

Análisis de preguntas de cuestionarios con Redes Neuronales

José de Jesús Zamarripa-Topete¹, Luz Evelia López-Chico¹ y Víctor Rafael Velázquez-Mejía¹

¹Universidad Autónoma de Baja California

Ensenada, B.C., México

jesuszamarripa@uabc.edu.mx, evelia@uabc.edu.mx, vvmejia@uabc.edu.mx

Paper received on 23/07/12, Accepted on 12/09/12.

Resumen. Se relacionan áreas conocimiento: 1. Diseño y Evaluación de Proyectos con cuestionario de Estudio Factibilidad; 2. Confiabilidad cuestionarios con alfa Cronbach; 3. Redes Neuronales para determinar preguntas a reelaborar. Alumnos en materias: Diseño y Evaluación Proyectos Electrónicos y Emprendedores, realizan cuestionario, pilotean y miden confiabilidad con alfa Cronbach, problema los menores de alfa 0.7. La Solución se integra: Módulo Estadístico y diez módulos Redes Neuronales. Estructura Redes Neuronales: capa entrada 10 neuronas, tres capas ocultas 50 neuronas, capa salida una neurona. Analizaron cuestionarios y Redes Neuronales establecieron preguntas con problemas, alumnos reelaboraron y en segundo piloteo lograron alfa mayor 0.7

Palabras Clave: Redes Neuronales, alfa de Cronbach.

1 Introducción

En este trabajo convergen tres áreas de conocimiento: 1. Diseño y Evaluación de Proyectos, en la Etapa de Inicio con el Estudio de Factibilidad del Proyecto y la elaboración de cuestionarios para determinar la opinión de los futuros usuarios del Proyecto; 2. El análisis estadístico de la confiabilidad de los cuestionarios con el alfa de Cronbach; 3. Las Redes Neuronales para establecer las preguntas a reelaborar en los cuestionarios que no alcanzan el alfa de Cronbach establecida.

La aportación principal del presente trabajo es la aplicación de las Redes Neuronales al analizar la información de la confiabilidad de los cuestionarios y establecer las preguntas que se deben hacer nuevamente, para mejorar el cuestionario y obtener un valor de alfa de Cronbach aceptable. La aportación secundaria es la propuesta de aplicar el alfa de Cronbach en el Diseño y Evaluación de Proyectos.

Esta aplicación se utilizó para apoyar a los estudiantes de Ingeniería en Electrónica del 9no semestre en las materias de Diseño y Evaluación de Proyectos Electrónicos y Emprendedores.

1.1 Diseño y Evaluación de Proyectos

Los alumnos proponen un Proyecto a desarrollar durante el ciclo escolar. Los Proyectos son: 1. Administrador inalámbrico de aparatos domésticos de bajo costo; 2. Tablero didáctico que determina el tipo de acentuación de palabras; 3. Cargador solar de celulares; 4. Control para un sistema doméstico de reciclaje de agua; 5. Llenador automático de garrafones de agua potable; 6. Casa inteligente.

En el Diseño y Evaluación de Proyectos, en su primera etapa se incluye el Estudio de Factibilidad para determinar la viabilidad del Proyecto. Las partes básicas del Estudio de Factibilidad son: 1. Análisis Técnico Operativo; 2. Análisis Económico Financiero; 3. Análisis de Normatividad Asociada; 5. Análisis de Mercado; 4. Análisis Social [1].

Parte del Análisis Social es la opinión del usuario con respecto al Proyecto: 1. Rechazo; 2. Poco Rechazo; 3. Indiferencia; 4 Poca Aceptación; 5. Aceptación.

Esto se realiza directamente con el usuario (fuente primaria), se emplean entrevista y cuestionarios, de diversos tipos [1].

Acotamiento del problema: 1. Los cuestionarios los realizan los alumnos. 2. La finalidad no es elaborar cuestionarios validados. 3. El tiempo disponible para realizar el cuestionario es de solo un par de semanas en el semestre escolar, por las fases de las materias que son: a) elegir el proyecto; b) determinar su factibilidad; c) diseñarlo; d) construirlo; e) probarlo; f) documentar el proceso; g) realizar una exposición del proyecto en funcionamiento [2].

1.2 Confiabilidad de cuestionarios con alfa de Cronbach

El cuestionario es el medio con el que se mide determinado comportamiento o conducta. La confiabilidad de un cuestionario es el grado en que sus resultados son consistentes y coherentes con el atributo a medir [6].

