

Optimización del geomarketing: Un enfoque en el preprocesamiento de datos, caso de estudio Ciudad de México

Norberto Carrillo García, Magdalena Saldana-Perez,
Marco Moreno-Ibarra, Jesús Saldaña-Pérez

Resumen—Este artículo se centra en el preprocesamiento de datos para el geomarketing, una disciplina que utiliza técnicas de geografía y marketing para analizar el comportamiento de los consumidores en función de su ubicación geográfica. Se propone un método de preprocesamiento de datos combina análisis de datos y ciencia geoespacial, que a su vez es útil para visualizar relaciones comerciales entre unidades económicas en la Ciudad de México. El método propuesto utiliza como base la distancia que hay entre unidades económicas tomando las “k” unidades más cercanas y se observa cómo el valor de “k” afecta las relaciones comerciales identificadas. Los resultados sugieren la existencia de valiosas relaciones comerciales basadas en la proximidad geográfica. Se plantean trabajos futuros que incluyen análisis espacio-temporal y la incorporación de datos adicionales.

Palabras clave—Geomarketing, ciencia geoespacial, preprocesamiento, conjunto de datos, aprendizaje automático.

Geomarketing Optimization: An Approach based on Data Preprocessing, Case Study of Mexico City

Abstract—This paper focuses on data preprocessing for geotargeting, a discipline that uses techniques from geography and marketing to analyze consumer behavior based on their geographic location. A data preprocessing method is proposed that combines data analysis and geospatial science, which is useful for visualizing business relationships between economic units in Mexico City. The proposed method is based on the distance between economic units, considering the 'k' closest units, and it examines how the value of 'k' affects the identified business relationships. The results suggest the existence of valuable business relationships based on geographic proximity. Future work is proposed, including space-time analysis and the incorporation of additional data.

Index Terms—Geomarketing, geospatial science, preprocessing, dataset, machine learning.

I. INTRODUCCIÓN

Una de las áreas del conocimiento que como muchas otras ha incluido a la tecnología en la mejora de sus actividades, es la administración. Algoritmos computacionales complejos que permiten un análisis rápido y profundo de grandes cantidades de datos son utilizados en diferentes actividades administrativas. Los estudios de mercado realizados para

aumentar ventas y transacciones entre empresas actualmente hacen uso de algoritmos computacionales para aumentar su eficiencia, tal es el caso del geo-mercadeo (geomarketing), un área de la administración, enfocada al análisis de indicadores económicos y su relación con el área geográfica en la que ocurren.

El geomarketing es una disciplina que combina técnicas de geografía y marketing para analizar y comprender el comportamiento de los consumidores en función de su ubicación geográfica. Su objetivo principal es utilizar datos espaciales para tomar decisiones estratégicas en marketing, como la selección de ubicaciones para tiendas, la segmentación de mercados o la planificación de campañas publicitarias, por otra parte, el geomarketing se aplica en diversas áreas, como la expansión de redes de tiendas, la evaluación del rendimiento de ubicaciones existentes, el análisis de competencia, la planificación de rutas de distribución, la segmentación de mercados y la personalización de mensajes publicitarios según la ubicación [1].

El geomarketing se basa en la idea de que la ubicación geográfica de los consumidores juega un papel importante en sus patrones de compra y comportamiento de consumo. Al analizar datos demográficos, socioeconómicos, de movilidad y otros indicadores relevantes, las empresas pueden identificar patrones y tendencias específicas en áreas geográficas específicas.

El uso de tecnologías de información geográfica (SIG), que se puede definir como una herramienta que permite recopilar, almacenar, manipular, analizar y visualizar datos geográficos [2], permite a las empresas visualizar y entender mejor la información espacial; además, les proporciona una ventaja competitiva al permitirles tomar decisiones más informadas y eficientes en sus estrategias de marketing y ventas.

El preprocesamiento es un paso fundamental en el geomarketing, pues permite analizar los datos que serán empleados en el proceso de geomarketing y relacionados con zonas geográficas determinadas. Es sabido que la salida de un sistema, y parte de su efectividad, dependen directamente de que los datos de entrada sean adecuados, y describan correctamente el problema que se desea resolver.

