

Identificación de la delimitación administrativa de la malaria usando redes neuronales artificiales

Boundary Delimitation of Malaria using Artificial Neural Networks

Identificação da Delimitação Administrativa da Malária usando Redes Neurais Artificiais

Fredy A. Salazar-Vasquez¹; Carlos Osorio-Serna²; María Alejandra Caicedo-Giraldo³; Wilfredo Alfonso-Morales⁴; Eduardo F. Caicedo-Bravo⁵

¹ Estudiante Ingeniería Electrónica, Grupo de investigación Percepción y Sistemas Inteligentes, escuela de ingeniería eléctrica y electrónica, Universidad del Valle, Cali, Colombia.

² Estudiante Ingeniería Electrónica, Grupo de investigación Percepción y Sistemas Inteligentes, escuela de ingeniería eléctrica y electrónica, Universidad del Valle, Cali, Colombia.

³ Médica y Cirujana. Grupo Centro de Investigación Científica Cauceseco. Estudiante de Especialización en Urología Universidad Nacional de Colombia, Bogotá, Colombia.

⁴ Ingeniero Electrónico, M.Sc, PhD. Profesor asistente, Grupo de investigación Percepción y Sistemas Inteligentes, escuela de ingeniería eléctrica y electrónica, Universidad del Valle, Cali, Colombia.

⁵ Ingeniero Eléctrico, M.Sc, PhD. Profesor titular, Director Grupo de investigación Percepción y Sistemas Inteligentes, escuela de ingeniería eléctrica y electrónica, Universidad del Valle, Cali, Colombia.

Email: fredy.salazar@correounivalle.edu.co

Recibido: febrero 23 de 2017

Aprobado: mayo 11 de 2017

Resumen

La metodología de clustering fue utilizada para agrupar tres barrios en Quibdó teniendo en cuenta factores que favorecen el desarrollo de la malaria. Los mapas auto-organizados de Kohonen fueron utilizados para el análisis de las características más significativas en la clasificación. Los clusters detectados fueron comparados con la clasificación geográfica de las casas, encontrando, que los mapas auto-organizados de Kohonen clasifican las casas por las condiciones ambientales propicias para el desarrollo del mosquito más que por la clasificación administrativa de la ciudad.

Palabras clave: Clustering, Malaria urbana, Mapa Auto Organizado de Kohonen, Red Neuronal Artificial.

Resumo

A Metodologia de Clustering foi usada para agrupar três bairros em Quibdo, Colômbia, levando em consideração fatores que favorecem o desenvolvimento da malária. Mapas auto-organizados de Kohonen foram utilizados para a análise das características mais significativas no agrupamento. Os Clusters detectados foram comparados com o agrupamento geo-

gráfico de casas, mostrando que os mapas auto-organizados de Kohonen agrupam as casas pelas condições ambientais favoráveis ao desenvolvimento do mosquito e não pelo agrupamento administrativo da cidade.

Palavras-chave: Clustering, Malária Urban, Mapa Auto Organizado Kohonen, Redes Neurais Artificiais.

Abstract

Clustering methodology was used to group three neighborhoods in Quibdó taking into account factors that favor the development of malaria. The Kohonen self-organizing maps were used for the analysis of the most significant features in the standings. The detected clusters were compared with the geographical classification of houses, finding that the Kohonen self-organizing maps households classified by environmental conditions conducive to development rather than the administrative classification of the city.

Keywords: Artificial Neural Networks, Clustering, Malaria, Self-organized map of Kohonen.

Introducción

La malaria es una enfermedad infecciosa, causada por parásitos de la especie *Plasmodium* y transmitida por vectores (mosquitos de la especie *Anopheles*), que genera una alta morbi-mortalidad en los países y regiones donde es endémica. En Colombia se produjeron aproximadamente 83 000 casos de malaria durante el año 2016, siendo Chocó el departamento con mayor carga de la enfermedad en el país, puesto que 56.4% de los casos se presentaron en esta región (SIVIGILA, 2016).

Históricamente, la malaria ha sido una enfermedad cuya transmisión predomina en zonas rurales, donde las condiciones ambientales favorecen la presencia y reproducción de los mosquitos vectores. Sin embargo, el Sistema Nacional de Vigilancia en Salud Pública (SIVIGILA), reporta cada vez más casos de malaria clasificados como provenientes de cabeceras municipales (zonas urbanas), pasando a representar el 30% del total de casos reportados durante el 2015 en el país.

Dado que el SIVIGILA obtiene los datos a partir de fichas epidemiológicas que son diligenciadas por el personal de salud en cada región, existe el riesgo de que haya una clasificación errada del origen de los casos de malaria. Esto debido a que la clasificación se basa en lo que el paciente refiere, información que no es verificada objetivamente y que no siempre es veraz. Adicionalmente, no existe un consenso en el concepto de urbano, peri-urbano o rural a nivel nacional e internacional; clasificación que se da desde el punto de vista geográfico y político administrativo, sin considerar los factores ambientales que favorecen la proliferación de los vectores y la transmisión. En el caso de Colombia, la clasificación de los municipios y barrios obedece a delimitaciones asignadas por los entes administrativos, que no necesariamente corresponden con los lineamientos establecidos por entidades como el Departamento Administrativo Nacional de Estadística (DANE, 2000), ni a las características reales de los territorios.

