

## Evaluación de preferencias de consumidores en la compra de computadores portátiles: Un estudio de segmentación y clasificación desde el marketing transformacional

*Assessment of consumer preferences in laptop purchases: A segmentation and classification study from a transformational marketing perspective*

Andrea C. Droguett<sup>1\*</sup>  <https://orcid.org/0000-0002-2827-4818>

Marcelo León Vargas<sup>1</sup>  <https://orcid.org/0000-0002-4149-8920>

Elías J. Bracho<sup>1</sup>  <https://orcid.org/0000-0002-9639-8687>

Recibido 12 de octubre de 2024, aceptado 06 de noviembre de 2024

*Received: October 12, 2024 Accepted: November 06, 2024*

### RESUMEN

En un entorno empresarial dinámico, la segmentación de consumidores es vital para el crecimiento y desarrollo. Este estudio propone un enfoque integrado utilizando aprendizaje automático para la segmentación posterior basada en los beneficios buscados por los consumidores, desde una perspectiva de marketing transformacional. La investigación utiliza algoritmos de máquinas de soporte vectorial (SVM) y árboles de clasificación (DTC) para efectuar la clasificación inicial, y el algoritmo de K-medias para segmentar a los consumidores según sus preferencias. La muestra consistió en 1.000 participantes. Los resultados revelaron grupos de consumidores con patrones de preferencias / comportamientos distintivos, proporcionando una visión detallada de la cartera de consumidores potenciales. En el marco del marketing transformacional, estos hallazgos permiten diseñar estrategias más allá de la simple promoción de productos, centrándose en experiencias personalizadas y comunicaciones dirigidas. La segmentación basada en beneficios identificó cuatro segmentos: el primero destacando economía y usabilidad, el segundo enfocado en rendimiento, el tercero mostrando preferencias en diseño, rendimiento y economía, y el cuarto resaltando diseño y rendimiento. El estudio demuestra que tanto el algoritmo SVM como DTC tienen un alto rendimiento en la clasificación de consumidores, proporcionando una base sólida para la aplicación práctica de estas técnicas. En resumen, este enfoque integrado utilizando aprendizaje automático ofrece una visión detallada y práctica para la segmentación de consumidores, abriendo oportunidades para estrategias de marketing más efectivas.

**Palabras clave:** Segmentación a posteriori, comportamiento del consumidor, segmentación por beneficio buscado, técnicas de clasificación en Machine Learning.

### ABSTRACT

*In a dynamic business environment, consumer segmentation is vital for growth and development. This study proposes an integrated approach using machine learning for post-hoc segmentation based on the benefits sought by consumers from a transformational marketing perspective. The research uses*

---

<sup>1</sup> Universidad Viña del Mar. Escuela de Ingeniería y Negocios. Viña del Mar, Chile.

E-mail: andrea.droguett@uvm.cl; marcelo.leon@uvm.cl; elias.bracho@uvm.cl

\* Autor de correspondencia: andrea.droguett@uvm.cl

*Support Vector Machine (SVM) and Decision Tree Classification (DTC) algorithms to perform the initial classification and the K-means algorithm to segment consumers according to their preferences. The sample consisted of 1,000 participants. The results revealed groups of consumers with distinctive preference/behavior patterns, providing a detailed view of the potential consumer portfolio. Within the framework of transformational marketing, these findings allow for the design of strategies beyond simple product promotion, focusing on personalized experiences and targeted communications. Benefit-based segmentation identified four segments: the first highlighting economy and usability, the second focusing on performance, the third showing preferences in design, performance, and economy, and the fourth highlighting design and performance. The study demonstrates that both the SVM and DTC algorithms have high performance in consumer classification, providing a solid foundation for the practical application of these techniques. In summary, this integrated machine learning approach offers a detailed and practical view of consumer segmentation, opening opportunities for more effective marketing strategies.*

**Keywords:** *Post-hoc segmentation, consumer behavior, benefit sought segmentation, classification techniques in Machine Learning.*

## INTRODUCCIÓN

La palabra “marketing” comenzó a usarse en Estados Unidos a principios del siglo XX, cuando la Universidad de Michigan lanzó un curso llamado “La Industria distributiva y reguladora en Estados Unidos”, utilizando por primera vez el término marketing en su folleto [1]. Desde entonces, el marketing ha evolucionado para adaptarse a las nuevas necesidades de los clientes y a los cambios en la industria. La llegada de Internet ha revolucionado el comercio electrónico, especialmente durante la pandemia, y ha llevado al desarrollo del marketing digital en los años 90. Este concepto inicialmente se centraba en la publicidad, pero con el surgimiento de herramientas sociales y móviles, se expandió para crear experiencias más interactivas entre el usuario y la marca [2].

En la actualidad, las empresas buscan segmentar sus mercados para entender mejor las preferencias y gustos de los consumidores, tanto en el ámbito tradicional como en el digital [3] y en redes sociales [4]. El marketing ha evolucionado hacia el uso intensivo de datos para comprender el comportamiento del consumidor, incorporando aspectos éticos y morales en la toma de decisiones de compra [5]. La inteligencia artificial ha permitido a los especialistas en marketing utilizar herramientas digitales y técnicas de minería de datos con machine learning para obtener información valiosa sobre consumidores, competencia y mercados [6], [7].

La recopilación de información sobre los consumidores facilita la interacción y el desarrollo de engagement,

fomentando una mayor identificación con la marca y convirtiendo a los clientes en embajadores de marketing [8].

La segmentación del mercado sigue siendo una preocupación central para los ejecutivos de marketing, ya que es fundamental para diseñar campañas exitosas. Tradicionalmente, se han utilizado criterios a priori como variables geográficas, demográficas y psicográficas. Sin embargo, con el cambio hacia consumidores altamente informados y con diversos requerimientos, estos criterios parecen menos efectivos para alcanzar los objetivos comerciales.

Desde una perspectiva analítico-comportamental, el comportamiento del consumidor se entiende como una intersección entre los antecedentes del consumo (historia de aprendizaje y ambiente) y los resultados del consumo (beneficios o pérdidas), según el modelo de perspectiva comportamental BPM (Behavior Perspective Model) [9]. Este modelo sugiere que tanto el entorno del consumidor como el comportamiento cambian constantemente debido a la interacción con su ambiente. Así, los mecanismos para analizar estos cambios ofrecen la posibilidad de ajustar y proporcionar experiencias de consumo adaptadas a las necesidades del consumidor [10].

Cambios en el ambiente, como los experimentados durante la pandemia de COVID-19 [11], demostraron cómo los antecedentes y consecuencias del consumo pueden variar y afectar diferentes tipos de consumo en el futuro [12]. Esto subraya la necesidad de buscar nuevos criterios de segmentación basados en lo que

realmente desean y necesitan los consumidores, utilizando técnicas de machine learning.

