



UNIVERSIDAD DEL BÍO-BÍO
FACULTAD DE CIENCIAS
EMPRESARIALES
INGENIERIA CIVIL EN INFORMATICA

Software para la Implementación de Mapas Auto-Organizados de Kohonen (SOM) para la clasificación de las acciones de la bolsa de Chile

JAVIER HUMBERTO ANDANA QUIJADA
RAMÓN ALEJANDRO DURAN FIGUEROA

Profesor Guía : Carola Andrea Figueroa Flores

Profesor Informante : María Angélica Caro Gutiérrez

Nota Final Proyecto de Título : _____

Memoria para optar al título de
Ingeniero Civil en Informática

CHILLÁN, 2015

Resumen

En este informe se describe la implementación en lenguaje JAVA de un tipo de red neuronal artificial llamada Mapas Auto-organizados de Kohonen (SOM) con el objetivo representar el comportamiento de las acciones de la Bolsa de Comercio de Santiago. Utilizando datos de los retornos accionarios de la bolsa de Santiago durante diferentes periodos de tiempo como datos de entrada para alimentar la red neuronal se buscó obtener clústers que permitan agrupar las acciones cuyo comportamiento de sus retornos sean similares. Lo anterior con el objetivo de establecer relaciones que ayuden a entender comportamiento de las acciones de mejor manera. En el informe se describe detalladamente el algoritmo de la red de Kohonen y su implementación, además, se muestra los resultados obtenidos al alimentar la red con los retornos de las acciones de la Bolsa de Comercio de Santiago durante los periodos 2009, 2010, 2011, 2012, 2013 y 2014. Finalmente los resultados obtenidos son contrastados con los resultados conseguidos por los profesores Carola Figueroa Flores y Edinson Cornejo Saavedra en el documento llamado “Estructura Topológica del Mercado Bursátil de Chile, período 2009 – 2013”.

Abstract

This report describes the implementation in JAVA language a type of artificial neural network called Self -organizing maps of Kohonen (SOM) in order to represent the behavior of the shares in the Santiago Stock Exchange . Using data from stock returns to the Santiago stock exchange during different time periods as input to feed the neural network sought to obtain clusters that allow group actions the behavior of their returns are similar. This with the aim of establishing relationships to help understand behavior of the better. The report describes in detail the algorithm of Kohonen network and its implementation also shows the results obtained by feeding the network with the returns of the shares of the Santiago Stock Exchange during the periods 2009, 2010, 2011, 2012, 2013 and 2014. Finally, the results obtained are compared with the results achieved by teachers Carola Figueroa Cornejo Flores and Edinson Saavedra in the document called “Topological Structure Stock Market in Chile, period 2009-2013 “.

Índice General

1	<u>INTRODUCCIÓN</u>	7
2	<u>MARCO TEÓRICO</u>	9
2.1	TÉCNICAS DE CLASIFICACIÓN EXISTENTES	9
2.1.1	REDES NEURONALES ARTIFICIALES	10
2.1.2	ÁRBOLES DE DECISIÓN.....	11
2.1.3	ALGORITMOS GENÉTICOS.....	12
2.1.4	DETECCIÓN DE CONGLOMERADOS (CLUSTERS).....	13
2.1.5	TÉCNICAS DE INFERENCIA ESTADÍSTICA.....	14
2.1.6	VISUALIZACIÓN	16
2.1.7	REGRESIÓN LINEAL.....	17
2.2	¿QUÉ SON LAS REDES NEURONALES?	18
2.2.1	TIPOS DE REDES NEURONALES	18
2.3	REDES NEURONALES ARTIFICIALES	19
2.3.1	REDES NEURONALES ARTIFICIALES SUPERVISADAS.....	19
2.3.2	REDES NEURONALES ARTIFICIALES NO SUPERVISADAS	19
2.4	MODELOS DE LAS REDES NEURONALES ARTIFICIALES	20
3	<u>DESCRIPCIÓN DEL PROBLEMA Y TRABAJOS RELACIONADOS</u>	21
3.1	PROBLEMA	21
3.2	TRABAJOS SIMILARES	21
3.2.1	LOS MÉTODOS NO LINEALES	21
3.2.2	LOS MODELOS LINEALES.....	24
3.2.3	PREDICCIÓN DE LA BOLSA A TRAVÉS DE MODELOS HÍBRIDOS.....	24
3.2.4	MODELOS BASADOS EN VARIABLES ECONÓMICAS	25
3.3	MOTIVACIÓN	26
4	<u>DEFINICIÓN PROYECTO</u>	27
4.1	OBJETIVO DEL PROYECTO	27
4.2	OBJETIVOS ESPECÍFICOS	27
4.3	DEFINICIONES, SIGLAS Y ABREVIACIONES	28
5	<u>DESCRIPCIÓN DE LA METODOLOGÍA Y MAPAS AUTO-ORGANIZADOS DE KOHONEN (SOM)</u>	29
5.1	MÉTODO	29
5.2	HERRAMIENTAS Y DATOS UTILIZADOS	29
5.3	MAPAS AUTO-ORGANIZADOS DE KOHONEN (SOM)	30
5.3.1	DESCRIPCIÓN DEL ALGORITMO DE KOHONEN	30
5.3.2	IMPLEMENTACIÓN ALGORITMO DE KOHONEN	37
5.3.3	DIAGRAMA DE CASOS DE USO.....	48
5.3.4	ESPECIFICACIÓN CASOS DE USO.....	49
5.3.5	DIAGRAMA DE CLASES SOM	52
5.3.6	FUNCIONAMIENTO DE LA APLICACIÓN	56
5.3.7	PRUEBA DE FUNCIONAMIENTO	60
6	<u>RESULTADOS Y CONCLUSIONES</u>	63

7	BIBLIOGRAFÍA.....	73
8	ANEXOS.....	75
8.1	CLASIFICACIÓN DE LAS ACCIONES DEL AÑO 2009 SEGÚN SU RUBRO	75
8.2	CLASIFICACIÓN DE LAS ACCIONES DEL AÑO 2010 SEGÚN SU RUBRO	77
8.3	CLASIFICACIÓN DE LAS ACCIONES DEL AÑO 2011 SEGÚN SU RUBRO	79
8.4	CLASIFICACIÓN DE LAS ACCIONES DEL AÑO 2012 SEGÚN SU RUBRO	81
8.5	CLASIFICACIÓN DE LAS ACCIONES DEL AÑO 2013 SEGÚN SU RUBRO	83
8.6	CLASIFICACIÓN DE LAS ACCIONES DEL AÑO 2014 SEGÚN SU RUBRO	85
8.7	RESULTADO INVESTIGACIÓN PROFESORES CAROLA-EDINSON AÑO 2009	87
8.8	RESULTADO INVESTIGACIÓN PROFESORES CAROLA-EDINSON AÑO 2010	88
8.9	RESULTADO INVESTIGACIÓN PROFESORES CAROLA-EDINSON AÑO 2011	89
8.10	RESULTADO INVESTIGACIÓN PROFESORES CAROLA-EDINSON AÑO 2012.....	90
8.11	RESULTADO INVESTIGACIÓN PROFESORES CAROLA-EDINSON AÑO 2013.....	91
8.12	RESULTADO INVESTIGACIÓN PROFESORES CAROLA-EDINSON AÑO 2014.....	92

Índice Tablas

Tabla 1. Caso de Uso Seleccionar Período.....	49
Tabla 2. Caso de Uso Entrenar SOM	50
Tabla 3. Caso de Uso Clasificar Acciones.....	51
Tabla 4. Tabla Atributos y Métodos SOM	56
Tabla 5. Alimentos con sus distintas especificaciones.....	60
Tabla 6. Color representativo por Mercado	63
Tabla 7. Grupo 1 y Grupo 2	64
Tabla 8. Grupo 3 y Grupo 4	65
Tabla 9. Grupo 5 y Grupo 6	66
Tabla 10. Grupo 7 y Grupo 8.....	67
Tabla 11. Grupo 9 y Grupo 10.....	68
Tabla 12. Clasificación Acciones año 2009 según su rubro	76
Tabla 13. Clasificación Acciones 2010 según su rubro.....	78
Tabla 14. Clasificación Acciones año 2011 según su rubro	80
Tabla 15. Clasificación Acciones año 2012 según su rubro	82
Tabla 16. Clasificación Acciones año 2013 según su rubro	84
Tabla 17. Clasificación Acciones año 2014 según su rubro	86

Índice Figuras

Figura 1. Red Neuronal Biológica vs Red Neuronal Artificial (Epistemowikia, 2014)	10
Figura 2. Ejemplo Simple Árbol de decisión (Morales & González, 2013).....	11
Figura 3 . Algoritmo Genético Simple (Baillard, 2012)	13
Figura 4. Clustering Jerárquico (Sanchez, 2013).....	14
Figura 5. Imagen 3D Visualización Previa (Orallo, 2004).....	17
Figura 6. Árbol de decisión Visualización Posterior (Granada, 2011).....	17
Figura 7. Ejemplo Regresión Lineal (GEOtutoriales, 2014).....	18
Figura 8. Ejemplo Mapa Auto-Organizado (Caparrini, 2014).....	31
Figura 9. Arquitectura Red SOM (Caparrini, 2014).....	32
Figura 10. Reducción Radio de Entorno de la BMU (Caparrini, 2014).....	34
Figura 11. Ejemplo SOM según Nivel de Pobreza (Caparrini, 2014)	36
Figura 12. Ejemplo SOM Proyección de países (Caparrini, 2014).....	37
Figura 13. Diagrama General SOM	40
Figura 14. Subproceso Entrenamiento SOM.....	41
Figura 15. Subproceso Clasificación SOM.....	47
Figura 16. Diagrama Casos de Uso SOM.....	48
Figura 17. Diagrama de Clases SOM.....	52
Figura 18. Menú Selección de Período	56
Figura 19. Matriz Inicial.....	57
Figura 20. Matriz Post- Entrenamiento.....	58
Figura 21. Matriz Post- Funcionamiento	59
Figura 22. Resultado SOM clasificación Alimentos (MetaQuotes, 2014).....	61
Figura 23. Resultado Alimento implementado con SOM del proyecto.....	62
Figura 24. Resultado Análisis Período 2011.....	69
Figura 25. Resultado Kruskal Acciones Bursátiles Año 2009 (Figueroa Flores & Cornejo, 2014)	87
Figura 26. Resultado Kruskal Acciones Bursátiles Año 2010 (Figueroa Flores & Cornejo, 2014)	88
Figura 27. Resultado Kruskal Acciones Bursátiles Año 2011 (Figueroa Flores & Cornejo, 2014)	89
Figura 28. Resultado Kruskal Acciones Bursátiles Año 2012 (Figueroa Flores & Cornejo, 2014)	90
Figura 29. Resultado Kruskal Acciones Bursátiles Año 2013 (Figueroa Flores & Cornejo, 2014)	91
Figura 30. Resultado Kruskal Acciones Bursátiles Año 2014 (Figueroa Flores & Cornejo, 2014)	92

1 INTRODUCCIÓN

Una de las mayores preocupaciones en el ámbito de las inversiones accionarias o de quienes toman decisiones gerenciales basándose en el posible comportamiento futuro del mercado, es disminuir al máximo los riesgos y la incertidumbre al momento de realizar la toma de decisiones. Para esto se busca encontrar métodos o modelos predictivos del comportamiento de las acciones que sean cada vez más exactos y que idealmente no sean dependientes de la opinión de un experto. Es en este aspecto en el cual utilizando datos históricos de la variable bajo estudio como acciones o índices bursátiles, se busca predecir los futuros comportamientos de los mercados a través de diferentes técnicas basadas en modelos matemáticos y estadísticos. En las últimas décadas con los grandes avances en el ámbito de la computación y la informática se han ido incorporando técnicas que incluyen redes neuronales artificiales y sistemas difusos.

Las redes neuronales artificiales poseen la propiedad de capturar las características no lineales de los índices de la bolsa debido a que pueden ser entrenadas con una cantidad suficiente de información para identificar dichas relaciones no lineales entre los valores de entrada y los de salida.

En este documento se utilizará un tipo de red neuronal artificial no supervisada llamada Mapas Auto-organizados (SOM) o también llamadas Redes o Mapas de Kohonen con el objetivo de encontrar relaciones entre los retornos accionarios de las acciones de la bolsa de Santiago para conformar grupos o clusters de acciones cuyos comportamientos de sus retornos sean similares.

Ahora se explicará un resumen sobre los distintos capítulos existentes en este informe:

Capítulo 1: Introducción con respecto al tema a desarrollar, junto con el resumen de cada capítulo.

Capítulo 2: Desarrollo de un Marco teórico acerca de las redes neuronales. Posteriormente se hará referencia a las diversas técnicas existentes últimamente sobre la clasificación de los datos.

Capítulo 3: Explicación sobre el problema existente a resolver. Además se explicarán trabajos similares sobre la clasificación de las acciones bursátiles.

Capítulo 4: Descripción sobre el proyecto a desarrollar. Aquí se explicarán los objetivos a cumplir, además de las diversas siglas a utilizar durante el desarrollo de este informe.

Capítulo 5: Explicación sobre la metodología utilizada en el desarrollo del proyecto. También se explicará sobre las herramientas y datos que se utilizaron en la creación del software. Por último se hablará sobre el Algoritmo de Kohonen. En esta parte se dividirá en 2 partes; la primera tratará sobre el funcionamiento del algoritmo, mostrando además algunos ejemplos en los que se ha utilizado. La segunda parte será sobre el código del algoritmo desarrollado, mostrando y explicando su funcionamiento a través de pedazos de código. Además se mostrará un ejemplo que corroborará la buena implementación del algoritmo.

Capítulo 6: Muestra de los resultados obtenidos por el software y conclusiones obtenidas a raíz de estos resultados.

2 MARCO TEÓRICO

En esta sección, se definirán las diferentes técnicas que existen hoy en día para la clasificación de datos empleados principalmente en la Minería de Datos, aparte de definir la técnica a utilizar para el desarrollo del proyecto.

2.1 Técnicas de Clasificación Existentes

Cuando se tiene una gran cantidad de información es necesario poder realizar alguna clasificación u ordenamiento de ésta, de manera de poder obtener y descubrir ciertas relaciones o patrones entre estos datos. Para esto, una de las herramienta más importantes, sobre todo en el ámbito de la informática, es la minería de datos la cual ocupa diversas técnicas de clasificación de datos (Moreno, 2007).

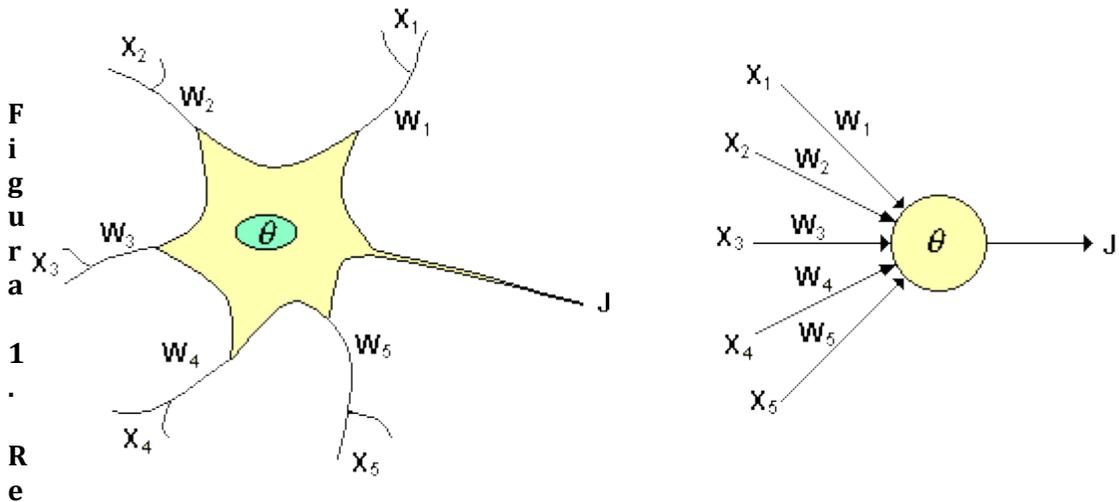
Las Técnicas más usadas son:

- Redes Neuronales
- Árboles de Decisión
- Algoritmos Genéticos
- Detección de Conglomerados (Clústers)
- Aprendizaje Automático
- Técnicas de inferencia estadística
- Visualización
- Regresión Lineal

En el capítulo 2.2 se detallará los tipos de redes neuronales y su funcionamiento.

2.1.1 Redes Neuronales Artificiales

Una red neuronal artificial es un sistema compuesto de muchos elementos procesadores operando en paralelo, cuya función es determinada por la estructura de la red, la fuerza en las conexiones y el procesamiento realizado por los nodos. Esta red está basada en el funcionamiento de las redes neuronales biológicas que se encuentran en el cerebro. En la Figura. 1 se muestra un paralelo entre una red neuronal biológica y una red neuronal artificial.



Red Neuronal Biológica vs Red Neuronal Artificial (Guadalajara, 2014)

En el **capítulo 2.2** se explicará con mayor detalle el funcionamiento de las redes neuronales y sus tipos.

2.1.2 Árboles de Decisión

Los árboles de decisión son una de las técnicas de aprendizaje inductivo supervisado no paramétrico. Se utiliza para la predicción y se emplea en el campo de inteligencia artificial, donde a partir de una base de datos, se construyen diagramas de construcción lógica muy similares a los sistemas de predicción basados en reglas. Estos sirven para representar y categorizar una serie de condiciones que ocurren en forma repetitiva para la solución de un problema. En la Figura 2 se muestra un árbol de decisión básico el cual busca determinar si determinadas condiciones climáticas son positivas (P) o negativas(N).

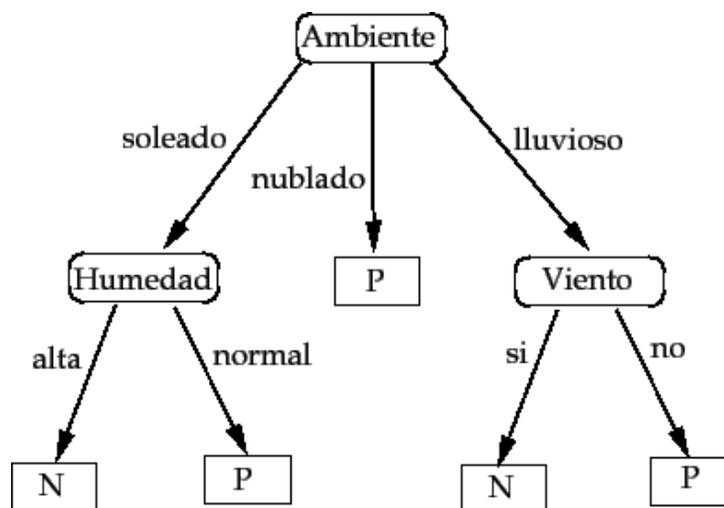


Figura 2. Ejemplo Simple Árbol de decisión (Morales & González, 2013)

Propiedades:

Una de las propiedades de esta técnica es que permite una organización eficiente de un conjunto de datos, debido a que los árboles son construidos a partir de la evaluación del primer nodo (raíz) y de acuerdo a su evaluación o valor tomado se va descendiendo en las ramas hasta llegar al final del camino (hojas del árbol), donde las hojas representan clases y el nodo raíz representa todos los patrones de entrenamiento los cuales se han de dividir en clases.

Los sistemas que implementan árboles de decisión tales como ID3 son muy utilizados en lo que se refiere a la extracción de reglas de dominio. Este método (ID3) se construye a partir del método de Hunt. La heurística de Hunt, consiste en escoger la característica más

discriminante del conjunto X, luego realizar divisiones recursivas del conjunto X, en varios subconjuntos disyuntos de acuerdo a un atributo seleccionado.

2.1.3 Algoritmos Genéticos

Los AGs son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Están basados en el proceso genético de los organismos vivos. A lo largo de las generaciones, las poblaciones evolucionan en la naturaleza de acorde con los principios de la selección natural y la supervivencia de los más fuertes, postulados por Darwin (1859).

Por imitación de este proceso, los Algoritmos Genéticos son capaces de ir creando soluciones para problemas del mundo real. La evolución de dichas soluciones hacia valores óptimos del problema depende en buena medida de una adecuada codificación de las mismas.

Un algoritmo genérico generalmente consta de una población inicial, la cual es sometida a un ciclo (iteración) que encadena hibridaciones, mutaciones y selecciones para crear nuevas generaciones de individuos. En cada nueva generación, se verifica si se ha logrado el objetivo antes de lanzar una nueva iteración.

En la imagen siguiente (Figura 3) se muestra de manera visual, el algoritmo genético Simple.