Esta problemática genera la necesidad de estudiar el origen real de los casos de malaria y las condiciones de los barrios de donde provienen, pues se sospecha que el aumento en la prevalencia de malaria urbana podría estar relacionado con errores en la clasificación del SIVIGILA (Padilla *et al.*, 2015).

Este estudio se realizó con el fin de caracterizar tres barrios de Quibdó, en los que históricamente se han reportado casos de malaria, utilizando técnicas de clustering basadas en Redes Neuronales Artificiales (RNA). Para el análisis de datos se utilizó el modelo de mapas auto-organizados de Kohonen, siguiendo una metodología para la selección y posterior procesamiento de las características, que busca identificar aquellas que tengan una mayor capacidad de discriminación. El resultado obtenido logró agrupar las casas pertenecientes a estos tres barrios por características similares y logró establecer un patrón diferencial entre los tres lugares, lo que permite identificar la delimitación administrativa de la malaria desde el punto de vista de los vectores que transmiten la enfermedad.

La Malaria y sus vectores

La malaria es una enfermedad transmitida por las hembras de mosquitos pertenecientes al género *Anopheles*. Estas pican a los humanos, pues la sangre que extraen es necesaria para la producción de huevos y propagación de su especie.

Cuando una de estas hembras pica a un humano infectado con el *Plasmodium*, ingiere las formas sexuales del parásito que se encuentran en el torrente sanguíneo de la persona enferma. Estas formas sexuales entran al intestino del mosquito donde, a su vez, realizan su ciclo de reproducción. Como resultado se obtienen las formas infectivas del parásito, llamadas esporozoítos, que se desplazan hacia las glándulas salivares del mosquito donde se almacenan. Finalmente, cuando el mosquito pica nuevamente, inyecta las formas infecti-

vas del parásito a otro ser humano, completando de esta manera el ciclo vital (Beier, 1998).

En Colombia, se han identificado entre 40 a 47 especies diferentes de *Anopheles*, 12 de las cuales transmiten malaria (Gonzalez y Carrejo, 2007). Sin embargo, se considera que solo tres de estas (*An. darlingi*, *An. nuñeztovari* y *An. albimanus*), cumplen con las características de un vector principal de malaria que incluyen: presentar una amplia distribución geográfica, tener una capacidad significativa de colonización y dispersión, alta adaptabilidad a medios modificados por el hombre, antropofilia y susceptibilidad de ser infectados por especies de *Plasmodium* que infectan humanos (Montoya-Lerma, 2011).

Hay varios factores que influyen en la presencia y reproducción de los mosquitos, entre ellos los más importantes son la humedad, la temperatura, el espesor de la vegetación, presencia de aguas estancadas o la existencia de otros sitios adecuados de cría. Por lo tanto, la cantidad de mosquitos en cierta zona podría ser un buen indicador del nivel de urbanización de la zona.

Sitio de estudio

Los barrios escogidos para la realización de este estudio están clasificados como pertenecientes al casco urbano de Quibdó, aunque en realidad son heterogéneos entre sí y presentan características ambientales diferentes. La Yesquita es un barrio situado en el centro de la ciudad, que se acoge a la definición de barrio urbano planteada por el DANE: las calles son pavimentadas, las casas están agrupadas en cuadras, estas cuentan con servicios públicos (electricidad, acueducto, alcantarillado) y no hay presencia de vegetación abundante; Casablanca y Cabí, son barrios situados en el extremo norte y sur de la ciudad respectivamente. Estos tienen menor cantidad de vías pavimentadas y las casas no presentan una organización definida en cuadras, con una dispersión mayor en el territorio. Adicionalmente, se encuentra gran cantidad de vegetación alrededor de las casas y colecciones de agua. En cuanto a servicios públicos, la mayoría de los hogares tienen electricidad, sin embargo, el abastecimiento de agua depende de la recolección de aguas lluvia y no se cuenta con alcantarillado.

Redes Neuronales y Clustering

Se puede definir el término clustering como la tarea de identificar y agrupar un conjunto de elementos con características similares en un grupo, llamado clúster; y que difiere de otro por las características de los ele-

mentos que lo conforman (Peña, 2002). Esta tarea es uno de los principales fundamentos en la minería de datos, y una técnica muy útil para el análisis de datos estadísticos, además de ser una de las principales y más comunes aplicaciones de las redes neuronales artificiales (RNA).

Las RNA surgen como una solución de problemas de clasificación basándose en una imitación del funcionamiento de las neuronas del cerebro humano. Son un sistema de interconexión de neuronas que transmiten una salida a partir de los estímulos percibidos de una entrada; este proceso es denominado sinapsis y ocurre de igual forma en las neuronas biológicas (Anderson, 2007).

