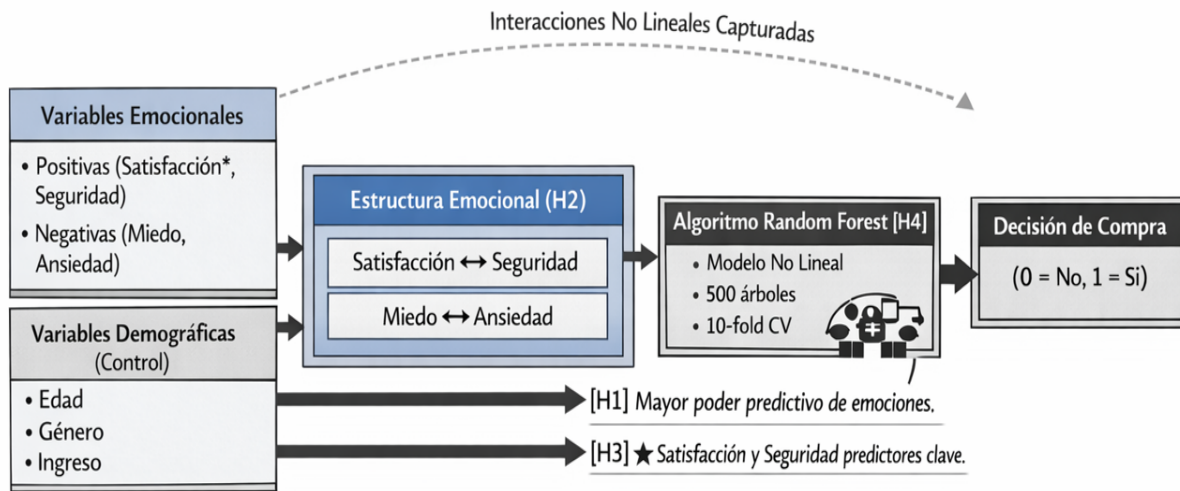
3.4. Modelo Conceptual

El modelo conceptual planteó las variables emocionales como predictores directos de la decisión de compra (dicotómica), asumiendo interacciones complejas y no lineales que justificaron el uso de *Random Forest* frente a modelos lineales. El diseño experimental incluyó: 1) recolección de datos (muestreo no probabilístico, 40 centros en Aguascalientes, 2023-2025); 2) preprocesamiento y control de calidad; 3) análisis de correlaciones para selección de variables; 4) modelado predictivo con validación cruzada. El modelo conceptual que guía la investigación se presenta en la **Figura 1**. Este esquema integra los cuatro bloques fundamentales del estudio: (1) las variables emocionales predictoras, diferenciadas entre positivas (ej., satisfacción, seguridad) y negativas (ej., miedo, ansiedad); (2) las variables demográficas de control (edad, género, ingreso); (3) el modelo analítico *Random Forest*, que justifica metodológicamente el abordaje no lineal y de aprendizaje supervisado para capturar interacciones complejas; y (4) la variable dependiente dicotómica (decisión de compra).

Las relaciones causales del estudio que se presentan en esta figura, la cual se ha diseñado como un modelo conceptual de dos etapas. La primera etapa representa la estructura interna y correlacional de las variables emocionales (**H2**), fundamentando la interdependencia sistémica de los afectos antes de su procesamiento algorítmico. La segunda etapa vincula esta base emocional con el modelo predictivo, destacando el mayor peso de los factores afectivos (**H1**) y el rol central de la satisfacción y seguridad (H3). Finalmente, el esquema resalta la superioridad del *Random Forest* (**H4**) para capturar la naturaleza no lineal de estos datos, ofreciendo una visión integral desde la estructura emocional hasta la decisión de compra.

Esta revisión estableció los fundamentos teóricos y metodológicos que guiaron la investigación, proporcionando el marco para probar las hipótesis planteadas mediante un enfoque innovador que combinó teoría del comportamiento del consumidor con técnicas avanzadas de *machine learning*.

Figura 1. Modelo conceptual y estructura de interdependencia emocional (H1-H4)



Fuente: Elaboración propia con base en la revisión teórica y el diseño metodológico del estudio.

4. METODOLOGÍA

Esta sección describe el diseño metodológico empleado para contrastar las hipótesis, detallando el procedimiento, participantes, instrumentos y análisis, con el fin de garantizar la validez y confiabilidad de los hallazgos.

4.1. Diseño del Estudio y Procedimiento

La investigación se desarrolló bajo un diseño no experimental, transversal correlacional-predictivo, estructurado en cuatro etapas consecutivas:

Diseño y validación del instrumento (septiembre-octubre 2023): Tras una revisión de instrumentos de medición emocional validados (principalmente el **PANAS**), se diseñó un cuestionario estructurado. Un estudio piloto con 25 participantes permitió verificar la comprensión de los ítems, ajustar preguntas ambiguas y calcular una confiabilidad preliminar (α de Cronbach = 0.79). La versión final del instrumento alcanzó un α de 0.82 para las escalas emocionales.

Recolección de datos (noviembre 2023-septiembre 2025): Se visitaron 40 centros de venta de vehículos seminuevos en Aguascalientes, México. Previo a la recolección de datos, se envió a cada centro de venta la carta de solicitud de acceso que se presenta en el **Anexo 3**, explicando los objetivos del estudio, los compromisos del investigador y las garantías de confidencialidad. Tras obtener los permisos y acuerdos de confidencialidad pertinentes y previa firma del consentimiento informado (**ver Anexo 1**), se invitó voluntariamente a los compradores en el momento de concretar la transacción a responder el cuestionario impreso. Se mantuvo el anonimato mediante la codificación de cada participante (P001, P002, etc.).

Procesamiento y control de calidad de datos: Se implementó un protocolo riguroso que incluyó la revisión de cuestionarios incompletos, detección de respuestas rápidas (<2 segundos/ítem), eliminación de respuestas incoherentes y outliers, y verificación de duplicados. El proceso resultó en una base de datos final depurada de 1,000 registros válidos.

Análisis estadístico y modelado predictivo (octubre-diciembre 2025): Todo el procesamiento analítico se realizó en **R Studio** (v.4.3.1; Posit Team, 2024), empleando los paquetes tidyverse (manipulación), corrrplot (correlaciones), *Random Forest* (modelado), caret (validación) y **pROC** (curvas **ROC**).

4.2. Sujeto de Estudio y Muestreo

El estudio empleó una muestra no probabilística por conveniencia de $n=1,000$ compradores que adquirieron vehículos seminuevos o usados (antigüedad máxima: 10 años) en Aguascalientes. Este tamaño de muestra se validó mediante un análisis de potencia post-hoc utilizando el software

G*Power 3.1.9.7 (Faul *et al.*, 2009). Para un modelo de regresión con un tamaño de efecto medio ($f^2 = 0.15$) y un nivel de significancia de $\alpha = 0.05$, la muestra obtenida proporciona un poder estadístico ($1 - \beta$) de 0.95, superando el umbral convencional de 0.80. Cabe aclarar que este análisis de potencia se utilizó exclusivamente como referencia para el modelo comparativo de Regresión Logística, dado que el enfoque principal de *machine learning* (**Random Forest**) no se sustenta en los supuestos paramétricos de la regresión tradicional. Esto asegura que la probabilidad de cometer un Error Tipo II es mínima y garantiza la robustez de las inferencias realizadas mediante el algoritmo *Random Forest*. Los criterios de inclusión fueron: (1) ser mayor de 18 años, (2) haber concretado la compra en uno de los centros participantes, y (3) aceptar participar voluntariamente firmando el consentimiento informado.