En el presente documento se presenta una propuesta de preprocesamiento de datos de entrada para un proceso de geomarketing, que combina el análisis de datos y la ciencia geoespacial. El área de estudio propuesta es la Ciudad de México. En la primera sección se presenta la introducción al

The authors are with the Instituto Politécnico Nacional (IPN), Centro de Investigación en Computación (CIC), México ({ncarrillog2020, amagdasaldana, mmorenoi, jsaldanap2023}@cic.ipn.mx).

Manuscript received on 15/04/2023, accepted for publication on 20/05/2023.

tema propuesta, en la sección 2 se describen trabajos relacionados con la investigación, en la sección 3 se presentan los métodos y materiales, en la sección 4 se presentan los resultados obtenidos a partir de la metodología propuesta, finalmente, en la sección 5 se presentan las conclusiones de la investigación.

II. TRABAJOS RELACIONADOS

A continuación, se presentan algunos de los trabajos relacionados con respecto al diagnóstico del TDAH mediante técnicas de inteligencia artificial.

En los últimos años, se ha producido un aumento significativo en el número de trabajos de investigación relacionados con el geomarketing. Este crecimiento se debe a varios factores, entre los que se incluyen: el crecimiento de nuevas tecnologías, el aumento en la disponibilidad de datos geoespaciales y el interés de las empresas en usar esta tecnología. Este crecimiento de la investigación en geomarketing ha dado lugar a una serie de avances significativos en el campo. Los investigadores han desarrollado nuevos métodos y técnicas para la recopilación, el análisis y la interpretación de datos geoespaciales. En este apartado se realiza y se describen los trabajos relacionados con la presente investigación.

Estos avances han permitido a los investigadores estudiar una amplia gama de temas relacionados con el geomarketing, como la ubicación de negocios, la segmentación de mercados y la asignación de recursos. En [3] se realiza una revisión sobre el concepto de geomarketing y 2 campos de conocimiento que lo componen, también, expone los principales campos de aplicación y las oportunidades que ofrece. Por otra parte, en [4] se realiza un proceso de geomarketing para evaluar las actividades y operaciones comerciales bancarias siendo el área de estudio la ciudad de Benin en este estudio se encontró que los bancos se posicionan cerca de los supermercados con alta actividad económica y tráfico de peatones, por último también se encontró que el 62 % de los bancos se ubican en los primeros 3 km del centro de la ciudad lo que sugiere que pueden existir zonas de oportunidad a las afueras de la ciudad de Benin para las entidades bancarias.

Existen otros trabajos que realizan revisiones sobre los modelos más utilizados en el geomarketing como es el caso de [5] en el cual se realiza una revisión de modelos gravitacionales y modelos de optimización. En [6] se examinan una gran variedad de técnicas y herramientas y un modelo que aborda la inteligencia de mercado geográfico para ver cómo esto puede beneficiar o contribuir a las campañas de geomarketing.

En [7] se propone una sistematización metodológica e informatizada para el proceso de geomarketing, esto, partiendo de la idea de que en el geomarketing no existe una metodología y también lo aplica en un caso práctico. En la literatura existen trabajos que proponen el uso del geomarketing como herramienta empresarial como es el caso de [8] donde de igual manera se propone el método deductivo y se utiliza la investigación descriptiva.

Por otra parte, con la creciente rama de la inteligencia artificial es inevitable que la rama de la cartografía, topografía y los servicios basados en localización se vean influenciados. En [9] se postula que las ramas relacionadas con la

geolocalización no desaparecerán, sino que se transformarán para dar paso a nuevas disciplinas. La prueba de esto último es [10]. En este trabajo se presenta una nueva estructura de datos 3D para administrar ubicaciones en un espacio tridimensional, el cual podría ser visto como edificios o más concreto en centros comerciales. La estructura de datos que se presenta está diseñada sobre la base de un grupo de clasificaciones y agrupamientos y se complementa con la información obtenida de su vecino más cercano.