Una RNA es un procesador paralelo masivamente distribuido e interconectado que tiene una facilidad natural para almacenar el conocimiento obtenido de la experiencia para luego hacerlo utilizable (Caicedo, 2009).

El método de aprendizaje de las redes neuronales, es una de las características más importantes en un proceso de clasificación de patrones, ya que es durante este aprendizaje que la red se configura internamente ante unas determinadas entradas para obtener una salida, de forma que la red pueda responder ante nuevos estímulos o situaciones diferentes a las aprendidas.

Existen dos tipos de aprendizaje, el primero es el aprendizaje supervisado, el cual es similar al caso de un tutor que enseña a un estudiante y evalúa si este aprende o no, entonces se presenta una entrada a la red y se compara la salida obtenida con la deseada, se evalúa el error y se realizan modificaciones a la red para corregir o disminuir este error; el segundo es el aprendizaje no supervisado, en la red no se evalúa una salida deseada y solo se tienen vectores de entrada, por lo que la red se configura para dar una salida consistente ante entradas similares. En este tipo de aprendizaje, la red extrae características, abstrayendo las propiedades colectivas subyacentes del conjunto de entrenamiento y agrupa por clases de similitudes (Gómez, Fernández, López y Díaz, 1994).

Mapas Auto Organizados de Kohonen

Uno de los algoritmos que utilizan las redes neuronales con aprendizaje no supervisado es el aprendizaje competitivo, en el cual las neuronas compiten unas con otras con el fin de realizar una tarea. Una de las redes neuronales que utilizan este tipo de algoritmos son los mapas auto organizados (Kohonen, 1984). El modelo de mapas auto organizados, propuesto por

Teuvo Kohonen (Kohonen, 1990), está totalmente inspirado en el comportamiento neurobiológico y muchos científicos investigadores lo utilizan en aplicaciones del mundo real, como clustering (Jain, Murty y Flynn, 1999), identificación de personas, clasificación multiespectral de imágenes, modelos de diagnósticos clínicos, etc. (Chelali, Djeradi y Cherabit, 2015), (Hamami, Mercier y Hamouda, 2014), (Jinliang, Caiqing, Jiansheng y Jianjing, 2010). La red de Kohonen apunta a usar el aprendizaje sin supervisión para ajustar los pesos de la capa competitiva y finalmente resultar en un clasificador de patrones.

El modelo SOM (Self Organizing Map, mapa auto organizado en inglés) tiene dos capas de neuronas o nodos, la capa de entrada y la capa salida. En todas las RNA, las neuronas de la misma capa están desconectadas entre sí, pero en la red de Kohonen las neuronas adyacentes en la capa de salida están conectadas entre sí por una relación de vecindad que representa la estructura del mapa, como se muestra en la figura 1.

La capa de entrada representa M señales o características de entrada que están totalmente conectadas a una capa de salida de dimensión N , la cual es elegida al azar, en esta capa a las neuronas se les asocia un peso sináptico. Los pesos de estas conexiones definen el vector $\omega_j(t)$ en el instante t , que tiene el mismo tamaño m que un vector de entrada $x(t) = (x_1, x_2, \dots, x_m)$. En el $\omega_j(t)$, $j = 1, 2, \dots, l$, j es el índice de neuronas, y l

es el número de neuronas en el mapa (Baracho, Melo y Coelho, 2016). El entrenamiento de este Mapa Auto-Organizado de Kohonen se realiza competitivamente y sin supervisión, repitiendo tres pasos básicos (muestreo, coincidencia por similitud y actualización) hasta que se logra un criterio mínimo de satisfacción.

En primer lugar, se eligen valores aleatorios para el vector de pesos $\omega_j(0)$, de manera que $\omega_j(0)$ son diferentes para cada j . El paso de muestreo toma un ejemplo de entrenamiento, o patrón, con una cierta probabilidad; Este ejemplo es entonces el vector de entrada x . La coincidencia por similitud es el procedimiento usado para encontrar la neurona con la mejor coincidencia (ganadora) usando la distancia euclidiana, de acuerdo con la ecuación:

$$i(x) = \arg \min_j \|x(n) - w_j\|, j = 1, 2, \dots, l$$

Así, $i(x)$ es el índice de la neurona ganadora para la entrada x . En el paso de actualización, uno tiene que ajustar los vectores de pesos sinápticos de todas las neuronas usando la ecuación:

$$w_j(n+1) = w_j(n) + \eta(n) \cdot h_{j,i(x)}(n) (x(n) - w_j(n))$$

Donde $\eta(n)$ representa la tasa de aprendizaje, y $h_{j,i}(x)(n)$ es la función de vecindad centrada en la neurona ganadora $i(x)$; Tanto $\eta(n)$ como $h_{j,i}(x)(n)$ son dinámicamente modificados durante el proceso de

