



### Mapas auto organizativos para la asistencia médica oftalmológica

Franyelit Suárez Carreño<sup>1</sup>, Jaime Sayago Heredia<sup>2</sup>

<sup>1</sup>Pontificia Universidad Católica del Ecuador, Esmeraldas, Ecuador, <sup>2</sup>Pontificia Universidad Católica del Ecuador, Esmeraldas, Ecuador

<sup>1</sup>frangelits@gmail.com, <sup>2</sup>jaimepsayago@gmail.com

Recibido: 18/10/2017

Aceptado: 15/01/2018

#### RESUMEN

En este trabajo se presenta un algoritmo inteligente para la clasificación de personas con lentes de contacto en un consultorio médico de oftalmología. Se desarrolló un sistema basado en mapas auto organizativos de MatLab®. Este tipo de herramientas permite utilizar las imágenes características de los pacientes y clasificarlas a través de un aprendizaje no supervisado. Los Mapas Auto Organizativos (SOM por sus siglas en inglés) son un tipo de red neuronal del tipo Kohonen, que utiliza las características visual y auditiva del cerebro humano. El sistema logró realizar tres categorías: personas con lentes de contacto duro, personas con lentes de contacto blando y personas sin lentes. Este sistema sirve de apoyo a la consulta médica oftalmológica y optometrista, pudiendo agilizar el proceso de diagnóstico en los pacientes. Se empleó el índice de agrupación de Davies-Boulding para la optimización del error.

**PALABRAS CLAVES:** Mapas auto organizativos, asistencia médica, sistema inteligente, clasificación, lentes de contacto.

#### ABSTRACT

In this paper an intelligent algorithm for the classification of people with contact lenses is presented in a medical ophthalmology office. A system based on self-organizing maps of MatLab © was developed. This type of tools allows the use of the patient's characteristic images and classifies them through unsupervised learning. Auto Organizational Maps (SOM) are a type of neuronal network of the Kohonen type, which uses the visual and auditory characteristics of the human brain. The system managed to realize three categories: people with hard contact lenses, people with soft contact lenses and people without glasses. This system serves as support for the ophthalmological and optometrist medical consultation, being able to speed up the process of diagnosis in patients. The Davies-Boulding grouping index was used to optimize the error.

**KEYWORDS:** Organizational maps, medical assistance, intelligent system, classification, contact lenses.

## 1. Introducción

Los lentes de contacto son hoy en día, unos lentes que corrigen los problemas de la visión de forma cosmética y sin que se noten, se aplican directamente sobre la córnea. Este tipo de lentes tuvo su origen en el año 1508 cuando Leonardo Da Vinci sumergió la cara en un recipiente con agua y observó que las imágenes en el fondo se distorsionaban, afectando considerablemente la visión. Más adelante en el año 1887, F.A Muller utilizó estas ideas para crear los primeros lentes de contacto, que consistían en una



superficie de vidrio que se colocaba en la córnea, pero eran de gran tamaño y no podían utilizarse por mucho tiempo[1].

En el año 1948, Kevin Tuohy diseña los primeros lentes de contacto de metacrilato, los que resultaban de más fácil utilización y de mayor tolerancia para el paciente. Pero fue en el año 1950 que Otto Wichterle y Dreifus llevan a cabo la fabricación de los lentes de contacto blando, que pueden ser utilizados por 8 o 9 horas continuas y que no causan molestias al paciente[2].

En este trabajo se presenta el diseño de un sistema basado en inteligencia artificial para la clasificación de pacientes con lentes de contacto. El sistema fue desarrollado utilizando mapas auto organizativos de MatLab®, ya que proporciona una herramienta útil en el procesado de imágenes empleando redes neuronales del tipo Kohonen.