La investigación se sustenta bajo un muestreo no probabilístico por conveniencia, la selección del grupo de estudio corresponde a aquellos sujetos registrados en la base de datos de compras disponible durante el año 2021, asumiendo un criterio de inclusión de individuos que compraron computadores durante el período. Dicha base de datos representa la información recogida de diversas tiendas de retail en la Región Metropolitana de Santiago de Chile.

Las preguntas de investigación que emergen son: ¿Cuáles son las características de los consumidores de computadores en la muestra? ¿Qué beneficios buscan al comprar un computador portátil? ¿Existen grupos de consumidores con preferencias distintas según sus características sociodemográficas? ¿Hay diferencias significativas en el desempeño entre las técnicas de segmentación más comunes?

Las hipótesis derivadas son: Es posible identificar grupos con preferencias distintas sobre los atributos de un portátil. Las preferencias de los consumidores estarán relacionadas con sus características sociodemográficas. Los algoritmos de clasificación presentarán diferentes grados de desempeño en el proceso.

## REVISIÓN DE LA LITERATURA

### Segmentación

La segmentación de mercados es un proceso mediante el cual el conjunto de consumidores se divide en grupos homogéneos de menor tamaño [13], en el entendido de que no es posible ofrecer un “único” producto que satisfaga las necesidades de todos los compradores, o a todos los compradores de igual modo [14], [15], [16]. Tómese un producto tan simple como la goma de mascar. Se notará que los mismos se venden en distintas cantidades, con diferentes sabores, con distintos contenidos de azúcar y con diferentes niveles de calorías. Esta técnica de dividir un mercado heterogéneo en grupos más pequeños no siempre fue la norma. Hasta por lo menos la primera mitad del siglo XX, la mayoría de la producción era masiva y se centraba en un único producto para bajar el costo total unitario y trasladar los ahorros al consumidor.

### Segmentación a priori versus a posteriori

No existe una manera única de segmentar un mercado [16]. Una decisión importante que debe tomarse en cuenta al efectuar una segmentación de mercados refiere al método para delinear las características de cada segmento [17]. Una aproximación consiste en dividir el mercado en grupos de consumidores en función de un criterio que se fija de antemano, utilizando una o más variables en combinación [14], [16], [17]. Por ejemplo, un gestor de marketing puede dividir a los compradores de un mercado de acuerdo a su frecuencia de compra [14]. La segmentación de mercados a priori es fuente de críticas dado que tiene poca capacidad de predecir el comportamiento de los compradores [17], [18].

Una segunda aproximación, denominada segmentación a posteriori o segmentación post hoc, agrupa a los individuos en segmentos en función de los resultados de un estudio de mercado [14]. La segmentación de mercados post hoc deriva a los grupos de consumidores en base a datos empíricos en función de que tan similares o disímiles son las respuestas de los individuos a un conjunto de características medibles.

Algunos documentos académicos de segmentación de mercados turísticos recurren a la segmentación a priori, dado que los segmentos se conocían con anterioridad. Los estudios a posteriori recurren a las técnicas de análisis factorial y de análisis clúster para determinar el tamaño y las características de los segmentos de turistas [18].

El objetivo del estudio es segmentar, a posteriori según el criterio del beneficio buscado, el mercado de consumidores de computadores portátiles, comparando el desempeño de dos técnicas de agrupación de machine learning sobre las características buscadas por los consumidores, para permitir a los directivos de negocios diseñar estrategias de marketing transformacional efectivas, adecuando la oferta de valor de la empresa [19], [20].

Los objetivos específicos son: describir las variables en estudio, tanto las de tipo sociodemográfico como las de beneficios buscados; aplicar las técnicas de clasificación SVM y DTC al set de datos; segmentar la muestra con base en sus preferencias en los atributos de un equipo mediante el algoritmo de K-medias; y determinar asociaciones y relaciones entre las

variables sociodemográficas y de preferencias en los atributos.

## MATERIALES Y MÉTODOS

### Materiales

Para la selección de la muestra de 1.000 usuarios de computadores portátiles, se utilizó el método de muestreo no probabilístico por conveniencia, aplicado en las regiones Metropolitana y de Valparaíso por medio de una plataforma de encuestas on-line. El set de datos recuperado contiene 35 variables de tipo sociodemográfico, de preferencia acerca de atributos buscados al comprar un computador y de hábitos de búsqueda de información de tecnología, como de canales de compra más frecuentes. En la Tabla 1, se especifican las variables.

Para la selección de la muestra de 1.000 usuarios de computadores portátiles, se utilizó el método de muestreo no probabilístico por conveniencia, aplicado en las regiones Metropolitana y de Valparaíso por medio de una plataforma de encuestas on-line. El set de datos recuperado contiene 35 variables de tipo sociodemográfico, de preferencia acerca de atributos buscados al comprar un computador y de hábitos de búsqueda de información de tecnología, como de canales de compra más frecuentes. En la Tabla 1, se especifican las variables.

Para un mejor entendimiento de los datos, se realizó un esquema estrella de la estructura de los datos, que se presenta en la Figura 1.

Para el caso de las variables de atributos preferidos por los potenciales consumidores, se utilizó una escala de respuesta entre 0 y 11 puntos que indica el grado de acuerdo con la importancia de la variable para la decisión de compra. La amplitud en el intervalo de respuesta utilizado obedece a que se busca maximizar la capacidad discriminante entre los encuestados. Para resumir la importancia de cada categoría de esta variable se tomó la media en la puntuación total de cada categoría, creando cuatro nuevas variables de resumen, estas son: beneficio promedio en la categoría usabilidad, beneficio promedio en la categoría diseño, beneficio promedio en la categoría performance y beneficio promedio en la categoría economía. Estas cuatro categorías resumidas aparecen en la tabla central del esquema estrella en las primeras posiciones.

Tabla 1. Descripción de las variables del estudio.

Variables sociodemográficas	<ul style="list-style-type: none"> <li>• Sexo</li> <li>• Edad</li> <li>• Nivel de educación</li> <li>• Ingreso familiar (\$M)</li> <li>• Ocupación</li> <li>• Estado civil</li> <li>• N° de hijos</li> <li>• Tipo vivienda</li> <li>• Urbano-rural</li> <li>• Conocimiento tecnológico</li> </ul>
Variables atributos buscados	<ul style="list-style-type: none"> <li>• Usabilidad</li> <li>• Fácil uso</li> <li>• Fácil configuración</li> <li>• Baja complejidad</li> <li>• Amplia conectividad</li> <li>• Capacitación del vendedor</li> <li>• Existencia servicio técnico</li> <li>• Servicio confiable</li> <li>• Diseño</li> <li>• Atractivo</li> <li>• Diferenciado</li> <li>• Monitor amplio</li> <li>• Estilo sofisticado</li> <li>• Performance</li> <li>• Procesamiento eficiente</li> <li>• Manejo de volumen de datos</li> <li>• Rapidez</li> <li>• Alta memoria</li> <li>• Alta capacidad gráfica</li> <li>• Última tecnología</li> <li>• Economía</li> <li>• Accesorios económicos</li> <li>• El menor precio de mercado</li> </ul>
Variables de hábitos de búsqueda de información tecnológica	<ul style="list-style-type: none"> <li>• Televisión</li> <li>• Internet</li> <li>• Revistas especializadas</li> </ul>
Variables de hábitos de canal de compra	<ul style="list-style-type: none"> <li>• Online</li> <li>• Tiendas especializadas</li> <li>• Retail</li> </ul>

### Métodos

El primer método supervisado de clasificación utilizado es el algoritmo Support Vector Machine (SVM). En este método de clasificación supervisada, los datos (p) son interpretados como un vector p-dimensional en un hiperplano de n dimensiones en el que cada uno de ellos pertenece a posibles categorías, de esta forma el algoritmo basado en SVM construye un modelo capaz de predecir si un punto nuevo, cuya categoría es desconocida, pertenece a una categoría o a la otra.