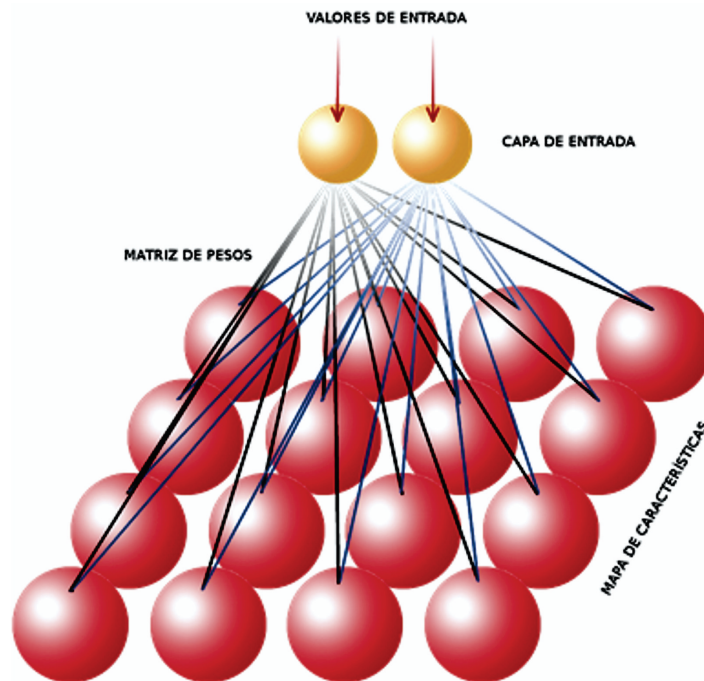


Figura 1. Mapa Auto Organizado de Kohonen.

aprendizaje para obtener mejores resultados. Los procedimientos de muestreo, coincidencia por similitud y actualización se repiten hasta que no se observan cambios significativos en el mapa (Baracho, Melo y Coelho, 2016).

El objetivo de este aprendizaje es categorizar los datos que se introducen en la red a través de un proceso de *clustering*; el objetivo del *clustering* es granular la información en mayor detalle, obteniendo una separación entre clases. De esta forma, informaciones similares son clasificadas para ser parte de la misma categoría, activando por tanto la misma neurona o grupo de neuronas vecinas de salida.

Este proyecto hace uso de un mapa de auto-organizado para realizar una tarea de *clustering* sobre una base de datos existente, la cual contiene información acerca de las características físicas y ambientales de un conjunto de viviendas en la ciudad de Quibdó, para identificar tres clústeres que, según variables representativas, permiten clasificar las viviendas entre urbanas, periurbanas y rurales. Dentro de los resultados se encontró una agrupación de las casas que presentaban características más afines entre sí, según los tres grupos mencionados y algunas casas que compartían características con aquellas de un barrio distinto al propio (Figura 2).

Desarrollo

El Centro de Investigación Científica Cauceseco suministró una base de datos la cual incluía información sociodemográfica, geográfica y entomológica de 15 casas ubicadas en 3 barrios de la ciudad de Quibdó, donde se identificaron casos de malaria, como parte de un proyecto piloto que busca documentar la trans-

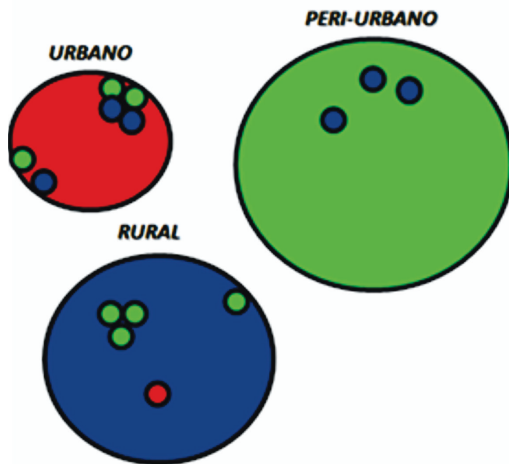


Figura 2. Representación de la salida esperada.

misión autóctona de malaria en estos barrios. Para el desarrollo de este proyecto se realizó un procesamiento de *clustering* de la información contenida en la base de datos descrita a través de un mapa auto-organizado de Kohonen, donde se procesaron 11 variables de 21 con el fin de obtener una caracterización y clasificación de las casas pertenecientes a cada barrio.

En la figura 3 se muestra el diagrama de flujo de todo el proceso.

En la primera parte del proceso, se realizó una poda de la base de datos con el objetivo de extraer las características más relevantes, es decir, aquellas que podían