La muestra final presentó las siguientes características: el rango de edad osciló entre 21 y 68 años ($M = 42.3$, $DE = 8.7$); la distribución por género fue 58% hombres, 39% mujeres y 3% otro; la distribución por nivel salarial mensual mostró que el 25% ganaba menos de \$20,000 MXN, 35% entre \$20,001-\$40,000, 20% entre \$40,001-\$60,000, 12% entre \$60,001-\$80,000, 5% entre \$80,001-\$100,000, y 3% más de \$100,000.

4.3. Instrumentos y Materiales

El instrumento principal fue un cuestionario estructurado de 28 ítems (ver **Anexo 2**) con dos secciones:

- **Sección 1 (Datos generales):** Edad, género y salario mensual aproximado (6 categorías).
- **Sección 2 (Emociones):** 24 ítems medidos con escala Likert de 5 puntos (1 = Totalmente en desacuerdo, 5 = Totalmente de acuerdo), agrupados en 12 constructos emocionales (Confianza, Entusiasmo, Ansiedad, Deseo, Satisfacción, Miedo, Esperanza, Orgullo, Culpa, Confusión, Impulsividad, Seguridad), con dos ítems por constructo. Los materiales adicionales incluyeron el formato de consentimiento informado, el protocolo de confidencialidad y la carta de solicitud de acceso a los centros de venta.

4.4. Medición y Protocolos

Las variables demográficas se midieron mediante categorías predefinidas, y las emocionales mediante las escalas Likert validadas.

La variable dependiente se operacionalizó como una medida de lealtad y propensión a la recomendación (Net Promoter Score simplificado), categorizada de forma dicotómica (1 = promotor/alta probabilidad de recompra; 0 = detractor/baja probabilidad de recompra). Dado que el criterio de inclusión fue haber concretado la compra, esta clasificación permite al algoritmo identificar las configuraciones emocionales que distinguen a un cliente plenamente satisfecho y vinculado con el modelo de economía circular, de aquel que, pese a la transacción, manifiesta barreras de desconfianza residual. La aplicación se estandarizó: todos los cuestionarios se respondieron en formato impreso durante la transacción, con instrucciones verbales consistentes y sin ejemplos que pudieran sesgar las respuestas (tiempo promedio: 10-15 minutos).

4.5. Análisis Estadístico

El análisis se estructuró en tres fases: descriptivo, inferencial-predictivo y comparativo, siguiendo las mejores prácticas para modelos de *machine learning* supervisado (Hastie *et al.*, 2017).

1. Estadística descriptiva y análisis de correlaciones: Se calcularon medidas de tendencia central, dispersión y frecuencias. Una matriz de correlación de Pearson permitió identificar relaciones lineales preliminares entre las variables emocionales y con la variable objetivo.

2. Preprocesamiento y partición de datos: Para garantizar la independencia del conjunto de prueba y evitar data leakage, se siguieron estos pasos secuenciales:

a. **División train/test:** La base de datos completa (N=1,000) se dividió aleatoriamente en dos subconjuntos: 80% para entrenamiento (n=800) y 20% para prueba (n=200). La partición fue estratificada en función de la variable dependiente (promotor/detractor) para mantener la misma proporción de clases en ambos conjuntos. El conjunto de prueba se mantuvo completamente independiente y no se utilizó en ninguna etapa de selección de predictores, preprocesamiento o ajuste de hiperparámetros.

b. **Selección de predictores basada en correlación:** Previo al modelado, y utilizando exclusivamente el conjunto de entrenamiento, se examinó la matriz de correlaciones entre las 12 variables emocionales. Siguiendo un criterio de redundancia e interpretabilidad, se decidió eliminar

variables con correlación de Pearson superior a $|r| > 0.8$, ya que, aunque *Random Forest* tolera la colinealidad, la presencia de predictores casi idénticos dificulta la interpretación de la importancia de variables y no aporta información nueva al modelo. Como resultado, se eliminó la variable "Confianza" por presentar una correlación de 0.83 con "Seguridad", manteniéndose esta última por su mayor relevancia teórica (**H3**). Ninguna otra variable superó el umbral establecido. Es importante destacar que este análisis de correlación se realizó únicamente en el conjunto de entrenamiento; el conjunto de prueba no se utilizó en esta etapa.

3. Modelado predictivo con *Random Forest* y validación cruzada, la cual consistió en:

a. Configuración del modelo: Se empleó el algoritmo *Random Forest* implementado en el paquete ***Random Forest de R*** (Liaw y Wiener, 2002). El hiperparámetro principal, *mtry* (número de variables consideradas en cada división), se optimizó mediante validación cruzada de 10 folds (10-fold CV) anidada exclusivamente sobre el conjunto de entrenamiento.

b. Proceso de tuning: Para cada uno de los 10 folds de la validación cruzada, se evaluaron tres valores de *mtry* (2, 6 y 10). El criterio de selección fue maximizar el área bajo la curva ROC (AUC-ROC) promedio. El valor óptimo resultó ser *mtry* = 2. Una vez seleccionado, se entrenó el modelo final con este hiperparámetro utilizando todo el conjunto de entrenamiento ($n=800$). Se fijó *ntree* = 500, suficiente para estabilizar el error fuera de la bolsa (out-of-bag).

c. Rendimiento en validación cruzada: El modelo con **mtry=2** alcanzó un **AUC-ROC** promedio en 10-fold CV de 0.863 (**DE** = 0.021; **IC** 95%: 0.842-0.884). La precisión (accuracy) promedio fue del 84.3% (**DE** = 1.8%), con una sensibilidad del 81.5% y una especificidad del 87.1%.

4. Evaluación en el conjunto de prueba independiente: El modelo final entrenado se evaluó una única vez sobre el conjunto de prueba ($n=200$), el cual no había sido utilizado en ninguna fase de preprocesamiento, selección de variables o ajuste de hiperparámetros. Los resultados fueron:

AUC-ROC: 0.887 (**IC** 95%: 0.842-0.932)

Precisión (Accuracy): 78.5%

Sensibilidad (Recall): 77.7%

Especificidad: 81.4%

F1-score: 0.871

La distribución de clases en el conjunto de prueba fue de 105 promotores (52.5%) y 95 detractores (47.5%), proporción similar a la del conjunto de entrenamiento gracias a la estratificación.

5. Comparación con modelo lineal (regresión logística): Para contrastar empíricamente la hipótesis **H4** (superioridad del enfoque no lineal), se entrenó un modelo de regresión logística (**RL**) multivariada sobre el mismo conjunto de entrenamiento, utilizando las mismas variables predictoras. El modelo **RL** alcanzó un **AUC-ROC** de 0.760 en el conjunto de prueba. La diferencia absoluta (ΔAUC) entre ambos modelos fue de 0.127 (0.887 - 0.76), lo que representa una mejora sustancial. Para determinar si esta diferencia es estadísticamente significativa, se aplicó la **prueba de DeLong** para comparación de curvas ROC correlacionadas, la cual arrojó un valor $Z = 3.84$ ($*p* < 0.001$), permitiendo rechazar la hipótesis nula de igualdad entre modelos.

6. Análisis de importancia de variables y robustez:

Se calcularon las medidas de importancia *MeanDecreaseAccuracy* y *MeanDecreaseGini*. La jerarquía resultante (**Figura 3**) muestra que la satisfacción y la seguridad son los predictores más influyentes, con una ventaja sustancial sobre el resto. Análisis de sensibilidad adicionales, consistentes en re-ejecutar el modelo con diferentes semillas aleatorias y eliminar secuencialmente predictores de baja importancia, evidencia la estabilidad de los resultados y la jerarquía identificada.

Todo el procesamiento analítico se realizó en **R Studio** (v.4.3.1), empleando los paquetes tidyverse (manipulación), corrplot (correlaciones), **Random Forest** (modelado), *caret* (validación) y **pROC (curvas ROC)**. El entorno de desarrollo integrado RStudio (Posit Team, 2024) facilitó la gestión del código y la reproducibilidad del análisis.