Existe una gran diversidad de algoritmos inteligentes para la clasificación de datos, este proceso consiste en clusterizar la información de acuerdo a características similares, tomando en cuenta algún criterio de determinación del error[3]. El proceso de clasificación se realiza siguiendo un patrón establecido, que permite generar una respuesta gráfica basada en símbolos, que es interpretada para luego ser procesada por otro algoritmo complementario.

Los mapas auto organizativos de Kohonen[4] son algoritmos diseñados para programar datos sobre un mapa bidimensional, de forma tal que distribuya sus características gradualmente y su implementación se realiza con el fin de producir una representación discreta del espacio de las muestras de entrada. Estos algoritmos han demostrado ser una herramienta muy útil para organizar la información, así también para visualizar los resultados, por lo que además facilita las búsquedas orientadas en gran contenido de información[5].

El desarrollo de un sistema inteligente para la asistencia médica, resulta de interés para dar apoyo al tratamiento de pacientes y al diagnóstico de posibles enfermedades. Este sistema pretende mejorar el tiempo de la consulta, facilitar el proceso de diagnóstico y contribuir al especialista médico. Se desarrolló la investigación bibliográfica, la obtención de datos, el procesado de información, la normalización de datos y finalmente la programación de la red neuronal para la clasificación.

## 2. Desarrollo

### Fundamentación teórica

Los mapas auto organizativos de Kohonen (SOM por sus siglas en inglés) son estructuras neuronales que utilizan aprendizaje competitivo para tratar de generar una relación espacial-topológica entre los vectores que caracterizan sus neuronas, a partir de un entrenamiento y en función de los vectores de entrada[6].

Pero para llevar a cabo la creación de un SOM se debe tener presente la inicialización de la red neuronal, definiendo el número de neuronas y especificando su topología, e inicializando el vector de pesos de cada neurona, algo que puede realizarse por comandos de configuración o a través de un constructor.

García et al[7] afirman que en este tipo de aprendizaje las neuronas compiten entre sí, activándose sólo una de ellas ante la presencia de un patrón o estímulo de entrada. El objetivo es, a partir de un proceso iterativo de comparación con la colección de datos de entrada, agrupar estos datos en base a su similitud. Para ello se presentan al mapa de vectores de entrada de igual dimensión que la de sus vectores característicos. Lo que significa que el conjunto de vectores de entrada debe tener características dimensionales iguales para su tratamiento.

La neurona ganadora establecerá el conjunto de neuronas cuyos vectores deben modificarse. Las neuronas están conectadas con sus vecinas mediante una relación impuesta por la estructura del SOM. El tamaño



del vecindario disminuirá a lo largo del entrenamiento; esa es la clave de la auto organización; reducir el vecindario para optimizar el resultado con el mínimo error. La función de actualización de pesos del mapa tiene la forma:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)] \quad (1)$$

donde  $t$  es el instante de tiempo discreto correspondiente a una iteración del algoritmo,  $x(t)$  es el vector de entrada en la iteración  $t$  y  $h_{ci}$  es la región de influencia que el vector de entrada tiene sobre el SOM, es decir es el núcleo de vecindad. Esta función es la que define la “rigidez” de la “red elástica” del SOM en el espacio de los datos[8].

$$h_{ci} = \alpha(t) \cdot \exp \frac{-||r_c - r_i||^2}{2\sigma^2(t)} \quad (2)$$

La función (2) que define el núcleo de vecindad  $h_{ci}$  puede ser de tipo gaussiano, lo que hará que la modificación de valores disminuya con la distancia en el vecindario, siendo  $\sigma(t)$  la función que define este radio de vecindario, que se va reduciendo después de cada iteración  $t$ . En los mapas bidimensionales las neuronas pueden ordenarse en un retículo rectangular o hexagonal, con lo que cada neurona tendrá 6 u 8 vecinos respectivamente.