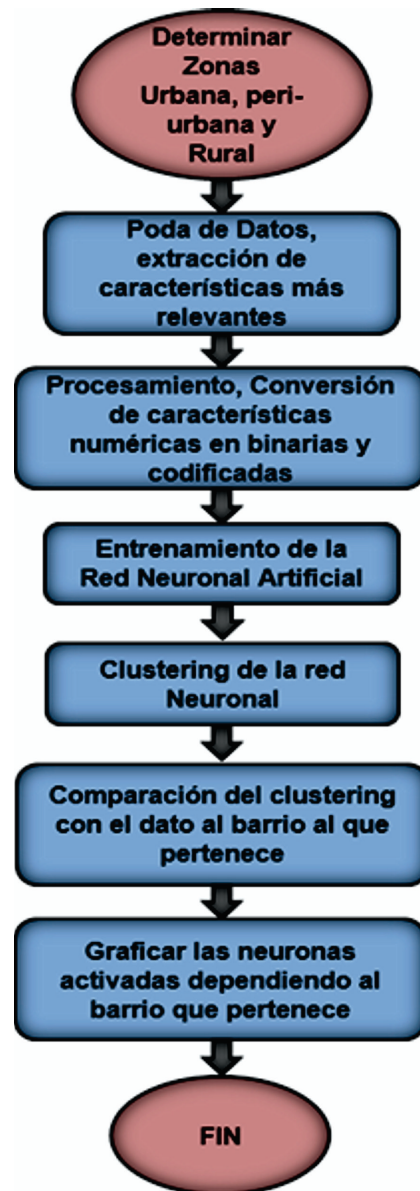


Figura 3. Diagrama de Flujo.

llegar a tener mayor influencia en la clasificación, y se eliminaron aquellas que no tenían ningún tipo de diferencia. Algunas características que dan soporte a la diferenciación de una clase a otra son: la altura, la temperatura máxima y mínima, la cantidad de mosquitos encontrados en cada casa y la variedad de especies que cohabitan en el lugar; cabe resaltar que en las muestras se consideró realizar un procesamiento de los datos del tipo de mosquitos, dando importancia a la existencia de estos y no a la cantidad de cada tipo, convirtiendo estas características en variables booleanas. También se codificaron algunas características dadas en texto asignándoles valores numéricos, y se descartaron aquellas que al momento de la recolección de los datos fueron omitidas.

Las 11 características para hacer *clustering* luego de la poda son:

1. nAlt: Altura sobre el nivel del mar.
2. Sdomi: Ubicación de la toma de la muestra dentro o fuera de la casa.
3. nTmin: Temperatura mínima medida en el intervalo de tiempo en que se tomó de la muestra.

4. Tmax: Temperatura máxima medida en el intervalo de tiempo en que se tomó la muestra.
5. nHumi: Nivel de humedad medido en el momento de la muestra.
6. nMosq: Cantidad total de mosquitos recolectados.
7. API: Especie de Mosquito recolectado.
8. DAR: Especie de Mosquito recolectado.
9. NUN: Especie de Mosquito recolectado.
10. TRI: Especie de Mosquito recolectado.
11. SP: Especie de Mosquito recolectado.

Luego de entrenar el mapa auto-organizado, la distribución de la información a través de los mapas de características se presenta en la figura 4.

Para la interpretación de los resultados se comparó la clasificación de la muestra tomada en el aforo respecto a la clasificación realizada por el mapa auto-organizado, obteniendo el mapa de vecindad de configuración hexagonal, el cual pinta cada neurona dependiendo del clúster en que se activó (ver Figura 5).