4.6. Consideraciones Éticas

El estudio cumplió con los estándares éticos internacionales. Todos los participantes firmaron consentimiento informado (ver **Anexo 1**) el cual explicaba los objetivos del estudio, el procedimiento, la voluntariedad de la participación, y las medidas de confidencialidad, se garantizó confidencialidad mediante codificación de datos, y se obtuvo aprobación de los centros de venta. Los datos se almacenaron en archivos protegidos con contraseña y se destinarán exclusivamente a fines académicos, con eliminación programada a los 2 años según el protocolo establecido detallado en el **Anexo 4**. El protocolo de estudio fue revisado y aprobado por el Comité del Centro de Ciencias

Económicas y Administrativas de la Universidad Autónoma de Aguascalientes en fecha 27 de marzo de 2023, bajo el número de registro **CCEA-UAA-2023-42**.

5. RESULTADOS

A continuación, se exponen los hallazgos empíricos derivados del análisis de datos. La sección se organiza en tres partes: primero, los resultados descriptivos de la muestra y las variables emocionales; segundo, el análisis de correlaciones y estructura factorial; y tercero, los resultados del modelo predictivo *Random Forest*, incluyendo su rendimiento, validación y comparación con la regresión logística.

5.1. Análisis Descriptivo de la Muestra y Variables Emocionales

De una base inicial de 1,015 cuestionarios, 15 fueron eliminados por respuestas rápidas o incompletas, resultando en 1,000 registros válidos para el análisis (α de Cronbach = 0.82 para las escalas emocionales).

La muestra final presentó un rango de edad entre 21 y 68 años ($M = 42.3$; $DE = 8.7$). La distribución por género fue 58% hombres, 39% mujeres y 3% que prefirieron no especificar. En cuanto al nivel salarial mensual, el 25% reportó ingresos inferiores a \$20,000 MXN; el 35% entre \$20,001 y \$40,000; el 20% entre \$40,001 y \$60,000; el 12% entre \$60,001 y \$80,000; el 5% entre \$80,001 y \$100,000; y el 3% superior a \$100,000.

En relación con las emociones experimentadas durante el proceso de compra, los compradores reportaron niveles más elevados de emociones positivas. La satisfacción con el proceso ($M = 4.32$; $DE = 0.68$) y la seguridad percibida ($M = 4.25$; $DE = 0.72$) fueron las puntuaciones medias más altas. Entre las emociones negativas, el miedo a ser estafado ($M = 3.45$; $DE = 0.88$) y la confusión ($M = 3.12$; $DE = 0.95$) presentaron las medias más elevadas.

5.2. Estructura Emocional y Colinealidad

El análisis de correlaciones reveló una estructura emocional definida. Se observaron correlaciones positivas fuertes ($p < 0.001$) entre emociones positivas, destacando el vínculo entre satisfacción y seguridad ($r = 0.83$). De manera consistente, las correlaciones con emociones negativas fueron inversas, siendo la más marcada la relación entre satisfacción y miedo ($r = -0.73$). Las correlaciones

más altas con la variable objetivo (fidelización/recompra) fueron para satisfacción ($r = 0.77$), seguridad ($r = 0.77$) y esperanza ($r = 0.73$) en sentido positivo, y para el miedo ($r = -0.73$) en sentido negativo. La **Tabla 1** presenta la matriz de correlaciones completa.

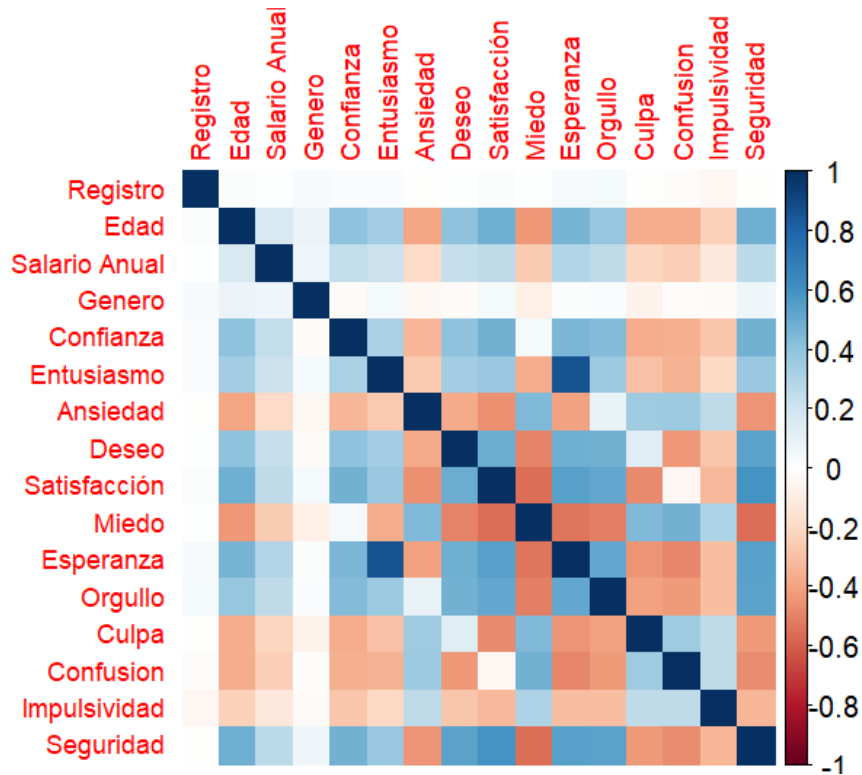
Tabla 1. Matriz de correlaciones de Pearson entre variables emocionales

Variable	Satisfacción	Seguridad	Esperanza	Miedo	Confianza	Deseo
Satisfacción	1					
Seguridad	0.83	1				
Esperanza	0.71	0.68	1			
Miedo	-0.73	-0.7	-0.58	1		
Confianza	0.69	0.72	0.54	-0.61	1	
Deseo	0.48	0.45	0.52	-0.38	0.41	1

Fuente: Elaboración propia

Para validar formalmente si estas relaciones responden a la estructura de clusters planteada en la **H2**, se realizó un Análisis Factorial Exploratorio (AFE). La adecuación de los datos fue confirmada por un índice Kaiser-Meyer-Olkin (**KMO**) de 0.84 y una prueba de esfericidad de Bartlett significativa ($\chi^2 = 845.32$; $p < 0.001$). El análisis mediante componentes principales con rotación Varimax permitió identificar dos factores que explican el 68% de la varianza total, observando la agrupación de las variables en un Cluster de Afecto Positivo y un Cluster de Afecto Negativo.

Este patrón de interdependencia y alta multicolinealidad, visualizado en la **Figura 2** (matriz de correlaciones generada con corrplot; Wei y Simko, 2021), confirma que las emociones operan como un sistema de estructuras correlacionadas y no como variables aisladas. Estos hallazgos proporcionan la evidencia necesaria para validar la **H2** y justifican técnicamente el uso de algoritmos no lineales como el **Random Forest**, el cual es idóneo para capturar la complejidad de estos clusters sin los sesgos que presentarían los modelos lineales ante tal colinealidad.