En [11] se propone desarrollar una herramienta cuantitativa basada en el aprendizaje automático y el geomarketing, con el objetivo de encontrar oportunidades de trabajo y contribuir al proceso de selección de ubicación en una red de franquicias. Existen trabajos en los cuales se plantea un trabajo de optimización de rutas como lo es [12] en el cual plantea métodos de aprendizaje no supervisado para el análisis de datos espaciales esto con el fin de identificar puntos de interés en la ruta de diferentes tipos de clientes.

En esta sección se realizó una revisión del estado del arte y si bien se puede concluir que para el geomarketing no existe un método definido debido a que este varía según el objetivo, el preprocesamiento es un aspecto en el que la mayoría de los trabajos no aborda.

III. MATERIALES Y MÉTODOS

En esta sección se presentan los métodos y los materiales necesarios para alcanzar el objetivo de la presente investigación, también, se propone un caso de estudio y área de estudio. El método que se propone en este trabajo se compone de 3 partes las cuales son las siguientes:

Conjunto de datos

En esta parte se realiza la selección del conjunto de datos a utilizar. A modo de implementación para este trabajo se utilizó el conjunto de datos que proporciona el DENUÉ INEGI el cual es un conjunto de datos que contiene información de todos los negocios o unidades económicas que se encuentran en México [13]. Este conjunto de datos está dividido por estados y cuenta con cerca de 4.5 millones de unidades económicas. Como se mencionó antes, el área de estudio de esta investigación es la Ciudad de México por lo que se realizó una selección de las unidades económicas de la Ciudad de México. La cardinalidad de esta selección es de 474,321 unidades económicas.

Selección de tipo de unidad económica de interés

En esta parte del método propuesto se selecciona la categoría o el giro del negocio de interés. Para este caso práctico se seleccionaron los laboratorios de análisis clínicos en el área de la Ciudad de México. Los laboratorios de análisis clínicos, en el conjunto de datos seleccionado, se encuentran en la categoría de servicios de salud y asistencia social y subcategoría de "Laboratorios médicos y de diagnóstico". En esta subcategoría se encuentran otras 2 subcategorías las cuales son sector público y sector privado en este caso práctico se seleccionaron tanto el sector público como el sector privado.

Partición del conjunto de datos

En esta parte de la metodología se realiza una partición en el conjunto de datos. El procedimiento que se realiza en esta parte

