

## LA COMPUTACIÓN INTELIGENTE PARA LA CLASIFICACIÓN DE SEMILLAS DE TRIGO

Andrés Hurtado Pasquel<sup>1</sup>, Wendy Estupiñan Estupiñan<sup>2</sup>, Franyelit Suárez Carreño<sup>3</sup>  
angel.andrew.hurtad@gmail.com, wendy.paola.estupinan@gmail.com, franyelit.suarez@pucese.edu.ec  
Pontificia Universidad Católica del Ecuador, Sede Esmeraldas-Ecuador

Recibido (06/06/17), aceptado (08/02/18)

---

**Resumen:** Los mapas auto-organizados (SOM) de Kohonen, son un tipo de redes neuronales artificiales de aprendizaje no supervisado, que permite la identificación de patrones entre datos midiendo las distancias entre ellos para formar clústers mediante el encuentro de un representante por cada grupo. En este artículo se presenta el algoritmo SOM para representar mapas de datos, se utilizará un programa en Matlab con la herramienta SOM toolbox para llevar a cabo la clasificación de tres tipos de semillas de trigo, a partir de los atributos de las mismas.

---

**Palabras Clave:** SOM, entrenamiento, red neuronal artificial, base de datos, clasificación.

## INTELLIGENT COMPUTING FOR THE CLASSIFICATION OF WHEAT SEEDS

---

**Abstract:** Kohonen's self-organized maps (SOM) are a type of artificial neural networks of unsupervised learning that allows the identification of patterns between data by measuring the distances between them to form clusters by meeting one representative for each group. In this article we present the SOM algorithm to represent data maps, we will use a program in Matlab with the tool SOM toolbox to carry out the classification of three types of wheat seeds, based on the attributes of the same.

---

**Key words:** Keywords - SOM, training, artificial neural network, database, classification.

## I. INTRODUCCIÓN

La inteligencia en la actualidad ha evolucionado, ésta ya no es solo para las personas, sino también las máquinas; el hombre ha creado máquinas electrónicas capaz de resolver problemas al igual que un ser humano, para comprobarlo se han realizado estudios y pruebas, entre ellos el Test de Turing, el cual describe lo siguiente: si una persona establece una conversación con una máquina que se encuentra en un cuarto oscuro, y al finalizar piensa que estuvo conversando con una persona, se considera que esa máquina es inteligente.

Otro gran paso en el avance tecnológico ha sido la Inteligencia Artificial (IA) la misma que pretende simular el funcionamiento del cerebro humano para aprender y resolver problemas.

El presente trabajo consiste en utilizar el algoritmo de clustering SOM, analizando su desempeño para distintos parámetros de entrenamiento. Para esto se utilizará una base de datos conformada por tres especies de semillas con siete atributos cada una [1][2]. Esta base de datos contiene siete atributos de mediciones de las propiedades geométricas de los núcleos pertenecientes a tres diferentes variedades de trigo para 210 muestras, las cuales son:

1. Zona (A)
2. Perímetro (P)
3. Compacidad ( $C = 4 * \pi * A / P^2$ )
4. Longitud de kernel
5. Anchura del kernel,
6. Coeficiente de asimetría
7. Longitud de la ranura del núcleo.

La base de datos contiene información de tres diferentes clases de trigos: Kama(1), Rosa(2) y canadiense(3), con 70 elementos cada uno, seleccionados al azar.

Mediante la herramienta Matlab se encontrará la correspondencia entre los tipos de trigos y las características de cada uno mediante una red SOM.

## II. FUNDAMENTACIÓN TEÓRICA

### A. Semillas de Trigo

El término trigo se emplea para identificar a otro tipo de cereal denominado Triticum, del cual existen diversas especies que han evolucionado a través del tiempo dependiendo el clima, la estación, región, temperatura, y más. Las semillas de trigo se caracterizan por ser de tipo endosperma (contenida en un solo saco embrionario); además de poder clasificarse por anchura, longitud, asimetría o zonas/regiones. [3]

### B. Redes Neuronales Artificiales

Las redes neuronales artificiales (RNA), consisten en un gran número de neuronas interconectadas que trabajan en paralelo hacia una solución específica, estas permiten el procesamiento de grandes volúmenes de datos a gran velocidad.

### C. Aprendizaje No Supervisado

Este tipo de aprendizaje es muy útil para la solución de problemas en los cuales se pretende buscar e identificar estructuras, tipo de organización o jerarquías de datos, sin contar con alguna forma de conocimiento previo para ejercer un entrenamiento.