SVM es un modelo que, partiendo de un conjunto de ejemplos de entrenamiento, los datos son

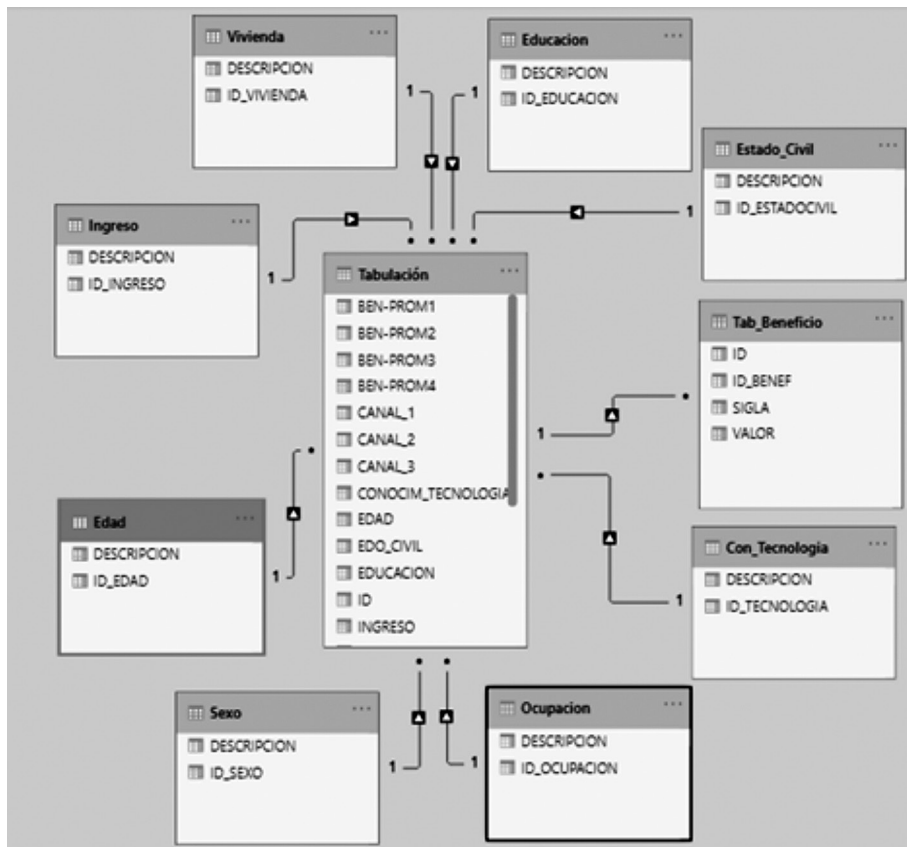


Figura 1. Esquema estrella de los datos.

etiquetados en diferentes clases y representarlas los puntos en el espacio para tratar de separar las diferentes clases mediante un espacio lo más amplio posible o separación óptima, para que cuando las nuevas muestras de los casos de prueba se pongan en correspondencia con dicho modelo puedan ser clasificadas correctamente en función de su proximidad [21].

El segundo método utilizado es el de Árboles de Decisión, técnica explicativa y descomposicional que utiliza un proceso de división secuencial, iterativo y descendente que, partiendo de una variable dependiente, forma grupos homogéneos definidos específicamente mediante combinaciones de variables independientes en las que se incluyen la totalidad de los casos recogidos en la muestra. En particular se eligió el algoritmo CART (Classification and Regression Trees) ya que genera árboles de decisión binarios y admite variables de entrada y de salida nominales, ordinales y continuas.

El tercer método utilizado es el de K-medias, es un algoritmo de clasificación no supervisada que agrupa objetos en  $k$  grupos basándose en sus características, en rigor es un algoritmo de optimización ya que el agrupamiento se realiza minimizando la suma de distancias entre cada objeto con respecto al centroide de su segmento o clúster. La principal ventaja que presenta es que es un método poco complejo y rápido, pero en qué se debe decidir previamente el número de segmentos para efectuar el agrupamiento de los datos [22].

## RESULTADOS Y DISCUSIÓN

### Resultados

Para el análisis de los datos y discusión de los resultados se completó el proceso descrito en la Figura 2.

La frecuencia de la variable edad se describe en la Tabla 2 y Figura 3. Se observa que la mayor proporción de la muestra está en el rango de edad

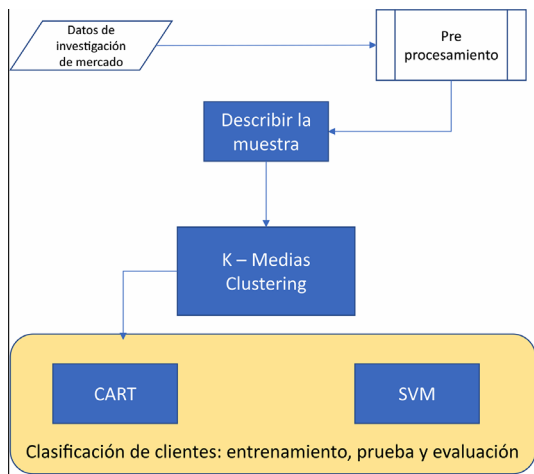


Figura 2. Procedimiento de análisis de datos.

Tabla 2. Edad.