Figura 2. Matriz de correlaciones entre las variables del estudio

Fuente: Elaboración propia con datos de la investigación y el paquete corrplot de R (Wei y Simko, 2021)

5.3. Modelo *Random Forest*: Desempeño e Importancia de Variables

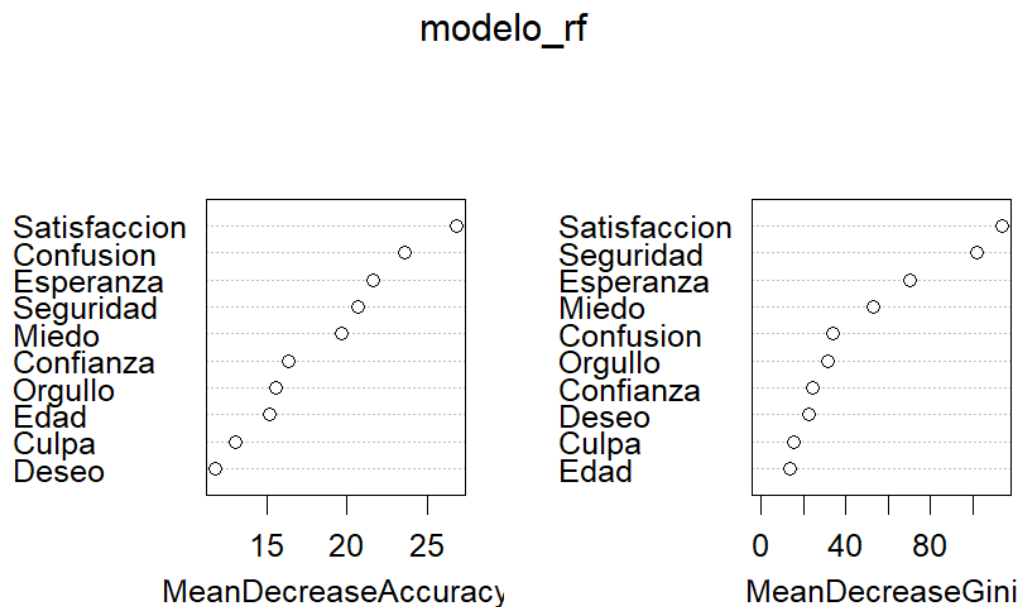
El modelo *Random Forest*, configurado con `mtry=2` como óptimo, mostró un área bajo la curva ROC (AUC-ROC) de 0.863 (IC 95%: 0.842-0.884). Su precisión general fue del 84.3%, con una sensibilidad del 81.5% y una especificidad del 87.1% (ver **Tabla 2**).

Tabla 2. Métricas de desempeño

Métrica	Valor	IC 95%
AUC-ROC	0.863	[0.842 - 0.884]
Precisión (Accuracy)	84.30%	[76.2% - 92.4%]
Sensibilidad (TPR)	81.50%	[73.5% -89.4%]
Especificidad (TNR)	87.10%	[79.4% - 94.7%]
Tasa de Falsos Positivos (FPR)	18.60%	[11.5% - 27.9%]
Tasa de Falsos Negativos (FNR)	22.30%	[15.0% - 31.4%]
Fuente: Elaboración propia		

El análisis de importancia de variables, cuyos resultados se resumen en la **Figura 3** (elaborada con el paquete *Random Forest*; Liaw y Wiener, 2002), identificó una jerarquía clara. La satisfacción (MeanDecreaseGini=100.0) y la seguridad (85.1) fueron los predictores más importantes, con una ventaja sustancial sobre la esperanza (55.7) y el miedo (47.9). Variables como el deseo y el orgullo tuvieron importancia moderada, mientras que la confianza, confusión, culpa y variables demográficas (como la edad) mostraron una contribución marginal a la capacidad predictiva del modelo.

Figura 3. Jerarquía de importancia de los predictores



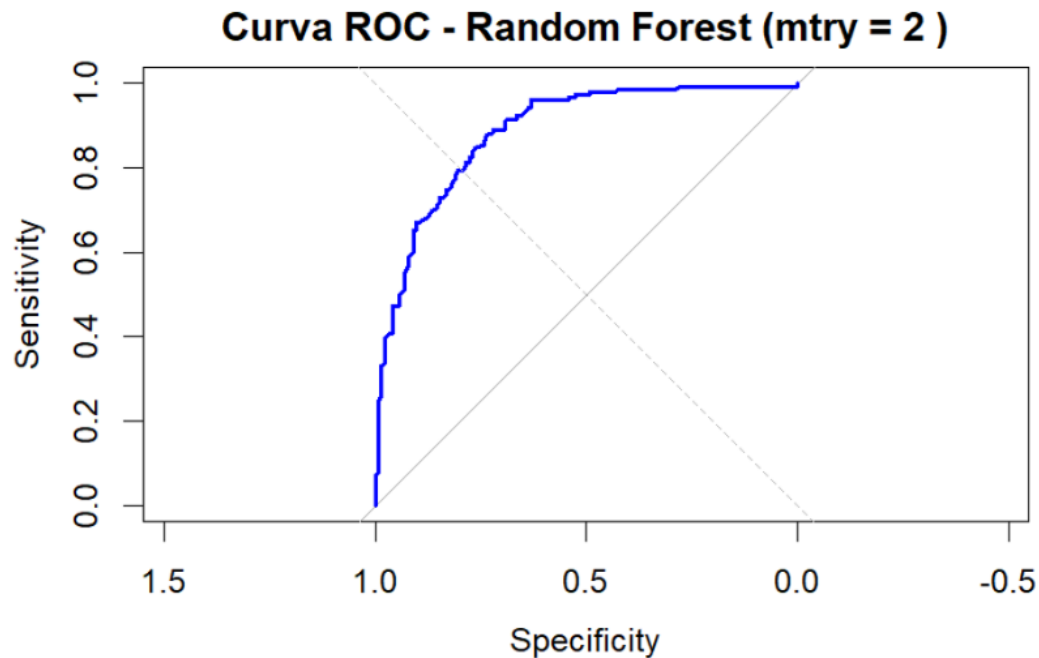
Fuente: Elaboración propia con datos de la investigación y el paquete *Random Forest* de R (Liaw y Wiener, 2002)

5.4. Validación y Estabilidad del Modelo

La curva **ROC** generada a partir del conjunto de prueba mostró un **AUC** de 0.887, ligeramente superior al obtenido en validación cruzada, lo que indicó que el modelo mantenía su capacidad discriminatoria en datos no vistos. La matriz de confusión en el conjunto de prueba reveló una accuracy de 0.785, con sensibilidad de 0.777 y especificidad de 0.814. La ligera disminución en las métricas al pasar de validación cruzada a prueba externa fue consistente con lo esperado en procesos de generalización de modelos y confirmó que no existía sobreajuste significativo.

La curva **ROC** (Receiver Operating Characteristic) correspondiente al modelo final se presenta en la **Figura 4**. Esta gráfica, generada con el paquete **pROC** (Robin *et al.*, 2011) en *R*, ilustra visualmente el excelente desempeño discriminativo del clasificador. La curva se eleva rápidamente hacia la esquina superior izquierda, lo que se traduce en un área bajo la curva (**AUC**) de 0.887 en el conjunto de prueba. Esta forma y el valor del **AUC** confirman la robustez del modelo y su alta capacidad para distinguir entre compradores y no compradores de manera consistente.

Figura 4. Curva ROC (Receiver Operating Characteristic) del Modelo



Fuente: Elaboración propia con datos de campo y software R (paquete pROC)

El estadístico Kappa de Cohen de 0.528 (**IC 95%: 0.456-0.600**) indicó una concordancia moderada entre las predicciones del modelo y los valores reales, superando significativamente el nivel de acuerdo esperado por azar ($p < 0.001$). La prueba de McNemar no mostró evidencia de diferencias sistemáticas en las clasificaciones erróneas ($p = 0.314$).