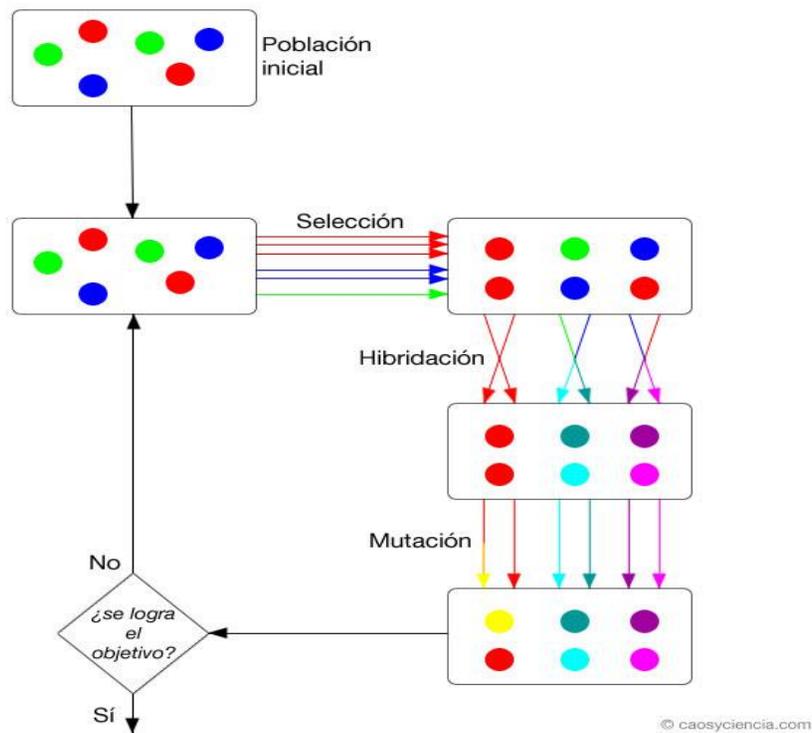


Figura 3 . Algoritmo Genético Simple (Baillard, 2012)

2.1.4 Detección de Conglomerados (Clusters)

El análisis de clúster o Clustering es el proceso de agrupar los datos en clases o en clústeres, de tal forma que, los datos de un mismo clúster tienen una alta similitud y a su vez, son muy diferentes de los de otro clúster. Un clúster de objetos puede ser tratado colectivamente como un grupo o ser considerado como una forma de compresión de datos

Uno de los problemas del análisis de clúster es que no existe una definición precisa de clúster. Lo cual ha originado el desarrollo de una gran cantidad de métodos, es así que los dos grandes grupos de métodos de clustering son: los jerárquicos y los no jerárquicos o particionales.

- **Clustering Jerárquicos**

La pertenencia a un grupo o clúster en un nivel de la jerarquía condiciona la pertenencia a grupos de un nivel superior (Ver Figura 4).

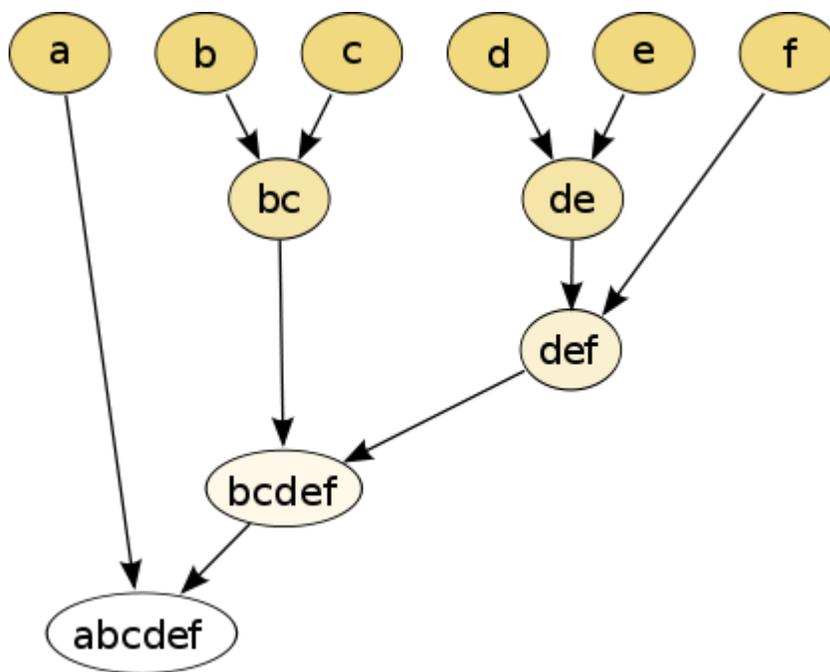


Figura 4. Clustering Jerárquico (Sanchez, 2013)

- **Clustering no Jerárquicos**

Los métodos particionales obtienen una única partición de los datos mediante la optimización de alguna función adecuada.

2.1.5 Técnicas de inferencia estadística

La inferencia estadística es el conjunto de métodos y técnicas que permiten inducir, a partir de la información empírica proporcionada por una muestra, la cuál es el comportamiento de una determinada población con un riesgo de error medible en términos de probabilidad.

Los métodos paramétricos de la inferencia estadística se pueden dividir, básicamente, en dos: métodos de estimación de parámetros y métodos de contraste de hipótesis. Ambos métodos se basan en el conocimiento teórico de la distribución de probabilidad del estadístico muestral que se utiliza como estimador de un parámetro.

- **Método estimación de Parámetros**

Consiste en asignar un valor concreto al parámetro o parámetros que caracterizan la distribución de probabilidad de la población.

Cuando se estima un parámetro poblacional, aunque el estimador que se utiliza posea todas las propiedades deseables, se comete un error de estimación que es la diferencia entre la estimación y el verdadero valor del parámetro.

El error de estimación es desconocido por lo cual es imposible saber en cada caso cual ha sido la magnitud o el signo del error; para valorar el grado de precisión asociado con una estimación puntual se parte de dicha estimación para construir un intervalo de confianza.

- **Método Contraste de Hipótesis**

Tienen como objetivo comprobar si determinado supuesto referido a un parámetro poblacional, o a parámetros análogos de dos o más poblaciones, es compatible con la evidencia empírica contenida en la muestra.

Los supuestos que se establecen respecto a los parámetros se llaman hipótesis paramétricas. Para cualquier hipótesis paramétrica, el contraste se basa en establecer un criterio de decisión, que depende en cada caso de la naturaleza de la población, de la distribución de probabilidad del estimador de dicho parámetro y del control que se desea fijar a priori sobre la probabilidad de rechazar la hipótesis contrastada en el caso de ser ésta cierta.

2.1.6 Visualización

La técnica de visualización de datos se utiliza fundamentalmente con dos objetivos:

- Aprovechar la gran capacidad humana de extraer patrones a partir de imágenes.
- Ayudar al usuario a comprender más rápidamente patrones descubiertos automáticamente por un sistema de KDD.

Estos dos objetivos marcan dos momentos diferentes del uso de la visualización de los datos (no excluyentes):

- **Visualización previa:** se utiliza para entender mejor los datos y sugerir posibles patrones o qué tipo de herramienta de KDD utilizar. En la Figura 5 se presenta un ejemplo sobre visualización previa.
- **Visualización posterior:** se utiliza para mostrar los patrones y entenderlos mejor. Un árbol de decisión es un ejemplo de visualización posterior. En la Figura 6 se presenta la forma de un árbol de decisión.

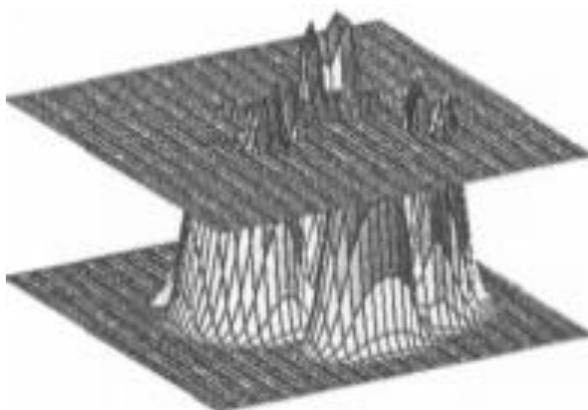


Figura 5. Imagen 3D Visualización Previa (Orallo, 2004)

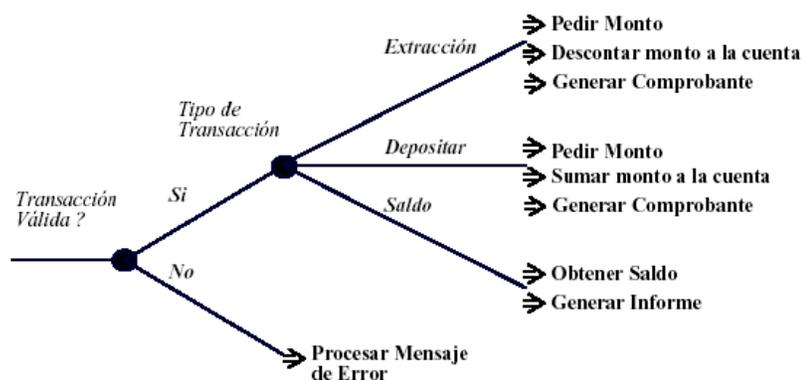


Figura 6. Árbol de decisión Visualización Posterior (Granada, 2011)

2.1.7 Regresión Lineal

La regresión lineal (Ver Figura 7) permite identificar relaciones entre variables numéricas y construir modelos de regresión: 1 variable salida y múltiples entradas numéricas. Se consideran relaciones de una variable de salida (dependiente) con múltiples variables de entrada (independientes).

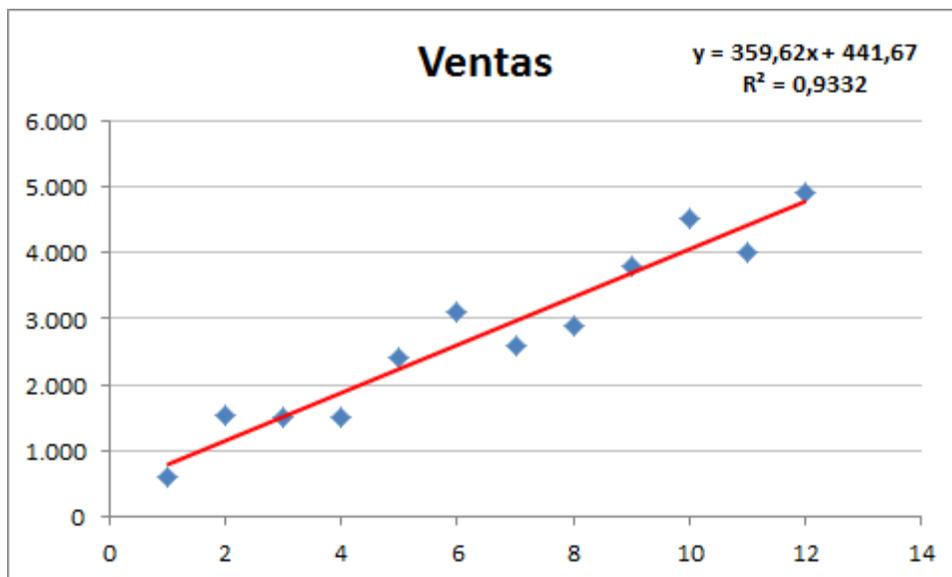


Figura 7. Ejemplo Regresión Lineal (GEOtutoriales, 2014)

2.2 ¿Qué son las Redes Neuronales?

Como se dijo anteriormente, una red neuronal (Fernandez, 2014) es un sistema compuesto de muchos elementos procesadores simples operando en paralelo, cuya función es determinada por la estructura de la red, la fuerza en las conexiones y el procesamiento realizado por los nodos.

2.2.1 Tipos de Redes Neuronales

Existen 2 tipos de Redes Neuronales:

- Redes Neuronales Biológicas
- Redes Neuronales Artificiales

En este trabajo, sólo se hará énfasis a las Redes Neuronales Artificiales.

2.3 Redes Neuronales Artificiales

Las redes neuronales artificiales (ANN) (Fernandez, 2014) imitan su funcionamiento a aquellas que se encuentran en el ámbito biológico. Son aptas para resolver problemas que no poseen un algoritmo claramente definido para transformar una entrada en una salida; aprenden, reconocen y aplican relaciones entre objetos.

Existen 2 tipos de ANN, las cuales son:

- Redes Neuronales Artificiales Supervisadas
- Redes Neuronales Artificiales No Supervisadas

2.3.1 Redes Neuronales Artificiales Supervisadas

Se le suministra a la red un conjunto de datos de entrada y la respuesta esperada. Se comparan los datos obtenidos por el sistema con los datos de entrada aportados, con la respuesta esperada y se modifican los pesos en función del error obtenido.

El conjunto de datos utilizados en el proceso de aprendizaje se denomina conjunto de entrenamiento (training set).

Si los vectores de entrada utilizados no contienen la información idónea la red puede llegar a no aprender.

2.3.2 Redes Neuronales Artificiales No Supervisadas

Sólo se aplica a la red neuronal los datos de entrada, sin indicarle la salida esperada. La red neuronal es capaz de reconocer algún tipo de estructura en el conjunto de datos de entrada (normalmente redundancia de datos) y de esta forma se produce auto-aprendizaje.

Durante el proceso de aprendizaje la red auto-organizada debe descubrir por sí misma rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, incorporarlos a su estructura interna de conexiones.

Se dice, por tanto, que las neuronas deben auto-organizarse en función de los estímulos (datos) procedentes del exterior.

2.4 Modelos de las Redes Neuronales Artificiales

Existe una serie de modelos que aparecen en la mayoría de estudios académicos y la bibliografía especializada, los cuales son:

- Perceptrón
- Adaline
- Perceptrón multicapa
- Memorias asociativas
- Máquina de Boltzmann
- Máquina de Cauchy
- Propagación hacia atrás
- Redes de Elman
- Redes de Hopfield
- Red de contrapropagación
- Redes de neuronas de base radial
- Redes de neuronas de aprendizaje competitivo
- Mapas Autoorganizados (RNA) (Redes de Kohonen)
- Crecimiento dinámico de células
- Gas Neuronal Creciente
- Redes ART

3 DESCRIPCIÓN DEL PROBLEMA Y TRABAJOS RELACIONADOS

3.1 Problema

Pronosticar el comportamiento de los mercados accionarios o patrones de comportamiento en los mercados financieros, ha sido un área de mucho interés tanto para personas dedicadas al área financiera como para académicos. Con el objetivo de poder disminuir la incertidumbre y los riesgos al momento de realizar una inversión financiera diferentes autores han buscado encontrar patrones de comportamiento en los mercados financieros para diversos activos financieros y países utilizando diferentes técnicas. Es en este contexto en el que este proyecto está orientado a representar la estructura del mercado bursátil chileno durante el período 2009-2014, a partir de las correlaciones entre los retornos de las acciones transadas en la Bolsa de Comercio de Santiago y el uso de Mapas Auto-organizados de Kohonen, para el desarrollo de una representación gráfica de las relaciones de dependencia mutua, clusters y acciones líderes que permitan ayudar a comprender el comportamiento de las acciones de la Bolsa de Comercio de Santiago.

3.2 Trabajos similares

Para el manejo de las acciones, se han utilizado diversos métodos o técnicas las cuales están orientadas principalmente a poder generar modelos que permitan predecir los cambios en un mercado bursátil, dentro de los métodos más usados se tienen los modelos lineales y los no lineales.

3.2.1 Los métodos no lineales

Dentro de los métodos no lineales están las redes neuronales artificiales las cuales poseen la propiedad de capturar las características no lineales de los índices de bolsa.

Dentro de los trabajos destacados se puede mencionar Doherty (2005) quien ofrece una clasificación intuitiva para grupos de empresas basándose en la actividad que describe con mayor precisión la naturaleza del negocio de la compañía. El análisis busca determinar si los precios de cierre de las acciones se mueven juntos, y si los grupos de acciones que no se mueven juntos son identificables en términos de actividad industrial. Usando TreeGNG, un

algoritmo de agrupamiento jerárquico, se verificará que una serie de precios de cierre de las acciones a través del tiempo se mantiene (Doherty, 2005).

Cupal, Deev & Linnertová (2012) buscaron identificar que tan interconectados estaban los mercados financieros europeos, es decir, cuanto un miembro de la comunidad europea influye en los demás y si eventualmente hay excepciones a esta situación. Para este estudio se utilizó una estructura topología de red y teoría de grafos basándose en la correlación de los mercados de la comunidad (Cupal, Deev, & Linnertová, 2012).

Giovanni Bonanno, Guido Caldarelli, Fabrizio Lillo, and Rosario N. Mantegna (2003) Compararon las propiedades topológicas del árbol de expansión mínima obtenida a partir de un gran grupo de acciones que se negocian en la Bolsa de Valores de Nueva York durante un período de comercio de 12 años con la obtenida a partir de datos simulados mediante modelos simples de mercado. Encontraron que el árbol real tiene características de una red compleja que no puede ser reproducida, ni siquiera como una primera aproximación, por un modelo de mercado al azar (Giovanni Bonanno, 2002).

Chen, Leung y Daouk (Chen, Leung, & Daouk, 2003) modelan y predicen el cambio en el índice de la Bolsa de Valores de Taiwán a través de otro modelo de red neuronal artificial denominado red neuronal probabilística (PNN); este método utiliza datos históricos del índice de la Bolsa de Valores de Taiwán, y los resultados muestran que las estrategias de inversión basadas en él obtienen mayores rendimientos que otras estrategias.

Toro, Molina y Garcés (Molina, Toro, & Garcés, 2005) presentaron un estudio comparativo entre la predicción de precios en bolsa de valores utilizando redes neuronales y neurodifusas. Se muestra una metodología aplicable a la predicción del comportamiento de cualquier tipo de acción, basada en el cálculo preliminar de la correlación entre el precio y otras variables de mercado.

Alonso y García (Alonso & García, 2009) emplean datos de alta frecuencia para encontrar un modelo estadístico que permita disminuir el grado de incertidumbre sobre el comportamiento del futuro inmediato (próximos diez minutos) de operadores del mercado accionario colombiano, específicamente del IGBC; para ello se usan los modelos Garch en

media (Garch-M), que tienen en cuenta los efectos Leverage, día de la semana, hora y día-hora.

George Szpiro (Szpiro, 1997), propone el uso de Algoritmos Genéticos para buscar dependencias entre conjuntos de datos para problemas en Finanzas y Economía. Se usan los algoritmos genéticos para problemas en minería de datos. Éstos se muestran idóneos para detectar e identificar relaciones ocultas en espacios multi-dimensionales que son caracterizados por la existencia de múltiples óptimos. De tal forma, se deja que los Algoritmos Genéticos descubran ecuaciones matemáticas que recrean, o al menos, imitan las fórmulas reales de generación de datos (Santana Luis Quintero Vicente, Diciembre de 2006.).

Eduardo Cruz, Jorge Restrepo y Pedro Medina utilizan redes neuronales para medir el comportamiento económico de la Bolsa de valores Colombiana a través del IGBC que es uno de los tres indicadores que permiten medir el comportamiento bursátil de la bolsa de ese país. A través de la metodología de las redes neuronales se buscaba darle al inversionista la posibilidad de proyectar el comportamiento del mercado bursátil colombiano de forma eficiente (Cruz, Restrepo, & Medina, 2009).

“Visa International” tiene actualmente en operación un sistema de detección de fraudes que usa redes neuronales llamado Cardholder Risk Identification Service (CHRIS). La red neuronal se entrena para detectar actividades fraudulentas, comparando transacciones legales con casos previos de fraude. Una vez que la red ha aprendido ciertos patrones de comportamiento "sospechosos", se le utiliza para alertar a un experto humano cuando una cierta transacción pudiese resultar fraudulenta (Castillo Tapia & Coello, 1999)

El Fuji Bank, en Tokio, usa un sistema que maneja lógica difusa para efectuar transacciones con bonos a corto plazo. Este sistema consta de unas 200 reglas difusas que se basan en estrategias financieras de uso común. Dado que en la vida real este tipo de reglas tienen un tiempo de vida limitado, las reglas difusas del sistema son también actualizadas regularmente por expertos humanos, usando como base datos históricos y su propia experiencia (Castillo Tapia & Coello, 1999).

3.2.2 Los modelos lineales

Los modelos lineales, por la fácil interpretación de sus elementos, tienen una considerable ventaja sobre otros, un ejemplo de estos ha sido la predicción de series de tiempo financieras. Según Alonso y García (Alonso & García, 2009) a principio de la década de 1980 los modelos empleados para modelar la media de un activo eran los Arima o Arimax, pero entrada la década de los noventa con los modelos ARCH y Garch se le da más importancia a la volatilidad que a la media de los datos.

3.2.3 Predicción de la bolsa a través de modelos híbridos

En la predicción de la bolsa de valores muchos autores han realizado combinaciones de métodos o modelos de predicción.

Kim y Han (Him & Han, 2000) expresan que existe una gran cantidad de estudios que incluyen modelos de redes neuronales artificiales (ANN), sin embargo las ANN tienen limitaciones en el aprendizaje de patrones debido a que los datos de la bolsa de valores tienen dimensionalidad compleja y gran ruido. Para solucionar esto se utilizan algoritmos genéticos (GA) que permiten discretizar los datos y así simplificar el proceso de aprendizaje para las ANN y reducir el ruido y los datos redundantes.

