



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

1 de 2

Neiva, 11 julio 2018

Señores

CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN

UNIVERSIDAD SURCOLOMBIANA

Ciudad

El (Los) suscrito(s):

Diego Mauricio Echeverri Suaza

con C.C. No. 7722852

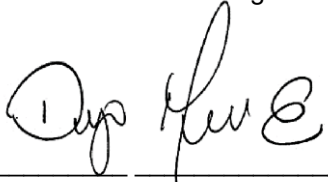
autor(es) de la tesis y/o trabajo de grado titulado "Estrategias Metodológicas Desde La Complejidad Para Integrar Las Redes Neuronales A Los Procesos Estocásticos Con Un Enfoque Interdisciplinario" presentado y aprobado en el año 2018 como requisito para optar al título de Magister En Estudios interdisciplinarios de la Complejidad.

Autorizo (amos) al CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN de la Universidad Surcolombiana para que con fines académicos, muestre al país y el exterior la producción intelectual de la Universidad Surcolombiana, a través de la visibilidad de su contenido de la siguiente manera:

- Los usuarios puedan consultar el contenido de este trabajo de grado en los sitios web que administra la Universidad, en bases de datos, repositorio digital, catálogos y en otros sitios web, redes y sistemas de información nacionales e internacionales "open access" y en las redes de información con las cuales tenga convenio la Institución.
- Permita la consulta, la reproducción y préstamo a los usuarios interesados en el contenido de este trabajo, para todos los usos que tengan finalidad académica, ya sea en formato Cd-Rom o digital desde internet, intranet, etc., y en general para cualquier formato conocido o por conocer, dentro de los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia.
- Continúo conservando los correspondientes derechos sin modificación o restricción alguna; puesto que de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación del derecho de autor y sus conexos.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

EL AUTOR/ESTUDIANTE: Diego Maurico Echeveri Suaza

Firma: 

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



**UNIVERSIDAD SURCOLOMBIANA
GESTIÓN SERVICIOS BIBLIOTECARIOS**



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

2 de 2

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	1 de 3
---------------	---------------------	----------------	----------	-----------------	-------------	---------------	---------------

TÍTULO COMPLETO DEL TRABAJO: Estrategias Metodológicas Desde La Complejidad Para Integrar Las Redes Neuronales A Los Procesos Estocásticos Con Un Enfoque Interdisciplinario

AUTOR O AUTORES:

Primero y Segundo Apellido	Primero y Segundo Nombre
Echeverri suaza	Diego Mauricio

DIRECTOR Y CODIRECTOR TESIS:

Primero y Segundo Apellido	Primero y Segundo Nombre
Londoño Betancourth	Gustavo

ASESOR (ES):

Primero y Segundo Apellido	Primero y Segundo Nombre

PARA OPTAR AL TÍTULO DE: Magister En Estudios Interdisciplinarios De La Complejidad

FACULTAD: Ciencias Exactas y Naturales

PROGRAMA O POSGRADO: Maestría en Estudios Interdisciplinarios de la Complejidad

CIUDAD: Neiva

AÑO DE PRESENTACIÓN: 2018 **NÚMERO DE PÁGINAS:**151

TIPO DE ILUSTRACIONES (Marcar con una X):

Diagramas_x_ Fotografías___ Grabaciones en discos___ Ilustraciones en general_x_ Grabados___
Láminas___ Litografías___ Mapas___ Música impresa___ Planos___ Retratos___ Sin ilustraciones___
Tablas o Cuadros_x_

SOFTWARE requerido y/o especializado para la lectura del documento: Lector PDF

MATERIAL ANEXO:

Vigilada mieducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO

CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	2 de 3
---------------	---------------------	----------------	----------	-----------------	-------------	---------------	---------------

PREMIO O DISTINCIÓN (En caso de ser LAUREADAS o Meritoria):

PALABRAS CLAVES EN ESPAÑOL E INGLÉS:

<u>Español</u>	<u>Inglés</u>	<u>Español</u>	<u>Inglés</u>
1. Estadística	Statistics	6. Complejidad	complexity
2. Modelización	modeling	7. Software	Software
3. RNA	RNA	8. Interdisciplinario	interdisciplinary
4. Bioestadística	Biostatistics	9. Sistemas Dinámicos	Dynamic systems
5. Regresión Lineal	Linear regression	10. Matemáticas Aplicadas	Applied mathematics

Estadística, modelización, Rna, Bioestadística, software, IBM Spss Statistic.