		Frecuencia	Porcentaje
Válido	17-25	232	23,2
	26-35	415	41,5
	36-50	250	25,0
	>50	103	10,3
	Total	1000	100,0

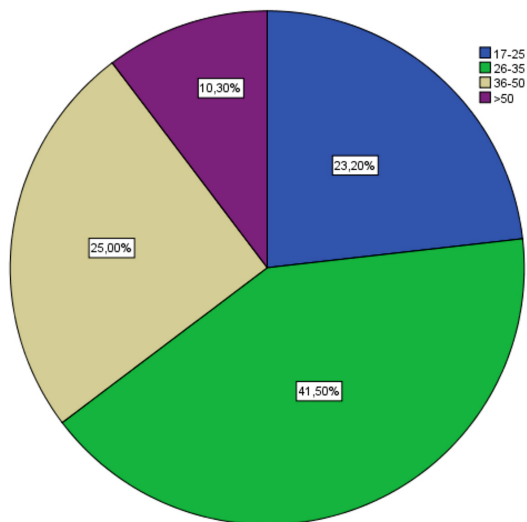


Figura 3. Edad.

entre 26 y 35 años que representa el 41,5% de los entrevistados. Seguido por el segmento entre 36 y 50 años que representa el 25% de la muestra.

La frecuencia de la variable género se describe en la Tabla 3 y Figura 4. Se aprecia que prácticamente hay igualdad entre los dos grupos compuesto por 52,5% de hombres y 47,5% de mujeres.

La frecuencia de la variable Educación se describe en la Tabla 4 y Figura 5. En la muestra estudiada la mayoría de los entrevistados tiene un nivel de educación universitaria incompleta que representa el 68,3% de los entrevistados.

La frecuencia de la variable ocupación se describe en la Tabla 5 y Figura 6. Se observa que la mayor

Tabla 3. Género.

		Frecuencia	Porcentaje
Válido	M	525	52,5
	F	475	47,5
	Total	1000	100,0

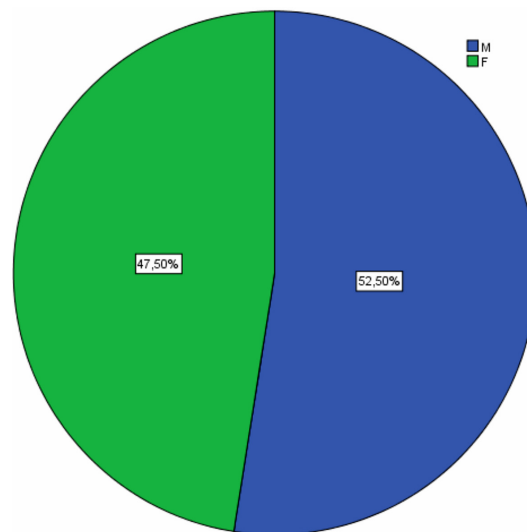


Figura 4. Género.

Tabla 4. Educación.

		Frecuencia	Porcentaje
Válido	Secundaria o menos	333	33,3
	Univ_incompleta	350	35,0
	Univ_completa	239	23,9
	Magíster o sup.	78	7,8
	Total	1000	100,0

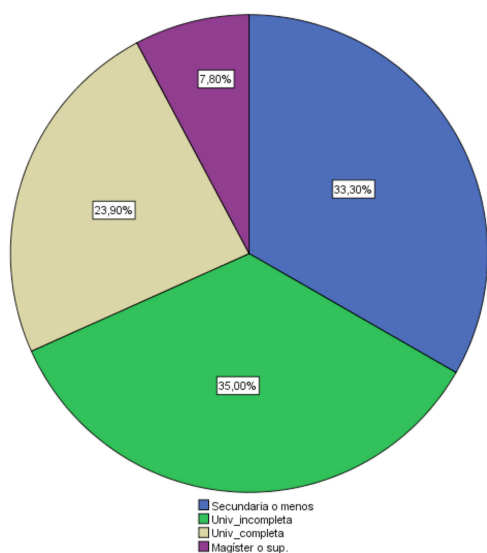


Figura 5. Educación.

Tabla 5. Ocupación.

		Frecuencia	Porcentaje
Válido	Obrero	124	12,4
	Empleado_of.	280	28,0
	Independiente	195	19,5
	Directivo	88	8,8
	Estudiante	187	18,7
	Trabajo_casa	126	12,6
	Total	1000	100,0

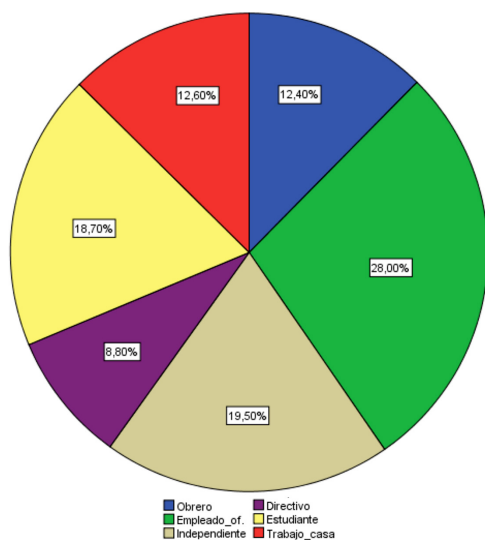


Figura 6. Ocupación.

proporción de la muestra tiene ocupación de empleado de oficina o estudiante que representan en conjunto 46,7% del total

La frecuencia de búsquedas en televisión se muestra en la Tabla 6. Claramente la televisión no es frecuentemente utilizada para adquirir información acerca de equipos portátiles. Sólo un 29,6% de la muestra la utiliza.

La frecuencia de búsquedas en Internet se muestra en la Tabla 7. En este caso sí presenta utilización importante ya que casi el 50% de la muestra la utiliza con frecuencia.

La frecuencia de búsquedas en revistas se muestra en la Tabla 8. Luego las revistas especializadas son una alternativa moderada a la hora de buscar información acerca del producto. Solo es utilizada por 35,3% de la muestra.

La frecuencia de sitios Web como canal de compra se muestra en la Tabla 9. Se aprecia entonces que es un canal con alta preferencia como punto de venta con 61,1% de la muestra.

Tabla 6. Televisión.

		Frecuencia	Porcentaje
Válido	No	704	70,4
	Si	296	29,6
	Total	1000	100,0

Tabla 7. Internet.

		Frecuencia	Porcentaje
Válido	No	514	51,4
	Si	486	48,6
	Total	1000	100,0

Tabla 8. Revistas.

		Frecuencia	Porcentaje
Válido	No	647	64,7
	Si	353	35,3
	Total	1000	100,0

Tabla 9. Web.

		Frecuencia	Porcentaje
Válido	No	389	38,9
	Si	611	61,1
	Total	1000	100,0

La frecuencia de uso de tiendas especializadas como canal de compra se muestra en la Tabla 10. Notoriamente este canal no es utilizado y sólo 23,1% las considera a la hora de efectuar una compra.

La frecuencia de búsquedas en retail se muestra en la Tabla 11. Igualmente, que las tiendas especializadas, el retail tampoco es una alternativa a considerar al momento de comprar un equipo, tan solo un 17,6% de los entrevistados lo considera una alternativa.

La agrupación de las variables género, edad, educación, ingreso y ocupación, se muestra en las Tablas 12-14 y Figuras 7-9, respectivamente.

En cuanto al género y edad, se aprecia tanto en la Tabla 12 como en la Figura 7 que la distribución

Tabla 10. Tiendas especializadas.