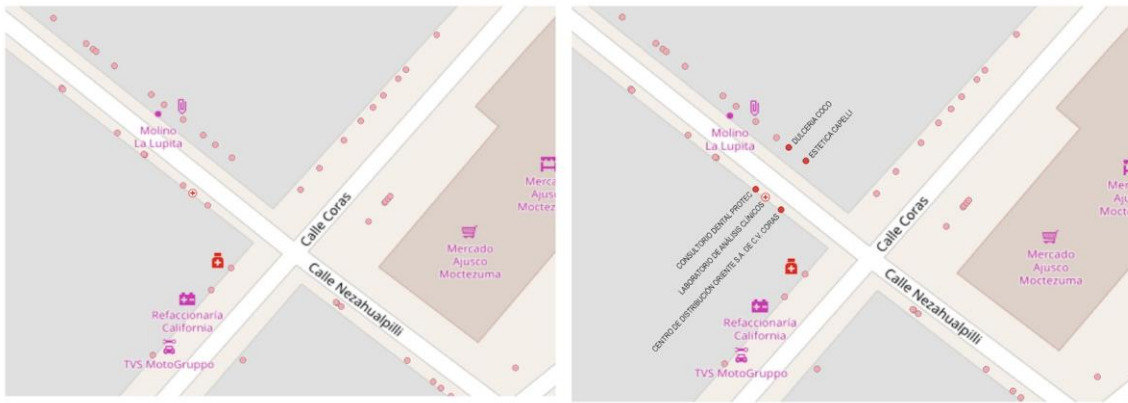


Fig. 1. Representación gráfica del procedimiento

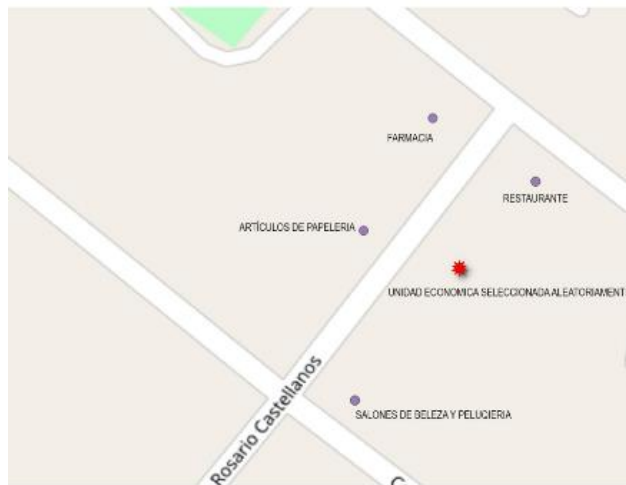


Fig. 2. Representación gráfica de la segunda parte del procedimiento

del método propuesto es extraer las unidades económicas de interés del conjunto de datos seleccionado. En este caso se extrajeron las unidades económicas de interés, en este caso son los laboratorios de análisis clínicos. Al final de este paso se obtienen 2 particiones del conjunto de datos original. En una partición se tienen las unidades económicas de interés, en este caso los laboratorios de análisis clínicos, y en la segunda partición con el resto de unidades económicas.

Algoritmo de cálculo de distancias.

En esta etapa del método se realiza un algoritmo de cálculo de distancias. Es importante mencionar que este algoritmo de cálculo de distancias está basado en el algoritmo K-NN (k - *Nearest Neighbor* por sus siglas en inglés) tal como se ve en algunos trabajos del estado del arte [10, 14, 15, 16].

La manera en la que se realiza este algoritmo es tomar las unidades económicas como una lista de listas en la cual cada lista tiene el nombre de cada unidad económica de interés y las coordenadas dado que los datos son referenciados. De igual manera el resto de las unidades económicas se guardan en otra lista de listas. El objetivo de esto es calcular las distancias que existen de las unidades económicas de interés al resto de unidades económicas:

$$d_E(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

En la fórmula (1) se describe el cálculo de la distancias donde P_1 representa la unidad económica de interés y los puntos x_1 y y_1 representan las coordenadas longitud y latitud respectivamente y P_2 representa otra unidad económica. El algoritmo en cuestión se presenta de la siguiente manera:

Algorithm 1 Algoritmo para generación de conjunto de datos

- 1: asignar valor a k
- 2: conjunto= []
- 3: **for** cada elemento en p1 **do**
- 4: vev-distancias= []
- 5: **for** cada elemento en p2 **do**
- 6: vec-distacias.append (Calcular distancia de p1[i], p2[j])
- 7: **end for**
- 8: Seleccionar las k negocios más cercanos a conjunto
- 9: Agregar negocios más cercanos a conjunto
- 10: **end for**
- 11: asignar etiqueta: optimo

En esta primera parte del procedimiento se obtiene la primera parte del conjunto final. Para obtener la segunda parte del conjunto de datos es necesario repetir el algoritmo realizando algunos ajustes. el primer ajuste es eliminar los k negocios más cercanos de cada elemento en P1 (Unidades económicas de interés). Seguido de esto se realiza una nueva partición con tamaño P1. Una vez obtenidas las 2 nuevas particiones se



Fig. 3. Representación geográfica del comportamiento del algoritmo con $k = 4$

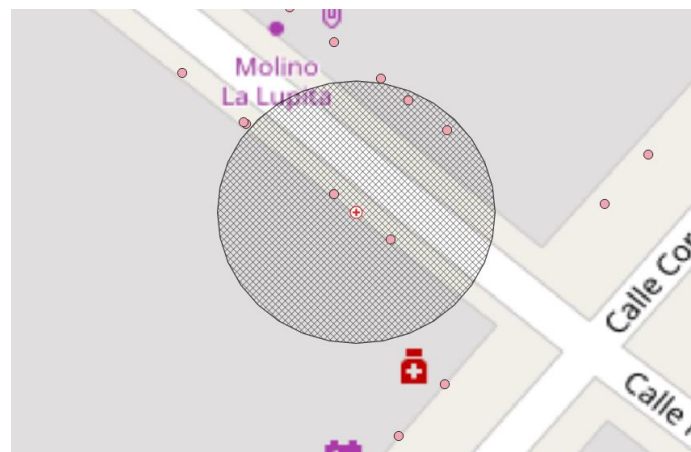


Fig. 4. Representación geográfica del radio de 19 metro creado por el algoritmo propuesto.

realiza el mismo procedimiento, pero en esta parte del proceso se cambia la etiqueta de óptimo a no. La primera parte del procedimiento se puede apreciar geográficamente en la Fig. 1.