Boyacioglu y Avci (Boyacioglu & Avci, 2010) confirman que los rendimientos del mercado bursátil se pueden predecir a través de rendimientos pasados y variables macroeconómicas y financieras. En este trabajo se utiliza el modelo Anfis, el cual combina la teoría de redes neuronales y la lógica difusa utilizando variables que comprenden otros índices de bolsa y variables macroeconómicas, el índice de precios del consumidor, la producción industrial, la tasa de cambio del dólar, etc. El modelo obtenido posee un R-cuadrado de 98%, lo que significa que Anfis predice con un alto grado de certeza el índice de la Bolsa de Valores de Estambul.

Hadavandi, Shavandi y Ghanbari (Hadavandi, Shavandi, & Ghanbari, 2010) presentan un enfoque integrado basado en sistemas difusos genéticos (GFS) y redes neuronales artificiales (RNA) en la construcción de sistemas expertos para el pronóstico del precio de las acciones, logrando precisión en las predicciones realizadas.

Wang, Cheng y Hsu (Wang, Cheng, & Hsu, 2010) incorporan la cadena de Markov en el modelo estocástico difuso y utiliza los datos del índice de la Bolsa de Taiwán. Dicho modelo es capaz de considerar simultáneamente las tasas de cambio, el aumento de las probabilidades y caídas de los índices bursátiles. De los 330 ensayos realizados durante los tres meses de duración del experimento el modelo resultó ser significativamente mejor en 298 de ellos.

3.2.4 Modelos Basados en variables económicas

En esta perspectiva se considera que los movimientos que se dan en la bolsa de valores están estrechamente ligados al comportamiento de la economía del país donde ella opera, entonces hay algunas variables que pueden llegar a incidir en el mercado público de valores: el crecimiento económico, las tasas de interés, el precio del dólar, la producción nacional, los precios del petróleo, los comportamientos de la inflación, el desempleo, entre otras variables de la economía macro de un país.

Cai, Chou y Li (Cai, Chou, & Li, 2009) publican un estudio en el cual se investigan las correlaciones dinámicas entre algunos índices bursátiles internacionales y la inflación. Se utiliza un modelo de series de tiempo autorregresivo denominado (DSTCC-CARR), que revela cómo las correlaciones son significativamente variables en el tiempo y la evolución de estas están relacionadas con las fluctuaciones cíclicas de las tasas de inflación y la volatilidad del mercado.

Majhi, Panda, Majhi y Sahoo (Majhi, Panda, Majhi, & Shao, 2009) proponen predecir índices bursátiles usando Adaptive Bacterial Foraging Optimization (ABFO) y BFO, la estructura utilizada en estos modelos de predicción es un combinador lineal simple. Los nuevos modelos son computacionalmente más eficientes, con predicción más precisa, y muestran una convergencia más rápida en comparación con otros modelos de computación evolutivos tales como el algoritmo genético.

3.3 Motivación

Las acciones de la bolsa se pueden agrupar de diversas formas según el tipo de estudio que se realice o el objetivo que se tenga al efectuar la clasificación. Una forma muy natural podría ser agruparlas de acuerdo a la naturaleza de su negocio, sería lógico plantear por ejemplo que el comportamiento de las acciones que pertenecen a compañías de un mismo tipo sea igual ya que se ven afectados por factores externos similares. Es interesante desde esta perspectiva analizar realmente que tan correlacionado está el comportamiento de las acciones de acuerdo a sus retornos accionarios para empresas del mismo sector económico y determinar si sus comportamientos son similares, sobre todo considerando que algunas pueden pertenecer a más de un sector económico o incluso si el perfil de una compañía puede cambiar a través del tiempo. Siguiendo lo dicho anteriormente, clasificar las acciones utilizando las correlaciones de sus retornos accionarios durante diversos periodos de tiempo usando la herramienta de Mapas Autorganizados de Kohonen, permitirá agrupar y generar una representación gráfica de estas relaciones. De esta forma, se podrán generar métricas que permitan establecer relaciones entre las acciones, generar modelos que faciliten entender el comportamiento de estas y eventualmente desarrollar técnicas de predicción.

4 DEFINICIÓN PROYECTO

4.1 Objetivo del proyecto

Este proyecto, tiene como objetivo el desarrollo de una aplicación utilizando para su creación los Mapas Auto-organizados de Kohonen y las correlaciones de las acciones de la Bolsa de Comercio de Santiago, con el fin de comparar los resultados con lo propuesto en el documento “Estructura Topológica del Mercado Bursátil de Chile, período 2009 – 2014” desarrollado por los profesores Carola Figueroa Flores y Edinson Cornejo Saavedra.

4.2 Objetivos específicos

- Calcular el coeficiente de correlación entre retornos accionarios, a partir de datos correspondientes al período 2009-2014, para las acciones que se transan en la Bolsa de Comercio de Santiago. Representar gráficamente dicho mercado durante el período 2009-2014.
- Analizar el mercado bursátil chileno, a partir de la estructura obtenida por medio del mapa auto-organizado de Kohonen del período 2009-2014, para distinguir relaciones de dependencia mutua, clústeres y acciones líderes.
- Implementar una red neuronal que permita clasificar las acciones de la bolsa de comercio de Santiago.
- Obtener conclusiones a partir de los resultados obtenidos por la red neuronal implementada que permitan comparar con lo propuesto en el documento “Estructura Topológica del Mercado Bursátil de Chile, período 2009 – 2014”.
- Obtener una representación gráfica o visual (a color) del mapa de Kohonen resultante.

4.3 Definiciones, Siglas y Abreviaciones

- SOM = Sigla en inglés “Self-Organizing Map” lo que a su traducción en español significa “Mapas Auto-Organizados”.
- RNA = Redes Neuronales Artificiales.
- AGs= Sigla en inglés “Algorithm Generics” lo que a su traducción en español significa “Algoritmos Genéricos”.
- ART= Sigla en inglés “Adaptative Resonance Theory” lo que a su traducción en español significa “Teoría de Resonancia Adaptativa”.
- KDD= Sigla en inglés “Knowledge Discovery in Databases” lo que a su traducción en español significa “Descubrimiento de conocimiento en Base de Datos”.

5 DESCRIPCIÓN DE LA METODOLOGÍA Y MAPAS AUTO-ORGANIZADOS DE KOHONEN (SOM)

5.1 Método

Durante el desarrollo del proyecto se determinó dividir las actividades principalmente en dos etapas. La primera está relacionada con el proceso de aprendizaje e investigación y captura de la información, la que tiene como objetivo principal, definir y determinar el marco teórico del problema y así manejar de mejor manera los conceptos relacionados con las Redes Neuronales, Algoritmo de Kohonen, SOM, procesamiento de información financiera y/o bursátil, etc. Como resultado de esta etapa se entregó un informe de marco teórico donde se sintetiza la información recolectada durante la etapa de investigación preliminar

La segunda etapa se realizó el desarrollo y construcción de la aplicación propiamente tal, la cual para su construcción, se utilizó la metodología iterativa incremental. Durante esta fase se crearon diferentes versiones del producto de manera de ir perfeccionando y mejorando la versión anterior y/o agregar nuevas funcionalidades en cada incremento.

Las actividades más frecuentes durante esta etapa fueron:

- Determinar uno o varios objetivos a cumplir durante la iteración como:
 - Corrección de errores del código.
 - Agregar funcionalidades que el cliente desee.
- Realizar una investigación (de ser necesaria) para determinar herramientas, librerías o técnicas que sean las más adecuadas para lograr alcanzar el objetivo.
- Codificar.
- Revisión de lo realizado.

5.2 Herramientas y datos utilizados

Para la realización de la aplicación se utilizó el lenguaje de programación Java. Para el diseño de la interfaz de usuario se usó principalmente sus librerías de Swing y para graficar la matriz del mapa auto-organizados de Kohonen las clases Graphics2D y BufferedImage de la librería AWT.

Los datos utilizados como fuente de alimentación para la red neuronal son suministrados mediante archivos de tipo Excel en formato “xls” los que son leídos utilizando la API de Excel de Java. Estos archivos contienen las correlaciones de las acciones de la bolsa de Santiago para los periodos 2009, 2010, 2011, 2012, 2013 y 2014 (Microsoft Excel, Versión 14.0.7116.5000), los cuales fueron calculados en base a la información obtenida utilizando la herramienta Económica (Económica, 1986) que permite recolectar información de los movimientos accionarios de las bolsas de diferentes países para su posterior análisis.

5.3 Mapas auto-organizados de Kohonen (SOM)

En esta sección se describirá el algoritmo de Kohonen utilizado en el desarrollo del proyecto y además se detallará como fue implementado en la creación de la aplicación.

5.3.1 Descripción del algoritmo de Kohonen

Los SOM, fueron inventados en 1982 por Teuvo Kohonen, profesor de la Academia de Finlandia. Los mapas auto organizados, también llamados redes de Kohonen, son un tipo de red neuronal no supervisada, competitiva, distribuida de forma regular en una rejilla de, normalmente, dos dimensiones. La técnica de Kohonen crea una red que almacena información de forma que las relaciones topológicas del conjunto de entrenamiento se mantienen.

Un ejemplo habitual que se usa para mostrar cómo funcionan los SOM se basa en la proyección de colores (asociados a vectores 3D formados a partir de, por ejemplo, sus 3 componentes RGB) en un espacio 2D.

En la imagen (Figura 8) muestra un ejemplo de SOM en donde se grafica las diferentes regiones del mapa utilizando como datos de entrenamiento los colores que se encuentran a la derecha del mapa.

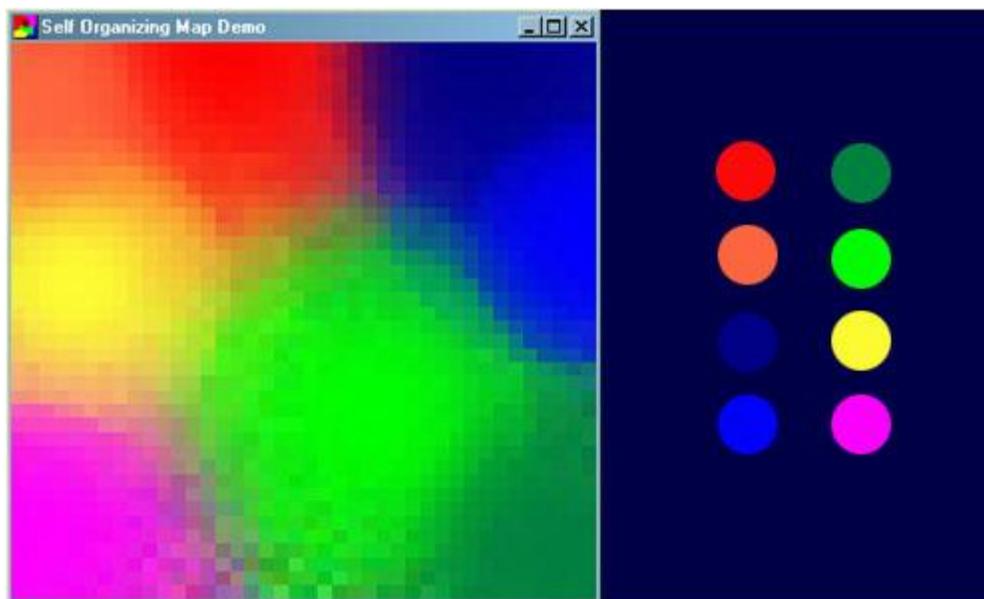


Figura 8. Ejemplo Mapa Auto-Organizado (Caparrini, 2014)

Uno de los aspectos más interesantes de los SOM es que aprenden a clasificar sin supervisión, lo que implica que no necesita un objetivo al cual aproximar, sino que genera la distribución a partir de la similitud entre los vectores.

5.3.1.1 Arquitectura en red

En general, el algoritmo SOM considera una arquitectura en 2 capas: por una parte se tiene una red de nodos de aprendizaje o matriz de salida junto con un conjunto de vectores de entrenamiento, de forma que todos los elementos de la primera capa están conectados con todos los elementos de la segunda capa en la Figura 9 , se puede ver gráficamente como sería la arquitectura de la red donde los vectores de entrada son de dos dimensiones y la matriz de entrenamiento es de 4 por 4 (la red de aprendizaje viene representada por nodos rojos, y los vectores de entrenamiento vienen representados en verde).

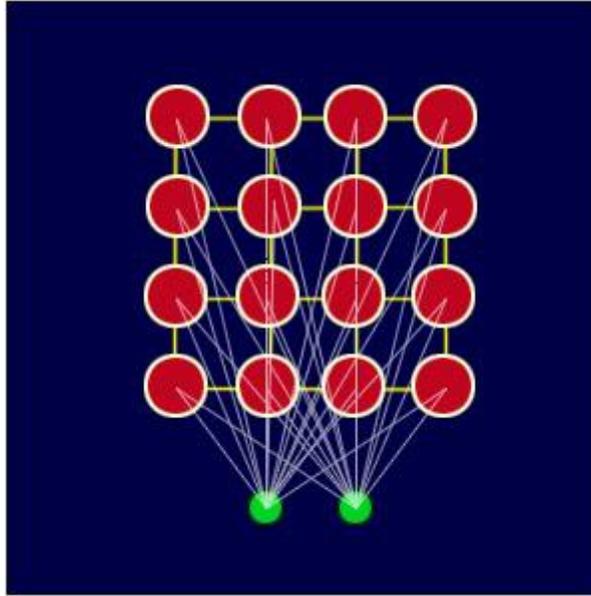


Figura 9. Arquitectura Red SOM (Caparrini, 2014)

Cada nodo tiene un posición topológica específica (que suele estar asociada a las coordenadas (x,y)) y contiene un vector de pesos de la misma dimensión que los vectores de entrenamiento. Es decir, si cada vector, V , de entrenamiento tiene dimensión n , cada nodo tendrá un vector de pesos, W , de dimensión n .

5.3.1.2 Aprendizaje

El aprendizaje en el algoritmo es de tipo OFF LINE, por lo que se distingue una etapa de aprendizaje y otra de funcionamiento. También utiliza un aprendizaje no supervisado de tipo competitivo. Sólo una neurona de la capa de salida se activa ante la entrada, ajustándose los pesos de las conexiones en función de la neurona que ha resultado vencedora.

Durante la etapa de entrenamiento, se presenta a la red un conjunto de informaciones de entrada para que ésta establezca, en función de la semejanza entre los datos, las diferentes categorías (una por neurona de salida) que servirán durante la fase de funcionamiento para realizar clasificaciones de nuevos datos que se presenten a la red. Los valores finales de los pesos de las conexiones entre cada neurona de la capa de salida con las de entrada se corresponderán con los valores de los componentes del vector de aprendizaje que consigue activar la neurona correspondiente.

El aprendizaje no concluye después de presentarle una vez todos los patrones de entrada, sino que habrá que repetir el proceso varias veces para refinar el mapa topológico de salida, de tal forma que cuantas más veces se presenten los datos, tanto más se reducirán las zonas de neuronas que se deben activar ante entradas parecidas, consiguiendo que la red pueda realizar una clasificación más selectiva.

5.3.1.3 Algoritmo

Sea N el número de neuronas de entrada y M el número de neuronas de salida.

1. Inicializar los pesos w_{ji} de la matriz de salida con valores aleatorios Normalmente, vectores en $[0,1]^n$
2. Presentar a la red una información de entrada en forma de vector $E_k = (e_1, \dots, e_N)$, cuyas componentes e_i sean valores continuos.
3. Determinar la neurona vencedora o BMU (Best Matching Unit) de la capa de salida: será aquella cuyo vector de pesos, W_j , sea el más parecido a la información de entrada E_k . Recordando que las componentes de W_j son los valores de los pesos de las conexiones entre esa neurona, j , y cada una de las neuronas de entrada. Para ello se calcula la distancia euclidiana entre los vectores E_k y W_j para cada neurona de salida. La expresión matemática sería la siguiente:

$$d_j = \sum_{i=1}^n (e_i - w_{ji})^2 \text{ para } 1 \leq j \leq M$$

Por motivos de eficiencia no será calculada la raíz cuadrada.

4. Se calcula el radio del entorno de BMU. Este radio comenzará siendo grande (como para cubrir la red completa) y se va reduciendo en cada iteración.

La fórmula que establece el radio en función de la iteración (que hace que vaya disminuyendo, pero no linealmente) es:

$$r(t) = r_0 e^{-\frac{t}{\lambda}}$$

Donde r_0 es el radio inicial (habitualmente, el radio de la red, es decir, uno suficiente para cubrir en el primer paso todos los nodos) y λ una constante que permite hacer que el radio sea muy pequeño cuando llegamos a la iteración máxima:

$$\lambda = \frac{\text{Tiempo_de_Entrenamiento}}{\ln r_0}$$

La Figura 10 muestra el efecto de ir reduciendo paulatinamente el radio del entorno, donde se marcan los nodos que se verían afectados si el nodo BMU es el nodo amarillo

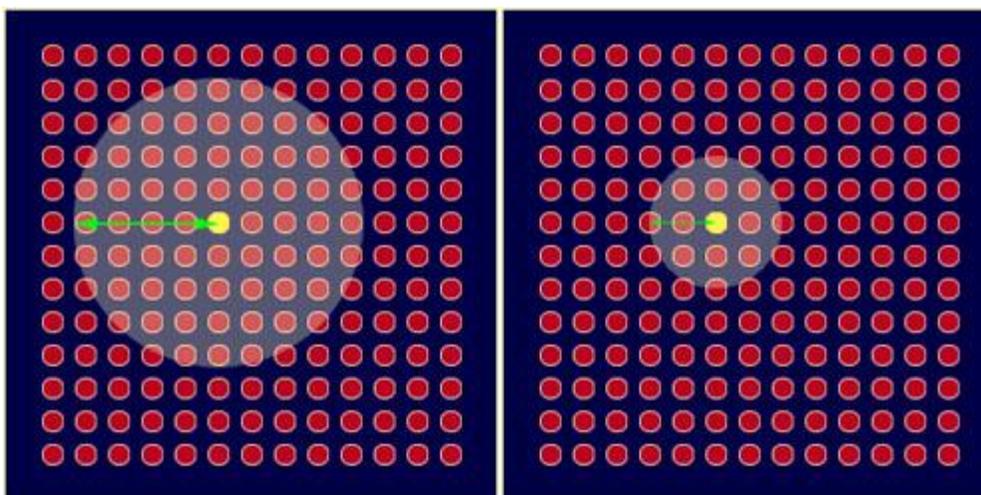


Figura 10. Reducción Radio de Entorno de la BMU (Caparrini, 2014)

5. Una vez localizada la neurona vencedora, se actualizan los pesos de sus conexiones de entrada y también los de las neuronas vecinas (las que pertenecen a su zona de vecindad). Lo que se consigue es asociar la información de entrada con una cierta zona de la capa de salida.

Para aproximar los pesos (W) de los nodos de la matriz de aprendizaje al vector de entrenamiento (V) se puede usar la siguiente fórmula:

$$W(t + 1) = W(t) + L(t)(V(t) - W(t))$$

6. El factor $L(t)$ se denomina tasa de aprendizaje, y permite aproximar W a V con el paso del tiempo. Como se quiere que su valor también disminuya a medida que el tiempo pasa, se puede usar una expresión similar a la del radio:

$$L(t) = L_0 e^{-\frac{t}{\lambda}}$$

El valor de L_0 se ajusta experimentalmente. Con el fin de que el efecto de aprendizaje sea más notable en las cercanías del BMU, se puede añadir un factor más al producto anterior, que hace que los nodos más cercanos al BMU se vean más afectados:

$$W(t + 1) = W(t) + D(t)L(t)(V(t) - W(t))$$

Por ejemplo, haciendo que $D(t)$ siga una gaussiana de la forma:

$$D(t) = e^{-\frac{d^2}{2r(t)^2}}$$

Donde d es la distancia del nodo que se está ajustando al BMU

7. El proceso se debe repetir, volviendo a presentar todo el juego de patrones de aprendizaje E_1, E_2, \dots , un mínimo de 500 veces ($t = 500$).
8. Después del aprendizaje se pasaría a la fase de funcionamiento de la red, durante la cual se pueden presentar nuevos puntos para la que la red clasifique de acuerdo a lo aprendido.

5.3.1.4 Ejemplo de Aplicación

Los SOM se usan habitualmente para proporcionar ayudas visuales, ya que permiten mostrar relaciones entre grandes cantidades de datos y que precisarían muchas más dimensiones (algo inviable para el ser humano) para ser mostradas adecuadamente. Con el fin de trabajar con una topología en los nodos que refleje un mayor número de conexiones entre ellos, pero sea realista desde un punto de vista 2D, es habitual trabajar con un teselado hexagonal del plano, identificando los hexágonos con los nodos de la red.

Por ejemplo en la Figura 11 se pueden ver los países organizados según su nivel de pobreza.