### D. Herramienta SOM Toolbox

SOM Toolbox es una biblioteca de funciones libres para MATLAB 5 que implementa el algoritmo SOM, visualización y análisis [4][5]. Esta herramienta permite el entrenamiento de redes SOM en diferentes tipologías y parámetros de aprendizaje, permite calcular diferentes errores, calidad y medidas para SOM. La visualización de esta se presenta a través de matrices, planos de componentes, codificación de colores por clústers, enlaces de colores entre SOM y más.

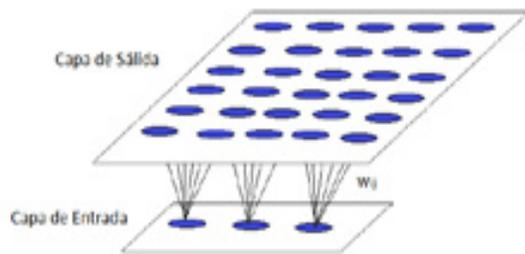
### E. Redes SOM

SOM por sus siglas en inglés corresponden a Self-Organizing Maps que significa Mapas auto-organizados o también llamados Mapas de Kohonen, es un tipo de red neuronal artificial con entrenamiento no supervisado, el mismo que permite agrupar los datos a través de un representante por cada grupo, para así visualizar un mapa con la descripción intuitiva de la similitud entre los datos.

#### Arquitectura del SOM

Kohonen [6] propuso el modelo neuronal de la red SOM con el objetivo de demostrar que a partir de estímulos externos o datos de entrada de espacios multidimensionales, la estructura propia de la red artificial puede describir y proyectar la información abstraída en los estímulos, organizando datos de salida en un mapa bidimensional. Un modelo SOM está compuesto por dos capas de neuronas [7]. La capa de entrada, formada por N neuronas, una por cada variable de entrada, se encarga de recibir y transmitir a la capa de salida la información procedente del exterior. La capa de salida, formada por M neuronas, es encargada de procesar la información y formar el mapa de rasgos [8].

En la Figura 1 se observa la estructura que presenta la red SOM, compuesta por la capa de entrada y salida.



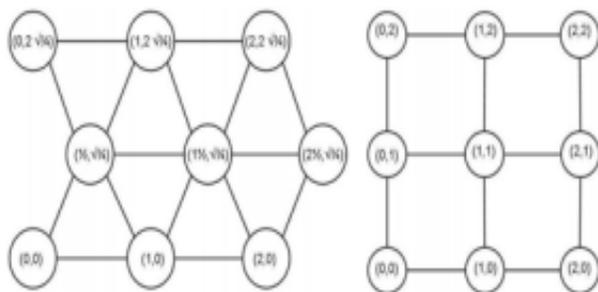
**Figura 1. Estructura de la Red SOM.**

Las Redes SOM se caracterizan por su aprendizaje secuencial que consiste en presentar ejemplos, buscar el resultado más cercano y posteriormente actualizar su posición en el espacio destinado para este [9].

La Arquitectura de las redes SOM cuentan con:

**Matriz de neuronas:** estas son un conjunto de neuronas de la red que se distribuyen por todo un campo de dos o tres dimensiones, pudiendo ser hexagonal, rectangular, entre otras, de tal manera que se identifican los grupos o clústeres [4][10].

En la Figura 2, se evidencia las matrices de neuronas con dimensión hexagonal y rectangular.



**Figura 2. Configuraciones hexagonal y rectangular en la redícula del SOM.**

**Espacio de entrada:** conjunto de datos de entrada (inputs) que se identifican como un vector de componentes por cada atributo y así se determina la dimensión del vector de pesos iniciales.

**Espacio de salida:** corresponde al conjunto de datos de salida (outputs/targets) de la red neuronal SOM de menor dimensión que el conjunto de entrada.

**Relación topológica entre neuronas:** entre las neuronas existe una relación de vecindad, la cual es subyacente al conjunto de datos debido a su estructura topológica natural. Lo más importante es definir la regla de asociación al inicio de la Red Neuronal SOM [8].

## F. Entrenamiento

Cada vez que se entrena la red neuronal SOM, a esta se le introduce un vector de datos de entrada y otro vector aleatorio correspondiente a los pesos de inicialización de cada neurona, se cuantifica la disimilitud entre ellos, y la neurona vencedora será aquella que se encuentre más cerca a esa medida, formando clústeres con las demás neuronas que se aproximen a la neurona vencedora [8].