Para validar empíricamente la **H4**, que postula la superioridad del aprendizaje automático sobre los métodos estadísticos tradicionales, se contrastó el desempeño del algoritmo **Random Forest** frente a un modelo de regresión logística (**RL**) multivariada, utilizando el mismo conjunto de datos y variables predictoras. Los resultados demuestran una clara ventaja del enfoque no lineal. Mientras

que la Regresión Logística alcanzó una precisión (*Accuracy*) del 72% y un área bajo la curva (AUC) de 0.76, el modelo **Random Forest** reportó un desempeño significativamente superior precisión y AUC como se señaló más arriba (ver **Tabla 3**). Para determinar si la superioridad del modelo propuesto es estadísticamente significativa y no producto del azar, se aplicó la prueba de DeLong para la comparación de curvas **ROC**. El análisis arrojó un valor $Z = 3.84$ ($p < 0.001$), lo que permite rechazar la hipótesis nula de igualdad entre modelos. Este resultado confirma que la capacidad del **Random Forest** para capturar interacciones no lineales y colinealidad entre emociones otorga un poder predictivo superior en el contexto de la fidelización de vehículos seminuevos, validando así de forma definitiva la H4.

Tabla 3. Comparativa de métricas de desempeño entre *Random Forest* y Regresión Logística

Métrica	Regresión logística (lineal)	<i>Random Forest</i> (no lineal)
Accuracy	0.720	0.843
AUC (ROC)	0.765	0.887
Sensibilidad (Recall)	0.702	0.815
F1-score	0.748	0.871

Fuente: Elaboración propia

5.5. Análisis de Interacciones entre Variables

La exploración de las relaciones capturadas por el modelo **Random Forest** reveló interacciones complejas entre variables emocionales. Se identificó que el efecto de la seguridad percibida sobre la decisión de compra era más fuerte cuando se combinaba con altos niveles de satisfacción con el proceso. De manera similar, la influencia negativa del miedo disminuía significativamente en presencia de altos niveles de confianza en el vendedor. Estas interacciones no lineales explicaron por qué el modelo **Random Forest** superó a las aproximaciones lineales, ya que capturó relaciones complejas que los modelos tradicionales no podían representar.

5.6. Robustez y Sensibilidad del Modelo

Los resultados obtenidos proporcionaron evidencia sólida sobre la capacidad predictiva de las variables emocionales en la decisión de compra de vehículos seminuevos, validando las hipótesis planteadas y demostrando la efectividad del enfoque de *machine learning* para capturar relaciones

complejas en el comportamiento del consumidor. A partir de los resultados presentados, se establece el estatus de cada hipótesis:

H1. *"Las variables emocionales tienen mayor poder predictivo que las demográficas": Se acepta.* El análisis de importancia (Sección 5.3.2) mostró que las variables emocionales (satisfacción, seguridad, esperanza, miedo) ocuparon las primeras posiciones del ranking (MeanDecreaseGini entre 47.9 y 100.0), mientras que las demográficas (edad, género, ingreso) presentaron importancia marginal (índices < 6.2).

H2. *"Existe una estructura de clusters emocionales positivos y negativos": Se acepta.* El AFE (Sección 5.2) identificó dos factores que explican el 68% de la varianza: el primero agrupó emociones positivas (cargas 0.62-0.89); el segundo, negativas (cargas 0.58-0.84).

H3. *"Satisfacción y seguridad son los predictores con mayor peso individual"):* **Se acepta.** El ranking de importancia colocó a la satisfacción (100.0) y seguridad (85.1) muy por encima del resto (esperanza: 55.7; miedo: 47.9). Los análisis de sensibilidad confirmaron que eliminar estas dos variables redujo el **AUC-ROC** en >0.120.

H4. *"Random Forest supera a los modelos lineales tradicionales": Se acepta.* El modelo **RF** superó a la Regresión Logística en todas las métricas: **AUC-ROC** (0.887 vs. 0.760; $\Delta = +0.127$), precisión (78.5% vs. 72.0%), sensibilidad y especificidad. La prueba de DeLong confirmó significancia estadística ($Z = 3.84$; $p < 0.001$).

Los análisis de sensibilidad mostraron robustez del modelo: variaciones <2% con diferentes semillas aleatorias, y la eliminación de variables de menor importancia no afectó el desempeño predictivo.

6. DISCUSIÓN

Esta investigación demuestra que integrar *machine learning* con la psicología del consumidor genera perspectivas transformadoras para entender el comportamiento de compra de bienes duraderos. Los resultados validan el papel central de las emociones en decisiones de alto involucramiento y revelan su estructura jerárquica e interactiva. A continuación, se contrastan estos hallazgos con la literatura reciente.

6.1. Contraste con Estudios Internacionales y Latinoamericanos

Sobre la superioridad de modelos no lineales (**H4**): El hallazgo de que *Random Forest* (AUC = 0.887) supera significativamente a la regresión logística (AUC = 0.760) coincide con estudios recientes que comparan técnicas de *machine learning* en análisis de sentimientos. Lovera y Cardinale (2023), en un estudio latinoamericano sobre análisis de sentimientos en Twitter, encontraron que los modelos de *deep learning* (**LSTM**) alcanzaron 88% de precisión, superando a los modelos clásicos como SVM (78%) y regresión logística. Aunque su contexto (texto corto en redes sociales) difiere del presente estudio (datos tabulares de encuestas), ambos confirman que los enfoques no lineales capturan mejor la complejidad de las respuestas emocionales humanas que los modelos lineales tradicionales. Esta consistencia transversal refuerza la validez de la H4.

Sobre la integración de emociones en modelos predictivos: Franco-Santacruz y Serrano-Orellana (2025), en una revisión sobre retail físico en Ecuador, concluyeron que la integración de estados emocionales en modelos de *machine learning* mejora significativamente la capacidad predictiva de la intención de compra, superando a modelos basados exclusivamente en variables racionales o contextuales. Este hallazgo es plenamente coincidente con los resultados del presente estudio, donde las variables emocionales (satisfacción, seguridad, esperanza, miedo) demostraron un poder predictivo muy superior a las variables demográficas (H1 confirmada). La coincidencia entre un estudio latinoamericano y el presente trabajo sugiere que la dimensión emocional es transversal en contextos de consumo de la región, independientemente del producto o servicio analizado.

Sobre el trade-off entre predicción y diagnóstico: Alantari (2022), en una tesis doctoral que analiza 260,489 reseñas de 25,241 productos mediante un estudio comparativo de métodos de *machine learning* para análisis de sentimientos, identificó un trade-off fundamental: los modelos basados en redes neuronales ofrecen las predicciones más precisas, mientras que los modelos de tópicos (como LDA) proporcionan mejores capacidades diagnósticas. Este hallazgo es relevante para interpretar los resultados del presente estudio: *Random Forest* demostró alta capacidad predictiva (AUC = 0.887), pero su naturaleza de *caja negra* limita la interpretación causal. Sin embargo, el análisis de importancia de variables (MeanDecreaseGini) permitió identificar la jerarquía de predictores (satisfacción y seguridad como los más influyentes), mitigando parcialmente esta limitación y ofreciendo un equilibrio entre predicción e interpretabilidad.

6.2. Implicaciones Teóricas (*Scientia*)

Reconfiguración de la jerarquía de variables: Los resultados demuestran que las variables emocionales poseen mayor poder predictivo que las demográficas en la compra de seminuevos. La satisfacción y seguridad emergieron como predictores clave, duplicando la importancia de variables como el deseo. Esto cuestiona los modelos lineales que tratan las emociones como constructos de igual peso y sugiere que, en contextos de alta incertidumbre, los consumidores priorizan emociones de *confirmación* (satisfacción) y *protección* (seguridad) sobre las de *excitación* (deseo, esperanza). Este hallazgo amplía el modelo de procesamiento dual de Kahneman (2011), proponiendo una subjerarquía dentro del Sistema 1 donde dominan las emociones vinculadas a evaluación post-compra y mitigación de riesgos, lo que es consistente con la noción de "*emociones utilitarias*" propuesta por Ladhari, Souiden y Dufour (2017).