La modificación de pesos depende también de la distancia entre una neurona  $n_i$  y la ganadora  $n_c$ , cuyos radios de vecindad  $r_i$  y  $r_c$ , tiende a cero según aumenta la malla o grid. La tasa de aprendizaje  $\alpha(t)$  es una función monótona decreciente respecto al tiempo  $t$  en el rango  $[0,1]$ . En otras palabras, tiende a cero a medida que se van procesando los vectores del conjunto de entrenamiento.

De este modo, una vez entrenado el mapa, después de un número fijo de iteraciones o cuando se cumpla la condición de parada establecida, ya se pueden introducir vectores clasificados manualmente que permiten etiquetar las zonas del mapa correspondientes.

Por tanto, los tres pasos fundamentales en la creación de un SOM corresponden a: inicialización, entrenamiento y ajuste. Una vez etiquetado, un SOM[9] puede usarse como un clasificador que asigna a cada vector de entrada la categoría con la que se haya etiquetado la neurona que se active en cada caso.

El algoritmo de agrupamiento de K-means es descrito en detalle por Hartigan[10] en 1975. Este afirma que la meta del algoritmo de K-means consiste en dividir  $M$  puntos en  $N$  dimensiones en  $K$  clúster, de tal forma que la suma dentro del clúster de cuadrados se minimiza.

La utilización de lentes de contacto es una práctica médica común para la corrección de problemas de la vista. Existen millones de usuarios de lentes de contacto, por lo que el conocimiento de la interacción de éstos con la superficie ocular es básico para el buen funcionamiento visual[11]. El uso de los lentes de contacto proporciona varios beneficios, pero también involucra algunos riesgos y complicaciones, que pueden desencadenar en procesos infecciosos o factores inmunológicos o tóxicos.

La herramienta MatLab® es ampliamente conocida y utilizada en universidades e institutos para el aprendizaje de cursos básicos y avanzados de matemáticas, ciencias y especialmente, ingeniería. Según Moler[12], MatLab® es una herramienta de software matemático que brinda un entorno de desarrollo integrado conocido como IDE con un lenguaje de programación propio denominado lenguaje M[13].

### 3. Metodología

El proceso metodológico para los mapas auto organizativos debe comprender en principio la evaluación de las variables involucradas en el proceso. La base de datos utilizada contiene 4 atributos y 24 instancias. La clasificación se lleva a cabo a través de tres categorías; 1) el paciente contiene lentes de contacto duro, 2) el paciente contiene lentes de contacto blandas, 3) el paciente no contiene lentes de contacto.



## 4. Diseño

Se desarrolló un algoritmo usando la herramienta toolbox de redes neuronales de MatLab®, ya que resulta didáctico y de fácil acceso para la clasificación de variables. Para asimilar la base de datos en ASCII se utilizó la función `som_read_data` que ofrece un proceso directo para el manejo de la información. Es importante resaltar que no se utilizó la función `rand` porque esta genera datos aleatorios que no son requeridos en este trabajo.

Los datos están contenidos en una matriz cuyas columnas constituyen los datos variables, mientras las filas corresponden a las muestras.

El mapa SOM fue construido a partir de la función `som-make` que ofrece una herramienta sencilla, versátil y con resultados óptimos. La figura 1 se presenta un diagrama de bloque que detalla las fases del sistema.

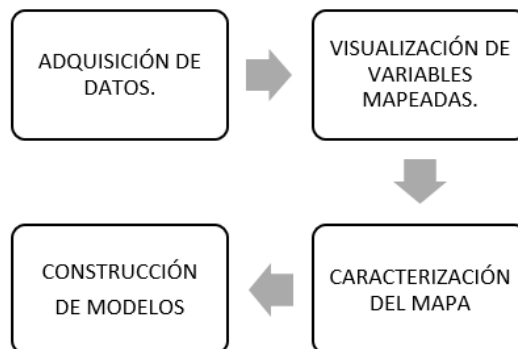


Figura 1: Diagrama del sistema auto organizativo.