## III. METODOLOGÍA

Al analizar el desempeño del algoritmo de clustering Red Neuronal SOM a través de distintos parámetros de entrenamiento, se evidencia la importancia de la utilización de dichas redes/algoritmos en la vida cotidiana. Día tras día se crean nuevas aplicaciones para encontrar patrones similares y agruparlos entre ellos, de tal manera que se entienda la relación que existe.

El presente trabajo explica detalladamente el procedimiento que se siguió para la creación, entrenamiento y evaluación de la red neuronal SOM, para lograr la clasificación efectiva de un grupo constituido por núcleos examinados (semillas) pertenecientes a tres diferentes tipos de trigo: Kama, Rosa y Canadiense. Se cuenta con una base de datos titulada “seed” en formato \*.data tomada de UC Irvine Machine Learning Repository, esta base de datos contiene 7 atributos y 70 elementos por cada tipo de semilla, siendo en total 210 instancias seleccionadas al azar para el experimento. Estos datos formarán la matriz de entrada.

Para obtener los datos agrupados de las diferentes semillas de trigo se instaló la herramienta SOM toolbox en la aplicación Matlab, procediendo así al ingreso de los datos de entrada y salida en un solo archivo con formato \*.data, la creación e inicialización del mapa, la normalización y entrenamiento de los datos. Adicionalmente se procedió a la creación del código en un archivo con formato \*.m con el cual una vez listo se puede llegar a visualizar los mapas que se requieran para el entendimiento y posterior análisis de los mismos.

Finalmente se realizó la clasificación de patrones teniendo en cuenta los diferentes pesos aleatorios generados, que permiten visualizar de manera intuitiva en un plano 2D la organización de los distintos clústeres de los datos suministrados, los cuales presentan un espectro de colores según la clase de semilla a la que pertenecen.

Finalmente se realizó la clasificación de patrones teniendo en cuenta los diferentes pesos aleatorios generados, que permiten visualizar de manera intuitiva en un plano 2D la organización de los distintos

clústeres de los datos suministrados, los cuales presentan un espectro de colores según la clase de semilla a la que pertenecen.

#### IV. RESULTADOS

Es necesario que la herramienta SOM sea capaz de agrupar en el mapa los tres tipos de trigos, proporcionándole únicamente los datos sobre las siete características citadas. Para ello se crearán las categorías de forma no supervisada por la propia red. Se inicializó la red con los datos de las semillas, tomando en cuenta las características propias de cada tipo

```
Rseed= som_read_data('C:\IA2\seed.data'); % Lee
datos en código ascii
Rseed = som_normalize(Rseed,'var'); % Normaliza el
conjunto de datos
Eseed = som_make(Rseed); % crea, inicializa y
entrena a la red SOM
Eseed = som_autolabel(Eseed,Rseed,'vote'); % ordena
las etiquetas automáticamente
som_
```

```
Una vez declarado podemos empezar a trabajar
con la herramienta, obteniendo por cada atributo
un código de colores que lo representa y que hace
énfasis en las características propias de cada semilla .
subplot(1,3,1)
som_cplane(Eseed,Code,DisMatriz);
hold on
som_grid(Eseed,'Label',cellstr(int2str(hits)),...
'Line','none','Marker','none','Labelcolor','k');
hold off
title('Codigo de Colores')
```

Una vez terminada la parte de codificación, es posible caracterizar los 7 clústeres asociados a los tipos de trigo. La Figura 3 muestra la matriz de todos los atributos de los tres tipos de semillas, matriz de pesos agrupados por atributos y la clasificación, En las matrices, los tonos de colores más oscuros representan una gran distancia entre las neuronas y los colores más claros significa que los vectores de los atributos de las semillas están próximos entre sí en el espacio de entrada.