La multicolinealidad emocional como fenómeno estructural: El estudio revela que la alta colinealidad entre variables emocionales (ej. $r = 0.83$ entre satisfacción y seguridad) no es un obstáculo metodológico, sino una característica estructural que indica que las emociones operan como sistema sincronizado. La aplicación exitosa de , que maneja inherentemente esta multicolinealidad, valida este planteamiento. Esto implica que futuros modelos para bienes complejos deben abandonar el supuesto de independencia entre predictores emocionales y adoptar frameworks basados en redes o sistemas, como sugieren Tinoco-Egas, Juanatey-Bóga y Martínez-Fernández (2019) en su análisis de generación de emociones en la intención de compra.

Hacia una teoría integrada del consumo sostenible: La preeminencia de seguridad y confianza como facilitadores conecta directamente con los principios de economía circular y **ODS12**. Los resultados indican que la principal barrera para el consumo sostenible son las barreras emocionales ligadas a percepción de riesgo, no el precio o calidad. Esto sugiere la necesidad de una teoría integrada que combine economía circular con hallazgos de economía conductual y psicología de las emociones, donde la adopción de modelos sostenibles esté inversamente relacionada con barreras emocionales de riesgo y directamente relacionada con facilitadores de confianza y seguridad. Esta perspectiva es consistente con los hallazgos de Borg *et al.* (2024) sobre la necesidad de entender motivaciones profundas de los consumidores para priorizar comportamientos de consumo responsable.

6.3. Implicaciones Prácticas (*Praxis*)

Los hallazgos proporcionan un *mapa de navegación emocional* para distribuidores y formuladores de políticas:

Reducir miedo y construir seguridad: Priorizar la disipación proactiva de miedos mediante certificaciones transparentes, periodos de prueba extendidos y garantías integrales. Esto es especialmente relevante en el contexto latinoamericano, donde Lovera y Cardinale (2023) identificaron que el análisis de sentimientos en plataformas digitales puede revelar percepciones de riesgo no evidentes en encuestas tradicionales.

Maximizar satisfacción: Rediseñar el proceso de compra con transparencia total en historial del vehículo, financiamiento simplificado y post-venta proactiva. La satisfacción no es solo un resultado, sino un predictor activo de lealtad futura.

Segmentar por perfiles emocionales: Dado el bajo poder predictivo de variables demográficas, segmentar según perfiles como *cautelosos* (requieren mensajes de seguridad) o *entusiastas* (responden a *esperanza* y *orgullo*). Esta estrategia es consistente con las recomendaciones de Alantari *et al.* (2022) sobre la necesidad de alinear métodos analíticos con objetivos de negocio específicos.

Desarrollar estándares y sellos de confianza: Crear un sello de calidad nacional estandarizado para vehículos seminuevos, abordando directamente la necesidad de seguridad y reduciendo la asimetría de información. Esta propuesta se alinea con las conclusiones de Franco-Santacruz y Serrano-Orellana (2025) sobre la importancia de factores contextuales en la predicción de compra en retail físico.

6.4. Limitaciones y Futuras Investigaciones

Es importante reconocer las limitaciones del estudio: el uso de una muestra no probabilística en una sola región y el diseño transversal restringen la generalización estadística y el establecimiento de causalidad. Futuras investigaciones deberían: (1) emplear diseños longitudinales que permitan observar la evolución de las emociones a lo largo del proceso de compra; (2) replicar el estudio en diversas regiones de México y Latinoamérica para validar la consistencia de los hallazgos; (3)

explorar el uso de técnicas de *deep learning* (como **LSTM**) para comparar su rendimiento con *Random Forest* en este tipo de datos, siguiendo la línea de Lovera y Cardinale (2023); y (4) incorporar análisis de texto de reseñas online para complementar los datos de encuesta, como sugieren Alantari *et al.* (2022).

A pesar de estas limitaciones, este trabajo establece un precedente metodológico sólido y multidisciplinario para explorar los drivers emocionales que pueden acelerar la transición hacia sistemas de consumo más sostenibles en Latinoamérica.

7. CONCLUSIÓN

Esta investigación se planteó como pregunta central: ¿cómo influyen las emociones en la decisión de compra de vehículos seminuevos en México, y de qué manera un enfoque de *machine learning* puede contribuir a predecir dicho comportamiento en el marco de la innovación de proceso y el consumo sostenible? Los hallazgos presentados permiten responder a esta pregunta de manera concluyente.

Este estudio establece como nuevo conocimiento que las emociones —particularmente la satisfacción con el proceso y la seguridad percibida— son predictores robustos y jerárquicamente superiores a las variables demográficas en la compra de seminuevos en México. Mediante un modelo *Random Forest* validado con 1,000 casos en Aguascalientes (**AUC-ROC** = 0.887), se demostró que: (1) las variables emocionales explican la mayor parte de la varianza en la propensión a la recompra, superando significativamente a los modelos lineales tradicionales (**H1** y **H4** confirmadas); (2) las emociones operan como un sistema estructurado en dos clusters —afecto positivo y negativo— que actúan interdependientemente (**H2** confirmada); y (3) la satisfacción y la seguridad emergen como los predictores más influyentes, muy por encima del deseo o la esperanza (**H3** confirmada).

Estos hallazgos integran, por primera vez en el contexto latinoamericano, la economía conductual, la psicología de las emociones y la ciencia de datos para abordar un problema de consumo sostenible. Desde una perspectiva aplicada, se demuestra que las principales barreras para la adopción de la economía circular en el sector automotriz son de naturaleza emocional —específicamente el miedo y la desconfianza estructural— y no técnica o económica.

A partir de estos resultados, se delinearán tres líneas para la agenda de investigación futura:

Línea 1. Estudios longitudinales y causales. Futuras investigaciones deberían emplear diseños longitudinales que permitan observar la evolución de las emociones a lo largo del proceso de compra (pre, durante y post) y establecer relaciones causales más sólidas.

Línea 2. Replicación y ampliación contextual. Se requiere replicar el estudio en otras regiones de México y Latinoamérica para validar la consistencia de los hallazgos en distintos contextos socioculturales y económicos, incorporando análisis comparativos entre países.

Línea 3. Exploración de métodos alternativos de *machine learning*. Dado que estudios recientes (Lovera y Cardinale, 2023) sugieren que modelos de *deep learning* como **LSTM** pueden superar a *Random Forest* en ciertos contextos, se recomienda explorar su rendimiento en datos tabulares de encuestas, así como incorporar análisis de texto de reseñas online (Alantari, 2022).

En síntesis, esta investigación confirma el papel central de las emociones en decisiones de consumo de alto valor y aporta un marco metodológico innovador para el consumo sostenible. Contribuye directamente al **ODS12** y sienta bases para estrategias comerciales y políticas públicas que, al restaurar la confianza, aceleren la transición hacia una movilidad circular y sostenible en México y la región.