El índice de confiabilidad es el alfa de Cronbach que mide la consistencia interna del cuestionario. Sus valores son de 0 a 1, los valores mínimos de aceptación son 0.6 a 0.7, los valores entre 0.9 a 1 implican que existen redundancias [4]. Para este trabajo los valores de aceptación son de 0.7 a 0.9.

El alfa de Cronbach se emplea ampliamente en la Psicología, Ciencias Humanas, Ciencias de la Educación, etc. Pero muy poco en Diseño y Evaluación de Proyectos.

En este trabajo se calculará el alfa de Cronbach de los cuestionarios elaborados por los alumnos de sus respectivos Proyectos.

1.3 Redes Neuronales

La neurona es la representación en programación de su contraparte biológica, con respecto a sus entradas, procesamiento y activación [3].

La neurona presentada en la Figura 1 ($neurona_i$), se constituye de las entradas (x_1, x_2, \dots, x_j), los pesos ($w_{i1}, w_{i2}, \dots, w_{ij}$) que se calculan con el entrenamiento, las entradas se multiplican por su respectivo peso ($x_1w_{i1}, x_2w_{i2}, \dots, x_jw_{ij}$), en el cuerpo de la neurona se suman las multiplicaciones ($\sum_j x_jw_{ij}$), la salida (y_i) es el resultado de la

función de activación, para este estudio es la función Sigmoidea ($1/[1+e^{-x}]$). Las Redes Neuronales son arreglos de neuronas interconectadas, que se acomodan por capas: capa de entrada, una o varias capas ocultas y una capa de salida. Se utilizan las redes neuronales para identificar patrones o características en la información [3].

En el presente trabajo se emplearán las Redes Neuronales para analizar la información de la confiabilidad de los cuestionarios elaborados por los estudiantes.

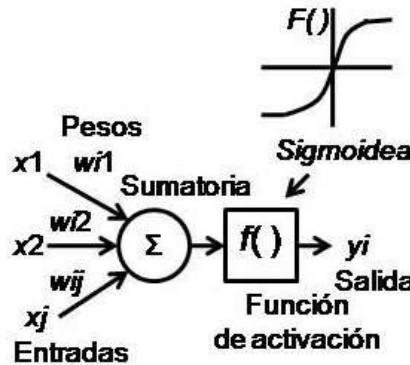


Figura 1. Neurona.

2. Desarrollo

2.1 Elaboración del cuestionario

Son seis equipos de alumnos que realizaron cada uno un cuestionario de su respectivo Proyecto. Por sus características los Proyectos pueden ser: 1. Mejorar un producto o servicio, que actualmente se ofrece; 2. Es nuevo y nadie lo ofrece [1].

Se define el perfil de los potenciales usuarios de los Proyectos:

1. Proyecto: Administrador inalámbrico de aparatos domésticos de bajo costo; Usuarios: con dispositivos WiFi.
2. Proyecto: Tablero didáctico que determina el tipo de acentuación de palabras; Usuarios: maestros de educación básica.
3. Proyecto: Cargador solar de celulares; Usuarios: con teléfono celular;
4. Proyecto: Control para un sistema doméstico de reciclaje de agua; Usuarios: con problema de suministro de agua.
5. Proyecto: Llenador automático de garrafones de agua potable; Usuarios: tiendas de autoservicio que vendan agua.
6. Proyecto: Casa inteligente; Usuarios: constructores o personas que deseen instrumentar viviendas.

A los alumnos se les entrega una plantilla para el desarrollo de preguntas del cuestionario (Tabla 1). Los alumnos elaboran diez preguntas con respuesta dicotómica de "sí" o "no", que cubren (en su caso) los indicadores, ejemplo de preguntas: ¿En el último año ha comprado un Smartphone?, ¿Compraría un cargador solar para su Smartphone? De esta forma se genera la versión 1.0 de su cuestionario.

Tabla 1. Plantilla para el desarrollo de preguntas del cuestionario.

Dimensión:	Indicadores:	Actividad:
Problema.	- Identificación del problema. - Importancia del problema.	
Uso de tecnología.	- Aceptación de la tecnología. - Uso de la tecnología.	
Gasto en tecnología.	- Gasto en el pasado en tecnología. - Intención de inversión en tecnología.	
Proveedor de producto o servicio (en su caso).	- Satisfacción con el proveedor actual. - Disposición a cambiar de proveedor.	
Aceptación del Proyecto.	- Intención de adquisición del Proyecto. - Intención de uso del Proyecto.	Elaborar diez preguntas dicotómicas que cubran los indicadores de su Proyecto.