		Frecuencia	Porcentaje
Válido	No	769	76,9
	Si	231	23,1
	Total	1000	100,0

Tabla 11. Retail.

		Frecuencia	Porcentaje
Válido	No	824	82,4
	Si	176	17,6
	Total	1000	100,0

Tabla 12. Género y edad.

Recuento						
		EDAD				Total
		17-25	26-35	36-50	>50	
Sexo	M	132	202	130	61	525
	F	100	213	120	42	475
Total		232	415	250	103	1000

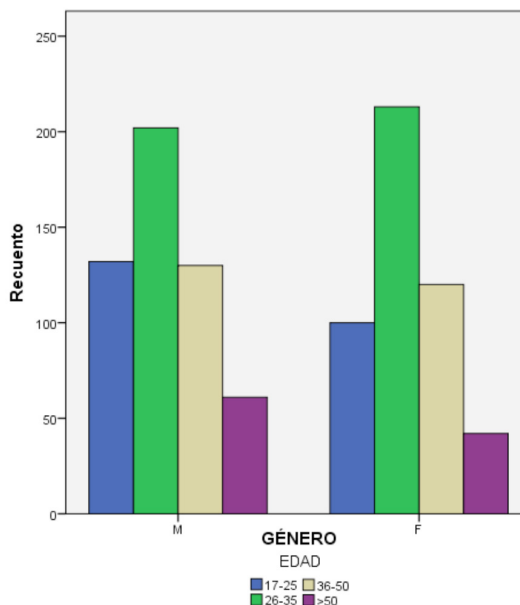


Figura 7. Género y edad.

por sexo en cada rango de edad es equitativa sin tendencia marcada hacia un sexo en particular.

A partir de la Tabla 13, un elemento a destacar es que la mayor parte de la muestra tiene un ingreso familiar ente \$MM 1 y \$MM 2, presenta educación universitaria incompleta y no hay mayor diferencia en la proporción por género.

Respecto al grado de conocimiento de tecnología, la Tabla 14 muestra que en un nivel básico se ubica el segmento de empleados de oficina, con nivel medio los segmentos que comparten la mayoría son los estudiantes y empleados de oficina y con un nivel avanzado también está representado por el grupo de empleados de oficina, explicado por que son el segmento con mayor proporción de la muestra alcanzando 28% del total. En cuanto al género no



Tabla 13. Género, educación e ingreso.

Recuento							
Ingreso FAM			Educacion				Total
			Secundaria o menos	Univ_incompleta	Univ_completa	Magíster o sup.	
<1000	Sexo	M	110	22	18	6	156
		F	93	23	16	4	136
	Total		203	45	34	10	292
Entre1000 y 2000	Sexo	M	39	96	69	15	219
		F	46	91	48	12	197
	Total		85	187	117	27	416
Entre2000 y 3000	Sexo	M	11	47	26	12	96
		F	19	39	31	10	99
	Total		30	86	57	22	195
Entre3000 y 5000	Sexo	M	8	19	17	10	54
		F	7	13	14	9	43
	Total		15	32	31	19	97
Total	Sexo	M	168	184	130	43	525
		F	165	166	109	35	475
	Total		333	350	239	78	1000

Tabla 14. Género, ocupación y dominio de tecnología.

Recuento									
Conocim_tecnologia			Ocupacion					Total	
			Obrero	Empleado_of	Independiente	Directivo	Estudiante		Trabajo_casa
Nulo	Sexo	M	13	23	6	5	15	11	73
		F	11	16	15	3	17	15	77
	Total		24	39	21	8	32	26	150
Básico	Sexo	M	23	60	48	7	28	22	188
		F	18	59	34	14	19	21	165
	Total		41	119	82	21	47	43	353
Medio	Sexo	M	16	43	32	19	33	22	165
		F	24	34	24	14	44	12	152
	Total		40	77	56	33	77	34	317
Avanzado	Sexo	M	10	24	21	14	14	16	99
		F	9	21	15	12	17	7	81
	Total		19	45	36	26	31	23	180
Total	Sexo	M	62	150	107	45	90	71	525
		F	62	130	88	43	97	55	475
	Total		124	280	195	88	187	126	1000

se aprecian mayores diferencias en los segmentos identificados.

## DISCUSIÓN

El estudio descriptivo en las características socio-demográficas da cuenta que la mayor parte de la muestra está en edad laboralmente activa, entre 26 y 50 años, presentando mayormente educación universitaria incompleta. Respecto a los medios de búsqueda de información tecnológica y canales de compra, Internet aparece liderando como medio de búsqueda de información, seguido por revistas especializadas, en cuanto a los canales de venta más utilizados por lejos la compra en sitios Web es la más recurrente, evidenciando el rezago de los puntos de venta presenciales como alternativa para efectuar una compra de equipos portátiles.

Al efectuar la agrupación de las variables, destaca que por cada segmento de edad estudiado no hay mayor diferencia de género, presentando la muestra un ingreso familiar más frecuente entre \$MM 1 y \$MM 2 en el grupo de educación universitaria incompleta. Respecto al conocimiento tecnológico, tanto en los niveles básico, medio y avanzado son los empleados de oficina los que presentan mayor representatividad.

### Segmentación de la muestra con algoritmo K-Medias

Para la selección del número óptimo de segmentos se usó el método gráfico del codo, que permite

identificar el punto de corte para el cual se minimiza la suma de cuadrados de las distancias desde el centroide a cada dato (WCSS) o dicho de otra forma se minimiza la varianza intragrupo, como se presenta en la Figura 8.

Se observa que el punto donde el error pasa a tener rendimientos marginales decrecientes es desde los cuatro grupos y por tanto en este punto se identifica el óptimo de segmentos a seleccionar en el algoritmo de K-Medias. La presentación de los segmentos se presenta en la Figura 9.

A partir del total de la muestra de 1.000 personas, se identifican cuatro segmentos según los beneficios buscados al comprar computadores portátiles, identificados desde el segmento 0 al segmento 3, la composición de estos es: 353 personas etiquetados en el segmento 0, 147 personas etiquetadas en el segmento 1, 298 personas etiquetadas en el segmento 2 y 202 personas etiquetadas en el segmento 3. Cabe señalar que los procedimientos de segmentación a posteriori (después de una investigación de mercado), actualmente se consideran mucho más efectivos que los procedimientos de segmentación a priori (a partir de variables conocidas), como son los realizados con variables socioeconómicas, demográficas, psicográficas, entre otras.