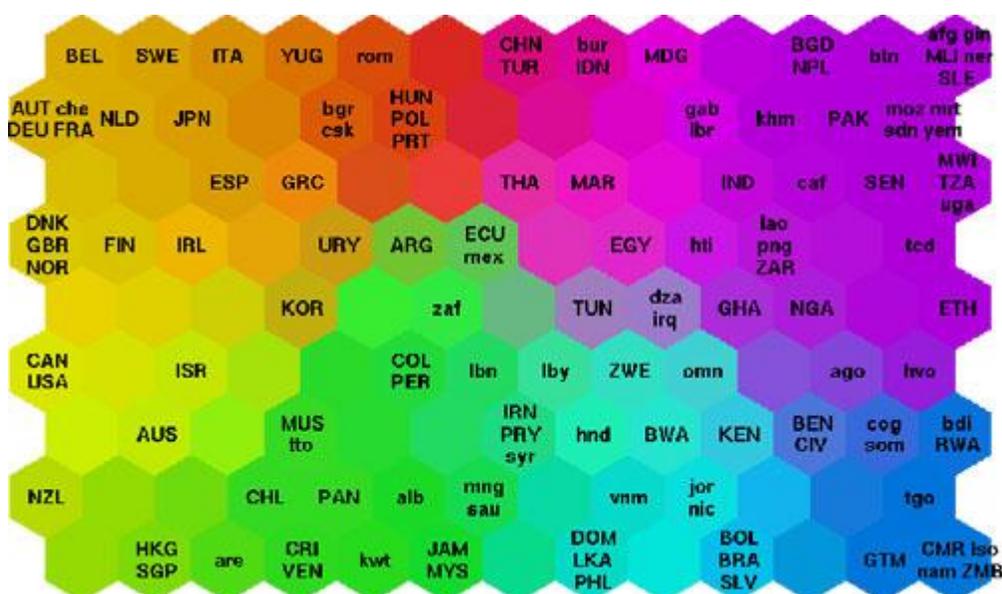


Figura 11. Ejemplo SOM según Nivel de Pobreza (Caparrini, 2014)

Junto a la representación anterior en la Figura 12 se puede volver a proyectar los países en un mapa estándar, de forma que visualmente se puede interpretar simultáneamente la información geográfica con la procedente de los datos anteriores:

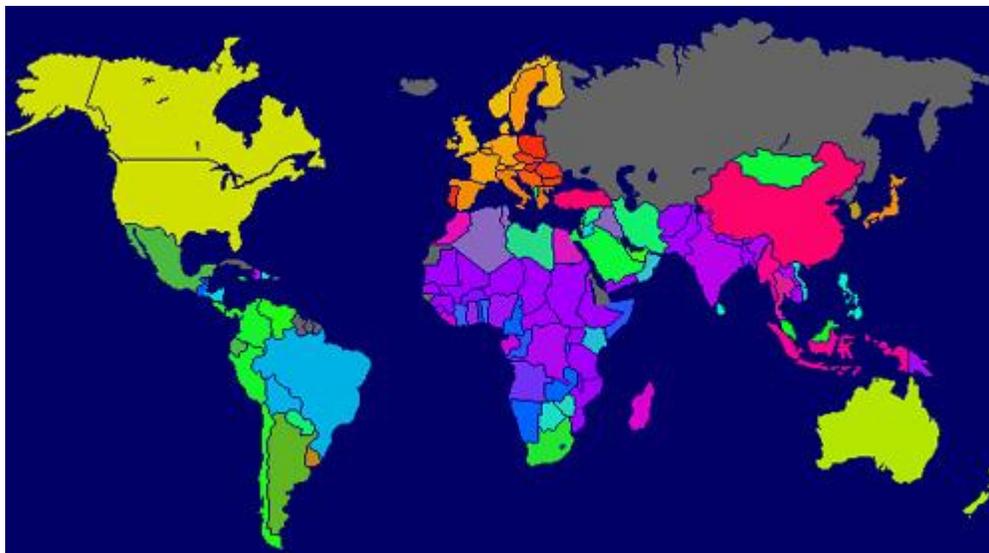


Figura 12. Ejemplo SOM Proyección de países (Caparrini, 2014)

Según los diversos factores que se usan para medir la calidad de vida de los países, podemos usar SOM para representar las agrupaciones que forman los diversos países en una red 2D.

5.3.2 Implementación Algoritmo de Kohonen

La implementación de este algoritmo se llevó a cabo utilizando el lenguaje de programación Java. Los datos que se usaron para la ejecución del algoritmo fueron las correlaciones de las acciones de la bolsa de Santiago para los periodos 2009, 2010, 2011, 2012, 2013 y 2014.

5.3.2.1 Rubros de las Acciones Bursátiles

Las acciones utilizadas fueron agrupadas en conjuntos según su rubro. De acuerdo a la página web de la Bolsa de Comercio de Santiago (Santiago, 2014), las acciones bursátiles se clasifican según los siguientes rubros:

- Comerciales y Distribuidoras
- Inversiones e Inmobiliarias
- Servicios Públicos
- Agropecuarias y Forestales
- Alimenticias y Bebidas
- Bancarias y Financieras
- Comerciales y Distribuidoras
- Industriales
- Marítimas y Navieras
- Mineras
- Pesqueras
- Productos Diversos
- Servicios Varios
- Productos Químicos

Los datos que se utilizaron para el entrenamiento del mapa SOM fueron las correlaciones de diez acciones las que fueron seleccionadas por representar cada una a uno de los rubros que agrupan la mayor cantidad de acciones de la bolsa de comercio de Santiago con el fin de lograr una muestra representativa de las acciones que posteriormente se clasificaran.

Las acciones utilizadas con su respectivo rubro se podrán observar en el ítem de Anexos desde el Anexo 8.1 al Anexo 8.6.

En la etapa de funcionamiento del algoritmo se utilizaron las correlaciones de las acciones de cada periodo con el objetivo de clasificar estas acciones de acuerdo a lo que se aprendió en la etapa de entrenamiento.

5.3.2.2 Implementación del algoritmo

Como se describió en el capítulo 5.3.1.3 el Algoritmo de Kohonen cuenta con una serie de actividades o pasos a realizar los cuales se deben ver reflejados en los procesos internos de la aplicación, estas actividades se ven graficadas en la Figura 13, que representa un diagrama de flujo general de los procesos del sistema SOM.

Las principales actividades realizadas en la aplicación son:

- Inicializar la matriz de nodos.
- Leer datos de entrenamiento.
- Realizar entrenamiento.
- Leer datos a clasificar
- Clasificar datos.

El entrenamiento y la clasificación de datos serán detalladas más adelante.

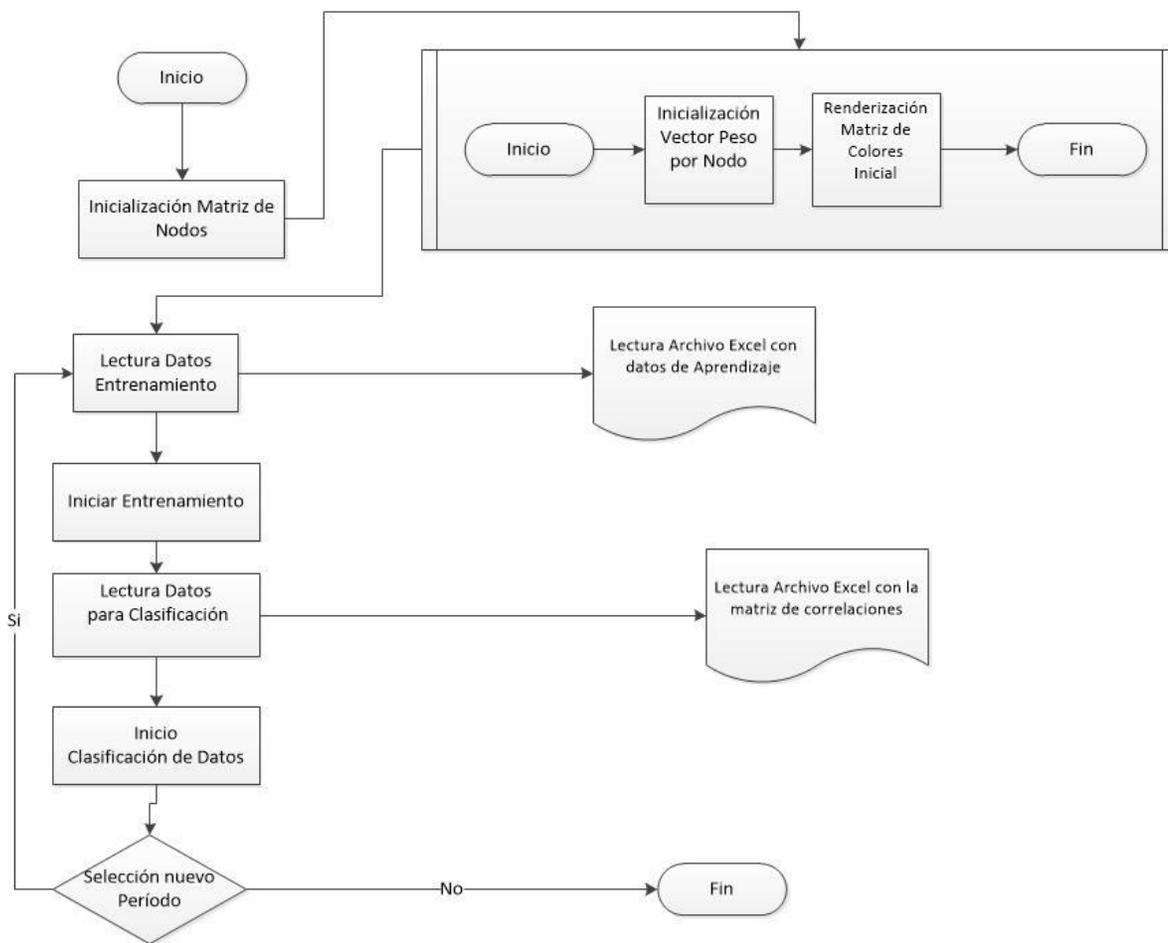


Figura 13. Diagrama General SOM

El entrenamiento de la matriz cuenta con una serie de pasos internos que se detallan en la Figura 14 que representa un diagrama de actividades del subproceso de entrenamiento.

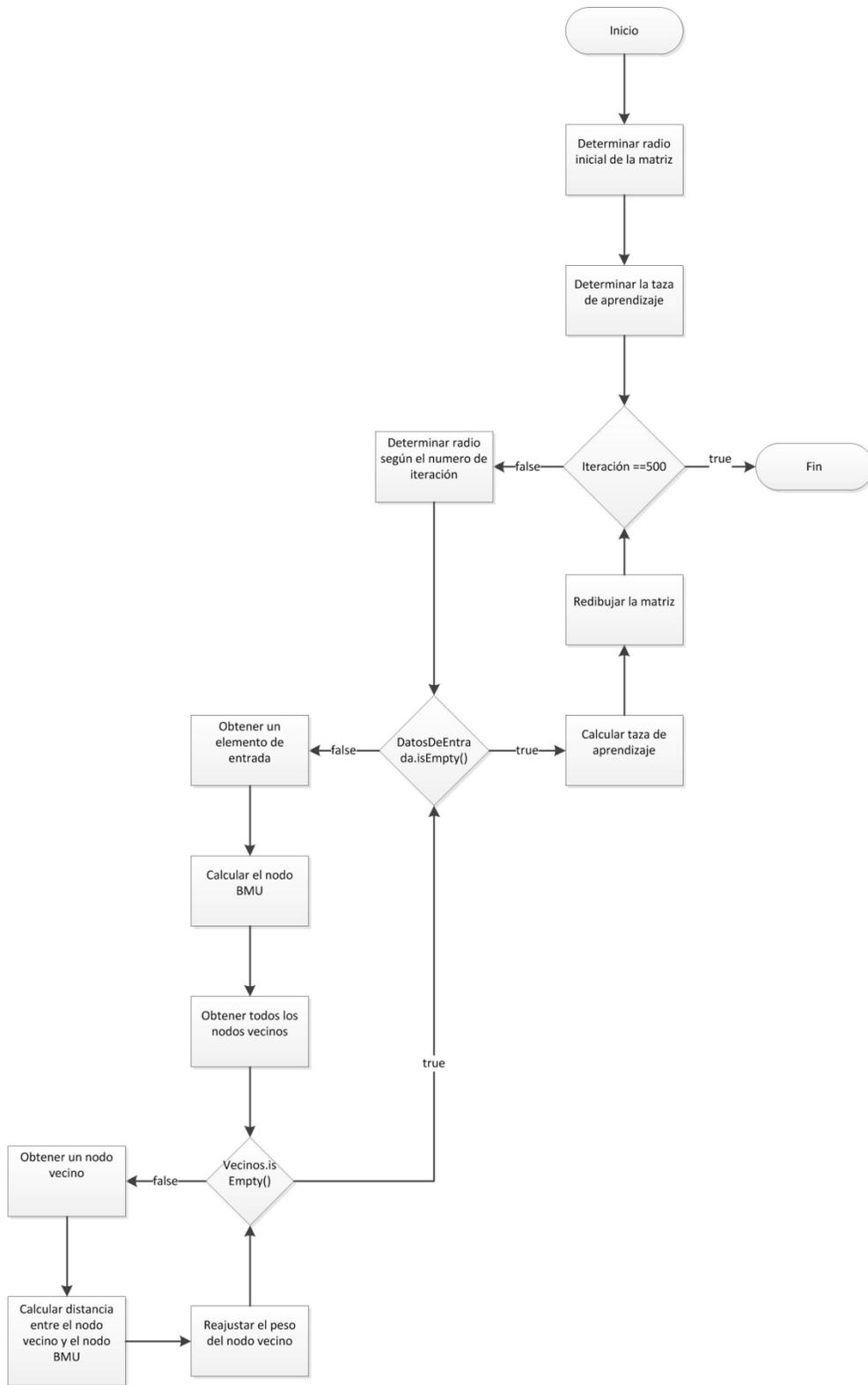


Figura 14. Subproceso Entrenamiento SOM

A continuación se detallarán las actividades realizadas durante el aprendizaje explicando sus funciones e ilustrando secciones del código fuente de la aplicación. Cabe señalar que muchas de la formulas en la que se basan los códigos expuestos a continuación fueron explicadas en el capítulo 5.3.1.

Paso 1: Determinar radio inicial de la matriz:

Primeramente se determina el radio de la matriz en base al tamaño de sus columnas y/o filas considerando el tamaño mayor de estas dos variables y dividiendo ese número en dos. Desde la perspectiva del código seria lo siguiente:

```
radio = Math.max(col, fil) / 2;
```

Paso 2: Definir tasa de aprendizaje:

La tasa de aprendizaje se define inicialmente como 0.05

Paso 3: Definir un radio Según el número de iteración r(t):

El tercer paso está definido dentro de un ciclo que se ejecutará por cada una de las iteraciones de la etapa de aprendizaje, el número de iteraciones está definido por defecto como 500. En este paso se define un radio el cual va variando respecto de en qué iteración se encuentre la ejecución del algoritmo dentro de la etapa de aprendizaje, de tal forma, de conseguir que al principio, cuando la matriz tiene pocos o nulos conocimientos el radio sea muy grande, prácticamente del tamaño del radio total de la matriz (obtenido en paso 1), y en las iteraciones finales cuando la matriz ya ha ingresado mucha información a su sistema el radio sea cercano a 1.

Desde la perspectiva de la programación la función es la siguiente:

```
r(t)=radio * Math.exp(-iteracion / constante);
```

Donde la constante es λ y se define de la siguiente forma en el código:

```
constante = NUM_ITERACIONES / Math.log(radio);
```

Paso 4: Obtener un elemento de entrada:

En esta etapa por cada elemento de entrada del vector de entrenamiento lo primero que se realiza es obtener el correspondiente elemento para ingresarlo a la matriz.

Paso 5: Calcular el nodo BMU:

Una vez que se obtiene un elemento de los vectores de entrada se calcula el nodo BMU dentro de la matriz SOM lo que en el código se realiza de la siguiente forma:

```
public Nodo getBMU(Vector inputVector) {

    Nodo bmu = matriz[0][0];
    double bestDist = distancia(bmu.getVector(), inputVector);
    double dist;

    for (int x = 0; x < columnas; x++) {
        for (int y = 0; y < filas; y++) {
            dist = distancia(matriz[x][y].getVector(), inputVector);
            if (dist < bestDist) {
                bmu = matriz[x][y];
                bestDist = dist;
            }
        }
    }

    return bmu;
}
```

Básicamente se calcula la distancia del vector entrante por cada uno de los nodos de la matriz y se va guardando el nodo que tenga el mejor valor y una vez terminado de recorrer todos los elementos de la matriz este nodo es retornado.

La función para calcular la distancia se define de la siguiente forma:

```

public double distancia(Vector<Double> v1, Vector<Double> v2) {
    double sumatoria = 0, temp;
    for (int i = 0; i < v2.size(); i++) {
        temp = v2.elementAt(i) - v1.elementAt(i);
        temp *= temp;
        sumatoria += temp;
    }

    return sumatoria;
}

```

Esta función recibe dos vectores y en base a estos se calcula la distancia Euclidiana entre ellos (no se calcula la raíz cuadrada por temas de eficiencia como se mencionó en el capítulo 5.3.1).

Paso 6: Obtener los vecinos:

Una vez obtenido el nodo BMU se obtienen los vecinos de éste para su posterior ajuste de valores. El vecindario corresponde a todos los nodos de la matriz que estén dentro del radio el cual es determinado en el paso 3 tomando como punto central el nodo BMU obtenido anteriormente.

Desde la perspectiva de la programación el código es el siguiente:

```

xstart = (int) (bmu.getX() - radioPorIteracion);
ystart = (int) (bmu.getY() - radioPorIteracion);
xend = (int) (xstart + (radioPorIteracion * 2));
yend = (int) (ystart + (radioPorIteracion * 2));
if (xend > col) {
    xend = col;
}
if (xstart < 0) {
    xstart = 0;
}
if (yend > fil) {
    yend = fil;
}
if (ystart < 0) {
    ystart = 0;
}

```

Lo que se busca lograr con esta sección de código es poder simular una sub matriz dentro de la matriz SOM la cual está compuesta por todos los vecinos del nodo BMU.

Para esto se obtiene un punto de inicio y fin tanto del eje X como del eje Y dentro de la matriz SOM. Los puntos de inicio se determinan en base a la posición del BMU restándole el radio por iteración (calculado en el paso 3) en ambos ejes.

Los puntos de fin se calculan con el punto de inicio sumando el diámetro para ambos ejes. Luego se ajustan los valores de inicio y fin para ambos ejes para que no sean mayor o menor que el ancho y el largo de la matriz SOM.

Finalmente se puede recorrer la sub matriz utilizando los puntos de inicio y fin de la siguiente manera:

```
for (int x = xstart; x < xend; x++) {
    for (int y = ystart; y < yend; y++) {
        // Realizar acción
    }
}
```

Paso 7: Obtener nodo vecino:

Por cada nodo vecino del BMU (obtenidos en el paso anterior) se obtiene uno de ellos para realizar el reajuste de pesos.

Paso 8: Calcular distancia entre el nodo vecino y el nodo BMU:

Para poder ajustar los pesos de los nodos vecinos se debe obtener la distancia de este con el BMU. En el código esto se realiza con la siguiente función:

```
dist = bmu.distanceTo(vecino);
```

La cual calcula la distancia Euclidiana entre ambos nodos de la siguiente forma:

```
public double distanceTo(Nodo n2) {
    int xleg, yleg;
    xleg = getX() - n2.getX();
    xleg *= xleg;
    yleg = getY() - n2.getY();
    yleg *= yleg;
    return xleg + yleg;
}
```

Paso 9: Reajustar pesos:

Se ajustan los pesos del vector con la siguiente función:

```
public void ajusteDePesos(Vector<Double> v, double L, double D) {
    double W, V;
    for (int i = 0; i < dataVector.size(); i++) {
        W = ((Double)dataVector.elementAt(i)).doubleValue();
        V = v.elementAt(i);
        W += D * L * (V - W);
        dataVector.setElementAt(new Double(W), i);
    }
}
```

Donde el vector v es el vector de entrenamiento, L es la tasa de aprendizaje y D es obtenida de la siguiente forma:

```
D= Math.exp(-(distSq) / (2 * (radioPorIter * radioPorIter)));
```

Donde $distSq$ es el valor obtenido en el **Paso 8**.

La fórmula para realizar el ajuste de peso se explicó en la sección **5.3.1**.

Paso 10: Calcular tasa de aprendizaje:

Una vez se terminó con una de las iteraciones de la etapa de aprendizaje se vuelve a calcular la tasa de aprendizaje en el código esto se realiza de la siguiente manera:

```
learningRate=TAZA_APRENDIZAJE_INICIAL*
    Math.exp((double)iteration/NUM_ITERACIONES);
```

Paso 11: Redibujar matriz

Finalmente se redibuja la matriz en la interfaz de usuario.

El proceso de clasificación de datos se puede ver graficada en la **Figura 15**, la cual muestra un diagrama de actividades con sus respectivas tareas.