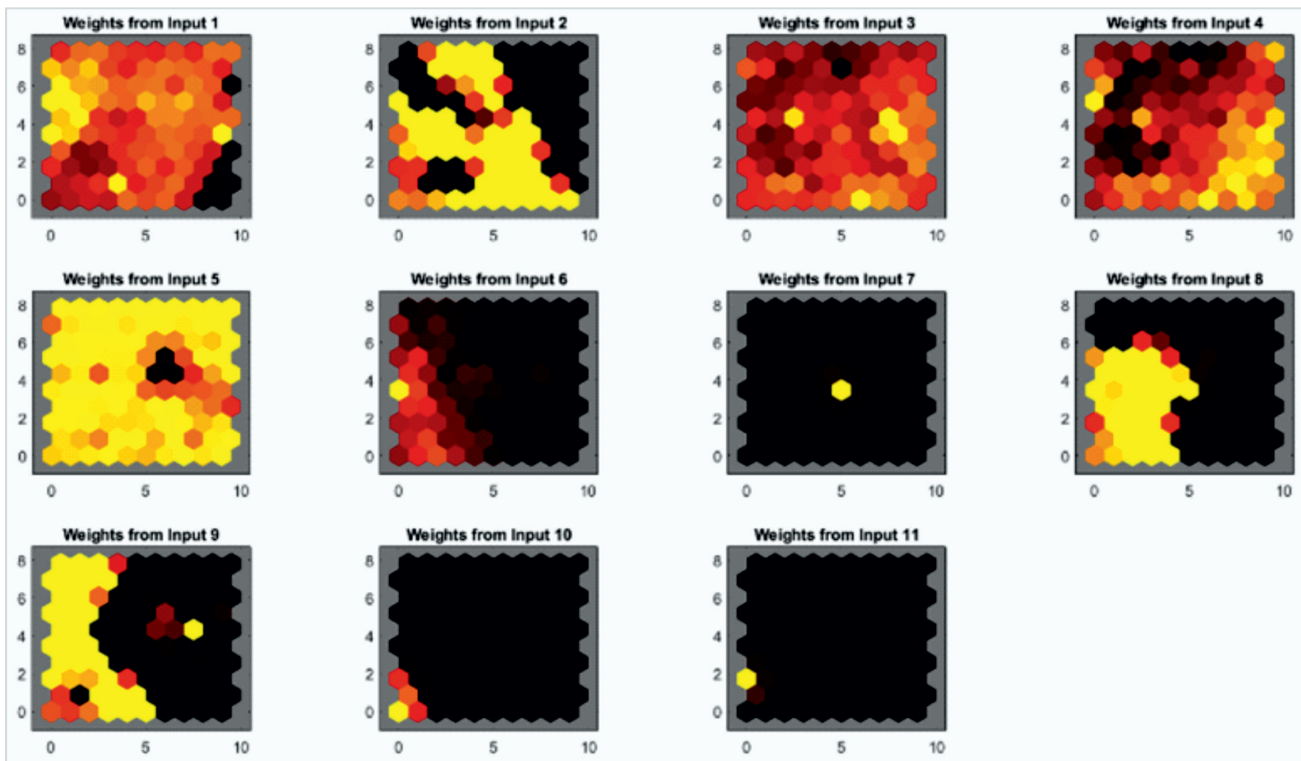


Figura 4. Comportamiento individual de las características

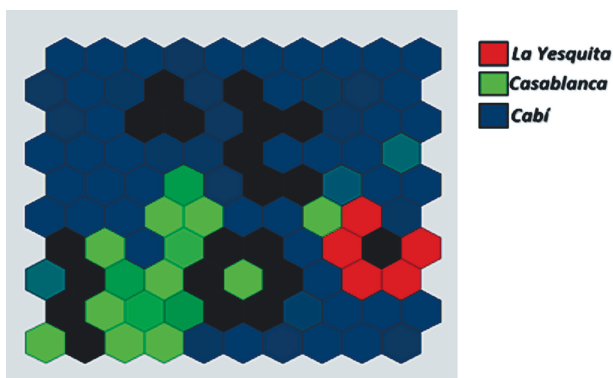


Figura 5. Neuronas activadas comparadas con la base de datos.

Para el análisis de los resultados se asignaron tres colores para distinguir las tres regiones consideradas en el caso de estudio: Rojo para la Yesquita, Verde para Casablanca, Azul para Cabí y el color negro corresponde a neuronas inactivas. Las combinaciones entre casas o sectores que compartían información son representadas por otros colores. En un panorama ideal, haciendo caso a las delimitaciones asignadas por los entes administrativos, el mapa debería presentar los tres colores sin combinaciones tal que se puedan identificar directamente las regiones; sin embargo, este mapa hace uso de las relaciones de las características comunes y nos presenta la distribución de la malaria en las regiones conforme al conjunto de sus características.

En otras palabras, las regiones que comparten características pueden ser más o menos vulnerables a los casos de malaria conforme a las características que describen las situaciones en las que se encuentran los habitantes y las condiciones reproductivas de los vectores más allá de las delimitaciones gubernamentales. Como una alternativa y haciendo uso de la separación en tres regiones, la figura 6 presenta la distribución y separación de clústeres que definen la malaria tal como se presentaron en Cabí (clúster superior), Casa-

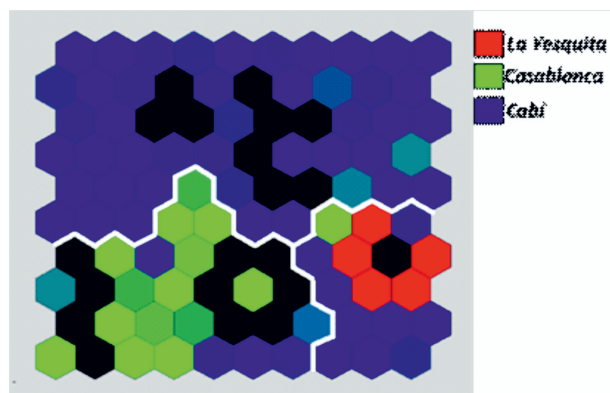


Figura 6. Regiones identificadas.

blanca (Inferior Izquierdo) y La Yesquita (Inferior Derecho).

Puede observarse que dentro de las regiones distinguidas hay neuronas que están resaltadas de otro color, esto es debido a que hay casas que están ubicadas en alguna de las regiones, pero presentan características que no son típicas de esta y que pueden ser de otra de las regiones. Asumiendo que cada región corresponde a una delimitación administrativa, siendo Cabí una zona rural, Casablanca una zona peri-urbana y la Yesquita una zona urbana, se puede observar como en la región que se identificó como peri-urbana, se observan neuronas activadas de color azul, lo que indica que una o varias casas de Cabí presentan características típicas de una región peri-urbana (figura 6).

Al concatenar ambos resultados se presenta una herramienta que es capaz de clasificar una casa o región como urbano, peri-urbano y rural; desde la perspectiva de la malaria, basándose en un conjunto de características dadas, tal como se muestra en la figura 7.