Por otra parte, en la Fig. 2 se muestra como se observaría la representación geográfica de la segunda parte del algoritmo.

Conjunto de datos

El resultado final del procedimiento propuesto es el conjunto de datos el cual está formado, para la clase óptima, todas las unidades económicas que rodean una unidad económica de interés, en este caso, el caso de estudio son los laboratorios de análisis clínicos. La segunda parte del conjunto de datos, la clase no óptima, son todas las unidades económicas que rodean un punto aleatorio donde no se encuentra ninguna unidad económica de interés y por lo tanto se supone que es un punto no óptimo. En la imagen x se muestra una ejemplificación del conjunto de datos.

IV. PRUEBAS Y RESULTADOS

En este apartado se describen las pruebas realizadas de la metodología planteada en este trabajo. Para las pruebas realizadas se tomaron valores de $k = 4, 10, 30, 60, 100$ y 150 .

Para la primera prueba con $k = 4$ se analizaron las 4 unidades económicas más cercanas a la unidad económica de interés, en este caso laboratorios de análisis clínicos.

En la Fig. 3 podemos observar de manera gráfica cómo se comporta el algoritmo, en este ejemplo tomado de la alcaldía Coyoacán se puede observar que con $k = 4$ los negocios cerca de un laboratorio de análisis clínicos son consultorio dental, dulcería, estética y un centro de distribución.

En la Fig. 4 se muestra la misma selección de $k = 4$ en un buffer cuyo radio es de 19 metros. haciendo un análisis de cuál es la unidad económica que se repite más se repite según un laboratorio de análisis clínicos se obtuvieron los siguientes resultados: en primer lugar, consultorios de medicina especializados, consultorios dentales, salones de belleza y peluquerías, servicio de alimentos o comida corrida y por último tiendas al por menor o tiendas misceláneas.

Por otra parte, para $k = 10$ se obtuvo que el agrupamiento resultante tiene un radio de entre 50 y 100 metros y las unidades económicas que más se repiten se muestran en la Tabla 1.

Cuando se considera un valor de k igual a 30, el agrupamiento resultante muestra un radio promedio de 100 a 115 metros. Las unidades económicas más frecuentes en estos grupos se ilustran en la Tabla 2.

Al establecer k en 60, los grupos generados muestran, en promedio, un radio de 120 a 150 metros. Las unidades económicas que se destacan en estos grupos se detallan en la Tabla 3.

TABLA I
DISTRIBUCIÓN DE UNIDADES ECONÓMICAS CON K = 10

Unidad económica	Nº
Consultorio de medicina especializada del sector privado	78
Salones de belleza y peluquerías	60
Consultorios dentales	55
Banca múltiple	37
locales de comida corrida	32

TABLA II
DISTRIBUCIÓN DE UNIDADES ECONÓMICAS CON K = 30

Unidad económica	Nº
Salones de belleza y peluquerías	144
Consultorios dentales	127
Consultorio de medicina especializada	125
Banca Múltiple	103
Locales de comida corrida	76

TABLA III
DISTRIBUCIÓN DE UNIDADES ECONÓMICAS CON K = 60

Unidad económica	Nº
Salones de belleza y peluquerías	250
Consultorios dentales	207
Banca Múltiple	200
Consultorio de medicina especializada	153
Tiendas al por menor o tiendas misceláneas	145

TABLA IV
DISTRIBUCIÓN DE UNIDADES ECONÓMICAS CON K = 100

Unidad económica	Nº
Salones de belleza y peluquerías	413
Banca Múltiple	384
Consultorios dentales	310
Tiendas al por menor o tiendas misceláneas	276
Locales de comida corrida	228

TABLA V
DISTRIBUCIÓN DE UNIDADES ECONÓMICAS CON K = 150

Unidad económica	Nº
Salones de belleza y peluquerías	601
Banca Múltiple	508
Tiendas al por menor o tiendas misceláneas	468
Consultorios dentales	379
Locales de comida corrida	343

Cuando se emplea un valor de k igual a 100, se observa en la tabla que el agrupamiento resultante presenta un radio promedio de 160 a 200 metros. En la Tabla 4 se observan las unidades económicas, más frecuentes en estos grupos.