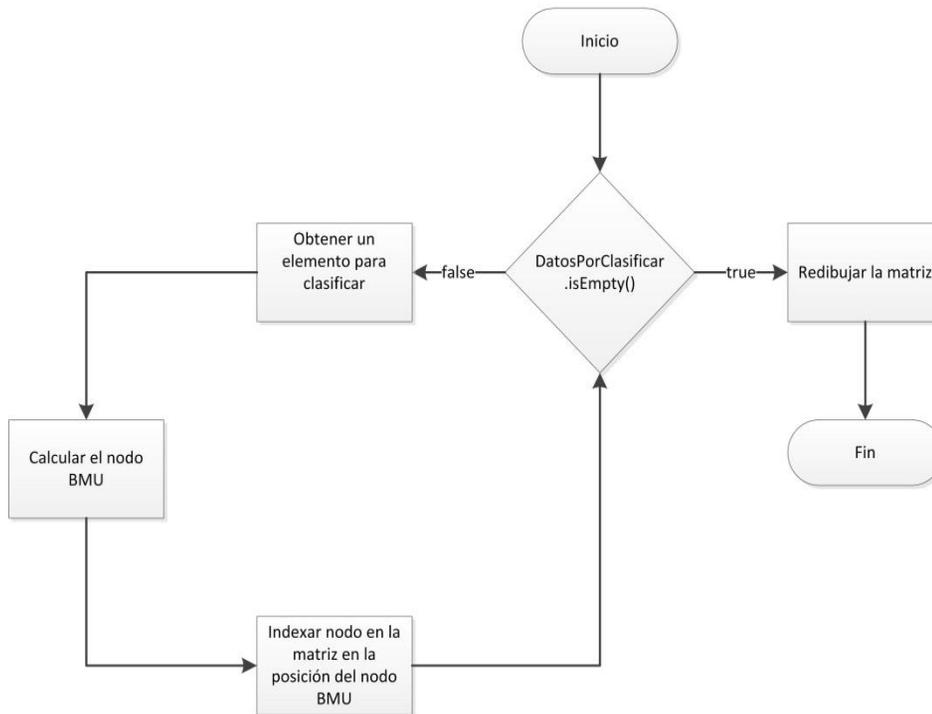


Figura 15. Subproceso Clasificación SOM

Paso 1: Obtener elemento a clasificar:

En el proceso de clasificación se realiza un ciclo por cada vector a clasificar primeramente se obtiene un elemento de la lista.

Paso 2: Calcular nodo BMU

El método para calcular el nodo BMU es el mismo que se utilizó en el proceso de aprendizaje.

Paso 3: Indexar nodo en la matriz en la posición del nodo BMU

En esta etapa una vez encontrado el nodo BMU se indexa el vector en la posición del BMU, es decir, que se asignan al BMU los valores del vector a clasificar. Desde la perspectiva de la programación la manera de realizar esto es la siguiente:

```

public void indexar(Vector v) {
    for (int i = 0; i < dataVector.size(); i++) {
        dataVector.setElementAt(v.elementAt(i), i);
    }
}
    
```

Donde el vector v es el vector a clasificar y el `dataVector` es el vector del nodo BMU de la matriz.

Paso 4: Redibujar matriz:

Una vez clasificado todos los vectores de entradas se redibuja la matriz en la interfaz de usuario.

5.3.3 Diagrama de casos de uso

El diagrama de casos de uso de la aplicación se puede ver en la imagen de la Figura 16.

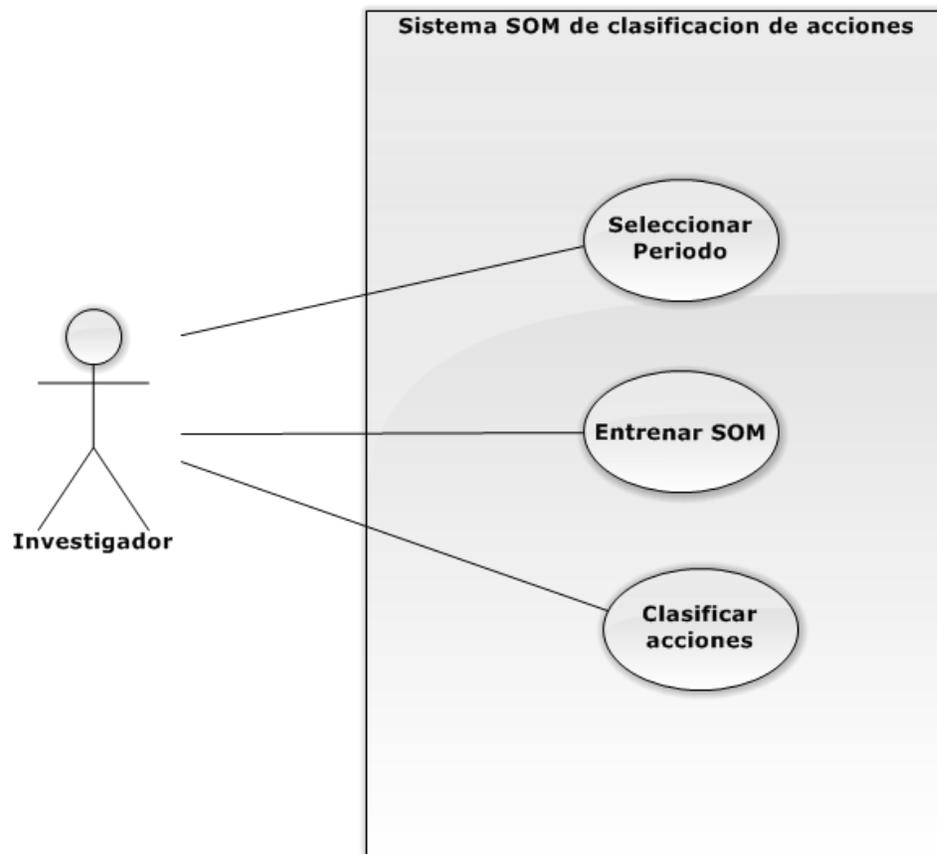


Figura 16. Diagrama Casos de Uso SOM

5.3.4 Especificación Casos de Uso

Caso de uso	Seleccionar periodo
Descripción	Permite al usuario seleccionar un periodo con el cual trabajar.
Pre-Condiciones	No existe.
Actores	Investigador.
Flujo normal	<ol style="list-style-type: none"> 1. El investigador inicia la ejecución. 2. El sistema muestra una lista con los periodos seleccionables. 3. El usuario selecciona un periodo y presiona aceptar.
Post condiciones	Los datos de entrenamiento y la matriz de correlaciones del periodo seleccionado deben estar correctamente cargados en el sistema.

Tabla 1. Caso de Uso Seleccionar Período

Caso de uso	Entrenar SOM
Descripción	El sistema entrenará la red neuronal utilizando los datos de entrenamiento.
Pre-Condiciones	Se debe seleccionar previamente un periodo con el cual trabajar.
Actores	Investigador.
Flujo normal	<ol style="list-style-type: none"> 1. El investigador inicia el entrenamiento. 2. La red neuronal realiza su etapa de entrenamiento o aprendizaje. 3. El sistema muestra de forma gráfica los resultados de su aprendizaje.
Post condiciones	La red neuronal estará preparada para clasificar datos en base a lo aprendido en su proceso de entrenamiento.

Tabla 2. Caso de Uso Entrenar SOM

Caso de uso	Clasificar acciones
Descripción	El sistema ordenará o clasificará las acciones basándose en lo aprendido en el entrenamiento.
Pre-Condiciones	El sistema debe haber sido entrenado previamente.
Actores	Investigador.
Flujo normal	<ol style="list-style-type: none"> 1. El investigador inicia la clasificación. 2. El sistema realiza la clasificación. 3. El sistema muestra de forma gráfica como han quedado ordenadas las acciones.
Post condiciones	El sistema estará listo para volver a iniciar su proceso de funcionamiento.

Tabla 3. Caso de Uso Clasificar Acciones

5.3.5 Diagrama de Clases SOM

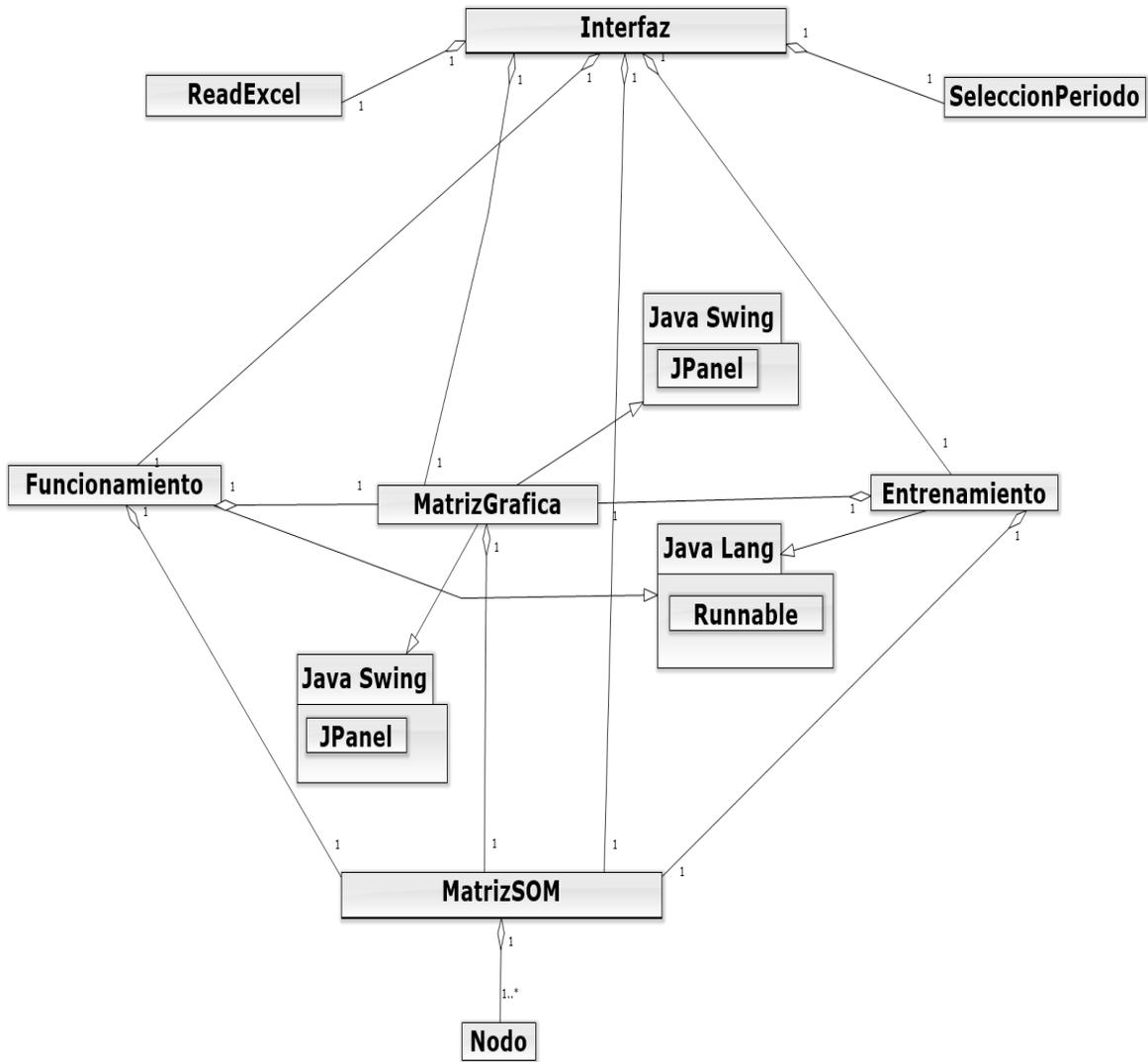


Figura 17. Diagrama de Clases SOM

En la Figura 17, se muestra el diagrama de clases resultante del software SOM. Además se observa las diferentes importaciones hacia otras clases utilizadas. Los atributos y métodos de las diversas clases se encuentran en la Tabla 4.

Nombre de la clase	Nombre Atributos	Constructor y Métodos
Nodo	- dataVector: Vector - x: int - y: int - nombre: String - isBMU: boolean = false	«Constructor» + Nodo (tamaño : int) + getNombre (): String + setNombre (nombre : String): void + getX (): int + setX (x : int): void + getY (): int + setY (y : int): void + distanceTo (n2 : Nodo): double + ajusteDePesos (v : Vector<Double>, L : double, D : double): void + ajusteDePesos2 (v : Vector, L : double, D : double): void + indexar (v : Vector): void + isIsBMU (): boolean + setBMU (isBMU : boolean): void
MatrizSOM	- columnas: int - filas: int - matriz: Nodo [][] - v: Vector<Nodo>	«Constructor» + MatrizSOM (columnas : int, filas : int, tamañoVector : int) + getBMU (inputVector : Vector): Nodo + getBMU2 (inputVector : Vector): Nodo + distancia (v1 : Vector<Double>, v2 : Vector<Double >): double + getNode (x : int, y : int): Nodo + getColumnas (): int + setColumnas (): void + getFilas (): int + setFilas (): int + getMatriz (): Nodo [][] + setMatriz (matriz : Nodo = [][]): void

Nombre de la clase	Nombre Atributos	Constructor y Métodos
MatrizGrafica	<ul style="list-style-type: none"> - img: BufferedImage = null - arialFont: Font - matriz: MatrizSOM - ready: boolean = false - entrenar: boolean = true 	<p>«Constructor» + MatrizGrafica () + paint (Graphics : g): void + ingresarMatriz (matriz : MatrizSOM): void + render (matriz : MatrizSOM): void + getImage (): BufferedImage + panelMouseMoved (evt : MouseEvent): void + isEntrenar (): boolean + setEntrenar (entrenar : boolean): void</p>
Funcionamiento	<ul style="list-style-type: none"> - <u>TAZA APRENDIZAJE INICIAL: double = 0.7</u> - <u>NUM ITERACIONES: int = 500</u> - <u>running: boolean</u> - radio: double - entrada: Vector - runner: Thread - renderer: MatrizGrafica - matriz: MatrizSOM - nombreAcciones: Vector - archivoNombres: String 	<p>«Constructor»+Funcionamiento (archivoNombresAcciones : String) + getD (distSq : double, radio : double): double + setDatos (matriz : MatrizSOM, entrada : Vector, latticeRender : MatrizGrafica): void + setDatos (matriz : MatrizSOM, entrada : Vector): void + start (): void + run (): void + stop (): void + setNombreArchivo (arch : String): void + llenarVectorNombres (v : Vector, arch : String): void + nombreArchivo (): String</p>
Entrenamiento	<ul style="list-style-type: none"> - <u>TAZA APRENDIZAJE INICIAL: double = 0.7</u> - <u>NUM ITERACIONES: int = 500</u> - <u>running: boolean</u> - radio: double - entrada: Vector - runner: Thread - renderer: MatrizGrafica - matriz: MatrizSOM - constante: double - running: boolean 	<p>«Constructor» + Entrenamiento () + radioSegunIteracion (iteracion : double): double + getD (distSq : double, radio : double): double + setDatos (matriz : MatrizSOM, entrada : Vector, latticeRender : MatrizGrafica): void + setDatos (matriz : MatrizSOM, entrada : Vector): void + start (): void + run (): void + stop (): void</p>

Nombre de la clase	Nombre Atributos	Constructor y Métodos
ReadExcel	- <u>dir</u> : String - <u>URL</u> : String	«Constructor» + ReadExcel () + leerDatos (nombreArchivo : String): void
Interfaz	- columnas: int = 40 - filas: int = 40 - datosEntrada: Vector - datosAClasificar: Vector - tamVector: int - matriz: MatrizSOM - trainer: Entrenamiento - func: Funcionamiento - matrizGrafica: MatrizGrafica - datosEntrenamiento: Vector - datosAClasificar: Vector - tamVector: int - seleccionPeriodo: SeleccionPeriodo - <u>periodo</u> : String - <u>matrizEntren</u> : String - <u>matrizAClasificar</u> : String - <u>archivoNombreAcciones</u> : String	«Constructor» + Interfaz (seleccionperiodo : SeleccionPeriodo) + JBEntrenarActionPerformed (evt : ActionEvent): void + JBFuncionarActionPerformed (evt : ActionEvent): void + cambiarPerdiodoActionPerformed (evt : ActionEvent): void + iniciar (): void + asignarColor (): void + getPeriodo (): String + setPerdiodo (perdiodo : String): void + getMatrizEntren (): String + setMatrizEntren (matrizEntren : String): void + getMatrizAClasificar (): String + setMatrizAClasificar (matrizAClasificar : String): void + setNombreArchivoAcciones(nombreArc hivo : String): void
SeleccionPeriodo	- <u>MATRIZ ENTREN 2009</u> : String = "MatrizEntrenamiento2009" - <u>MATRIZ ENTREN 2010</u> : String = "MatrizEntrenamiento2010" - <u>MATRIZ ENTREN 2011</u> : String = "MatrizEntrenamiento2011" - <u>MATRIZ ENTREN 2012</u> : String = "MatrizEntrenamiento2012" - <u>MATRIZ ENTREN 2013</u> : String = "MatrizEntrenamiento2013" - <u>MATRIZ ENTREN 2014</u> : String = "MatrizEntrenamiento2014" - <u>MAT 2009</u> : String = "Matriz2009" - <u>MAT 2010</u> : String = "Matriz2010" - <u>MAT 2011</u> : String = "Matriz2011" - <u>MAT 2012</u> : String = "Matriz2012" - <u>MAT 2013</u> : String = "Matriz2013" - <u>MAT 2014</u> : String = "Matriz2014" - <u>ACC 2009</u> : String = "nombre_acciones_2009" - <u>ACC 2010</u> : String = "nombre_acciones_2010" - <u>ACC 2011</u> : String = "nombre_acciones_2011" - <u>ACC 2012</u> : String = "nombre_acciones_2012" - <u>ACC 2013</u> : String = "nombre_acciones_2013" - <u>ACC 2014</u> : String = "nombre_acciones_2014"	+ «Constructor» + SeleccionPeriodo () + IniciarActionPerformed (evt : ActionEvent): void + <u>main</u> (args : String): void

Tabla 4. Tabla Atributos y Métodos SOM

5.3.6 Funcionamiento de la aplicación

A continuación se describe el funcionamiento de la aplicación desarrollada.

Al iniciar la aplicación se despliega un menú el cual permite seleccionar algunos de los periodos a estudiar como se ve en la Figura 18.



Figura 18. Menú Selección de Período

Al seleccionar un periodo y presionar el botón comenzar se despliega la matriz con neurona inicializadas con valores aleatorios.

Posteriormente a esto, se visualizará una matriz similar a lo que se muestra en la Figura 19.

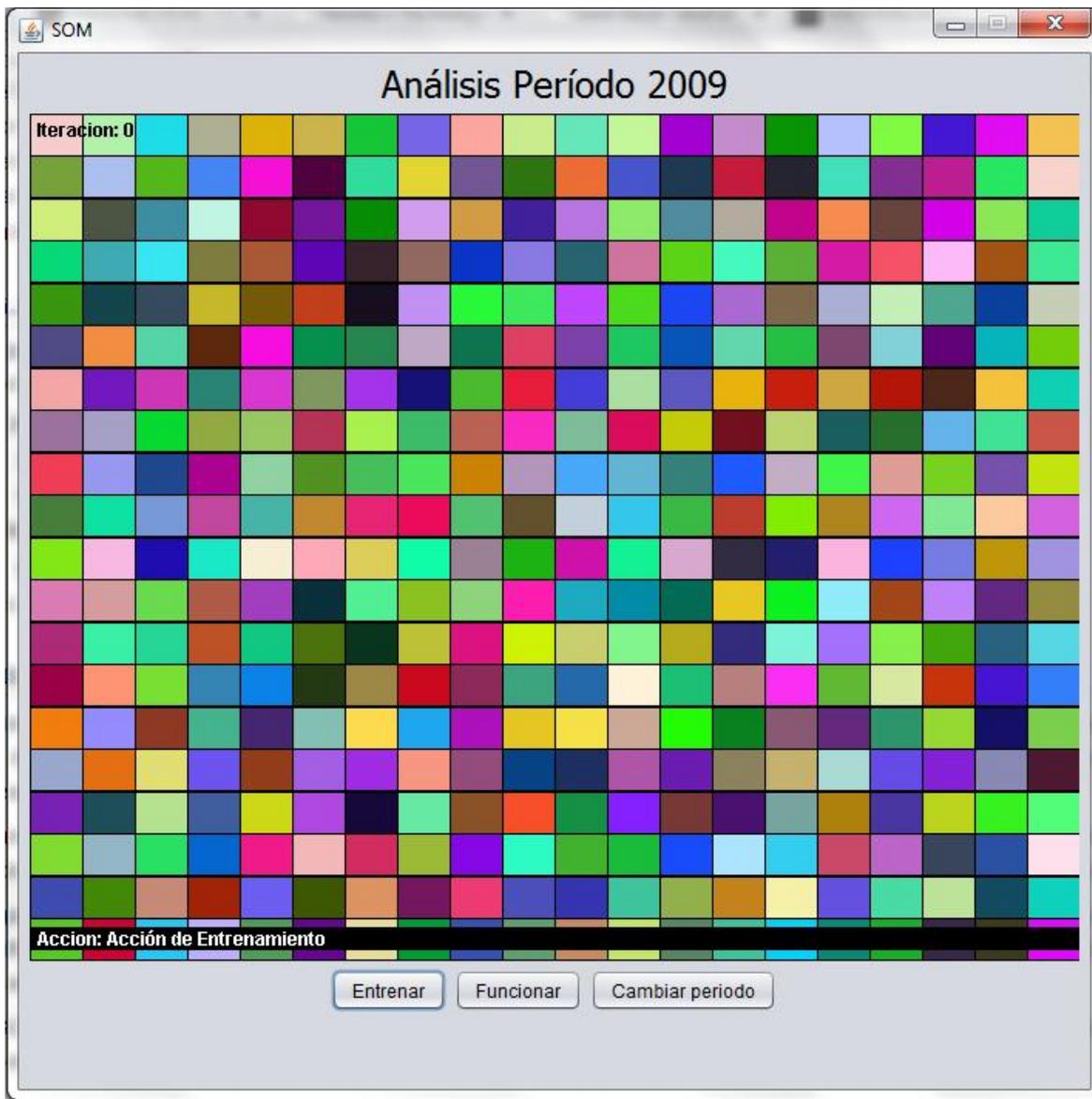


Figura 19. Matriz Inicial

Posteriormente al presionar el botón entrenar comenzara a leer los datos de entrenamiento y la red comenzara su etapa de aprendizaje en el cual ejecutará quinientas iteraciones las que se pueden visualizar en la esquina superior derecha. En la Figura 20 se puede ver el estado de las neuronas una vez terminado el entrenamiento.