Consideraciones finales

Mediante el uso de redes neuronales se obtuvo una clasificación de las casas ubicadas en los sitios de estudio que permite distinguir tres grupos fácilmente diferenciables en el mapa auto organizado de Kohonen. Al contrastar los resultados suministrados por dicho mapa, con la ubicación real de las casas en cada uno de los tres barrios estudiados, se confirmó que estas casas fueron agrupadas correctamente por la red neuronal. Esto significa que existen diferencias significativas entre los barrios estudiados, aunque todos estén clasificados como pertenecientes a la cabecera municipal. En la figura 8 se presenta una herramienta que facilita la interpretación y comparación de lo que podría ser denominado como urbano, peri-urbano y rural desde la perspectiva de la malaria en la ciudad de Quibdó.

Se puede observar que la mayoría de los casos y por lo tanto de las casas, provenía del barrio Cabí, en el que se identificaron características más compatibles con el área rural. Esto quiere decir que, aunque se reporta un amplio número de casos de malaria en las cabeceras municipales, es probable que la mayoría de estos se presenten en realidad en barrios con características rurales. Por lo tanto, esto sugiere que para obtener una clasificación más acertada del origen los casos de malaria, se deberían incluir en el análisis las características ambientales que rodean la casa donde se produce el caso, así como la presencia de factores que propician el desarrollo del vector y, en consecuencia, la transmi-

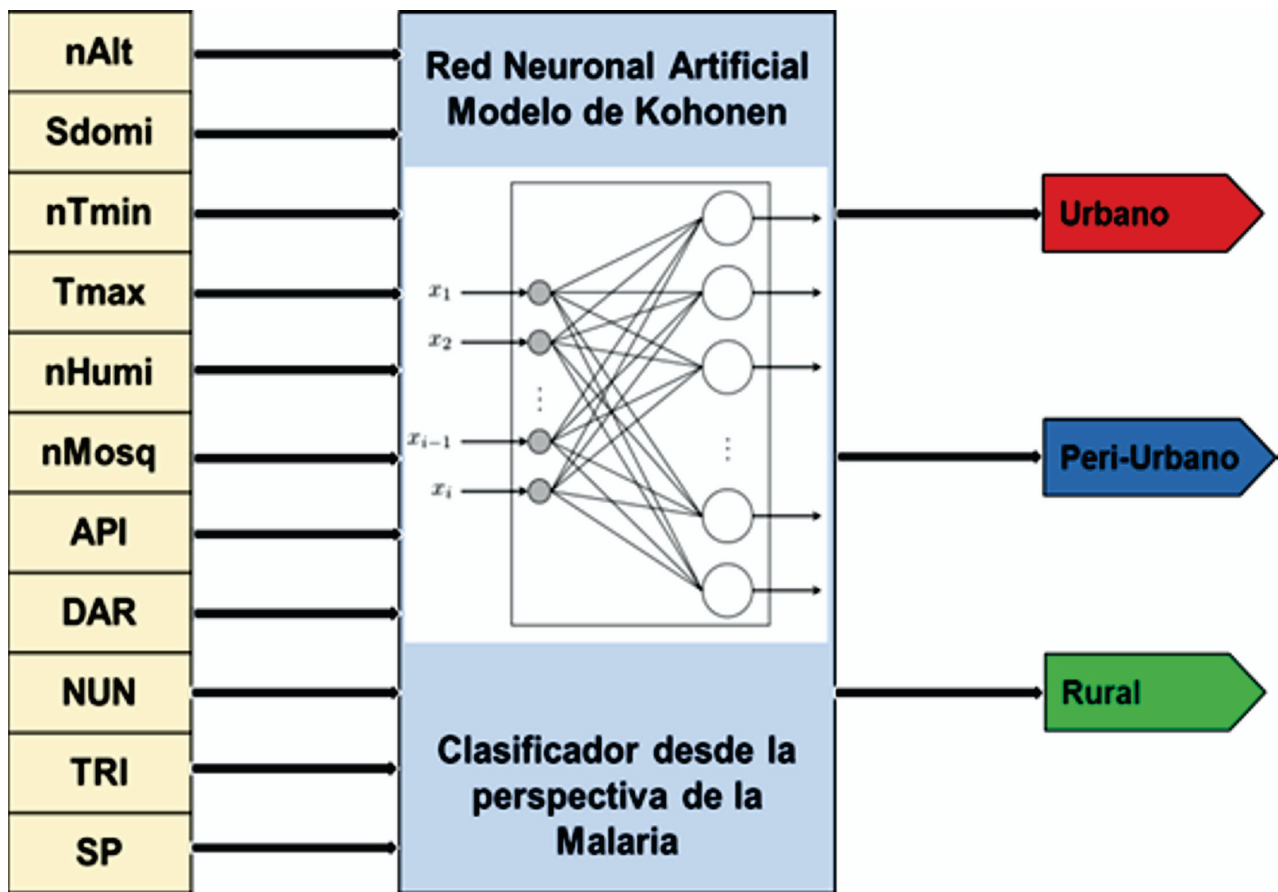


Figura 7 Clasificador de casos de Malaria

sión del parásito, y no solamente basarse en las delimitaciones administrativas existentes.