## 5. Resultados

En este trabajo se presentan dos tipos de resultados; 1) la generación de clústeres que representan la categorización de los lentes de contacto, 2) la representación gráfica de la clasificación.

### 1. Generación de clúster

Para la generación de clústeres se ha considerado el desarrollo de un código en Matlab®, donde se han cargado los datos y se ha creado el SOM, para más adelante ser normalizados.

```
sD = som_read_data('lenses.data');  
sD = som_normalize(sD,'var');  
sM = som_make(sD);  
sM = som_autolabel(sM,sD,'vote');
```

Para lograr un agrupamiento en el mapa, se realizó un proceso de optimización. Este proceso se llevó a cabo a través del índice de agrupación Davies-Boulding. Pero antes se utilizó la función “`kmeans_clusters`” para obtener un particionamiento inicial.

```
[c,p,err,ind] = kmeans_clusters(sM, 7);  
[dummy,i] = min(ind)  
subplot(1,3,2)  
som_cplane(sM,Code,Dm)  
subplot(1,3,3)  
som_cplane(sM,cl)
```

Una vez establecido el agrupamiento de las variables, se procedió a crear modelos sobre el mapa basados en vecinos más cercanos. Luego se usó la función “`som_supervised`” para crear un clasificador para el conjunto de datos de la base de datos lentes.



La figura 2 muestra la matriz de peso, las etiquetas de cómo están agrupados los tipos de lentes y la correlación entre las variables, donde se puede observar que entre el atributo astigmático y la prescripción existe una correlación inversa notable.

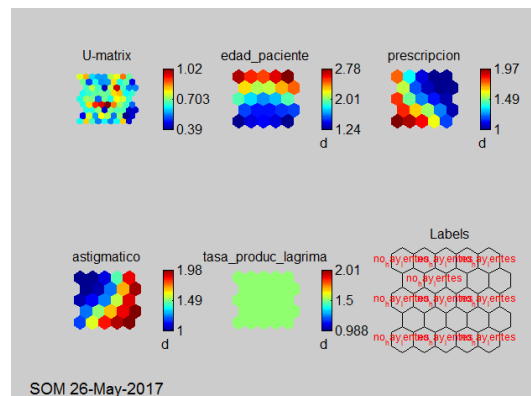


Figura 2: Inspección visual del mapa.

En la figura 3, se presentan la proyección, la misma que confirma la existencia de un clúster, es decir las unidades interpolativas indican que hay un grupo formado por la etiqueta “no\_hay\_lentes”. Esta etiqueta hace referencia a los pacientes que no necesitan usar lentes.

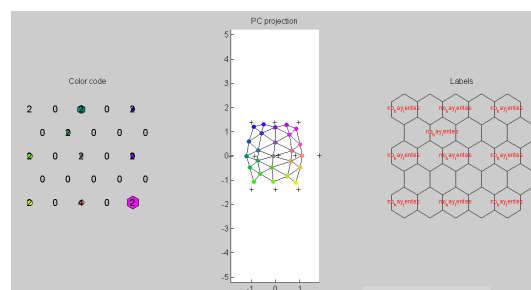


Figura 3: Proyección.

La figura 4 muestra los resultados del diagrama de dispersión simple e histograma de todas las variables. La información de los tipos de lentes se codifica como una quinta variable: 1) paciente con lentes de contacto duro, 2) para paciente con lentes de contacto blandas y 3) para paciente que no necesitan lentes. Los puntos de datos originales se encuentran en el triángulo superior, los valores del prototipo del mapa en el triángulo inferior y los histogramas en la diagonal: el color negro identifica el conjunto de datos y el rojo a los valores del prototipo del mapa.