RESUMEN DEL CONTENIDO: (Máximo 250 palabras)

Se presenta el diseño de una propuesta metodológica para el programa de matemática aplicada en la facultad de ciencias exactas y naturales de la universidad Surcolombiana que, desde una perspectiva interdisciplinaria y partiendo de un análisis tanto de la estructura curricular como de los requerimientos planteados en los contextos internacional, nacional e institucional, permite incorporar las redes neuronales en el plan de estudios, enfocado hacia el perfil profesional: “El Matemático egresado de este programa curricular es un profesional con conocimientos bien fundamentados de la matemática universal y aplicaciones en: Biomatemáticas, Economía Matemática, Estadística y Sistemas Dinámicos; es un docente e investigador con capacidad de utilizar la modelización matemática, los sistemas de computación simbólica y TIC en la resolución de problemas interdisciplinarios identificados a través de la gestión y ejecución de proyectos”. Se referencia la situación actual del plan curricular y se plantea la necesidad de fomentar el trabajo interdisciplinario enfocado a la solución de problemas de complejidad creciente. Una de las estrategias consiste en cursar asignaturas de naturaleza interdisciplinaria como lo son las redes neuronales cuyo microdiseño curricular se elabora. Teniendo como referentes conceptuales: la formación del pensamiento matemático y en particular el estocástico, la incorporación de las TICs al currículo de matemáticas y la propuesta de matemáticas para la creatividad, se hace un análisis detallado de los conceptos involucrados en las redes neuronales como modelos estocásticos para la modelación y simulación de sistemas complejos. Mediante una metodología con enfoque cualitativo, se elabora la propuesta señalando los contenidos a abordar.

ABSTRACT: (Máximo 250 palabras)

The design of a methodological proposal for the program of applied mathematics in the Faculty of Exact and Natural Sciences of the Surcolombiana University is presented, from an interdisciplinary perspective and starting from an analysis of both the curricular structure and the requirements raised in the contexts international, national and institutional, allows to incorporate neural networks in the curriculum, focused on



CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	3 de 3
---------------	---------------------	----------------	----------	-----------------	-------------	---------------	---------------

the professional profile: "The mathematician graduated from this curricular program is a professional with well-grounded knowledge of universal mathematics and applications in: Biomathematics, Mathematical Economics , Statistics and Dynamic Systems; is a teacher and researcher with the ability to use mathematical modeling, symbolic computer systems and ICT in solving interdisciplinary problems identified through the management and execution of projects ". The current situation of the curricular plan is referenced and the need to promote interdisciplinary work focused on the solution of problems of increasing complexity. One of the strategies consists of taking subjects of an interdisciplinary nature, such as neural networks whose curriculum micro-design is elaborated. Taking as conceptual referents: the formation of mathematical thinking and in particular Stochastic, the incorporation of ICTs into the mathematics curriculum and the proposal of mathematics for creativity, a detailed analysis of the concepts involved in neural networks is done as stochastic models for the modeling and simulation of complex systems. Through a methodology with a qualitative approach, the proposal is drawn up, indicating the contents to be addressed.