2.2 Prueba Piloto de los cuestionarios versión 1.0

En el Diseño y Evaluación de Proyectos, el muestreo probabilístico para la aplicación de cuestionarios es poco usado [1]. En este trabajo se emplea el Muestreo de Conveniencia de Sitio, el cual reúne a los usuarios que cumplen con el perfil predefinido. Para algunos Proyectos se eligen tiendas que ofrezcan los últimos modelos tecnológicos relacionados al Proyecto.

Al ser una actividad académica y por los tiempos escolares (solo dos semanas), se proponen a diez personas por Proyecto para aplicarles la prueba piloto. Con respecto a la representatividad y evitar el sesgo, se solicita a los alumnos que cada persona elegida sea de diferente fraccionamiento, escuela o zona de la ciudad.

2.3 Codificación por cuestionario

Las personas se codificaron como “*Persona1*” a “*Persona10*” y las preguntas de “X1” a “X10”. A las respuestas se les asigna para el “*sí*” valor de “1” y para “*no*” el “0”.

El ejemplo de un cuestionario del proyecto 4 (control para un sistema doméstico de reciclaje de agua) versión 1.0 contestado se vacía en el formato de la Tabla 2.

Tabla 2. Formato Cuestionario de Proyecto 4 contestado.

Personas:	Preguntas:									
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Persona1	1	0	0	0	0	1	0	0	1	0
Persona2	1	1	0	1	1	1	1	0	0	0
Persona3	0	0	0	0	0	0	0	0	1	0
Persona4	1	1	1	1	1	0	0	1	1	0
Persona5	0	0	0	0	0	0	1	1	1	1
Persona6	1	1	1	0	1	1	1	1	1	1
Persona7	0	0	1	1	1	0	0	0	0	1
Persona8	0	0	1	1	1	1	1	1	1	1
Persona9	0	1	0	1	1	0	0	0	0	1
Persona10	1	1	0	0	1	1	1	1	1	1

2.4 Confiabilidad del cuestionario

Para determinar la confiabilidad del cuestionario se emplea el alfa de Cronbach (α) [5]:

Paso 1. Se obtiene la correlación entre las preguntas (r_{ij}) con la Ecuación 1 (r de Pearson). Requiere: la pregunta número i (X_i); la pregunta número j (X_j); el número de preguntas (n).

$$r_{ij} = \frac{\sum X_i X_j - \frac{(\sum X_i)(\sum X_j)}{n}}{\sqrt{\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n} \right] \left[\sum X_j^2 - \frac{(\sum X_j)^2}{n} \right]}} \quad (1)$$

Paso 2. Calcular la correlación promedio (\bar{r}) con la Ecuación 2. Los términos son: la correlación entre las preguntas (r_{ij}); el número de correlaciones (N).

$$\bar{r} = \frac{\sum r_{ij}}{N} \quad (2)$$

Paso 3. Obtener el alfa de Cronbach (α) de la Ecuación 3. Los elementos son: número de preguntas (n); correlación promedio (\bar{r}).

$$\alpha = \frac{n(\bar{r})}{1 + \bar{r}(n - 1)} \quad (3)$$

Los cuestionarios de los Proyectos obtuvieron los valores de alfa presentados en la Tabla 3. Para determinar que un cuestionario es confiable el valor de alfa debe ser entre 0.7 a 0.9, solo el "Cuestionario Proyecto 2" cumple (0.714) y puede ser aplicado a su público meta. El resto de los cuestionarios no cumplen y deben ser reelaborados [4]. El siguiente paso es determinar las preguntas a corregir.

Tabla 3. Alfa de los cuestionarios.

	Alfa de Cronbach (α):	$0.7 < \alpha < 0.9$
Cuestionario Proyecto 1	0.634	No
Cuestionario Proyecto 2	0.714	Si
Cuestionario Proyecto 3	0.230	No
Cuestionario Proyecto 4	0.661	No
Cuestionario Proyecto 5	0.595	No
Cuestionario Proyecto 6	0.290	No

2.5 Establecer las preguntas con problemas

Para el cuestionario en su versión 1.0 pilotado que no cumple con el alfa de Cronbach, se utiliza la Ecuación 1 para obtener la matriz de correlación inter-elemento, en la Tabla 4 se presenta el Cuestionario Proyecto 4. Esta matriz se procesa con redes neuronales para establecer las preguntas que tienen problemas.