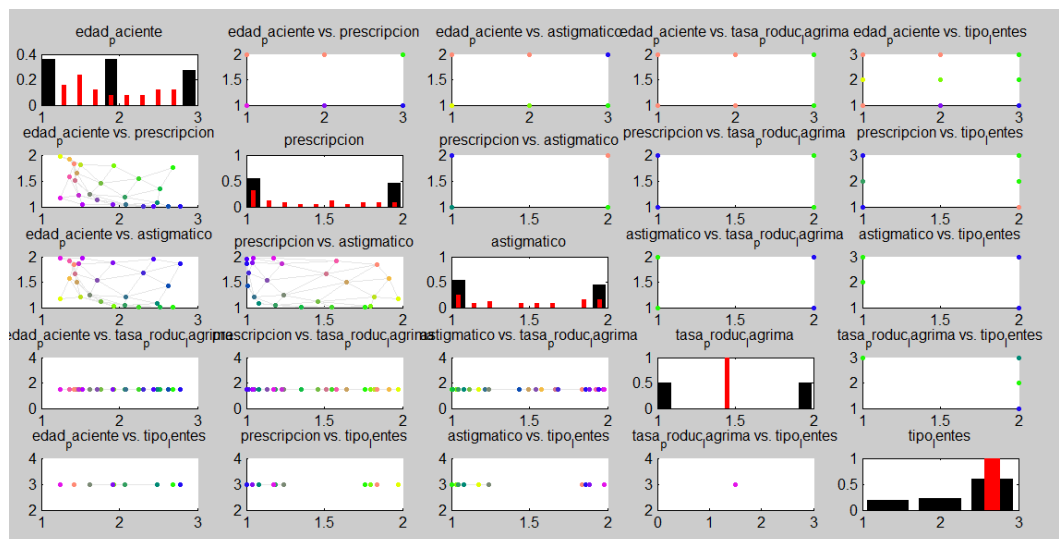


Figura 4: Diagrama de dispersión simple e histograma.

A partir de la Figura 5 se confirma que hay un solo grupo conformado por los pacientes que no necesitan lentes de ningún tipo a través de las distribuciones.

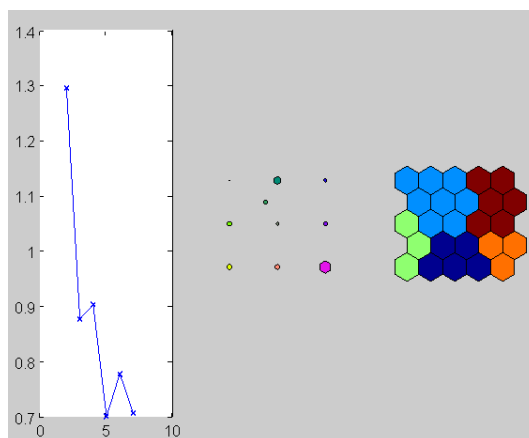


Figura 5: Distribuciones.

## 2. Representación gráfica de la clasificación

La figura 6 presenta el índice de agrupación Davies-Boulding, que se encarga de minimizar el error con la mejor agrupación.

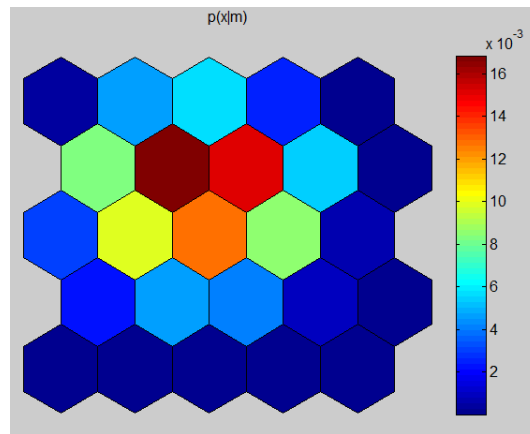


Figura 6: Categorización de grupos.

El índice Davies-Bouldin indica que hay un grupo en el mapa. Además este índice permite optimizar la búsqueda.