APROBACION DE LA TESIS

Nombre Presidente Jurado:

Firma:

Nombre Jurado: Mauro Montealegre Cárdenas

Firma:

Nombre Jurado: Edgar Montealegre Cárdenas

Firma:

UNIVERSIDAD SURCOLOMBIANA
FACULTAD CIENCIAS EXACTAS Y NATURALES
MAESTRÍA EN ESTUDIOS INTERDISCIPLINARIOS DE LA COMPLEJIDAD



UNIVERSIDAD
SURCOLOMBIANA

NIT: 891180084-2

ESTRATEGIAS METODOLÓGICAS DESDE DA COMPLEJIDAD PARA INTEGRAR LAS
REDES NEURONALES A LOS PROCESOS ESTOCÁSTICOS CON UN ENFOQUE
INTERDISCIPLINARIO

AUTOR

Esp. DIEGO MAURICIO ECHEVERRI SUAZA

TRABAJO DIRIGIDO POR:

M.Sc. GUSTAVO LONDOÑO BETANCOURTH

NEIVA- HUILA, 18 DE JULIO DE 2018

A mi familia mis hijos, mi esposa que con su paciencia y apoyo contribuyo en salir adelante con esta propuesta.

A mis padres, y a mi abuela quienes con su dedicación y esfuerzo durante toda la vida han hecho posible este logro y a quienes debo todo lo que soy.

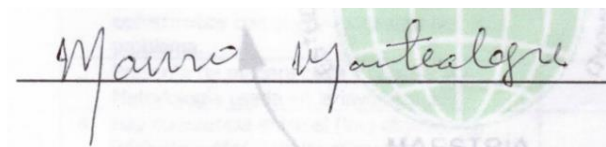
A la Corporación Universitaria del Huila CORHUILA por su apoyo.

A la vida que me ha dado otra oportunidad de seguir queriendo.

Nota de aceptación:

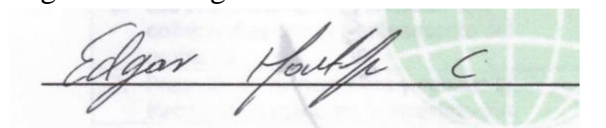
Aprobado por el Comité de Grado en cumplimiento de los requisitos exigidos por la Universidad Surcolombiana para optar al título de Magister en Estudios Interdisciplinarios de la Complejidad

Mauro Montealegre Cárdenas

A handwritten signature in black ink that reads "Mauro Montealegre". The signature is written over a horizontal line. In the background, there is a faint green circular logo with a globe-like pattern and the word "MAGISTERIA" visible at the bottom.

Jurado

Edgar Montealegre Cárdenas

A handwritten signature in black ink that reads "Edgar Montealegre C". The signature is written over a horizontal line. In the background, there is a faint green circular logo with a globe-like pattern and the word "MAGISTERIA" visible at the bottom.

Jurado

Neiva – Huila, 15 de junio de 2018

Agradecimientos

El autor expresa su agradecimiento a:

MAURO MONTEALEGRE CARDENAS, Doctor en Matemáticas. Profesor de la facultad de Ciencias Exactas y Naturales, Profesor de la facultad de Ciencias Exactas y Naturales, por sus aportes y conocimientos en el área.

GUSTAVO LONDOÑO BETANCOURTH, Magister En Matemáticas. director del presente proyecto por toda su colaboración paciencia y apoyo.

JASMIDT VERA CUENCA, Magister en Matemáticas, Profesora de la facultad de Ciencias Exactas y Naturales, por sus aportes y conocimientos en el área.

Tabla de contenido

Agradecimientos	III
Tabla de contenido	IV
Lista de tablas	VII
Lista de figuras.....	VIII
Resumen.....	1
Abstract.....	2
1. INTRODUCCIÓN	3
2. PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN.....	5
3. ANTECEDENTES Y JUSTIFICACIÓN	13
3.1. Antecedentes	13
3.1.1. Historia.....	13
3.1.2. Matemáticas para la creatividad.....	14
3.1.3. Modelización estadística con redes neuronales.	14
3.1.4. Redes Neuronales Artificiales aplicadas al análisis de datos.....	14
3.1.4.1. El proceso de autoevaluación del programa de matemática aplicada.	15
3.1.4.2. Condición 3: Contenido curricular.....	15
3.1.4.2.1. Fortalezas	15
3.1.4.2.2. Debilidades	16
3.1.4.2.3. Estrategias de mejoramiento	16
3.1.4.3. Condición 4: Organización de las actividades académicas.....	16
3.1.4.3.1. Fortalezas	16
3.1.4.3.2. Debilidades	16
3.1.4.3.3. Estrategias de mejoramiento	17
3.1.4.4. Condición 5: Investigación	17
3.1.4.4.1. Fortalezas	17
3.1.4.4.2. Debilidades	17
3.1.4.4.3. Estrategia de mejoramiento.....	17
3.2. Justificación	18
4. MARCO TEÓRICO.....	20
4.1. CARACTERÍSTICAS PRINCIPALES DE LAS REDES NEURONALES.....	20