Tabla 4. Matriz de correlación inter-elemento “ r_{ij} ” Cuestionario Proyecto 4.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
X1	$r_{1\ 1}$ 1.00	0.60	0.00	-0.20	0.22	0.60	0.20	0.20	0.22	-0.41
X2	0.60	1.00	0.00	0.20	0.66	0.20	0.20	0.20	-0.22	0.00
X3	0.00	0.00	1.00	0.41	0.54	0.00	0.00	0.41	0.09	0.25
X4	-0.20	0.20	0.41	1.00	0.66	-0.20	-0.20	-0.20	-0.66	0.00
X5	$r_{5\ 1}$ 0.22	0.66	0.54	0.66	$r_{5\ 5}$ 1.00	0.22	0.22	0.22	-0.43	$r_{5\ 10}$ 0.37
X6	0.60	0.20	0.00	-0.20	0.22	1.00	0.60	0.20	0.22	0.00
X7	0.20	0.20	0.00	-0.20	0.22	0.60	1.00	0.60	0.22	0.41
X8	0.20	0.20	0.41	-0.20	0.22	0.20	0.60	1.00	0.66	0.41
X9	0.22	-0.22	0.09	-0.66	-0.43	0.22	0.22	0.66	1.00	-0.09
X10	$r_{10\ 1}$ -0.41	0.00	0.25	0.00	$r_{10\ 5}$ 0.36	0.00	0.41	0.41	-0.09	$r_{10\ 10}$ 1.00

2.6 Integración de la solución

La solución presentada en la Figura 2, se integra de un Módulo Estadístico y diez Redes Neuronales.

La información de la Tabla 2 entra en el Módulo Estadístico y este entrega la matriz de correlación inter-elemento de la Tabla 4, las correlaciones son de -1 a 1, los valores negativos presentan las inconsistencias de la pregunta con respecto al cuestionario y los positivos la consistencia de la pregunta con el cuestionario.

Son diez Redes Neuronales por facilidad de su entrenamiento [8], cada Red Neuronal tiene la siguiente estructura (Figura 3): diez neuronas en la capa de entrada (1 a 10); 30 neuronas en la primera capa oculta (11 a 40); diez neuronas en la segunda capa oculta (41 a 50); 11 neuronas en la tercera capa oculta (51 a 61); una neurona en la capa de salida (62).

La información que entra a cada red neuronal corresponde a diez elementos de la matriz de correlación inter-elemento: $r_{1\ 1} - r_{1\ 10}$; $r_{2\ 1} - r_{2\ 10}$; ...; $r_{9\ 1} - r_{9\ 10}$; $r_{10\ 1} - r_{10\ 10}$. La salida de la Red Neuronal ($OutX1 \dots OutX10$) entrega el grado de problema (inconsistencia de la pregunta con respecto al cuestionario) de la pregunta ($X1, X2, \dots, X9, X10$) en el intervalo de 0.0 (sin problema) a 0.9 (problema grave) [8], con el resultado mayor a 0.3 (0.4 a 0.9) se debe reelaborar la pregunta. La solución se programó en C++ [7].

2.7 Entrenamiento de las Redes Neuronales

Las Redes Neuronales se entrenaron con aprendizaje supervisado de retropropagación, tolerancia 0.005, termino de momemtum 0.95 y razón de aprendizaje 0.5 [7].

Se realiza un entrenamiento segmentado [8]. El entrenamiento segmento 1: neurona 1 (capa entrada) y las neuronas 11 a 13 (capa oculta 1). Se tiene una entrada con 201 posibilidades (-1.0, -0.99...-0.01, 0, 0.01...0.99, 1.0) y 2 salidas (1, 0). Por lo tanto se tienen 201 vectores de entrenamiento. Se repite con las restantes (2 a 40).