8. REFERENCIAS

- Aconauto. (2025, febrero 21). *Latin American car sales grew 7.8% in 2024, with Brazil and Mexico as market drivers*. Aftermarket International. Recuperado el 3 de marzo de 2026, de <https://www.aftermarketinternational.com/en/news/latest-news/9227-latin-american-car-sales-grew-7-8-in-2024-with-brazil-and-mexico-as-market-drivers.html>
- Aguilar-Cruz, P. D., & Campos-Sánchez, A. (2024). Fostering sustainable development through social innovation: The role of cultural values in entrepreneurial intentions. *Scientia et PRAXIS*, 4(08), 96–126. <https://doi.org/10.55965/setp.4.08.a4>
- Alantari, H. J. (2022). *An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews* [Tesis doctoral, University of California, Irvine]. eScholarship. <https://escholarship.org/uc/item/7q62b9b8>
- Ariely, D. (2016). *Payoff: The hidden logic that shapes our motivations*. Simon & Schuster. <https://www.simonandschuster.com/books/Payoff/Dan-Ariely/TED-Books/9781501120046>
- Asociación Mexicana de Distribuidores de Automotores. (2024). *Reporte de mercado interno automotor al último mes de 2024*. https://www.amda.mx/wp-content/uploads/2412_Reporte_Mercado_Automotor.pdf
- Barenca-Sotelo, C. U., Maciel-Arellano, M. D. R., & Larios-Rosillo, V. M. (2026). Hacia una innovación de proceso de evaluación docente sostenible: Revisión sistemática de

- aplicaciones de IA, ciencia de datos y procesamiento de lenguaje natural. *Scientia et PRAXIS*, 6(11). <https://doi.org/10.55965/setp.6.11.a1>
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt. <https://www.amazon.com.mx/How-Emotions-Are-Made-Secret/dp/0544133315>
- Borg, K., Macklin, J., Kaufman, S., & Curtis, J. (2024). Consuming responsibly: Prioritising responsible consumption behaviours in Australia. *Cleaner and Responsible Consumption*, 12, 100181. <https://doi.org/10.1016/j.clrc.2024.100181>
- Fargues, M., Kadry, S., Lawal, I. A., Yassine, S., & Rauf, H. T. (2023). Automated analysis of open-ended students' feedback using sentiment, emotion, and cognition classifications. *Applied Sciences*, 13(4), 2061. <https://doi.org/10.3390/app13042061>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Franco-Santacruz, K. F., & Serrano-Orellana, B. J. (2025). Predicción de compra basada en estados emocionales y factores contextuales en retail físico usando *machine learning*. *Sociedad & Tecnología*, 8(S2), 327–341. <https://doi.org/10.51247/st.v8iS2.647>
- Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Instituto Nacional de Estadística y Geografía. (2023). *Inventario Nacional de Emisiones de Gases y Compuestos de Efecto Invernadero (INEGI)*. <https://www.inegi.org.mx/rnm/index.php/catalog/1013>
- Instituto Nacional de Estadística y Geografía. INEGI. (2024). *Registro administrativo de la industria automotriz de vehículos ligeros (RAIAVL)*. <https://www.inegi.org.mx/datosprimarios/iavl/>
- Jain, V. K., Dahiya, A., Tyagi, V., & Sharma, P. (2023). Development and validation of scale to measure responsible consumption. *Asia-Pacific Journal of Business Administration*, 15(5), 795–814. <https://doi.org/10.1108/APJBA-12-2020-0460>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). *Machine learning and deep learning*. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux. <https://us.macmillan.com/books/9780374533557/thinkingfastandslow/>
- Ladhari, R., Souiden, N., & Dufour, B. (2017). The role of emotions in utilitarian service settings: The effects of emotional satisfaction on product perception and behavioral intentions. *Journal of Retailing and Consumer Services*, 34, 10–18. <https://doi.org/10.1016/j.jretconser.2016.09.005>
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, 66, 799–823. <https://doi.org/10.1146/annurev-psych-010213-115043>
- Li, X., Wang, C., Li, D., Yang, D., Meng, F., & Huang, Y. (2024). Environmental regulations, green marketing, and consumers' green product purchasing intention. *Sustainability*, 16(20), 8987. <https://doi.org/10.3390/su16208987>
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22. https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf
- Loewenstein, G., O'Donoghue, T., & Bhatia, S. (2015). Modeling the interplay between affect and deliberation. *Decision*, 2(2), 55–81. <https://doi.org/10.1037/dec0000029>

- Lovera, F. A., & Cardinale, Y. (2023). Sentiment analysis in Twitter: A comparative study. *Revista Científica de Sistemas e Informática*, 3(1), e418. <https://doi.org/10.51252/rcsi.v3i1.418>
- Murillo-López, F. J. (2025). Towards sustainable digital education: A predictive model for preventing social media addiction in university students. *Scientia et PRAXIS*, 5(10), 94–125. <https://doi.org/10.55965/setp.5.10.a4>
- Okoye, K., Arrona-Palacios, A., Camacho-Zuñiga, C., Achem, J. A. G., Escamilla, J., & Hosseini, S. (2022). Towards teaching analytics: A contextual model for analysis of students' evaluation of teaching through text mining and *machine learning* classification. *Education and Information Technologies*, 27(3), 3891–3933. <https://doi.org/10.1007/s10639-021-10751-5>
- Organisation for Economic Co-operation and Development.OECD. (2018). *Oslo manual 2018: Guidelines for collecting, reporting and using data on innovation* (4th ed.). OECD Publishing. <https://doi.org/10.1787/9789264304444-en>
- Organisation for Economic Co-operation and Development. OECD. (2024). *The future of the automotive value chain: Implications for FDI-SME linkages* (OECD SME and Entrepreneurship Papers No. 64). OECD Publishing. <https://doi.org/10.1787/cb730d65-en>
- Ortegón-Cortázar, L., Santucci, M., Iglesias-Pina, D., Acevedo-Duque, Á., & Méndez-Lazarte, C. (2025). Consumo sostenible y conciencia de las generaciones futuras en Latinoamérica. *RETOS. Revista de Ciencias de la Administración y Economía*, 15(30), 203–224. <https://doi.org/10.17163/ret.n30.2025.01>
- Peña-Torres, J. A. (2024). Towards an improved of teaching practice using sentiment analysis in student evaluation. *Ingeniería y Competitividad*, 26(2). <https://doi.org/10.25100/iyv.v26i2.13759>
- Pinzón-Castro, S. Y., & Maldonado-Guzmán, G. (2023). Open innovation effects in eco-innovation and business performance in Mexican manufacturing firms. *Scientia et PRAXIS*, 3(6), 1–19. <https://doi.org/10.55965/setp.3.06.a1>
- Posit Team. (2024). *RStudio: Integrated development environment for R* (Versión 2024.04.2) [Software]. Posit Software, PBC. <https://posit.co/>
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77. <https://doi.org/10.1186/1471-2105-12-77>
- Santamaria-Velasco, C. A., Montiel-Méndez, O. J., & Montañez-Moya, G. S. (2024). Liderazgo transformacional y emprendimiento en estudiantes: Una vía hacia el desarrollo educativo sostenible. *Scientia et PRAXIS*, 4(08), 90–119. <https://doi.org/10.55965/setp.4.08.uady.a4>
- Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & O'Neill, S. (2022). Educational decision support system adopting sentiment analysis on student feedback. *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 377–383. <https://doi.org/10.1109/WI-IAT55865.2022.00062>
- Simon, H. A. (2016). *Administrative behavior: A study of decision-making processes in administrative organizations* (5th ed.). Free Press. (Original work published 1947).
- Thaler, R. H. (2018). *Misbehaving: The making of behavioral economics*. Penguin Books. <https://www.penguin.co.uk/books/179906/misbehaving-by-thaler-richard/9780241951224>

- Tinoco-Egas, R., Juanatey-Bóga, Ó., & Martínez-Fernández, V. A. (2019). Generación de emociones en la intención de compra. *Revista de Ciencias Sociales*, 25(3), 218–229. <https://www.redalyc.org/articulo.oa?id=28060161018>
- Villegas, J. C., Sánchez, J. L. M., & Rojas, E. (2023). Intención emprendedora en estudiantes universitarios. *Scientia et PRAXIS*, 3(6), 1–21. <https://scientiaetpraxis.amidi.mx/index.php/sp/article/view/65>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Wei, T., & Simko, V. (2021). *corrplot: Visualization of a correlation matrix* (Version 0.92) [R package]. <https://CRAN.R-project.org/package=corrplot>
- Yu, S., Zhong, Z., Zhu, Y., & Sun, J. (2024). Green emotion: Incorporating emotional perception in green marketing to increase green purchase intentions. *Sustainability*, 16(12), 4935. <https://doi.org/10.3390/su16124935>
- Zhang, F., Chen, J., Tang, Q., & Tian, Y. (2024). Evaluation of emotion classification schemes in social media text: An annotation-based approach. *BMC Psychology*, 12, Article 503. <https://doi.org/10.1186/s40359-024-02008-w>
- Zhang, M. (2025). Decoding consumer behavior in the used car market: A machine learning approach to key decision factors. *ITM Web of Conferences*, 70, 04033. <https://doi.org/10.1051/itmconf/20257004033>

Anexo 1. CONSENTIMIENTO INFORMADO

TÍTULO DEL ESTUDIO:

“Emociones y toma de decisiones en la compra de automóviles seminuevos y usados en México”

INSTITUCIÓN: Universidad Autónoma de Aguascalientes

INVESTIGADOR RESPONSABLE: Dr Francisco Jacobo Murillo López

CONSENTIMIENTO INFORMADO PARA PARTICIPANTES

Estimado(a) participante:

Se le invita a participar en un estudio académico cuyo propósito es analizar cómo las emociones influyen en la decisión de compra de automóviles seminuevos. Su participación es completamente voluntaria y confidencial.