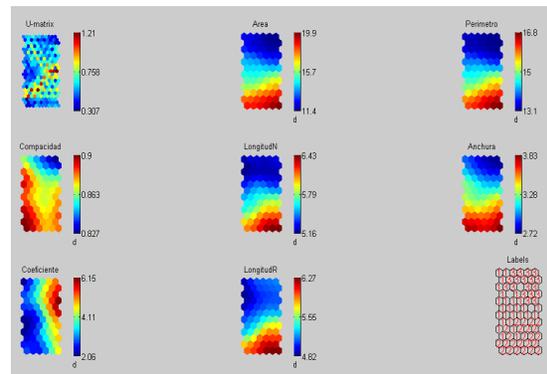


Figura3. Gráfica en Hexagonal de cada atributo Agrupado

Las proyecciones pueden también observarse en la Figura 4, en la cual se presenta la clasificación de los tres tipos de semillas, obsérvese que existen dos mapas hexagonales cada uno con las etiquetas propias de cada semilla agrupadas en tres clústeres diferentes, y un gráfico de proyecciones. En los tres casos se puede diferenciar como se encuentran separados los grupos de semilla.

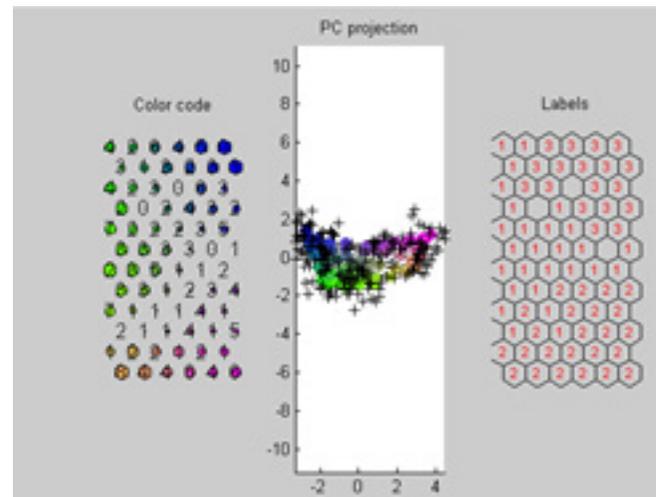


Figura 4. Gráfica de las Proyecciones

La Figura 5 se muestran los datos agrupados por atributos de dos en dos y la relación existente entre cada uno de ellos. Para lo cual, un atributo se coloca en el eje de las x y el siguiente en el eje de las y determinando lo dispersos o cercanos que se encuentran los datos.