La figura 7 ofrece los resultados de la creación de modelos. Donde se observa que el color amarillo representa los pacientes con lentes duros, el verde claro los pacientes con lentes blandos y el celeste los pacientes que no necesitan lentes.

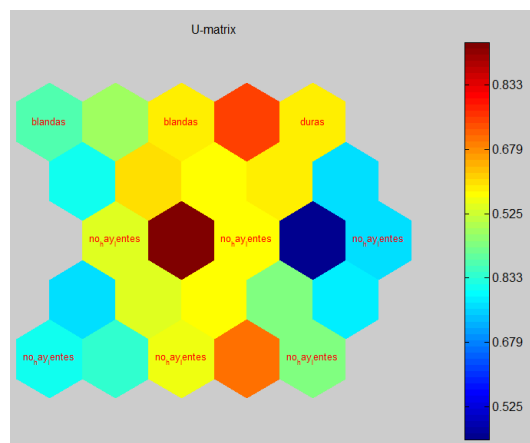


Figura 7: Modelación.

## 6. Conclusiones

Los lentes duros fueron clasificados a partir de un patrón que determina si el paciente no presenta astigmatismo, su tasa de producción de lágrima es reducida, su prescripción indica miopía y si es un paciente joven.

Los lentes blandos por su parte se clasificaron determinando si el paciente presenta astigmatismo, presbicia y su tasa de producción de lágrima es normal.

La clasificación para personas sin lentes se hizo a partir de quince datos de pacientes que padecen presbicia.



Los algoritmos de clasificación son posibles a través de las redes neuronales artificiales, las cuales presentan una gran robustez para el manejo de datos. Estos algoritmos permiten la clasificación de objetos en diferentes grupos.

## 7. Referencias

### Referencias

- [1] Juan A. Durán de la Colina. *Complicaciones de las lentes de contacto*. Ediciones Díaz de Santos, 1998.
- [2] Fernando Ballesteros. “De Da Vinci a nuestros días”. En: *Ciencia & Tecnología para la Salud Visual y Ocular* 7 (2006), págs. 107-115.
- [3] Teuvo Kohonen. “The self-organizing map”. En: *Proceedings of the IEEE* 78.9 (1990), págs. 1464-1480.
- [4] Ben Russell, Hujun Yin y Nigel M. Allinson. “Document clustering using the 1+ 1 dimensional self-organising map”. En: *International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 2002, págs. 154-160.
- [5] Miguel Ángel Navarro Huerga. “Análisis de medidas de potencia en interiores para su aplicación en sistemas de localización basados en la técnica del fingerprinting”. En: (2010).
- [6] Alberto P. García-Plaza, Víctor Fresno y Raquel Martínez. “Web page clustering using a fuzzy logic based representation and self-organizing maps”. En: *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT'08. IEEE/WIC/ACM International Conference on*. Vol. 1. IEEE, 2008, págs. 851-854.
- [7] Teuvo Kohonen y col. “Som pak: The self-organizing map program package”. En: *Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science* (1996).
- [8] John A. Hartigan y Manchek A. Wong. “Algorithm AS 136: A k-means clustering algorithm”. En: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1 (1979), págs. 100-108.
- [9] American Optometric Association. “Optometric Clinical Practice Guideline: Care of the contact lens patients.” En: 94.2 (2006), págs. 103-109.
- [10] R. Gorrochotegui y col. “Lentes de Contacto: Historia, Tipos y Complicaciones de su Uso.” En: *Informe Medico* 11.2 (2009).
- [11] A. Serrano y col. “Superficie ocular y lentes de contacto”. En: *Rev Mex Oftalmol* 82.6 (2008), págs. 352-65.
- [12] J Wiley. *MATLAB: An introduction With Applications*. Hoboken (NJ): USA: REVERTE S.A., 2005.
- [13] Richard Goering. “Matlab edges closer to electronic design automation world.” En: *Electronic Engineering Times* 1341 (2004), págs. 4-5.