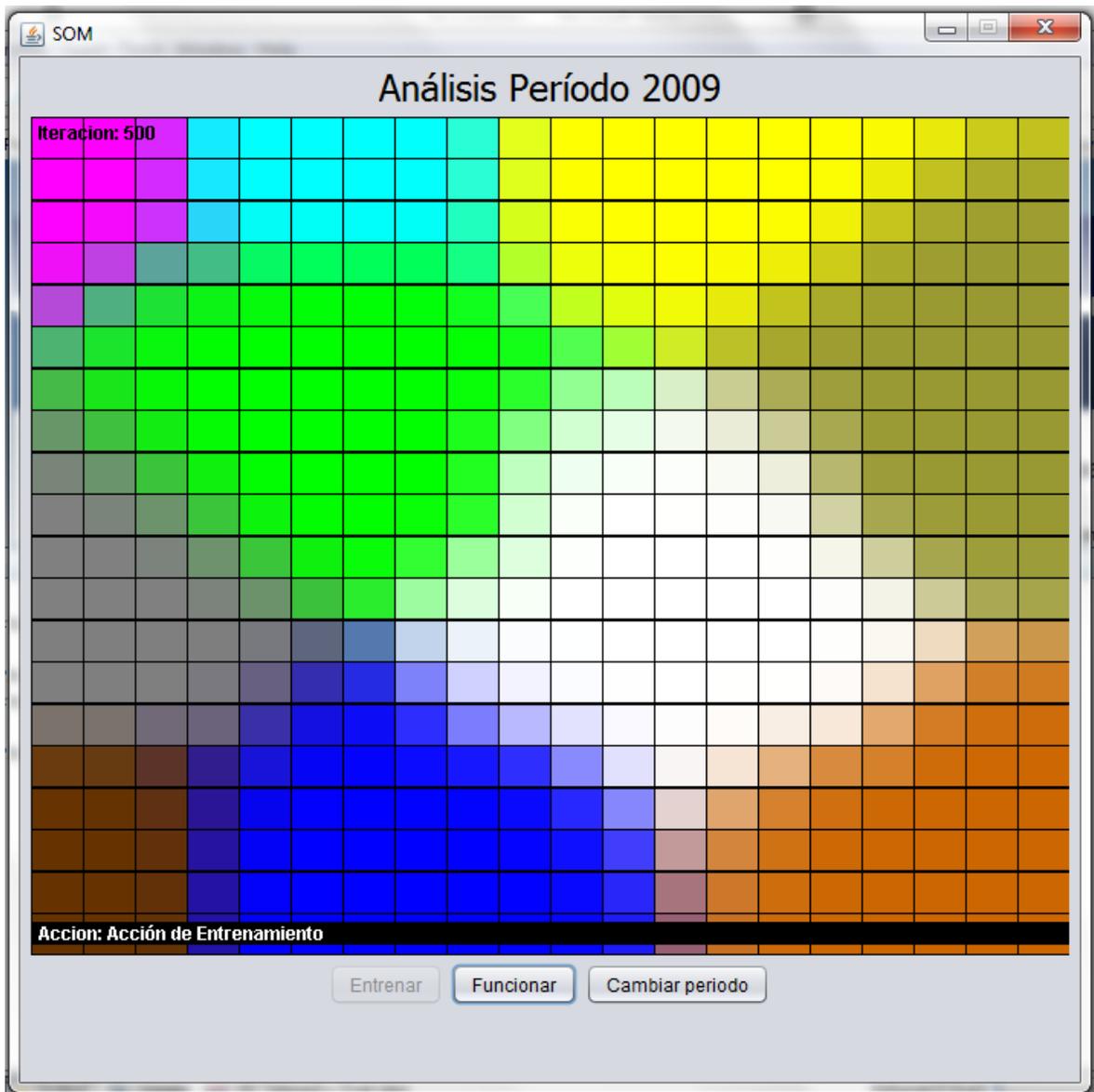


Figura 20. Matriz Post- Entrenamiento

Para finalizar al presionar el botón funcionar la aplicación comenzara a leer los datos de las correlaciones de las acciones del correspondiente periodo y la red realizará su clasificación en base a lo aprendido en la etapa de entrenamiento. En la Figura 21 se puede ver cómo quedan las acciones agrupadas por la red.

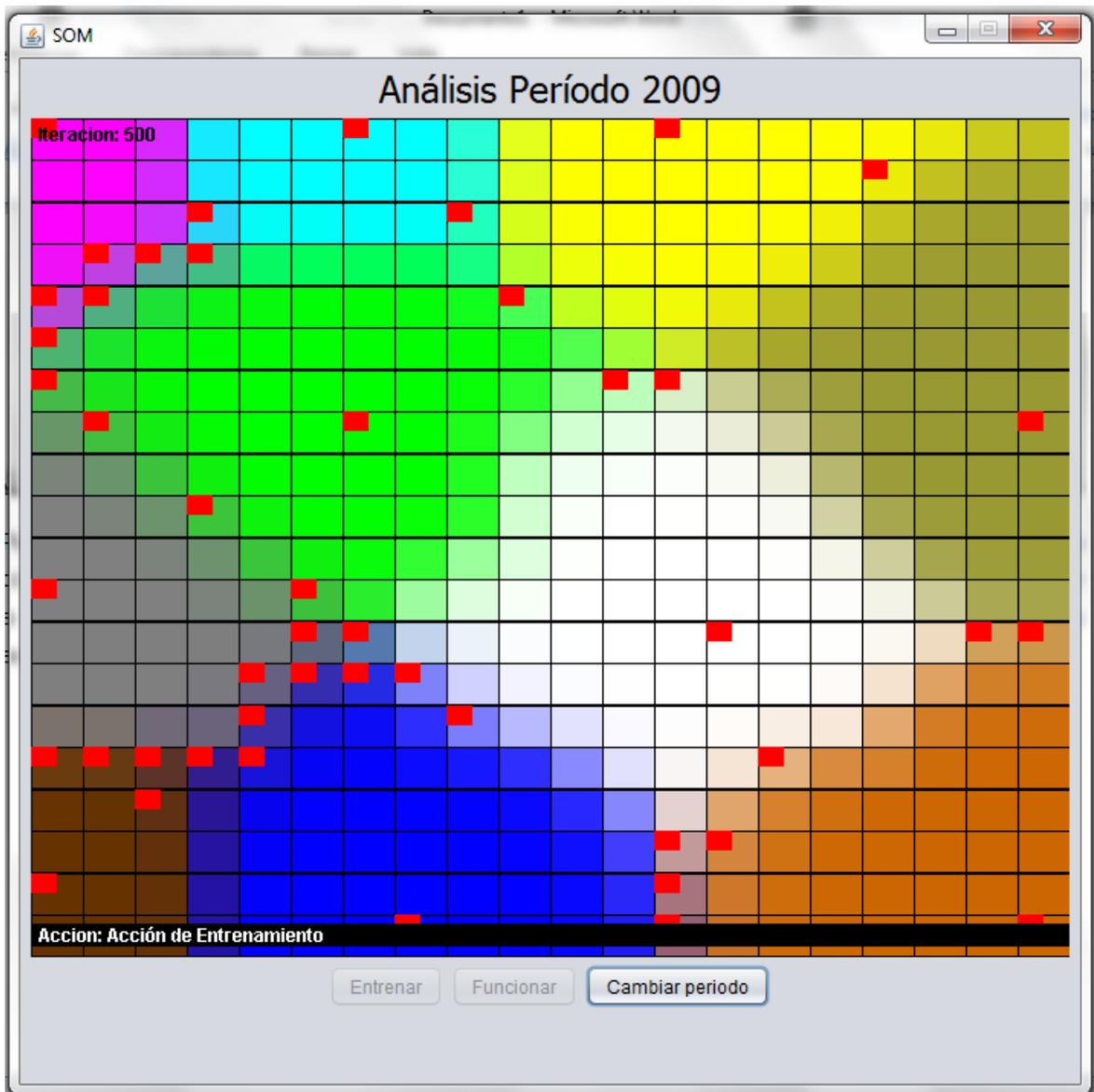


Figura 21. Matriz Post- Funcionamiento

Cada recuadro de la matriz que posee en su interior otro recuadro color rojo representa que una de las acciones se ha clasificado en esa posición, para saber qué acción es, se debe posicionar el puntero sobre el recuadro correspondiente y en la barra negra del inferior se mostrará la o las acciones que han quedado en esa posición.

5.3.7 Prueba de Funcionamiento

Para comprobar que la implementación del algoritmo fuese correcta se realizó una comparación de resultados con un trabajo de terceros. Para esto se utilizó un ejemplo obtenido en la página web MQL5 (MetaQuotes, 2014) en el cual se usan Mapas de Kohonen para clasificar diferentes alimentos en base a sus proteínas, carbohidratos y grasa. Los alimentos utilizados en este trabajo se encuentran en la tabla 5.

	Alimento	Proteína	Carbohidratos	Grasa
1	Manzanas	0.4	11.8	0.1
2	Aguacate	1.9	1.9	19.5
3	Plátanos	1.2	23.2	0.3
4	Filete de ternera	20.9	0	7.9
5	Hamburguesa Big Mac	13	19	11
6	Nueces de Brasil	15.5	2.9	68.3
7	Pan	10.5	37	3.2
8	Mantequilla	1	0	81
9	Queso	25	0.1	34.4
10	Tarta de queso	6.4	28.2	22.7
11	Galletas	5.7	58.7	29.3
12	Copos de maíz	7	84	0.9
13	Huevos	12.5	0	10.8
14	Pollo frito	17	7	20
15	Patatas fritas	3	36	13
16	Chocolate caliente	3.8	19.4	10.2
17	Pepperoni	20.9	5.1	38.3
18	Pizza	12.5	30	11
19	Pastel de carne porcina (Pork Pie)	10.1	27.3	24.2
20	Patatas	1.7	16.1	0.3
21	Arroz	6.9	74	2.8
22	Pollo asado	26.1	0.3	5.8
23	Azúcar	0	95.1	0
24	Filete de atún	25.6	0	0.5
25	Agua	0	0	0

Tabla 5. Alimentos con sus distintas especificaciones

El resultado obtenido por la web MQL5 se puede ver en la Figura 22. En esta figura se muestran los resultados que se obtuvo al ingresar los alimentos de la tabla 5 y clasificarlos utilizando la implementación de Mapas de Kohonen realizada en este proyecto

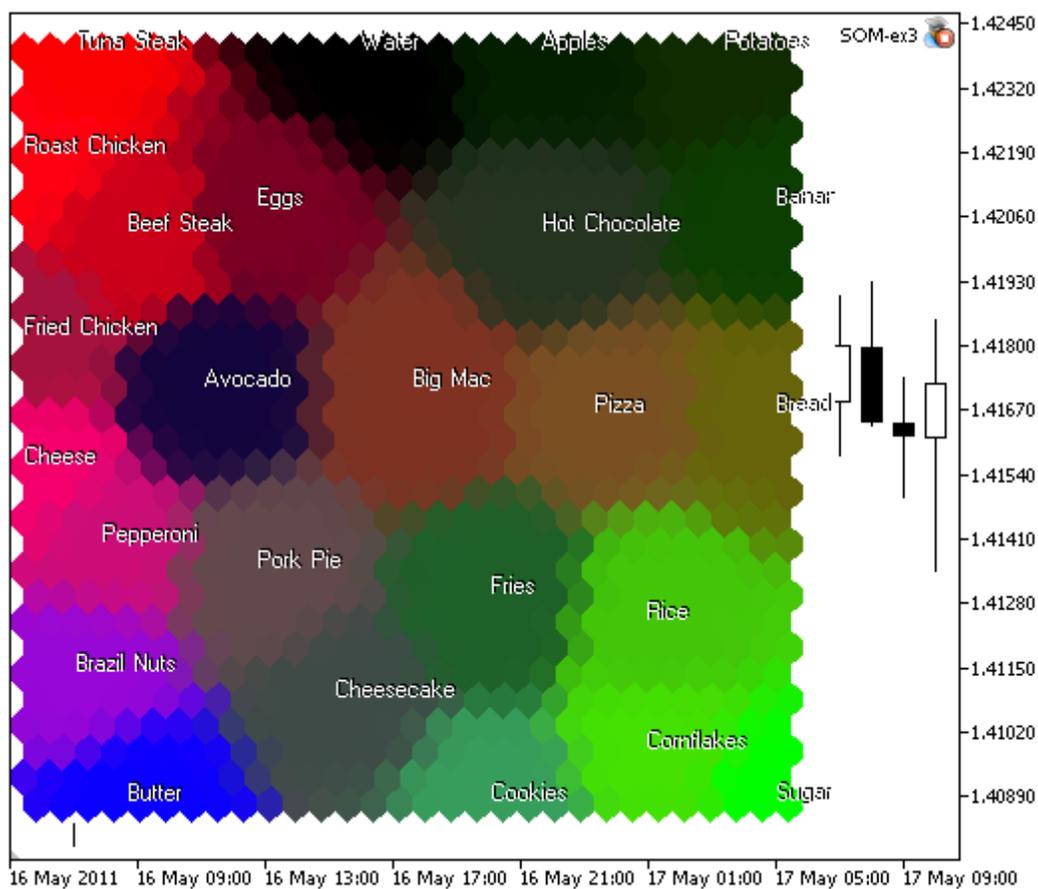


Figura 22. Resultado SOM clasificación Alimentos (MetaQuotes, 2014)



Figura 23. Resultado Alimento implementado con SOM del proyecto

Al comparar los dos resultados se puede ver que la distribución de los alimentos en relación a sus respectivos colores en el mapa es bastante similar, por ejemplo, el arroz, la azúcar y las galletas se agrupan en las regiones de color verde en ambos mapas, también la mantequilla y las nueces brasileñas se mantiene en la región azul en ambos mapas y se puede ver que las similitudes siguen en otros alimentos. Estos resultados equivalentes indican que la implementación de la red se realizó de manera adecuada.

6 RESULTADOS Y CONCLUSIONES

En esta sección se mostrarán los resultados obtenidos utilizando la aplicación de Mapa de Kohonen utilizando los datos de las correlaciones de los retornos de las acciones de la bolsa de Santiago para los períodos 2009 al 2014.

Utilizando su cercanía en relación al color de cada neurona se identificaron 10 clúster de acciones. La Tabla 5 muestra los distintos colores en la cual se representará cada rubro o mercado perteneciente cada acción.

Color	Mercado o Rubro
Plomo	Inversiones e Inmobiliarias
Verde	Agropecuarias y Forestales
Azul	Alimenticias y Bebidas
Blanco	Bancarias y Financieras
Amarillo	Comerciales y Distribuidoras
Rosado	Industriales
Celeste	Metalmecánicas
Café	Pesqueras
Naranja	Servicios Públicos
Verde Musgo	Servicios Varios

Tabla 6. Color representativo por Mercado

Las Tablas siguientes muestran los diferentes grupos que se pueden obtener al visualizar los mapas obtenidos post clasificación de las correlaciones de las acciones por cada período.

Tabla de Resultados Obtenidos

Grupo / Año	2009	2010	2011	2012	2013	2014
Grupo 1						
	BCI	BCI	BCI	BCI	BCI	BCI
	BSANTANDER	SM-CHILE B	RIPLEY	BSANTANDER	BSANTANDER	SM-CHILE B
	CHILE	CHILE	CORPBANCA		BESALCO	ECL
	CCU	CORPBANCA				COLBUN
	SONDA	SM-CHILE B				
	FALABELLA					
Grupo 2						
	CENCOSUD	CENCOSUD	CENCOSUD	CENCOSUD	CENCOSUD	CENCOSUD
	CGE	CALICHERAA	FALABELLA	RIPLEY		ENTEL
	NUEVAPOLAR	NORTEGRAN	CHILE			HITES
	RIPLEY	ORO BLANCO	SM-CHILE B			NORTEGRAN
	COLO COLO	SK	CONCHATORO			AESGENER
		IAM	SALFACORP			
		FASA				
		CAMPOS				
		IANSA				
		MARINSA				
		PILMAIQUEN				

Tabla 7. Grupo 1 y Grupo 2

Grupo / Año	2009	2010	2011	2012	2013	2014
Grupo 3						
	MASISA	MASISA	CCU	MASISA	IANSA	MASISA
	CORPBANCA	GASCO	PILMAIQUEN	COLO COLO	SMSAAM	VAPORES
		CEMENTOS	ALMENDRAL	SCHWAGER	SK	PARAUCO
		CINTAC	ENTEL	CAMPOS	SALFACORP	
		CIC	CAMPOS	SQM-B	ILC	
		BESALCO	IANSA	VAPORES	MASISA	
		SALFACORP	SONDA	CRUZADOS	ENTEL	
		SOCOYESA	MASISA	ENDESA	FALABELLA	
				ENERSIS	SECURITY	
				MOLYMET	ENJOY	
				PAZ	AQUACHILE	
				SM-CHILE B	CONCHATORO	
				QUINENCO		
				FALABELLA		
				PARAUCO		
				CHILE		
				PILMAIQUEN		
			CONCHATORO			
			LAS CONDES			
Grupo 4						
	VAPORES	MULTIFOODS	INVERMAR	MULTIFOODS	MULTIFOODS	MULTIFOODS
	MULTIFOODS	CRISTALES	MULTIFOODS	AUSTRALIS	NORTEGRAND	CAMANCHA
	PAZ	AGUAS-A	PARAUCO	INVERMAR		ENJOY
	SALFACORP	CRUZADOS	COLBUN	ALMENDRAL		SMSAAM
	SOCOYESA	INVERMAR	ENDESA			AQUACHILE
	SCHWAGER	COLO COLO	ENERSIS			
	WMTCL	CURAUMA	CAMANCHACA			
	PILMAIQUEN	PEHUENCHE				
	INVERMAR	PUCOBRE-A				

Tabla 8. Grupo 3 y Grupo 4

Grupo/Año	2009	2010	2011	2012	2013	2014
Grupo 5						
	LAN	LAN	CHILE	LAN	LAN	SK
	COLBUN	AESGENER	BESALCO	SOCOYESA	ENDESA	HABITAT
	ECL	ENTEL	SQM-B	BELSALCO	SQM-B	LAN
		INDISA	MULTIFOODS	SALFACORP	SOCOYESA	ILC
		QUINENCO	INVERMAR	SONDA	PAZ	QUINENCO
		FALABELLA	INVEXANS		CRUZBLANCA	
		SECURITY	CAMANCHACA		CORPBANCA	
		SQM-B	VAPORES			
		ECL				
		NUEVAPOLAR				
		VAPORES				
Grupo 6						
	FORUS	FORUS	ENJOY	BANMEDICA	FORUS	FORUS
	CUPRUM	CUPRUM	FORUS	ZOFRI	AZUL AZUL	EMBONOR-B
	SK	CGE	IAM	CFR	COLO COLO	TECHPACK
	IANSÁ	ALMENDRAL	SECURITY	GASCO	CCU	SALFACORP
	BANMEDICA		SK	INVEXANS	EMBONOR-B	SOCOYESA
	AGUAS-A			HABITAT	INVERMAR	NUEVAPOLAR
	PEHUENCHE			IAM	INVEXANS	SQM-B
				HITES	MOLYMET	BELSALCO
				FORUS	LAS CONDES	GASCO
					ALMENDRAL	
					AGUAS-A	
					IAM	
					BANMEDICA	
					CGE	
					GASCO	
				AESGENER		