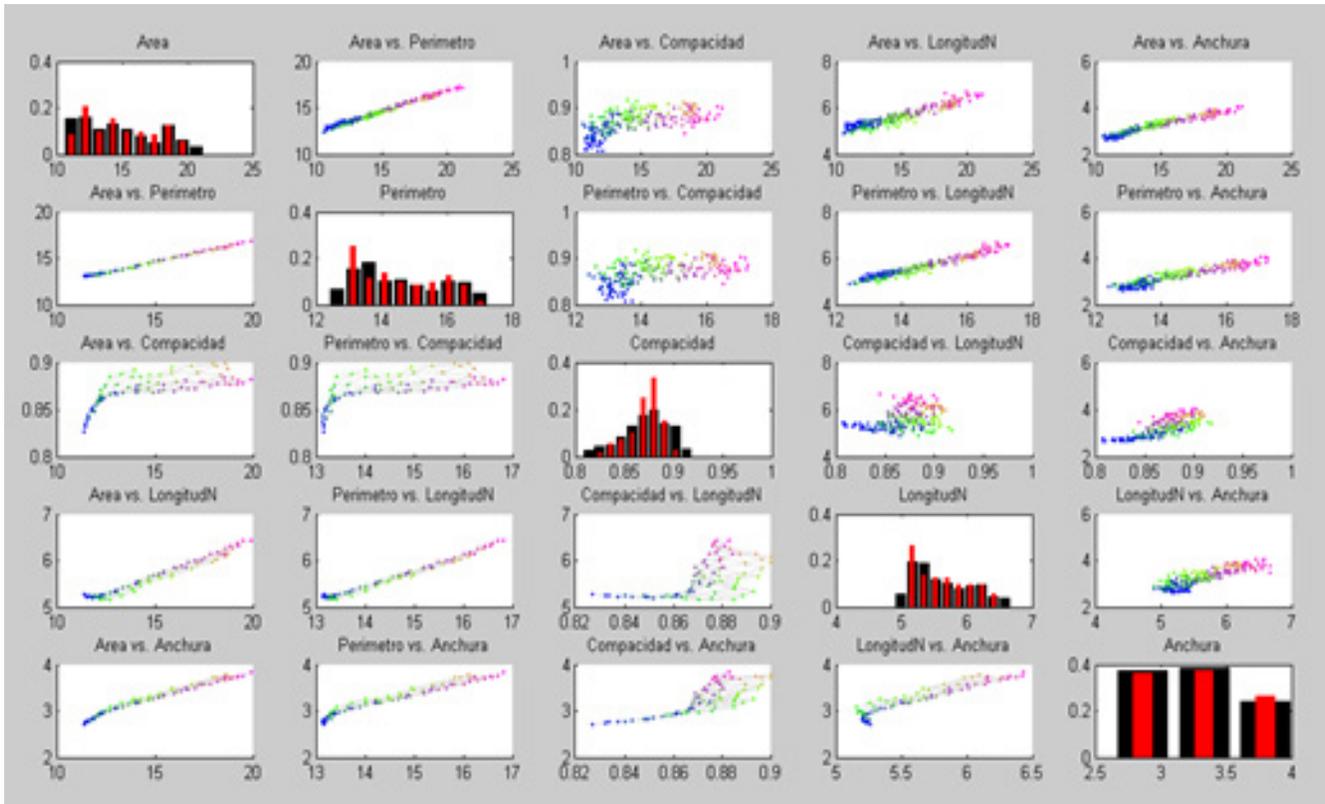


Figura 5. Gráfica de relaciones entre Atributos

Otra manera de analizar los datos a través de gráficas se encuentra en la Figura 6. En la imagen encontramos 3 gráficas diferentes, la del lado izquierdo muestra un plano que identifica el peso de las neuronas por atributos (eje x atributos, eje y pesos neuronas) y los siguientes gráficos muestran el agrupamiento de las matrices de pesos.

no tiene nada que ver uno con el otro, ya que existe un espacio significativo entre las neuronas (entre más diferente sea el color de una neurona y otra, más espacio existirá).

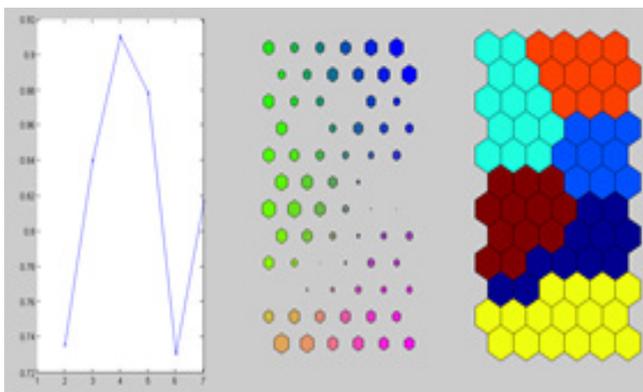


Figura 6. Agrupamiento de Valores por Atributos

Para finalizar, la Figura 7 muestra la matriz completa con los datos una vez ya agrupados correctamente luego del entrenamiento de la red. Se puede observar claramente que existen neuronas de color amarillo junto con otras de color azul, esto quiere decir que estos datos

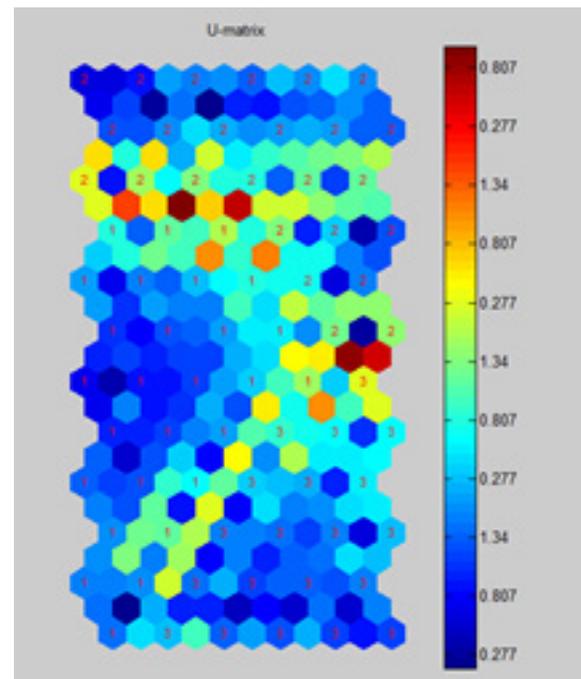


Figura 7. Matriz Hexagonal con las Neuronas de Entrenamientos (Datos Agrupados por clase de Semilla de Trigos)

Para llevar a cabo el agrupamiento, fue de suma importancia la utilización de los siguientes comandos de la herramienta SOM Toolbox: som\_read\_data permite la lectura del archivo con los datos de las semillas de trigo en formato .data, som\_normalize permite normalizar los datos con sus respectivos atributos para poder ser inicializado y creado el mapa con som\_make y por último som\_show o som\_show\_add que permiten la visualización de los mapas auto-organizativos, mostrando los tres tipos de semillas agrupados.