4.1.1.	INTRODUCCIÓN A LAS REDES NEURONALES	20
4.1.2.	FUNCIONAMIENTO DE UNA NEURONA BIOLÓGICA.....	21
4.2.	CARACTERÍSTICAS DE UNA RED NEURONAL ARTIFICIAL.....	25
4.2.1.	Notación.....	27
4.2.2.	Funciones de Transferencia.....	28
4.2.2.1.	Limitador fuerte (Hardlim).....	29
4.2.2.2.	Función de transferencia lineal (purelin):.....	30
4.2.2.3.	Función de transferencia sigmoideal (logsig).....	30
4.2.3.	Topología de una Red.	31
4.2.3.1.	Capa de entrada	33
4.2.3.2.	Capas ocultas	33
4.2.3.3.	Capa de salida	33
4.3.	Teoría de la Complejidad (interdisciplinarietàad)	37
4.4.	Estadística y Probabilidad.....	39
4.4.1.	Ideas estocásticas fundamentales	39
4.4.2.	Análisis exploratorio de datos.....	43
4.4.2.1.	Características educativas del análisis exploratorio de datos	44
4.4.2.2.	Asociación.....	45
4.4.2.2.1.	Importancia en estadística.....	45
4.4.2.2.2.	Asociación y Correlación.....	45
4.4.2.2.3.	Papel de la asociación y regresión en los métodos estadísticos.....	45
4.4.2.2.4.	Asociación y causalidad.....	46
4.4.3.	Modelos de regresión.....	46
4.4.3.1.	Modelo de regresión lineal simple.....	46
4.4.3.1.1.	Métodos de Mínimos Cuadrados	47
4.4.3.1.2.	Valores estimados y residuales del modelo	47
4.4.3.1.3.	Supuestos del modelo de regresión.....	48
4.4.3.1.4.	Coefficiente de determinación	48
4.4.4.	Otros modelos de regresión	48
4.5.	NRA y modelos estadísticos	49
5.	OBJETIVOS	54

5.1.	Objetivo general.....	54
5.2.	Objetivos Específicos.....	54
6.	METODOLOGÍA.....	55
6.1.	Tipo y enfoque de la investigación	55
6.2.	Universo de estudio, población y muestra	56
6.3.	Estrategias Metodológicas	56
6.4.	Técnicas e instrumento de Investigación	57
7.	RESULTADOS Y DISCUSIÓN	58
7.1.	Análisis de la evaluación de las condiciones 3, 4, 5	58
7.2.	Análisis del plan de estudios.....	61
7.3.	Propuesta Curricular	73
7.4.	Estrategias curriculares	75
7.4.1.	Ejemplo 1: Perceptrón Multicapa Usando IBM Spss Statistics v22.....	75
7.4.1.1.	Usar un Perceptrón multicapa para evaluar el riesgo de crédito.....	75
7.4.1.2.	Preparación de los datos para su análisis	76
7.4.2.	Ejemplo 2: Función de base radial con IBM Spss Statistics.....	82
7.4.2.1.	Uso de la función de base radial para la clasificación de clientes de telecomunicaciones 82	
7.4.2.2.	Preparación de los datos para su análisis	83
8.	CONCLUSIONES	90
9.	RECOMENDACIONES.....	91
10.	BIBLIOGRAFÍA	92
11.	ANEXOS	96
	Anexo A: Microdiseño curricular fundamentos de las redes neuronales artificiales.....	96
	Anexo B Tutorial de Introducción a las Redes Neuronales con IBM Spss Statistics.....	103