Los segmentos identificados presentan las puntuaciones promedio indicadas en la Tabla 15, en relación con los beneficios que buscan en un computador

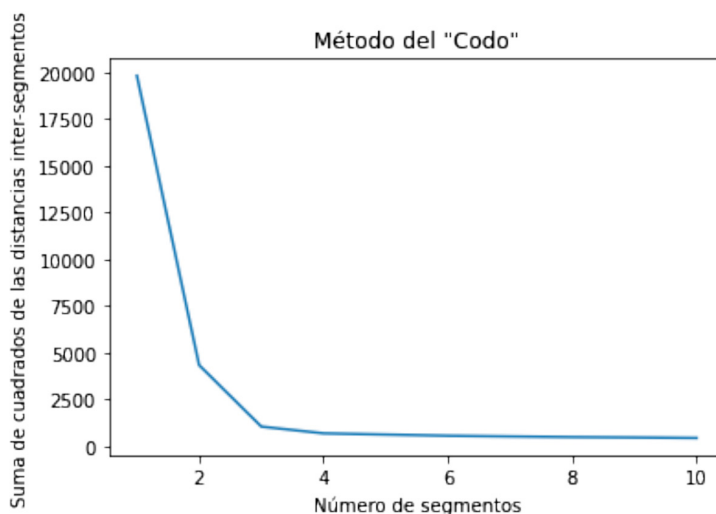


Figura 8. Selección óptima de segmentos.

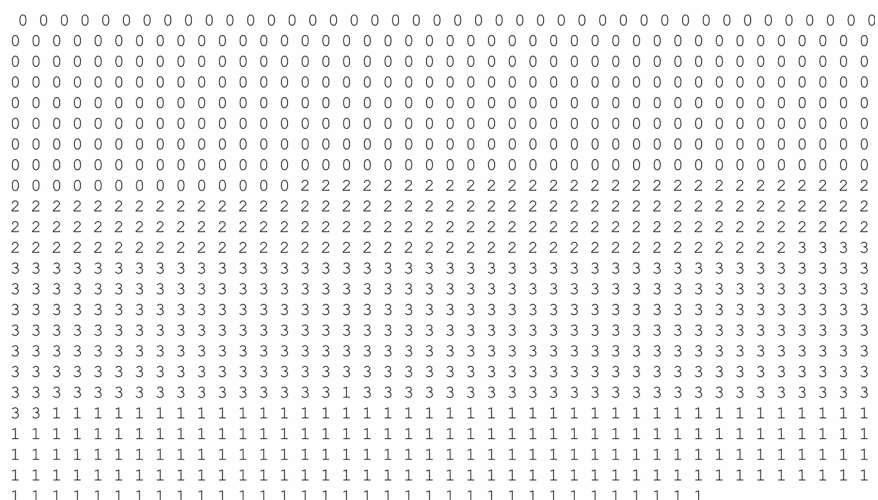


Figura 9. Presentación de los segmentos en mercados de computadores.

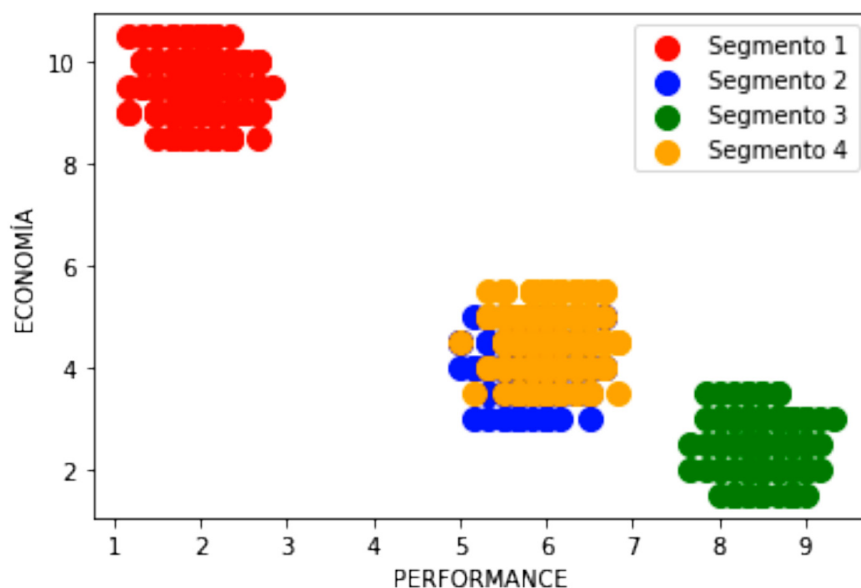


Figura 10. Representación gráfica de los segmentos en mercados de computadores.

portátil, considerando que las opciones de valoración eran de 1 a 11 puntos, dónde 1 era muy poco importante y 11 de máxima importancia.

Dónde las dimensiones son, Ben\_Prom1: Usabilidad. Ben\_Prom2: Diseño, Ben\_Prom3: Performance y Ben\_Prom4: Economía

Se aprecia que el segmento 0 (segmento 1 en la Figura 11) presenta marcada preferencia con los

beneficios asociados a Economía y Usabilidad. El segmento 1 (segmento 2 en la Figura 11) marcada preferencia con los beneficios asociados a Performance. El segmento 2 (segmento 3 en la Figura 11) marcada preferencia con los beneficios asociados a Usabilidad, Diseño y Performance y el segmento 3 (segmento 4 en la Figura 11) marcada preferencia con los beneficios asociados a Diseño y Performance. En la Figura 11 se corrobora las preferencias por segmentos.

Tabla 15. Puntuaciones por cada segmento.

S0				
	BEN-PROM1	BEN-PROM2	BEN-PROM3	BEN-PROM4
Media	7,854	2,218	1,994	9,497
S1				
	BEN-PROM1	BEN-PROM2	BEN-PROM3	BEN-PROM4
Media	3,229	2,53	8,514	2,443
S2				
	BEN-PROM1	BEN-PROM2	BEN-PROM3	BEN-PROM4
Media	6,422	6,017	6,002	4,513
S3				
	BEN-PROM1	BEN-PROM2	BEN-PROM3	BEN-PROM4
Media	4,766	5,961	5,851	4,052

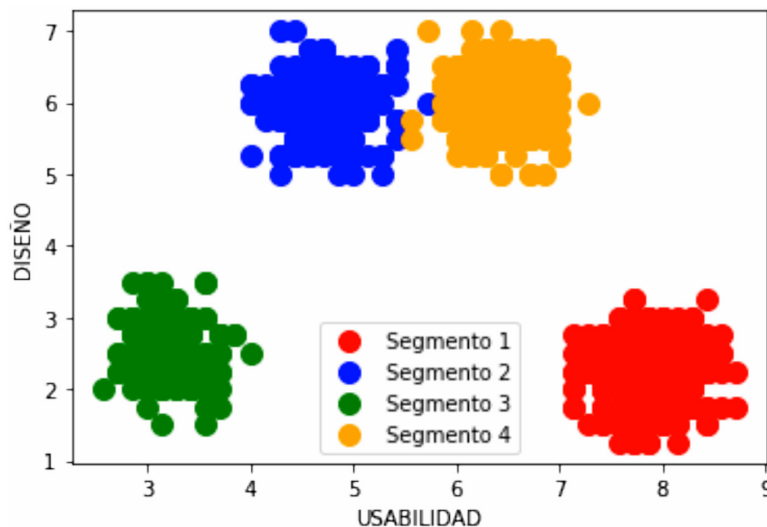


Figura 11. Esquema de segmentos y beneficios buscados en mercados de computadores.