## V. CONCLUSIONES

Los estudios se llevaron a cabo utilizando la combinación de grano de trigo cosechado procedente de yacimientos experimentales, explorado en el Instituto de Agrofísica de la Academia de Ciencias de Polonia en Lublín. El grupo constituido por núcleos examinados pertenecientes a tres diferentes variedades de trigo: Kama, Rosa y canadiense, 70 elementos cada uno, seleccionados al azar para el experimento [11]. Los cuales ofrecen una importante información para llevar la selección y clasificación de las variables.

Las redes SOM permiten visualizar de manera intuitiva en un plano 2D la organización de los distintos clústeres de los datos, los cuales pueden ser coloreados según la clase a la que pertenecen, haciendo más fácil la visualización de los vectores prototipo asociados a cada clúster.

La Red Neuronal SOM se utiliza para clasificar patrones teniendo diferentes pesos aleatorios, consume bastantes recursos cuando hay un número alto de neuronas, así como las iteraciones que realiza y el número de vecindad que posee. En la actualidad las redes neuronales SOMs se utiliza para clasificar y determinar grupos en los cuales hay muchos patrones. [1]

Finalmente, se concluye que fue posible caracterizar los 7 clústeres asociados a los tipos de trigo, en cuanto se logró una clasificación utilizando el mapa SOM con un alto porcentaje de datos clasificados correctamente.

## VIII.- REFERENCIAS

[1] T. York, A. Ukpong, S. Mylvaganam and Y. Ru, «Parameter estimation from tomographic data using self-organising maps», 2012 IEEE International

Conference on Imaging Systems and Techniques Proceedings, 2012.

[2] F. González, N. Obregón, «Self-organizing maps of Kohonen as a river clustering tool within the methodology for determining regional ecological flows ELOHA», Ingeniería y Universidad [en línea] 2013, 17 (Julio-Diciembre) : [Fecha de consulta: 28 de mayo de 2017] Disponible en: <<http://oai.redalyc.org/articulo.oa?id=47728826003>> ISSN 0123-2126.

[3] A. Yasar, E. Kaya y I. Saritas, «Classification of Wheat Types by Artificial Neural Network,» International Journal of Intelligent Systems and Applications in Engineering, vol. 4, n° 1, pp. 12-15, 2016.

[4] M. Chang, Y. Horng and H. Jia-Sheng, «Evolutionary self-organizing map», 1998 IEEE International Joint Conference on Neural Networks Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98CH36227).

[5] A. Akhterov, O. Lezina and A. Shastina, «DIAGNOSTICS OF DEVELOPMENT OF ORGANIZATIONAL AND MANAGERIAL COMPETENCIES OF THE ENGINEERS USING THE SELF-ORGANIZING KOHONEN MAPS», Automation and Control in Technical Systems, no. 42, 2013.

[6] E. Atenogenes, «Laboratorio de Dinámica no Lineal,» [En línea]. Available: [www.dynamics.unam.edu](http://www.dynamics.unam.edu). [Último acceso: 20 05 2017].

[7] I. Callejas, J. Píneros, J. Rocha, F. Hernández and F. Delgado, «Implementación de una red neuronal artificial tipo SOM en una FPGA para la resolución de trayectorias tipo laberinto», 2013 II International Congress of Engineering Mechatronics and Automation (CIIMA), 2013.

[8] A. Gómez, «Redes neuronales artificiales: The Self-Organizing Maps (SOM) para el reconocimiento de patrones.,» Institución Universitaria Los Libertadores , vol. 1, n° 1, 17 diciembre 2013.

[9] S. Gálvez, «SELF-ORGANIZING MAPS,» Universidad de Chile Facultad de Ciencias Físicas y Matemáticas Departamento de Ingeniería Eléctrica, Santiago, 2014.

[10] J. Vesanto, J. Himberg, E. Alhoniemi y J. Parhankangas, «Laboratory of Computer and Information Science Adaptive Informatics Research