Finalmente, al configurar k en 150, el agrupamiento resultante muestra un radio promedio de 200 a 250 metros. En la Tabla 5 se detallan las unidades económicas que se identificaron.

A partir de estos resultados, podemos observar cómo se comporta la distribución de los negocios en función de los

diferentes valores que asignamos a "k" en el algoritmo. Como se puede apreciar en las pruebas realizadas en la alcaldía Coyoacán, Ciudad de México, y centrándonos en los laboratorios de análisis clínicos, las unidades relacionadas en un radio de aproximadamente 20 metros incluyen consultorios de medicina especializada, salones de belleza y consultorios dentales.

Sin embargo, al ampliar el radio a 100 metros, notamos que las unidades económicas cambian. En otras palabras, mientras que en un radio de 20 metros el tipo de negocio más común es el consultorio de medicina especializada, en un radio de 100 metros, los salones de belleza y peluquerías se convierten en las unidades económicas predominantes, seguidos de los consultorios dentales, y en tercer lugar, los consultorios de medicina especializada.

A medida que aumentamos el valor de "k" y, por lo tanto, el tamaño del grupo de datos, seguimos encontrando relaciones interesantes. Por ejemplo, cuando "k" es igual a 60, con un radio promedio de 130 metros, observamos que la distribución de unidades económicas se mantiene casi idéntica, excepto que ahora la banca múltiple ocupa el tercer lugar.

En un radio promedio de 170 metros (con "k" igual a 100), encontramos que la banca múltiple se ubica en el segundo lugar. Finalmente, en un radio de aproximadamente 250 metros con "k" igual a 150, la distribución de las unidades económicas apenas cambia.

Este último agrupamiento revela que las tiendas misceláneas al por menor se vuelven más frecuentes. A partir de este análisis, podemos identificar posibles relaciones entre las unidades económicas. Utilizando un valor de "k" igual a 150, se destacan conexiones entre salones de belleza y peluquerías, bancos o cajeros automáticos (banca múltiple), tiendas de abarrotes, consultorios dentales y locales de comida corrida.

V. CONCLUSIONES

En esta sección, se presentan las conclusiones derivadas de la aplicación del método propuesto y la obtención de resultados. Como se mencionó anteriormente, el objetivo de este trabajo es proponer un método de preprocesamiento de datos que combine la ciencia de datos y la ciencia geoespacial.

Este método nos permite visualizar las relaciones comerciales y, dado que el resultado final es un conjunto de datos, es factible aplicar el aprendizaje automático para llevar a cabo una clasificación de zonas comerciales en relación a unidades económicas de interés.

Como resultado, podemos identificar diversas relaciones comerciales entre unidades económicas. Una conclusión relevante que se desprende es que a medida que aumenta el valor de "k," se puede realizar un análisis más detallado de las relaciones. En las pruebas realizadas, se mantuvo el valor de "k" en 150, ya que con este número de "k," el radio promedio se mantiene en 250 metros, lo cual puede considerarse una distancia relativamente corta.

Es importante destacar que las relaciones comerciales obtenidas en ocasiones pueden carecer de sentido. Esto puede deberse a varios factores. Entre los resultados, se encuentra el ejemplo de las relaciones entre laboratorios de análisis clínicos y salones de belleza. En estas dos unidades económicas, no es posible establecer una relación comercial clara.