Tabla 9. Grupo 5 y Grupo 6

Grupo/Año	2009	2010	2011	2012	2013	2014
Grupo 7						
	PROVIDA	PROVIDA	PROVIDA	PROVIDA	PROVIDA	PROVIDA
	SM-CHILE B	HITES	ZOFRI	CUPRUM	ZOFRI	SECURITY
		BANMEDICA	CGE		PEHUENCHE	
		ENJOY	CRUZADOS		SCHWAGER	
		HABITAT	SCHWAGER			
Grupo 8						
	CAP	CAP	CAP	CAP	COLBUN	CMPC
	ALMENDRAL	PARAUCO	HITES	COLBUN	PARAUCO	CAP
	ENTEL	WMTCL	PAZ	SK	ANTARCHILE	FALABELLA
	CMPC	INVERCAP	SOCOYESA	ENTEL	CAP	PAZ
	INVERCAP	CONCHATORO	BESALCO	ECL	SONDA	AGUAS-A
	ORO BLANCO	PAZ	SQM-B	AESGENER	VAPORES	ALMENDRAL
	FASA	SCHWAGER	ANTARCHILE	CORPBANCA	RIPLEY	
	IAM	RIPLEY	QUINENCO	AQUACHILE	QUINENCO	
	INVEXANS	SAN PEDRO		CAMANCHACA	HABITAT	
	NORTEGRAN	SONDA		IANSA	CFR	
	ZOFRI			NORTEGRAN	HITES	
	CALICHERAA			OROBLANCO	ENERSIS	
	PARAUCO			CRUZBLANCA		
	SQM-B			SECURITY		

Tabla 10. Grupo 7 y Grupo 8

Grupo/Año	2009	2010	2011	2012	2013	2014
Grupo 9						
	COPEC	INVEXANS	BSANTANDER	CMPC	COPEC	IAM
	ANTARCHILE	AXXION	INVERCAP	ANTARCHILE	CHILE	SONDA
		ZOFRI	COPEC	COPEC	SM-CHILE B	COPEC
		CMPC	CMPC		CMPC	ANTARCHILE
		COLBUN			ECL	CGE
		ENDESA				CHILE
		ENERSIS				ENDESA
		ANTARCHILE				ENERSIS
		COPEC				CORPBANCA
						RIPLEY
					BSANTANDER	
Grupo 10						
	INDISA	ANDINA-A	AESGENER	ANDINA-A	ANDINA-A	CONCHATORO
	ANDINA-A	ANDINA-B	ORO BLANCO	ANDINA-B	ANDINA-B	LAS CONDES
	AESGENER	CCU	BANMEDICA	EMBNOR -B	NUEVAPOLAR	SCHWAGER
	CONCHATORO		CUPRUM	CCU		INVEXANS
	ANDINA-B		EXESVAL-C	NUEVAPOLAR		MINERA
	ENDESA		HABITAT	AGUAS-A		AZUL AZUL
	ENERSIS		VAPORES	AZUL AZUL		COLO COLO
	QUINENCO		MOLYMET	CGE		CFR
			ANDINA-B	ENJOY		BANMEDICA
			NORTEGRAN			CRUZBLANCA
			ANDINA-A			CCU
						IANSA
						ANDINA-A

Tabla 11. Grupo 9 y Grupo 10

Al analizar los resultados obtenidos utilizando el Mapa de Kohonen lo primero que se puede apreciar al ver el estado final de la red (Ver Figura 22) es que la mayoría de las acciones tienden a caer en neuronas que están ubicadas en la zona donde convergen los colores y solo algunas acciones se posicionan más cercanas al centro de cada color lo que indica que las acciones que están ubicadas en sectores más céntricos son las más representativas de cada grupo.

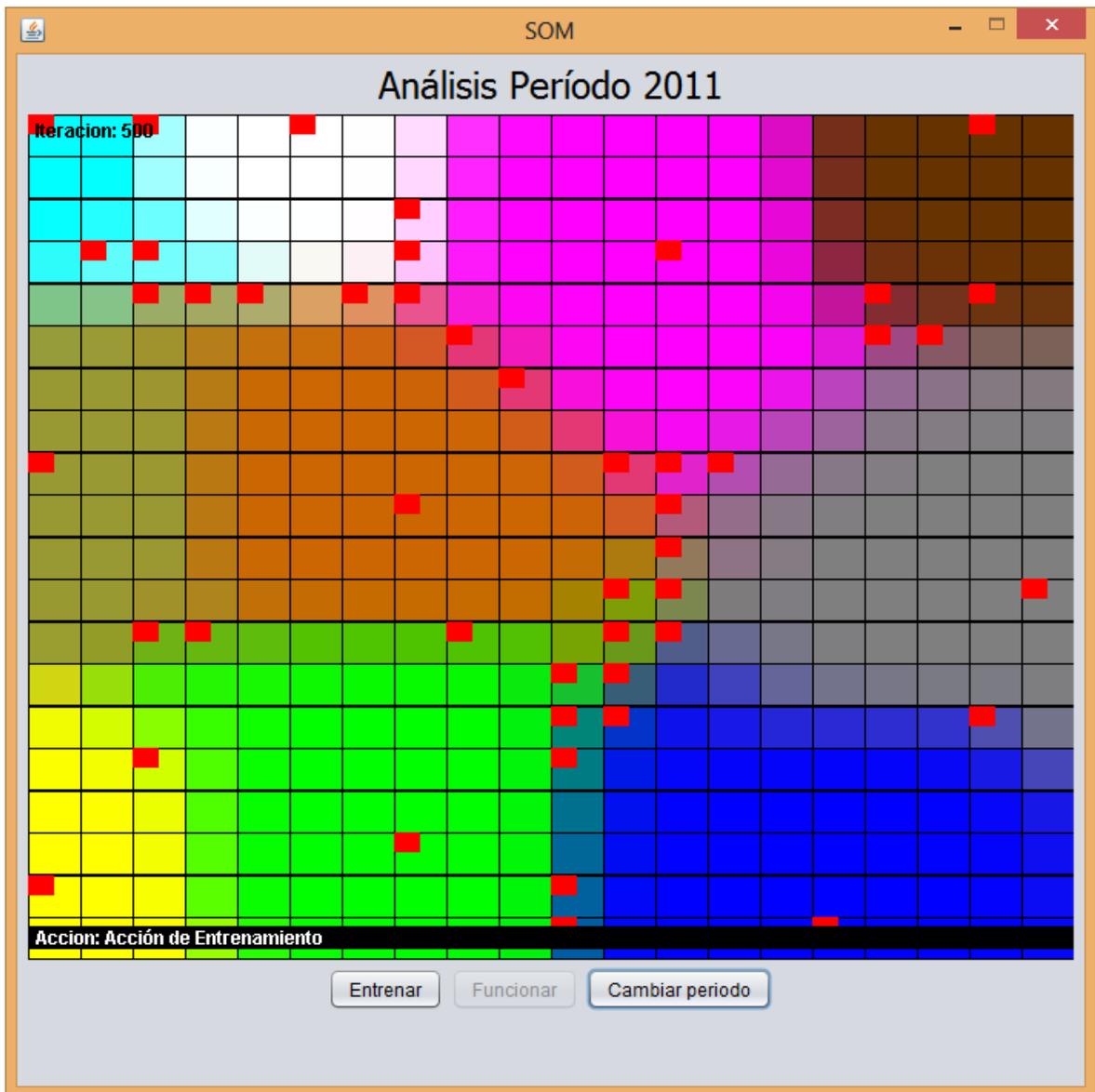


Figura 24. Resultado Análisis Período 2011

Al ejecutar el algoritmo reiteradamente se ve que las acciones que están en un mismo grupo tienden a mantenerse, aunque, en cada ejecución tiende a existir pequeñas variaciones en la clasificación final de los datos, esto se puede atribuir a que el estado inicial de la red es diferente en cada ejecución, ya que las neuronas se inicializan con valores aleatorios, lo que produce pequeñas variaciones en el estado final, aun así, a grandes rasgos las relaciones o grupos se mantienen en el tiempo. Otro aspecto que se identificó al realizar múltiples ejecuciones es que la distancia entre cada acción al ser clasificada no necesariamente indica que estas sean más cercanas o que pertenezcan al mismo grupo, por ejemplo, al ver la Figura 21 de Matriz Post- Funcionamiento se observa que hay un grupo de acciones que están ubicadas en neuronas que convergen entre el color plomo con el azul, a primera impresión se podría suponer que estas acciones forman un clúster ya que están muy juntas unas de otras sin embargo al realizar diversas ejecuciones se observó que los colores de la red se tienden a organizar en diferentes posiciones del mapa en cada ejecución de manera que, por ejemplo, en una siguiente ocasión el color azul con el plomo no necesariamente van a converger en cuyo caso las acciones que se encuentran muy cercanas en la figura 21 no se encontrarán en posiciones de vecindad. Esto llevo a la conclusión de que los grupos o clúster que se generan deben ser identificados principalmente utilizando el criterio de color es decir, siguiendo con el ejemplo anterior las acciones que están ubicadas en una neurona que tiene por color una variante del azul tienden a seguir juntas en todas las ejecuciones y las que esta ubicadas en una variante del color plomo seguirán perteneciendo al mismo grupo.

Se llevó a cabo un análisis comparativo entre los resultados obtenidos utilizando Mapa de Kohonen y los resultados conseguidos por los profesores Carola Figueroa Flores y Edinson Cornejo Saavedra en el documento llamado “Estructura Topológica del Mercado Bursátil de Chile, período 2009 – 2014”. Este estudio buscó modelar la estructura topológica del mercado bursátil chileno durante el período 2009-2014, a partir de las correlaciones entre los retornos de las acciones transadas en la Bolsa de Comercio de Santiago utilizando el algoritmo de Kruskal basado en la teoría de grafos, para el desarrollo de una representación gráfica de las relaciones de dependencia mutua. Los resultados de este trabajo se pueden ver en el ítem Anexos desde el Anexo 8.7 al Anexo 8.12.

Al comparar los resultados de ambos trabajos se encontró que la distribución de las acciones resultaba distinta, esto debido principalmente a los tipos de algoritmos. El algoritmo de Kruskal tiene como objetivo construir un árbol (subgrafo sin ciclos) formado por aristas sucesivamente seleccionadas de mínimo peso.

En el caso de lo que ocurre con las correlaciones de las acciones de la Bolsa de Comercio de Santiago, una diferencia sustancial que puede llevar a que los resultados sean distintos, es que cuando se utiliza el Mapa de Kohonen al momento de alimentar la red en su etapa de entrenamiento con los datos de acciones que sean representativas de los distintos mercados, se genera una tendencia de agrupamiento de acuerdo a estas acciones. A diferencia del algoritmo de Kruskal, el cual no depende de lo descrito anteriormente, sino que solo depende de las distancias del grafo que se forma, es decir, no sigue ninguna tendencia o predisposición. Por otra parte al clasificar las acciones con ambos métodos se encontró un punto común. Al clasificar las acciones en base a sus correlaciones de retornos accionarios se busca obtener clúster de acciones cuyos movimientos en el mercado sean similares, para así poder determinar patrones que permitan predecir sus movimientos accionarios en el futuro. Considerando esto se pensaba que los resultados finales es decir, los clúster que se conformarían, tuvieran relaciones con sus rubros o sectores económicos, es decir, que las empresas de industrias iguales o relacionadas tienda a estar en un mismo clúster. Una vez analizado los resultados finales se pudo observar que esto no necesariamente se producía. Esta tendencia también fue observada en el trabajo de los profesores Carola y Edinson (Figueroa Flores & Cornejo, 2014) en el cual tampoco se encontró que las acciones se agruparan en base a sus mercados, es decir, que esta tendencia se mantiene en ambos trabajos de análisis del mercado de la bolsa de Santiago y demostraría que no necesariamente las acciones de un mismo mercado se mueven en conjunto si no que por el contrario sus retornos tienden a ser independientes entre ellos. Esta tendencia podría explicarse considerando que hoy en día los mercados y las empresas están todas estrechamente relacionadas entre sí es decir una empresa del sector bancario y financiero está estrechamente relacionada como fuente de financiamiento en diversas compañías de distintos rubros y por ende sus fluctuaciones en el mercado también afectan a estas aunque la naturaleza de sus negocios sea completamente distinta. Además de esto actualmente la economía global los movimientos accionarios también se ven influenciados por factores económicos que afectan a gran escala como lo son cambios demográficos, cambios tecnológicos, energía y recursos naturales, inversión pública, política internacional

(Comercio, 2013), etc. Aun así considerando lo anterior se puede ver en este trabajo que hay una excepción a esto, la cual tiene relación con las acciones del rubro de Pesqueras las cuales tienden a agruparse todas o en su defecto la mayoría de estas en un solo clúster de acciones, esto podría deberse a que quizás en este sector económico es muy independiente de los demás es decir podría ser un rubro aislado del resto y sus retornos solo se ven afectados por los cambios propios de ese mercado.

Finalmente sería interesante de analizar en trabajos futuros cómo se comporta el mercado Bursátil de la bolsa de comercio de Santiago utilizando alguna técnica diferente de red neuronal para determinar las diferencias o ventajas y desventajas de las diferentes técnicas. También sería de interés utilizar la técnica de Mapas de Kohonen para clasificar los diferentes mercados pero a nivel de naciones de manera de poder determinar cómo las economías de los diferentes países se relacionan entre sí y analizar como las grandes potencias económicas influyen en el resto de los países.

7 BIBLIOGRAFÍA

- (s.f.).
- Alonso, J. C., & García, J. C. (2009). *¿ qué tan buenos son los patrones del IGBC para predecir su comportamiento?*
- Baillard, A. (24 de Enero de 2012). *caosy ciencia*. Recuperado el 21 de Septiembre de 2014, de http://www.caosyciencia.com/visual/imagen.php?id_img=475
- Boyacioglu, M. A., & Avci, D. (2010). *An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction*.
- Cai, Y., Chou, Y., & Li, D. (2009). *Explaining international stock correlations with CPI fluctuations*.
- Caparrini, F. S. (10 de Enero de 2014). *Dpto. de Ciencias de la Computación e Inteligencia Artificial*. Recuperado el 21 de Septiembre de 2014, de <http://www.cs.us.es/~fsancho/?e=76%20>
- Carmona, S. (2014). *IIB*. Recuperado el 21 de Septiembre de 2014, de <http://genoma.unsam.edu.ar/trac/docencia/wiki/Bioinformatica/Guias/DataMining>
- Castillo Tapia, M. G., & Coello, C. (1999). *Departamento de Computación*. Recuperado el 2015 de Febrero de 24, de <http://delta.cs.cinvestav.mx/~ccoello/journals/finanzas.pdf>
- Chen, S. A., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index. *Computers & Operations Research*, 901-923.
- Comercio, O. M. (2013). *C.Factores económicos fundamentales que afectan al comercio internacional*.
- Cruz, E., Restrepo, J., & Medina, P. (2009). *Universidad Tecnológica de Pereira*. Recuperado el 24 de Febrero de 2015, de <http://revistas.utp.edu.co/index.php/revistaciencia/article/view/2889/1603>
- Cupal, M., Deev, O., & Linnertová, D. (2012). *Network structures of the European stock markets*.
- Doherty, K. (2005). *ieeexplore*. Recuperado el 10 de Septiembre de 2014, de http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1662299&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D1662299
- Economática. (1986). *Economática*. Recuperado el 29 de Septiembre de 2014, de <https://economica.com/ES/>
- Epistemowikia. (Septiembre de 2014). *Cala*. Recuperado el 20 de Septiembre de 2014, de http://cala.unex.es/cala/epistemowikia/index.php?title=Comparativa_redes_neuronales_biol%C3%B3gicas_y_artificiales
- Fernandez, J. R. (20 de Septiembre de 2014). *Universidad San Martín de Porres*. Recuperado el 20 de Septiembre de 2014, de <http://www.usmp.edu.pe/publicaciones/boletin/fia/info32/pag4.htm>
- Figueroa Flores, C. A., & Cornejo, E. (2014). *Estructura Topológica del Mercado Bursátil de Chile, período 2009 - 2014*. Chillán.
- García, M. C., Jalal, A. M., Garzón, L. A., & López, J. M. (29 de Octubre de 2013). *Publicaciones Universidad EAFIT*. Recuperado el 22 de Septiembre de 2014, de <http://publicaciones.eafit.edu.co/index.php/ecos-economia/article/view/2298/2249>
- GEOtutoriales. (22 de Febrero de 2014). *gestion de operaciones*. Recuperado el 21 de Septiembre de 2014, de <http://www.gestiondeoperaciones.net/proyeccion-de-demanda/como-utilizar-una-regresion-lineal-para-realizar-un-pronostico-de-demanda/>
- Giovanni Bonanno, G. C. (2002). *Topology of correlation based minimal spanning trees in real and model markets*. Palermo.
- Granada, A. 8.-A. (08 de Abril de 2011). *Admon 8-A Diana Granada*. Recuperado el 21 de Septiembre de 2014, de http://admon8adianagranada.blogspot.com/2011_04_01_archive.html
- Guadalajara, U. d. (2014). Recuperado el 24 de Febrero de 2015, de <http://proton.ucting.udg.mx/posgrado/cursos/idc/neuronales2/RNArtificial.htm>

- Hadavandi, E., Shavandi, H., & Ghanbari, A. (2010). *Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. Knowledge-Based Systems.*
- Him, K., & Han, I. (2000). *Genetic algorithms approach to feature discretization in artificial neural.*
- Majhi, R., Panda, G., Majhi, B., & Shaoo, G. (2009). *Efficient prediction of stock market indices using adaptive bacterial foraging.*
- MetaQuotes. (31 de Marzo de 2014). *MQL5*. Recuperado el 18 de Febrero de 2015, de <http://www.mql5.com/es/articles/283>
- Molina, A., Toro, M. E., & Garcés, A. (2005). *Red de Revistas Científicas de América Latina y el Caribe*. Recuperado el 24 de Febrero de 2015, de <http://www.redalyc.org/pdf/2570/257021033007.pdf>
- Morales, E., & González, J. (Enero de 2013). *Instituto Nacional de Astrofísica, óptica y Electrónica*. Recuperado el 20 de Septiembre de 2014, de <http://ccc.inaoep.mx/~emorales/Cursos/NvoAprend/Acetatos/sbl.pdf>
- Moreno, G. (03 de Octubre de 2007). *Gimena Moreno WordPress*. Recuperado el 21 de Septiembre de 2014, de <http://gamoreno.wordpress.com/2007/10/03/tecnicas-mas-usadas-en-la-mineria-de-datos/>
- Ocampo, E., Molina, A., & Garcés, A. (3 de Agosto de 2005). *Universidad Distrial Francisco José de Caldas*. Recuperado el 24 de Febrero de 2015, de <http://tecatura.udistrital.edu.co/ojs/index.php/revista/article/view/186>
- Orallo, J. H. (2004). *Extracción Automática de Conocimiento en Bases de Datos e Ingeniería del Software*. Recuperado el 21 de Septiembre de 2014, de <http://users.dsic.upv.es/~jorallo/docent/doctorat/t2b.pdf>
- Porres, U. S. (s.f.). *USMP*. Recuperado el 20 de Septiembre de 2014, de <http://www.usmp.edu.pe/publicaciones/boletin/fia/info32/pag4.htm>
- Sanchez, F. (15 de 12 de 2013). *francescsanchezbi*. Recuperado el 13 de 10 de 2014, de El algoritmo de clústeres de Microsoft: <http://francescsanchezbi.webnode.es/news/el-algoritmo-de-clusteres-de-microsoft/>
- Santana Luis Quintero Vicente, C. C. (Diciembre de 2006). *Universidad de Olavide*. Recuperado el 21 de Noviembre de 2014, de <http://www.upo.es/RevMetCuant/pdf/vol2/art4.pdf>
- Santiago, B. d. (02 de Diciembre de 2014). *Bolsa de Comercio de Santiago*. Recuperado el 02 de Diciembre de 2014, de <http://www.bolsadesantiago.com/index.aspx>
- Szpiro, G. G. (1997). *A search for hidden relationships: data mining with genetic algorithms. Computational Economics.*
- Wang, Y. F., Cheng, S., & Hsu, M. H. (2010). *Incorporating the Markov chain concept into fuzzy stochastic prediction of stock indexes.*

8 ANEXOS

8.1 Clasificación de las Acciones del año 2009 según su rubro

Acción	Rubro
WMTCL	Comerciales y Distribuidoras
PROVIDA	Inversiones e Inmobiliarias
QUINENCO	Inversiones e Inmobiliarias
SALFACORP	Inversiones e Inmobiliarias
MASISA	Agropecuarias y Forestales
SCHWAGER	Agropecuarias y Forestales
ANDINA-A	Alimenticias y Bebidas
ANDINA-B	Alimenticias y Bebidas
CCU	Alimenticias y Bebidas
CONCHATORO	Alimenticias y Bebidas
IANSA	Alimenticias y Bebidas
BCI	Bancarias y Financieras
BSANTANDER	Bancarias y Financieras
CHILE	Bancarias y Financieras
CORPBANCA	Bancarias y Financieras
SM-CHILE B	Bancarias y Financieras
COPEC	Comerciales y Distribuidoras
FALABELLA	Comerciales y Distribuidoras
FASA	Comerciales y Distribuidoras
NUEVAPOLAR	Comerciales y Distribuidoras
RIPLEY	Comerciales y Distribuidoras
ZOFRI	Comerciales y Distribuidoras
FORUS	Industriales
CALICHERAA	Inversiones e Inmobiliarias
CUPRUM	Inversiones e Inmobiliarias
INDISA	Inversiones e Inmobiliarias
INVERCAP	Inversiones e Inmobiliarias
INVEXANS	Inversiones e Inmobiliarias
NORTEGRAN	Inversiones e Inmobiliarias