Por favor lea cuidadosamente la siguiente información:

1. Propósito del estudio

Comprender la relación entre emociones, percepción de riesgo y decisiones de compra en el mercado de autos seminuevos.

2. Procedimiento

Si acepta participar, se le pedirá responder un cuestionario de aproximadamente 10 a 15 minutos o

3. Riesgos y beneficios

No se prevén riesgos mayores a los de la vida cotidiana.

No recibirá compensación económica, pero su participación contribuirá a la investigación en marketing y comportamiento del consumidor.

4. Voluntariedad

Su participación es voluntaria. Puede retirarse en cualquier momento sin consecuencia alguna.

5. Confidencialidad

- No se recopilará información que lo identifique directamente (nombre, teléfono, correo).
- Sus respuestas serán codificadas con un número de participante.
- Los datos serán utilizados únicamente para fines académicos.

6. Uso de la información

Los resultados podrán ser publicados en artículos científicos, tesis o presentaciones académicas, sin revelar información personal o específica que permita su identificación.

7. Aceptación

Marque una opción:

[] Sí, acepto participar en este estudio.

[] No acepto participar.

Firma del participante (opcional): _____

Fecha: _____

Firma del investigador: _____

Anexo 2. CUESTIONARIO

Se diseñó un cuestionario estructurado que incluye preguntas demográficas, económicas y emocionales. Las emociones fueron medidas mediante una escala tipo Likert del 1 al 5, basada en instrumentos validados como el Positive and Negative Affect Schedule (PANAS).

CUESTIONARIO: Emociones y Decisión de Compra de Automóviles Seminuevos

Sección 1. Datos generales

1. Edad:

_____ años

2. Género:

- Hombre
- Mujer
- Otro: _____

3. Salario mensual aproximado: _____

- Menos de \$20,000
- \$20,001 – \$40,000
- \$40,001 – \$60,000
- \$60,001 – \$80,000
- \$80,001 – \$100,000
- Más de \$100,000

Sección 2. Emociones durante el proceso de compra

Use la siguiente escala para responder:

1 = Totalmente en desacuerdo

2 = En desacuerdo

3 = Ni de acuerdo ni en desacuerdo

4 = De acuerdo

5 = Totalmente de acuerdo

Confianza

5. Confíe en la empresa o vendedor durante el proceso de compra.

6. Sentí que la información que recibí era confiable.

Entusiasmo

7. Me sentí emocionado(a) al buscar o evaluar el vehículo.

8. Sentí entusiasmo cuando encontré opciones que me interesaban.

Ansiedad

9. Me sentí ansioso(a) mientras evaluaba diferentes opciones.

10. La posibilidad de equivocarme me generó ansiedad.

Deseo

11. Sentía un fuerte deseo de obtener el auto que estaba viendo.

12. Quería avanzar rápidamente con la compra porque el auto me gustaba mucho.

Satisfacción

13. Me sentí satisfecho(a) con el proceso de compra.

14. Al finalizar la compra, sentí satisfacción con mi decisión.

Miedo

15. Tenía miedo de ser estafado(a) o engañado(a).

16. Sentí temor de comprar un vehículo con fallas ocultas.

Esperanza

17. Tenía expectativas positivas sobre el desempeño del vehículo.

18. Esperaba que mi compra mejorara mi vida o movilidad.

Orgullo

19. Sentí orgullo al considerar o comprar el automóvil.

20. La compra del vehículo me hizo sentir bien conmigo mismo(a).

Culpa

21. Me sentí culpable por gastar dinero en esta compra.

22. Experimenté culpa pensando en otros gastos o responsabilidades.

Confusión

23. Me sentí confundido(a) por la cantidad de información disponible.

24. Tuve dificultades para comparar opciones o entender las diferencias.

Impulsividad

25. Tomé decisiones rápidas durante la compra.

26. Avancé sin analizar todos los detalles que debía considerar.

Seguridad

27. Sentí seguridad en el vehículo que evalué o compré.

28. Me sentí seguro(a) de haber tomado una buena decisión.

Anexo 3. CARTA PARA SOLICITAR ACCESO A EMPRESAS DE AUTOS SEMINUEVOS

[Aguascalientes, Ags. Méx.], [Fecha]

A quien corresponda:

Nombre de la empresa

Dirección

Ciudad

Estimados(as) señores(as):

Mi nombre es _____, profesor/investigador de _____.

Me permito solicitar su amable autorización para realizar una investigación titulada:

“Emociones y toma de decisiones en la compra de automóviles seminuevos en México”.

El estudio tiene como objetivo analizar cómo factores emocionales, cognitivos y de percepción de riesgo influyen en la adquisición de autos seminuevos. La información será obtenida de manera responsable mediante:

- Cuestionarios anónimos a clientes que hayan comprado un automóvil seminuevo en su establecimiento, o
- Entrevistas cortas a compradores recientes (opcional y si la empresa lo autoriza).

Compromisos del investigador:

1. No interferir en la operación o experiencia de venta.
2. No solicitar información personal de clientes.
3. No registrar datos sensibles o confidenciales.
4. Entregar resultados generales a la empresa si así lo solicita.
5. Garantizar anonimato total de participantes y de la empresa.

Agradezco de antemano su colaboración y quedo atento a resolver cualquier duda.

Atentamente,

Nombre

Teléfono

Correo electrónico

3. Protocolo de confidencialidad

Anexo 4. PROTOCOLO DE CONFIDENCIALIDAD DE DATOS

Estudio: *“Emociones y toma de decisiones en la compra de automóviles seminuevos”*

El presente documento establece las medidas que se adoptarán para garantizar la confidencialidad, seguridad y protección de los datos recabados en este estudio.

1. Naturaleza de los datos recolectados

Los datos consisten únicamente en:

- Variables sociodemográficas generales (edad, género, ingreso aproximado).
- Percepciones y emociones autoinformadas.
- Experiencia de compra.
- No se recogerán nombres, correos, teléfonos ni identificadores personales.

2. Codificación

Cada participante será identificado mediante un código alfanumérico (ej.: P001, P002).

Ningún archivo contendrá información personal asociada al código.

3. Almacenamiento seguro

- Los datos serán almacenados en archivos protegidos con contraseña.
- Solo el investigador responsable tendrá acceso.
- No se compartirán bases de datos sin anonimización.

4. Uso de datos

Los datos solo se utilizarán con fines académicos:

- artículos científicos
- reportes académicos
- presentaciones

Nunca se utilizarán para fines comerciales.

5. Eliminación de información

Una vez concluido el estudio, los datos serán resguardados por un periodo máximo de 2 años y posteriormente eliminados de forma definitiva.

Firma del investigador responsable: _____

Fecha: _____



This is an open access article distributed under the terms of the CC BY-NC license (<https://creativecommons.org/licenses/by-nc/4.0/>)