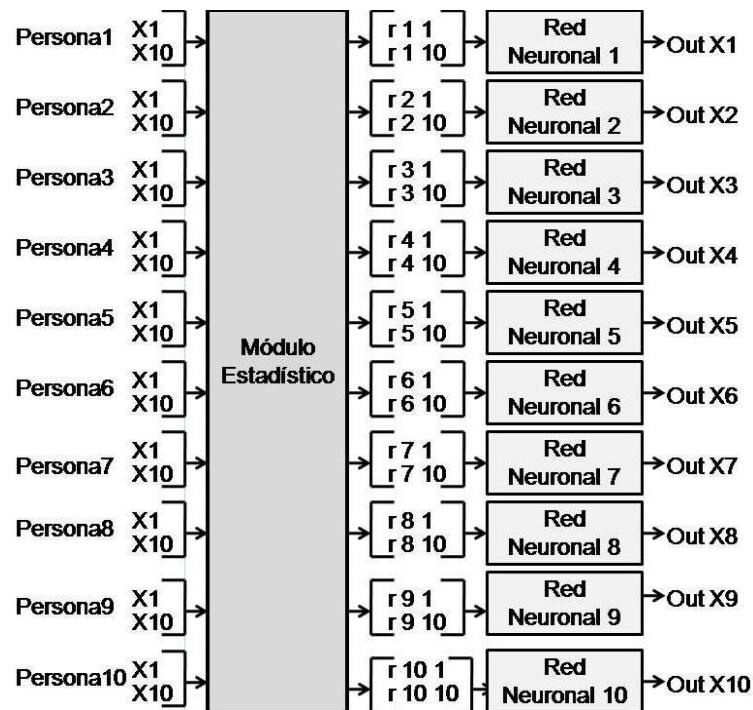


Figura 2. Integración de la solución.

El entrenamiento segmento 2: neuronas 41 a 50 (capa oculta 2), las neuronas 51 a 61 (capa oculta 3) y la neurona 62 (capa salida). Este segmento en cada entrada puede tener dos posibles valores (0 o 1) y la neurona de salida diez valores (0 a 0.9). Los vectores de entrenamiento requerido son: entradas (10) elevadas a las posibilidades (2), $10^2=100$ vectores de entrenamiento.

Se integran las Redes Neuronales entrenadas de forma segmentada, se prueban con vectores de entrada y salida conocidos para validar su entrenamiento.

3 Resultados

Con varias Redes Neuronales y su entrenamiento segmentado, el aprendizaje de las mismas se alcanzó en poco tiempo. Las Redes Neuronales con muchos elemen-

tos o con un número grande de vectores de entrenamiento, su aprendizaje puede demorar varios días o no ocurrir.

Los Cuestionarios de Proyecto que no cumplieron con el alfa de Cronbach (Tabla 3), se procesaron y se obtuvo la información de la Tabla 5. De los resultados se recomienda a los estudiantes que en su respectivo cuestionario reelaboren las siguientes preguntas con valores mayores de 0.3 (0.4 a 0.9): *Cuestionario Proyecto 1* la pregunta X4; *Cuestionario Proyecto 3* las preguntas X2, X3, X7; *Cuestionario Proyecto 4* la pregunta X4, X9; *Cuestionario Proyecto 5* las preguntas X4, X5; *Cuestionario Proyecto 6* las preguntas X2, X3, X6, X9, X10.

Los alumnos corrigieron las preguntas indicadas para generar la versión 2.0 de su cuestionario y realizaron una segunda prueba piloto con otros usuarios. Los cuestionarios contestados del segundo pilotaje se les calculó el alfa de Cronbach (Tabla 6), la totalidad cumplen con el valor de alfa y proceden a aplicarlos a sus respectivos públicos meta. Al utilizar la solución (programa) se sistematiza el análisis a los cuestionarios con problemas y permite a los alumnos trabajar en el tiempo asignado (dos semanas).