Acción	Rubro
ORO BLANCO	Inversiones e Inmobiliarias
PARAUCO	Inversiones e Inmobiliarias
SOCOVELSA	Inversiones e Inmobiliarias
ANTARCHILE	Inversiones e Inmobiliarias
BANMEDICA	Inversiones e Inmobiliarias
VAPORES	Marítimas y Navieras
CAP	Metalmecánicas
SQM-B	Mineras
INVERMAR	Pesqueras
MULTIFOODS	Pesqueras
CMPC	Productos Diversos
PEHUENCHE	Servicios Públicos
PILMAIQUEN	Servicios Públicos
LAN	Servicios Públicos
AESGENER	Servicios Públicos
AGUAS-A	Servicios Públicos
ALMENDRAL	Servicios Públicos
CGE	Servicios Públicos
COLBUN	Servicios Públicos
ECL	Servicios Públicos
ENDESA	Servicios Públicos
ENERSIS	Servicios Públicos
ENTEL	Servicios Públicos
CENCOSUD	Servicios Varios
COLO COLO	Servicios Varios
IAM	Servicios Varios
PAZ	Servicios Varios
SK	Servicios Varios
SONDA	Servicios Varios

Tabla 12. Clasificación Acciones año 2009 según su rubro

8.2 Clasificación de las Acciones del año 2010 según su rubro

Acción	Rubro
WMTCL	Comerciales y Distribuidoras
BESALCO	Construcción
CEMENTOS	Construccion
PROVIDA	Inversiones e Inmobiliarias
QUINENCO	Inversiones e Inmobiliarias
SALFACORP	Inversiones e Inmobiliarias
SECURITY	Inversiones e Inmobiliarias
MASISA	Agropecuarias y Forestales
SCHWAGER	Agropecuarias y Forestales
ANDINA-A	Alimenticias y Bebidas
ANDINA-B	Alimenticias y Bebidas
CCU	Alimenticias y Bebidas
CONCHATORO	Alimenticias y Bebidas
IANSA	Alimenticias y Bebidas
SAN PEDRO	Alimenticias y Bebidas
BCI	Bancarias y Financieras
BSANTANDER	Bancarias y Financieras
CHILE	Bancarias y Financieras
CORPBANCA	Bancarias y Financieras
SM-CHILE B	Bancarias y Financieras
COPEC	Comerciales y Distribuidoras
FALABELLA	Comerciales y Distribuidoras
FASA	Comerciales y Distribuidoras
HITES	Comerciales y Distribuidoras
NUEVAPOLAR	Comerciales y Distribuidoras
RIPLEY	Comerciales y Distribuidoras
ZOFRI	Comerciales y Distribuidoras
FORUS	Industriales
CALICHERAA	Inversiones e Inmobiliarias
CAMPOS	Inversiones e Inmobiliarias
CUPRUM	Inversiones e Inmobiliarias
CURAUMA	Inversiones e Inmobiliarias
HABITAT	Inversiones e Inmobiliarias
INDISA	Inversiones e Inmobiliarias
INVERCAP	Inversiones e Inmobiliarias
INVEXANS	Inversiones e Inmobiliarias
MARINSA	Inversiones e Inmobiliarias

Acción	Rubro
NORTEGRAN	Inversiones e Inmobiliarias
ORO BLANCO	Inversiones e Inmobiliarias
PARAUCO	Inversiones e Inmobiliarias
SOCOVESA	Inversiones e Inmobiliarias
ANTARCHILE	Inversiones e Inmobiliarias
AXXION	Inversiones e Inmobiliarias
BANMEDICA	Inversiones e Inmobiliarias
VAPORES	Marítimas y Navieras
CAP	Metalmecánicas
CINTAC	Metalmecánicas
PUCOBRE-A	Mineras
SQM-B	Mineras
INVERMAR	Pesqueras
MULTIFOODS	Pesqueras
CIC	Productos Diversos
CMPC	Productos Diversos
CRISTALES	Productos Diversos
LAN	Servicios Públicos
PEHUENCHE	Servicios Públicos
PILMAIQUEN	Servicios Públicos
AESGENER	Servicios Públicos
AGUAS-A	Servicios Públicos
ALMENDRAL	Servicios Públicos
CGE	Servicios Públicos
COLBUN	Servicios Públicos
ECL	Servicios Públicos
ENDESA	Servicios Públicos
ENERSIS	Servicios Públicos
ENTEL	Servicios Públicos
GASCO	Servicios Públicos
CENCOSUD	Servicios Varios
COLO COLO	Servicios Varios
CRUZADOS	Servicios Varios
ENJOY	Servicios Varios
IAM	Servicios Varios
PAZ	Servicios Varios
SK	Servicios Varios
SONDA	Servicios Varios

Tabla 13. Clasificación Acciones 2010 según su rubro

8.3 Clasificación de las Acciones del año 2011 según su rubro

Acción	Rubro
BESALCO	Construcción
PROVIDA	Inversiones e Inmobiliarias
QUINENCO	Inversiones e Inmobiliarias
SALFACORP	Inversiones e Inmobiliarias
SECURITY	Inversiones e Inmobiliarias
MASISA	Agropecuarias y Forestales
SCHWAGER	Agropecuarias y Forestales
ANDINA-A	Alimenticias y Bebidas
ANDINA-B	Alimenticias y Bebidas
CCU	Alimenticias y Bebidas
CONCHATORO	Alimenticias y Bebidas
IANSA	Alimenticias y Bebidas
BCI	Bancarias y Financieras
BSANTANDER	Bancarias y Financieras
CHILE	Bancarias y Financieras
CORPBANCA	Bancarias y Financieras
SM-CHILE B	Bancarias y Financieras
COPEC	Comerciales y Distribuidoras
FALABELLA	Comerciales y Distribuidoras
HITES	Comerciales y Distribuidoras
NUEVAPOLAR	Comerciales y Distribuidoras
RIPLEY	Comerciales y Distribuidoras
ZOFRI	Comerciales y Distribuidoras
FORUS	Industriales
CALICHERAA	Inversiones e Inmobiliarias
CAMPOS	Inversiones e Inmobiliarias
CUPRUM	Inversiones e Inmobiliarias
CURAUMA	Inversiones e Inmobiliarias
HABITAT	Inversiones e Inmobiliarias
INVERCAP	Inversiones e Inmobiliarias
INVEXANS	Inversiones e Inmobiliarias
NORTEGRAN	Inversiones e Inmobiliarias
ORO BLANCO	Inversiones e Inmobiliarias
PARAUCO	Inversiones e Inmobiliarias

Acción	Rubro
SOCOVELSA	Inversiones e Inmobiliarias
ANTARCHILE	Inversiones e Inmobiliarias
BANMEDICA	Inversiones e Inmobiliarias
VAPORES	Marítimas y Navieras
CAP	Metalmecánicas
MOLYMET	Metalmecánicas
SQM-B	Mineras
CAMANCHACA	Pesqueras
INVERMAR	Pesqueras
MULTIFOODS	Pesqueras
CMPC	Productos Diversos
PILMAIQUEN	Servicios Públicos
AESGENER	Servicios Públicos
AGUAS-A	Servicios Públicos
ALMENDRAL	Servicios Públicos
CGE	Servicios Públicos
COLBUN	Servicios Públicos
ECL	Servicios Públicos
ENDESA	Servicios Públicos
ENERSIS	Servicios Públicos
ENTEL	Servicios Públicos
EXESVAL-C	Servicios Públicos
LAN	Servicios Públicos
AZUL AZUL	Servicios Varios
CENCOSUD	Servicios Varios
COLO COLO	Servicios Varios
CRUZADOS	Servicios Varios
EJOY	Servicios Varios
IAM	Servicios Varios
PAZ	Servicios Varios
SK	Servicios Varios
SONDA	Servicios Varios

Tabla 14. Clasificación Acciones año 2011 según su rubro

8.4 Clasificación de las Acciones del año 2012 según su rubro

Acción	Rubro
BESALCO	Construcción
PROVIDA	Inversiones e Inmobiliarias
QUINENCO	Inversiones e Inmobiliarias
SALFACORP	Inversiones e Inmobiliarias
SECURITY	Inversiones e Inmobiliarias
MASISA	Agropecuarias y Forestales
SCHWAGER	Agropecuarias y Forestales
ANDINA-A	Alimenticias y Bebidas
ANDINA-B	Alimenticias y Bebidas
CCU	Alimenticias y Bebidas
CONCHATORO	Alimenticias y Bebidas
EMBONOR-B	Alimenticias y Bebidas
IANSÁ	Alimenticias y Bebidas
BCI	Bancarias y Financieras
BSANTANDER	Bancarias y Financieras
CHILE	Bancarias y Financieras
CORPBANCA	Bancarias y Financieras
SM-CHILE B	Bancarias y Financieras
COPEC	Comerciales y Distribuidoras
FALABELLA	Comerciales y Distribuidoras
HITES	Comerciales y Distribuidoras
NUEVAPOLAR	Comerciales y Distribuidoras
RIPLEY	Comerciales y Distribuidoras
ZOFRI	Comerciales y Distribuidoras
FORUS	Industriales
CAMPOS	Inversiones e Inmobiliarias
CUPRUM	Inversiones e Inmobiliarias
HABITAT	Inversiones e Inmobiliarias
INVERMAR	Inversiones e Inmobiliarias
LAS CONDES	Inversiones e Inmobiliarias
NORTEGRAN	Inversiones e Inmobiliarias
ORO BLANCO	Inversiones e Inmobiliarias
PARAUCO	Inversiones e Inmobiliarias
SOCOYESA	Inversiones e Inmobiliarias
ANTARCHILE	Inversiones e Inmobiliarias

Acción	Rubro
BANMEDICA	Inversiones e Inmobiliarias
VAPORES	Marítimas y Navieras
CAP	Metalmecánicas
MOLYMET	Metalmecánicas
SQM-B	Mineras
AQUACHILE	Pesqueras
AUSTRALIS	Pesqueras
CAMANCHACA	Pesqueras
INVEXANS	Pesqueras
MULTIFOODS	Pesqueras
CMPC	Productos Diversos
CFR	Productos Químicos
PILMAIQUEN	Servicios Públicos
AESGENER	Servicios Públicos
AGUAS-A	Servicios Públicos
ALMENDRAL	Servicios Públicos
CGE	Servicios Públicos
COLBUN	Servicios Públicos
ECL	Servicios Públicos
ENDESA	Servicios Públicos
ENERSIS	Servicios Públicos
ENTEL	Servicios Públicos
GASCO	Servicios Públicos
LAN	Servicios Públicos
AZUL AZUL	Servicios Varios
CENCOSUD	Servicios Varios
COLO COLO	Servicios Varios
CRUZADOS	Servicios Varios
CRUZBLANCA	Servicios Varios
ENJOY	Servicios Varios
IAM	Servicios Varios
PAZ	Servicios Varios
SK	Servicios Varios
SONDA	Servicios Varios

Tabla 15. Clasificación Acciones año 2012 según su rubro

8.5 Clasificación de las Acciones del año 2013 según su rubro

Acción	Rubro
BESALCO	Construcción
PROVIDA	Inversiones e Inmobiliarias
QUINENCO	Inversiones e Inmobiliarias
SALFACORP	Inversiones e Inmobiliarias
SECURITY	Inversiones e Inmobiliarias
MASISA	Agropecuarias y Forestales
SCHWAGER	Agropecuarias y Forestales
ANDINA-A	Alimenticias y Bebidas
ANDINA-B	Alimenticias y Bebidas
CCU	Alimenticias y Bebidas
CONCHATORO	Alimenticias y Bebidas
EMBONOR-B	Alimenticias y Bebidas
IANSÁ	Alimenticias y Bebidas
BCI	Bancarias y Financieras
BSANTANDER	Bancarias y Financieras
CHILE	Bancarias y Financieras
CORPBANCA	Bancarias y Financieras
SM-CHILE B	Bancarias y Financieras
COPEC	Comerciales y Distribuidoras
FALABELLA	Comerciales y Distribuidoras
HITES	Comerciales y Distribuidoras
NUEVAPOLAR	Comerciales y Distribuidoras
RIPLEY	Comerciales y Distribuidoras
ZOFRI	Comerciales y Distribuidoras
FORUS	Industriales
HABITAT	Inversiones e Inmobiliarias
ILC	Inversiones e Inmobiliarias
INVERMAR	Inversiones e Inmobiliarias
LAS CONDES	Inversiones e Inmobiliarias
NORTEGRAN	Inversiones e Inmobiliarias
PARAUCO	Inversiones e Inmobiliarias
SOCOVESA	Inversiones e Inmobiliarias

Acción	Rubro
ANTARCHILE	Inversiones e inmobiliarias
BANMEDICA	Inversiones e inmobiliarias
SMSAAM	Marítimas y Navieras
VAPORES	Marítimas y Navieras
CAP	Metalmecánicas
MOLYMET	Metalmecánicas
SQM-B	Mineras
AQUACHILE	Pesqueras
INVEXANS	Pesqueras
MULTIFOODS	Pesqueras
CMPC	Productos Diversos
CFR	Productos Químicos
PEHUENCHE	Servicios Públicos
AESGENER	Servicios Públicos
AGUAS-A	Servicios Públicos
ALMENDRAL	Servicios Públicos
CGE	Servicios Públicos
COLBUN	Servicios Públicos
ECL	Servicios Públicos
ENDESA	Servicios Públicos
ENERSIS	Servicios Públicos
ENTEL	Servicios Públicos
GASCO	Servicios Públicos
LAN	Servicios Públicos
AZUL AZUL	Servicios Varios
CENCOSUD	Servicios Varios
COLO COLO	Servicios Varios
CRUZBLANCA	Servicios Varios
ENJOY	Servicios Varios
IAM	Servicios Varios
PAZ	Servicios Varios
SK	Servicios Varios
SONDA	Servicios Varios

Tabla 16. Clasificación Acciones año 2013 según su rubro

8.6 Clasificación de las Acciones del año 2014 según su rubro

Acción	Rubro
BESALCO	Construcción
PROVIDA	Inversiones e Inmobiliarias
QUINENCO	Inversiones e Inmobiliarias
SALFACORP	Inversiones e Inmobiliarias
SECURITY	Inversiones e Inmobiliarias
MASISA	Agropecuarias y Forestales
SCHWAGER	Agropecuarias y Forestales
ANDINA-A	Alimenticias y Bebidas
ANDINA-B	Alimenticias y Bebidas
CCU	Alimenticias y Bebidas
CONCHATORO	Alimenticias y Bebidas
EMBONOR-B	Alimenticias y Bebidas
IANSA	Alimenticias y Bebidas
BCI	Bancarias y Financieras
BSANTANDER	Bancarias y Financieras
CHILE	Bancarias y Financieras
CORPBANCA	Bancarias y Financieras
SM-CHILE B	Bancarias y Financieras
COPEC	Comerciales y Distribuidoras
FALABELLA	Comerciales y Distribuidoras
HITES	Comerciales y Distribuidoras
NUEVAPOLAR	Comerciales y Distribuidoras
RIPLEY	Comerciales y Distribuidoras
FORUS	Industriales
TECHPACK	Industriales
HABITAT	Inversiones e Inmobiliarias
ILC	Inversiones e Inmobiliarias
LAS CONDES	Inversiones e Inmobiliarias
MINERA	Inversiones e Inmobiliarias
NORTEGRAN	Inversiones e Inmobiliarias
PARAUCO	Inversiones e Inmobiliarias
SOCOVELSA	Inversiones e Inmobiliarias

ANTARCHILE	Inversiones e inmobiliarias
BANMEDICA	Inversiones e inmobiliarias
SMSAAM	Marítimas y Navieras
VAPORES	Marítimas y Navieras
CAP	Metalmecánicas
SQM-B	Mineras
AQUACHILE	Pesqueras
CAMANCHACA	Pesqueras
INVEXANS	Pesqueras
MULTIFOODS	Pesqueras
CMPC	Productos Diversos
CFR	Productos Químicos
AESGENER	Servicios Públicos
AGUAS-A	Servicios Públicos
ALMENDRAL	Servicios Públicos
CGE	Servicios Públicos
COLBUN	Servicios Públicos
ECL	Servicios Públicos
ENDESA	Servicios Públicos
ENERSIS	Servicios Públicos
ENTEL	Servicios Públicos
GASCO	Servicios Públicos
LAN	Servicios Públicos
AZUL AZUL	Servicios Varios
CENCOSUD	Servicios Varios
COLO COLO	Servicios Varios
CRUZBLANCA	Servicios Varios
ENJOY	Servicios Varios
IAM	Servicios Varios
PAZ	Servicios Varios
SK	Servicios Varios
SONDA	Servicios Varios

Tabla 17. Clasificación Acciones año 2014 según su rubro

8.7 Resultado Investigación Profesores Carola-Edinson Año 2009

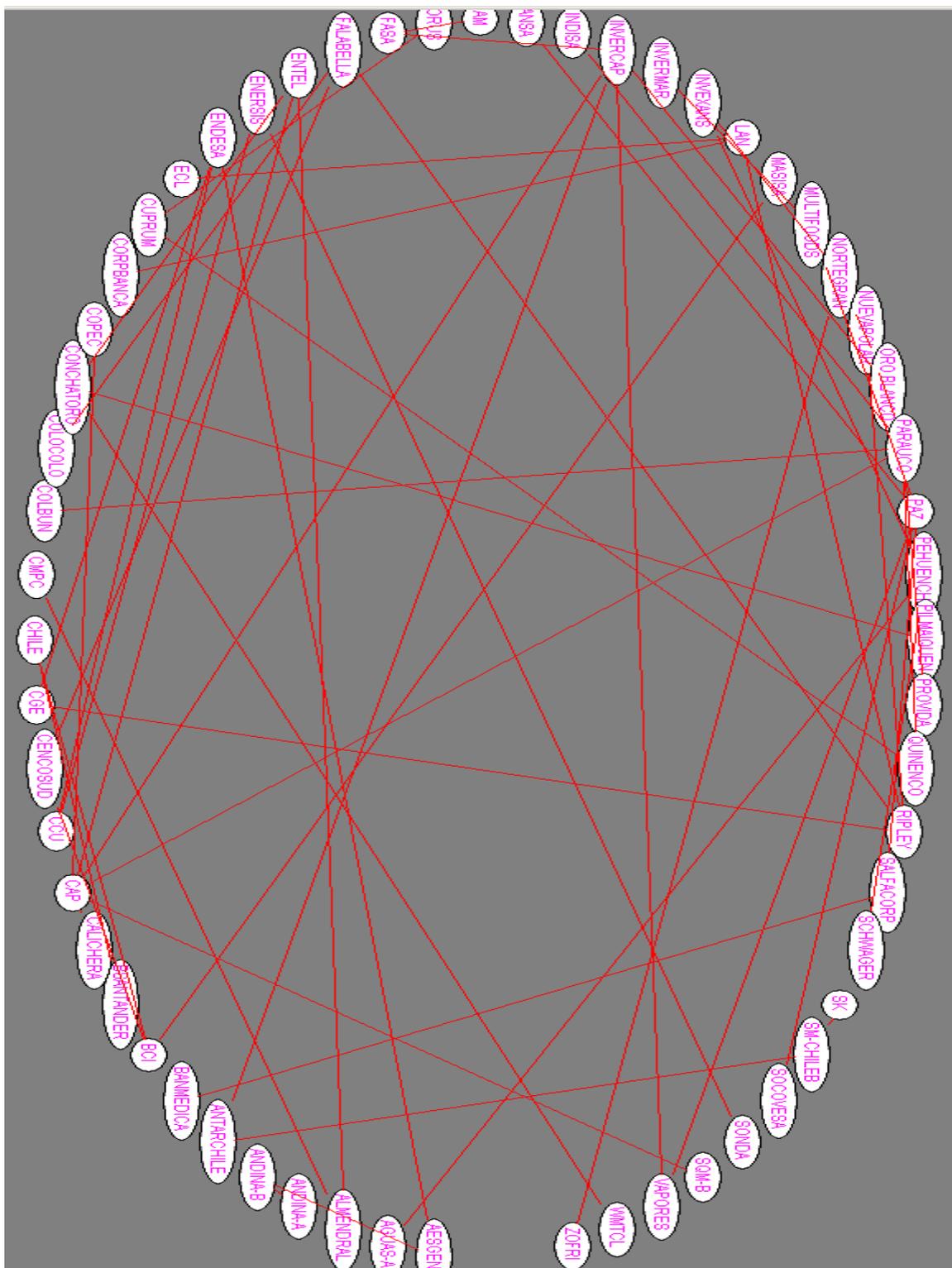


Figura 25. Resultado Kruskal Acciones Bursátiles Año 2009 (Figuerola Flores & Cornejo, 2014)