Se aprecia que el segmento 0 (segmento 1 en la Figura 10 y Figura 11) presenta marcada preferencia con los beneficios asociados a Economía y Usabilidad. El segmento 1 (segmento 2 en la Figura 10 y Figura 11) marcada preferencia con los beneficios asociados a Performance y Diseño. El segmento 2 (segmento 3 en la Figura 10) marcada preferencia con los beneficios asociados a Performance y el segmento 3 (segmento 4 en la Figuras10 y Figura 11) marcada preferencia con los beneficios asociados a Diseño, Usabilidad y Performance. En la Figura 10 y Figura 11 se corroboran las preferencias por segmentos.

Se aprecia que el primer segmento identificado (segmento 0) corresponde al color rojo que presenta (segmento 1 en la Figura 11) marcada preferencia por economía y usabilidad. El segundo segmento (segmento 1) corresponde al color verde (segmento 3 en la Figura 11) que presenta marcada preferencia por performance. El tercer segmento (segmento 2) corresponde al color amarillo (segmento 4 en la Figura 11) que presenta marcada preferencia por diseño, performance y economía. Finalmente, el cuarto segmento (segmento3) corresponde al color azul (segmento 2 en la Figura 11) que presenta marcada preferencia por diseño y performance.

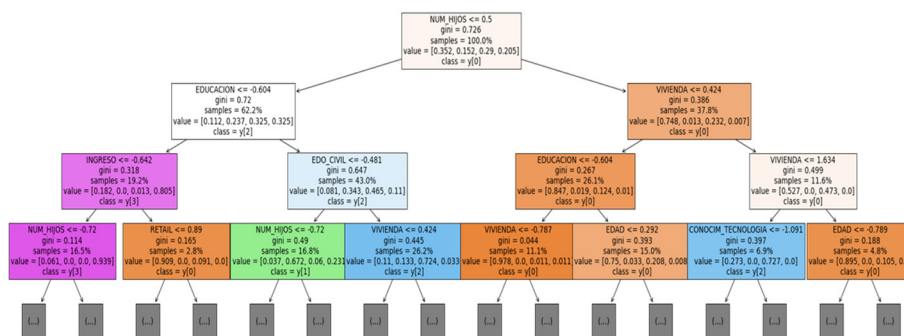


Figura 12. Árbol de clasificación.

**Clasificación con Algoritmo Árbol de Clasificación**

La Figura 12 muestra el Árbol de Clasificación. En esta visualización se aprecia que en el primer nodo se ubica el 100% de la muestra con las proporciones de entrevistado en cada segmento en el vector values, se ve que el mayor representante es el segmento 0 con 35,2% de la muestra el menor el segmento 1 con 15,2% de la muestra, esto con el índice de impureza de Gini de 0,726 que indica que hay más de dos clases, mientras mayor el índice, mayor cantidad de clase existirá. Seguidamente el criterio de segmentación más influyente para cada uno de los 4 segmentos es el número de hijos de la familia, en el punto de corte de 2,5 hijos aproximadamente. Luego para las familias con menor número de hijos desde el punto de corte, la variable que más explica es el nivel de educación, representando al 62,2% de la muestra y para las familias con mayor número de hijo desde el punto de corte será la variable tipo de vivienda representando al 37,8% de la muestra. De esta manera se puede seguir descendiendo en los niveles de clasificación, siguiendo los mismos criterios antes mencionados.

Las medidas de desempeño de este algoritmo supervisado se señalan en la Tabla 16, desde una matriz de confusión multiclase, partiendo de la premisa que para efectos de entrenamiento se ocupó el 80% de la muestra y para la prueba el 20% de los datos.

Lo que indica que el algoritmo tiene alto poder discriminante para detectar a los falsos negativos y

Tabla 16. Medidas desempeño Árbol de Decisión.

Precisión	0,94
Sensibilidad (Recall)	0,90
Especificidad	0,95

a los falsos positivos, vale decir aquellos sujetos mal clasificados en cada uno de los cuatro segmentos.

En la implementación de este algoritmo se utilizó Kernel de base radial que permite el tratamiento con datos no lineales Las medidas de desempeño de este algoritmo supervisado se señalan en la Tabla 17, desde una matriz de confusión multiclase, partiendo de la premisa que para efectos de entrenamiento se ocupó el 80% de la muestra y para la prueba el 20% de los datos.

Lo que indica que el algoritmo tiene alto poder discriminante para detectar a los falsos negativos y a los falsos positivos, vale decir aquellos sujetos mal clasificados en cada uno de los cuatro segmentos. En consecuencia, se comprueba que tanto el algoritmo de árboles de decisión y SVM tiene alto grado de desempeño en cuanto a su poder discriminante.

**CONCLUSIONES**

Los resultados del estudio permiten concluir que es posible identificar grupos de consumidores con distintas preferencias en relación con los atributos de un equipo portátil. Esta conclusión valida la primera hipótesis y demuestra la eficacia de los algoritmos de segmentación en la identificación de estos grupos. La capacidad de diferenciar grupos basados en sus preferencias específicas permite a las empresas desarrollar estrategias más precisas y efectivas.

Tabla 17. Medidas desempeño SVM.

Precisión	0,96
Sensibilidad (Recall)	0,92
Especificidad	0,98

Las preferencias de los consumidores están relacionadas con sus características transformacionales y sociodemográficas. Este hallazgo confirma la segunda hipótesis y resalta la importancia de considerar estos factores al segmentar el mercado. La inclusión de características transformacionales y sociodemográficas en los análisis de segmentación proporciona una visión más completa del comportamiento del consumidor, permitiendo a las empresas adaptar sus ofertas de manera más adecuada a las necesidades y deseos de diferentes segmentos de mercado.

Por su parte, los algoritmos utilizados en la clasificación transformacional de clientes presentaron diferentes grados de desempeño en el proceso. Este resultado refuta la tercera hipótesis y sugiere que la elección del algoritmo puede tener un impacto significativo en los resultados de la segmentación. La variabilidad en el desempeño de los algoritmos indica la necesidad de una selección cuidadosa y posiblemente una combinación de diferentes enfoques para obtener los mejores resultados en la segmentación del mercado.

Las clasificaciones en cada segmento de consumidores tienen alta confiabilidad y permiten elaborar propuestas de valor diferenciadas para cada uno de ellos. Esto implica que los atributos de producto diferenciados de acuerdo con las tendencias a corto plazo, niveles de precios por segmento, canales de distribución combinados y estrategias comunicacionales personalizadas pueden ser efectivamente implementados. La capacidad de desarrollar ofertas personalizadas mejora la relevancia y efectividad de las estrategias de marketing, lo que puede llevar a una mayor satisfacción del cliente y lealtad a la marca.