Lista de tablas

Tabla 1. Funciones de Transferencia	31
Tabla 2. Equivalencia en la terminología usada en estadística y de redes neuronales.	49
Tabla 3. Equivalencia entre modelos estadísticos y modelos de red neuronal	50
Tabla 4. Análisis de las condiciones encontradas en el proceso de autoevaluación.....	58
Tabla 5. Análisis de los cursos del componente básico del programa aplicada de naturaleza matemático.....	61
Tabla 6. Contenidos de redes neuronales a incorporar en cursos del plan de estudios y sugerencias metodológicas	74
Tabla 7 Resumen de procesamiento de casos	78
Tabla 8 Información de red.....	79
Tabla 9 Resumen del modelo.....	81
Tabla 10 Clasificación	81
Tabla 11. Resumen de procesamiento de casos	85
Tabla 12. Información de red.....	86
Tabla 13. Resumen del modelo.....	88
Tabla 14. Clasificación	89
Tabla 1 Iconos de nivel de medición	107
Tabla 6 <i>Iconos de nivel de medición</i>	129

Lista de figuras

<i>Figura 1</i> Neuronas Biológicas	22
Figura 2 Cambios asociativos de las fuerzas sinápticas durante el aprendizaje	23
Figura 3 Comunicación entre neuronas	24
Figura 4. Proceso químico de una sinápsis	25
Figura 5 Neurona Artificial.....	26
Figura 6. De la neurona biológica a la neurona artificial.....	26
Figura 7. Proceso de una red neuronal.....	27
Figura 8. Neurona de una sola entrada.....	28
Figura 9. Función de transferencia Hardlim	29
Figura 10. Función de transferencia Hardlims.....	29
Figura 11. Función de transferencia lineal.....	30
Figura 12.Función de transferencia sigmoideal.....	30
Figura 13. Neurona con múltiples entradas	32
Figura 14. Neurona con múltiples entradas, notación abreviada	32
Figura 15. Capa de S neuronas	33
Figura 16. Capa de S neuronas con notación abreviada	34
Figura 17. Red de tres capas	34
Figura 18. Redes Recurrentes	35
Figura 19. Bloque de retardo	35
Figura 20. Bloque integrador	36
Figura 21. Red neuronal dinámica recurrente.....	36
Figura 22. Clasificación de las Redes Neuronales.....	37

Figura 23. Perceptrón simple con función lineal = Modelo de regresión lineal	51
Figura 24. Red Neuronal para la Regresión Lineal Múltiple Multidimensional	52
Figura 25. Red Neuronal para la Regresion Polinomial.....	53
Figura 26 Cuadro de diálogo Generadores de números aleatorios	76
Figura 27 Perceptrón multicapa: Pestaña Variables fuente	77
Figura 28 Perceptrón multicapa: Pestaña Particiones	78
Figura 29 Modelo simulación de entrenamiento	80
Figura 30 Cuadro de diálogo Generadores de números aleatorios	83
Figura 31 Función de base radial: Pestaña Variables	84
Figura 32 Función de base radial: Pestaña Resultados	85
Figura 33. modelo de red neuronal artificial.....	87
Figura 1 Arquitectura Feedforward con una capa oculta.....	105
Figura 2 Perceptrón Multicapa En Spss ruta: Analizar/redes Neuronales	110
Figura 3 pestaña Particiones	112
Figura 4 Pestaña Arquitectura.....	116
Figura 5 Pestaña Entrenamiento	119
Figura 6 Pestaña Resultado.....	122
Figura 7 Pestaña Guardar.....	124
Figura 8 Pestaña Exportar.....	125
Figura 9 Pestaña Opciones.....	127