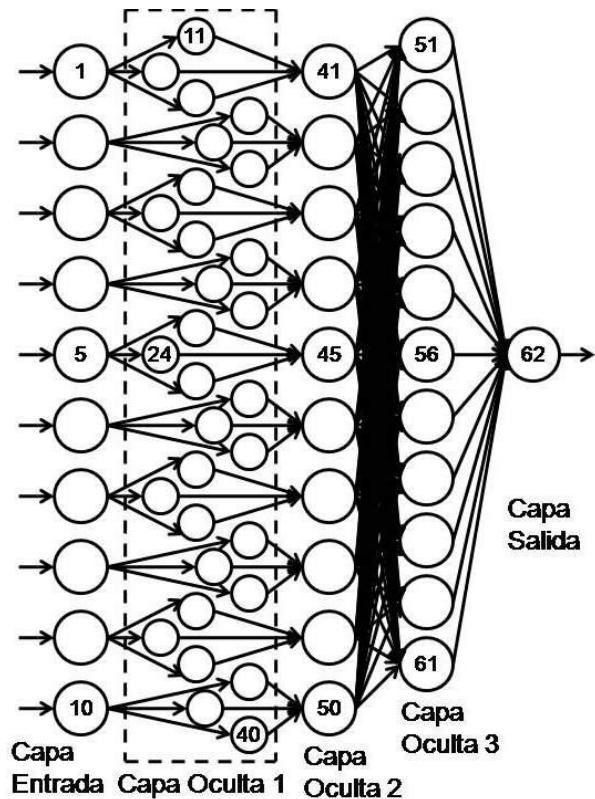


Figura. 3. Red Neuronal.

Tabla 5. Grado de problema de las preguntas.

	Pregunta:									
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Cuestionario Proyecto 1	0.2	0.2	0.1	0.4	0.2	0.0	0.0	0.0	0.3	0.2
Cuestionario Proyecto 3	0.2	0.7	0.6	0.3	0.3	0.1	0.6	0.3	0.2	0.3
Cuestionario Proyecto 4	0.2	0.1	0.0	0.5	0.1	0.1	0.1	0.1	0.4	0.2
Cuestionario Proyecto 5	0.2	0.0	0.1	0.6	0.5	0.2	0.1	0.2	0.2	0.1
Cuestionario Proyecto 6	0.3	0.4	0.5	0.1	0.3	0.5	0.1	0.1	0.6	0.5

Tabla 6. Alfa de Cronbach segundo piloteo.

	Alfa de Cronbach (α):	$0.7 < \alpha < 0.9$
Cuestionario Proyecto 1	0.722	Si
Cuestionario Proyecto 3	0.719	Si
Cuestionario Proyecto 4	0.751	Si
Cuestionario Proyecto 5	0.733	Si
Cuestionario Proyecto 6	0.729	Si

4 Conclusiones

Al tener la solución implementada se sistematiza el análisis de los cuestionarios con problemas. Con las Redes Neuronales se puede establecer cuáles son las preguntas que deben ser reelaboradas. Con esto se apoya a los alumnos para que realicen mejores cuestionarios con respecto a sus Proyectos y así su Estudio de Factibilidad se fortalece. De esta forma los alumnos cumplen en esta parte del desarrollo de proyectos en tiempo y forma. Al ser una solución en programa se puede instalar en un servidor y analizar los cuestionarios recibidos por correos electrónicos o contestados en línea. Al trabajar la solución con varias Redes Neuronales (diez) y estas con varias capas ocultas (tres) se facilitó el proceso de entrenamiento segmentado de las mismas. Las Redes Neuronales pueden ser una herramienta de apoyo para analizar información en las diversas fases del desarrollo de los Proyectos.

Referencias

1. Baca, G.: Evaluación de proyectos. 6^a Edición. Mc Graw Hill. ISBN 978-607-15-0260-5, (2010)

2. Cleand, D., King, W.: Manual para la administración de proyectos. Grupo Editorial Patria. ISBN 968-26-1151-2, (2007)
3. Martín del Brío, B., Sanz A.: Redes neuronales y sistemas borrosos. 3^a Edición. Alfaomega Ra-Ma. ISBN: 978-970-15-1250-0, (2007)
4. Oviedo, H. C., Campos, A.: Aproximación al uso del coeficiente alfa de Cronbach. Revista colombiana de psiquiatría. Vol XXXIV, número 004. ISSN 0034-7450, pp. 572-580, (2005)
5. Pagano, R.: Estadística para las ciencias del comportamiento. 7^a Edición. Thomson, ISBN 0-534-61767-0, (2006)
6. Pérez, H. E.: Estadística para las Ciencias Sociales, del Comportamiento y de la Salud. Cengage Learning. ISBN 978-970-686-491-1, (2007)
7. Rogers, J.: Object-Oriented Neural Networks in C++. Academic Press. ISBN 0125931158, (1997)
8. Sharkey, A. J.: Improving the expert networks of a modular multi-net system for pattern recognition. Springer. ISBN 978-3-540-38625-4, (2006)
9. Van Der Zwaag, B. J., Slump, C. H., Spaanenburg, L.: Process Identification Through Modular Neural Networks and Rule Extraction. World Scientific. ISBN 981-238-066-3, (2002)