Estos resultados apoyan la hipótesis de que hay un error en la forma en la que se clasifica la malaria desde la óptica administrativa de las regiones endémicas y permiten definir una metodología que ofrece una medida objetiva para diferenciar y clasificar los diferentes barrios de una ciudad, teniendo en cuenta características que favorecen la presencia de vectores y la transmisión del *Plasmodium*.

En próximos proyectos, se podría relacionar la salida que se obtiene mediante el uso de las RNA, con un mapa geográfico de la ciudad, departamento o región donde se haya ejecutado el estudio, según las coordenadas de cada casa, generando un mapa donde se identifiquen las zonas de más alto riesgo para la transmisión de la malaria. Esta información podría ser utilizada por entidades gubernamentales y de vigilancia en salud pública para desarrollar estrategias de control vectorial, tales como fumigación, eliminación de criaderos y edu-

cación a la comunidad para evitar comportamientos que perpetúen la presencia de mosquitos, conociendo de antemano las características ecológicas y de urbanización de los barrios donde van a ser aplicadas.

Por otro lado, esta información podría ser útil para predecir dónde pueden surgir nuevos casos de malaria, dado el mapa de zonas de riesgo y de esa forma realizar actividades para lograr un diagnóstico y tratamiento oportunos en la comunidad. Además, se podría aplicar esta herramienta en el estudio y clarificación de otras enfermedades de alto contagio y de interés epidemiológico en el país tales como la infección por el virus del dengue, chikungunya, entre otros.

Agradecimientos

Los autores agradecen al Centro de Investigación Científica Caucaseco por facilitar la base de datos con la información que fue recolectada en la ciudad de Quibdó, que permitió la realización de este trabajo de investigación.

Referencias

- SIVIGILA (2016) Boletín Epidemiológico Semanal. Semana epidemiológica número 52 de 2016 (25 al 31 de diciembre).
- Padilla JC, Chaparro PE, Molina K, Arevalo-Herrera M, Herrera S (2015) Is there malaria transmission in urban settings in Colombia? *Malaria journal* 14: 453
- DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA, DANE. División Político-administrativa de Colombia, años 2000, 1997, 1992, 1988, 1983, y 1970.
- Beier, J. C. (1998). Malaria parasite development in mosquitoes. *Annual review of entomology*, 43(1), 519-543.
- González, R.; Carrejo. (2007) N. Claves y notas de distribución. Universidad del Valle; Cali. Introducción al estudio taxonómico de Anopheles de Colombia; p. 237
- Montoya-Lerma, J., Solarte, Y. A., Giraldo-Calderón, G. I., Quiñones, M. L., Ruiz-López, F., Wilkerson, R. C., & González, R. (2011). Malaria vector species in Colombia: a review. *Memórias do Instituto Oswaldo Cruz*, 106, 223-238.
- Peña Daniel, (2002). Análisis de datos multivariados.
- Anderson, J. (2007). *Redes Neuronales*. Alfaomega.
- Caicedo, Eduardo y López, Jesús (2009). Una aproximación práctica a las redes neuronales artificiales.
- Gómez, Fernández, López y Díaz. *Aprendizaje con Redes Neuronales Artificiales*.
- Flórez, Raquel y Fernández, José (2008). *LAS REDES NEURONALES ARTIFICIALES, FUNDAMENTOS TEORICOS Y PLICACIONES PRACTICAS*.
- T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464-1480, 1990. [2] T Kohonen, "Self-Organizing Maps," Berlin/Heidelberg, Germany: Springer, 30, 1995.
- A.K. Jain, M.N. Murty, P. 1. Flynn, "Data clustering: A review," *ACM Computing Surveys*, vol. 31, pp. 264-323, September 1999.
- F. Z. Chelali, A. Djeradi and N. Cherabit, "Investigation of DCT/PCA combined with Kohonen classifier for human identification," 2015 4th International Conference on Electrical Engineering (ICEE), Boumerdes, 2015, pp. 1-7.
- I. Hammami, G. Mercier and A. Hamouda, "The Kohonen map for credal classification of large multispectral images," 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, 2014, pp. 3706-3709.
- H. Jinliang, Y. Caiqing, L. Jiansheng and S. Jianjing, "Research on Clinical Diagnostic Models of IPF Syndromes in TCM Based on Dynamic Kohonen Network and Decision Tree," 2010 Second WRI Global Congress on Intelligent Systems, Wuhan, 2010, pp. 107-110.
- S. F. Baracho, V. V. d. Melo and R. C. Coelho, "Automated Left Ventricle Posterior Wall Segmentation Using Kohonen Self-Organizing Map," 2016 5th Brazilian Conference on Intelligent Systems (BRACIS), Recife, 2016, pp. 456-461.
- I. Hammami, G. Mercies and A. Hamouda, "The Kohonen map for credal fusion of heterogeneous data," 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, 2015, pp. 2947-2950.