VI. TRABAJOS FUTUROS

En este apartado se discuten los trabajos futuros para la presente investigación. Uno de los trabajos futuros que pueden mejorar el presente trabajo es el de realizar un análisis espacio temporal, es posible debido a que el portal de donde estos datos fueron obtenidos cuenta con un histórico de censos que van desde el año 2010 al 2023. Como trabajo futuro es posible realizar el procedimiento utilizando 2 conjuntos de datos, esto con el fin de analizar sitios en los que se encontraban unidades económicas de interés y en el presente ya no se encuentran. Al realizar este procedimiento se podría incluso encontrar posibles relaciones comerciales que no son compatibles o posibles razones por las que no funcionan ciertas unidades económicas en determinadas zonas. Por otra parte, otro posible trabajo futuro es agregar al proceso más conjunto de datos como lo pueden ser conjuntos de datos del transporte colectivo de la ciudad como lo pueden ser estaciones de Metro y Metrobús, conjuntos como áreas verdes en la ciudad.

REFERENCIAS

- [1]. G. Cliquet, "Geomarketing Methods and Strategies in Spatial Marketing," <http://ci.nii.ac.jp/ncid/BA78279883>, (2013).
- [2]. M. Duckham, Q.C. Sun, and M.F. Worboys, "GIS: A computing perspective", *CRC press*, (2023).
- [3]. J. F. Rodríguez-Martínez, "Geomarketing una oportunidad para la analítica empresarial", *Revista Investigación y Negocios*, vol. 13, no. 22, pp. 103–110, (2020).
- [4]. T.F. Balogun and O.O. Ogbeide, "A geomarketing perspective of commercial banks distribution in benin city," *Ethiopian Journal of Environmental Studies & Management*, vol. 13, no. 1, (2020).
- [5]. J.A. Montejano-Escamilla and G. Cruz Bello, "Modelos de localización para geomarketing" (2018).
- [6]. T. Guarda, M.F. Augusto, and I. Lopes, "Geographic market intelligence as a competitive advantage," In: *2019 14th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1–5, (2019).
- [7]. E.A., Albornoz-Del Valle, F.J., Núñez-Cerda, and C. Mena-Frau, "Geomarketing: Desde una visión comercial a una aplicación social, en contextos metropolitanos", *Revista de Geografía Norte Grande*, vol. 76, pp. 143–167, (2020).
- [8]. E.V.R., Farrónán, C.A.V. Salazar, and J.E.V. Sandoval, "El geomarketing como estrategia para el posicionamiento de una óptica en Chiclayo 2018", *Tzhoecoen*, vol. 11, no. 1, pp. 24–33. (2019).
- [9]. J. Liu and K. Gao, "Challenges and opportunities for mapping and surveying and location based service in the age of intelligence", *武汉大学学报 (信息科学版)*, vol. 42, no. 11, pp. 1506–1517, (2017).
- [10]. S. Azri, U. Ujang, and A.A. Rahman, "Voronoi classified and clustered data constellation: A new 3D data structure for geomarketing strategies", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, pp. 1–16, (2020).
- [11]. M.F.F.D. Oliveira, P.H.M. Albuquerque, P.Y. Hao, and P.A. Henrique, "Mapping regional business opportunities using geomarketing and machine learning", *Gestão & Produção*, vol. 27, no. e4158, (2020).
- [12]. A. Festi, "Spatial unsupervised learning–applications of market basket analysis in geomarketing", In: *Applied Spatial Statistics and Econometrics*, pp. 517–539, Routledge, (2020).
- [13]. INEGI, (2023, 5 de octubre), DENUÉ, <https://www.inegi.org.mx/app/mapa/denué/default.aspx>.
- [14]. D. Zaim, A. Benomar, and M. Bellafkih, "Clustering The Customers By Using K-Means Algorithm", In: *2020 International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, Morocco, pp. 1–6, (2020). DOI: 10.1109/ISAECT50560.2020.9523683.
- [15]. S.C. Tékouabou-Koumétio and H. Toulmi, "Improving knn model for direct marketing prediction in smart cities", In: *Machine Intelligence and Data Analytics for Sustainable Future Smart Cities*", Springer International Publishing, pp. 107–118, (2021). DOI: 10.1007/978-3-030-72065-0_7.
- [16]. Y. Li, Y. Qin, and H. Wang, "K-nearest neighbor privacy protection query for distributed storage in location-based service", *Wireless Personal Communications*, vol. 121, pp. 1509–1532, (2021). DOI: 10.1007/s11277-021-08682-y.