En cuanto a las futuras líneas de investigación, se podrían explorar varios aspectos importantes. La evaluación de otros algoritmos de segmentación y la comparación de su desempeño con los utilizados en este estudio podría proporcionar información valiosa sobre las mejores prácticas en segmentación de mercados. Un análisis más profundo de la relación entre las características sociodemográficas de los consumidores y sus preferencias también sería beneficioso, proporcionando una comprensión más detallada de los factores que influyen en el comportamiento de compra. Además, el desarrollo

de estrategias de marketing transformacional basadas en los resultados de la segmentación podría ayudar a las empresas a implementar de manera efectiva los hallazgos de este estudio.

Finalmente, la aplicación de los métodos de segmentación a otros mercados o tipos de productos podría ampliar la utilidad de estos enfoques y proporcionar insights adicionales sobre la segmentación de mercados en diferentes contextos.

## REFERENCIAS

- [1] A. Martines, C. Ruiz y J. Escrivá, *Marketing en la Actividad Comercial*, Madrid, España: McGraw-Hill, 2014.
- [2] J. Rey, "Marketing digital, big data y publicidad programática; el caso de Walmart México", repository.usta.edu.co, 2020. [En línea]. Disponible en: <https://repository.usta.edu.co/handle/11634/23121> (Accedido: 3 de abril 2024).
- [3] S. Oporta y D. Torres, "Marketing Estratégico: Segmentación de Mercados, mercado meta y posicionamiento en el mercado", Tesis de Grado, Ciencias Económicas y Administrativas. Universidad Nacional Autónoma de Nicaragua, Managua, Nicaragua 2018.
- [4] K. Bricio, J. Calle y M. Zambrano, "El marketing digital como herramienta en el desempeño laboral en el entorno ecuatoriano: estudio de caso de los egresados de la Universidad de Guayaquil", *Revista Universidad y Sociedad*, vol. 10, no. 4, pp. 103-109, 2018. [En línea]. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_abstract&pid=S2218-36202018000400103](http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S2218-36202018000400103)
- [5] N. Casariego, "Metodología de análisis y segmentación de clientes usando secuencias de comportamiento", Tesis de Master, Universidad Internacional de la Rioja, La Rioja, España, 2019.
- [6] E. Olarte, M. Panizzi y R. Bertone, "Segmentación de mercado usando técnicas de minería de datos en redes sociales", en *XXIV Congreso Argentino de Ciencias de la Computación*, La Plata, Argentina, 2018, pp. 471-480. [En línea]. Disponible en: <https://sedici.unlp.edu.ar/handle/10915/73226>

- [7] K. Kalicanin, M. Colovic, A. Njegus, and V. Mitic, "Benefits of artificial intelligence and machine learning in marketing," in *Sinteza 2019 - International Scientific Conference on Information Technology and Data Related Research*, Belgrade, Serbia, 2019, pp. 472-477, doi:10.15308/Sinteza-2019-472-477.
- [8] J. Mejía, "Relación entre el principio de reciprocidad y el compromiso del consumidor", *Retos*, vol. 10, no. 20, pp. 249-264, 2020, doi: 10.17163/ret.n20.2020.04.
- [9] I. Medina y M. Sandoval, "Modelo de la perspectiva conductual: análisis y revisión conceptual", *Revista Latinoamericana de Psicología*, vol. 43, no. 3, pp. 429-441, 2011. [En línea]. Disponible en: [http://www.scielo.org.co/scielo.php?pid=S0120-05342011000300004&script=sci\\_abstract&tlng=es](http://www.scielo.org.co/scielo.php?pid=S0120-05342011000300004&script=sci_abstract&tlng=es)
- [10] M. Espejo, J. Niño y R. Rodríguez, "Un estudio de segmentación con estudiantes según su estilo de vida como consumidores de un servicio educativo universitario", *Revista Internacional de Estudios en Educación*, vol. 23, no. 1, 2023, doi: 10.37354/rie.2022.227.
- [11] A. Holmstrom, "Applying information security to the operational technology environment and the challenges it brings," *Journal of Information Security*, vol. 10, no. 2, pp. 100-110, 2021, doi: 10.4236/jis.2021.102008.
- [12] A. Jamett e L. Schweizer, "Cambios en el comportamiento del consumidor en los escenarios post COVID-19: el consumidor brasileño en 2025", *Revista Psicologia Organizações e Trabalho*, 2021, vol. 21, no. 1, pp. 1318-1326, doi: 10.5935/rpot/2021.1.20553.
- [13] L. Salgado, "Segmentación de los consumidores de alimentos orgánicos según sus actitudes, valores y creencias ambientales", *Contaduría y Administración*, vol. 64, no. 2, 2019, doi: 10.22201/fca.24488410e.2018.1491.
- [14] J. Peter and J. Donnelly, *A preface to marketing management*, 15th ed., New York, USA: McGraw Hill, 2019.
- [15] P. Kotler and G. Armstrong, *Principios de Marketing*, 12va ed., Madrid, España: Pearson, 2008.
- [16] P. Kotler, J. Bowen y J. Makens, *Marketing turístico*, 5ta ed., Madrid, España: Pearson, 2013.
- [17] A. Tan and A. Lo, "A benefit-based approach to market segmentation: a case study of an-American specialty coffeehouse chain in Hong Kong," *Journal of Hospitality & Tourism Research*, vol. 32, no. 21, pp. 342-362, 2008, doi:10.1177/1096348008317388.
- [18] L. Loker and R. Perdue, "A benefit-based segmentation of a nonresident summer travel market," *Journal of Travel Research*, vol. 31, no. 1, 1992, doi: 10.1177/004728759203100107.
- [19] T. Bieger and C. Laesser, "Market segmentation by motivation: The case of Switzerland", *Journal of Travel Research*, vol. 41, no. 1, pp. 68-76, 2002, doi: 10.1177/0047287502041001008.
- [20] P. Condori, "Segmentación de mercado: Machine Learning en marketing en contextos de covid-19", *Sistemas e Informática*, vol. 26, no. 1, doi: 10.15381/idata.v26i1.23623.
- [21] M. Callau-Zori, "Algoritmos de agrupación para flujos de datos en entornos centralizados y distribuidos", Tesis Doctoral, Facultad de Informática, Universidad Politécnica de Madrid, España 2013, doi: 10.20868/UPM.thesis.22467.
- [22] G. Venegas, "Método de clúster jerárquicos para datos funcionales multivariados: una aplicación en mercadeo", *bibliotecavirtualoducal.uc*, 2018. [En línea]. Disponible en: <http://bibliotecavirtualoducal.uc.cl/vufind/Record/oai:localhost:123456789-1512011> (Accedido: 12 de mayo 2024).