Resumen

Se presenta el diseño de una propuesta metodológica para el programa de matemática aplicada en la facultad de ciencias exactas y naturales de la universidad Surcolombiana que, desde una perspectiva interdisciplinaria y partiendo de un análisis tanto de la estructura curricular como de los requerimientos planteados en los contextos internacional, nacional e institucional, permite incorporar las redes neuronales en el plan de estudios, enfocado hacia el perfil profesional: “El Matemático egresado de este programa curricular es un profesional con conocimientos bien fundamentados de la matemática universal y aplicaciones en: Biomatemáticas, Economía Matemática, Estadística y Sistemas Dinámicos; es un docente e investigador con capacidad de utilizar la modelización matemática, los sistemas de computación simbólica y TIC en la resolución de problemas interdisciplinarios identificados a través de la gestión y ejecución de proyectos”(Documento maestro, 2015).

Se referencia la situación actual del plan curricular y se plantea la necesidad de fomentar el trabajo interdisciplinario enfocado a la solución de problemas de complejidad creciente. Una de las estrategias consiste en cursar asignaturas de naturaleza interdisciplinaria como lo son las redes neuronales cuyo microdiseño curricular se elabora. Teniendo como referentes conceptuales: la formación del pensamiento matemático y en particular el estocástico, la incorporación de las TICs al currículo de matemáticas y la propuesta de matemáticas para la creatividad, se hace un análisis detallado de los conceptos involucrados en las redes neuronales como modelos estocásticos para la modelación y simulación de sistemas complejos. Mediante una metodología con enfoque cualitativo, se elabora la propuesta señalando los contenidos a abordar, relaciones con otras asignaturas del currículo y sugerencias metodológicas.

Palabras Claves: Estadística, modelización, Rna, Bioestadística, software, IBM Spss Statistic.

Abstract

The design of a methodological proposal for the program of applied mathematics in the Faculty of Exact and Natural Sciences of the Surcolombiana University is presented, from an interdisciplinary perspective and starting from an analysis of both the curricular structure and the requirements raised in the contexts international, national and institutional, allows to incorporate neural networks in the curriculum, focused on the professional profile: “The mathematician graduated from this curricular program is a professional with well-grounded knowledge of universal mathematics and applications in: Biomathematics, Mathematical Economics , Statistics and Dynamic Systems; is a teacher and researcher with the ability to use mathematical modeling, symbolic computer systems and ICT in solving interdisciplinary problems identified through the management and execution of projects “(Master document, 2015).

The current situation of the curricular plan is referenced and the need to promote interdisciplinary work focused on the solution of problems of increasing complexity. One of the strategies consists of taking subjects of an interdisciplinary nature, such as neural networks whose curriculum micro-design is elaborated. Taking as conceptual referents: the formation of mathematical thinking and in particular Stochastic, the incorporation of ICTs into the mathematics curriculum and the proposal of mathematics for creativity, a detailed analysis of the concepts involved in neural networks is done as stochastic models for the modeling and simulation of complex systems. Through a methodology with a qualitative approach, the proposal is drawn up, indicating the contents to be addressed, relations with other subjects of the curriculum and methodological suggestions.

Keywords: Statistics, modeling, Rna, Biostatistics, software, IBM Spss Statistic.

1. INTRODUCCIÓN

Este trabajo pretende ser un primer paso práctico hacia una plena comprensión de las relaciones entre los modelos biológico y artificial de las redes neuronales, dada su importancia en muchos campos de aplicación, en particular para la identificación de sistemas dinámicos no lineales.

Está basado en los siguientes trabajos: Un documento de la Facultad de Ingeniería de la Universidad Tecnológica de Pereira¹. sobre aplicaciones de las redes neuronales a problemas de Ingeniería Eléctrica; Tesis de Maestría de Jasmidt Vera², donde se estudia la dinámica del modelo de Fitzhugh-Nagumo y su relación con las redes de Hopfield; y el Toolbox de IBM Spss Statistic sobre redes neuronales y el trabajo de tesis de especialización de Diego Mauricio Echeverri Suaza, Tutorial sobre redes neuronales artificiales y algunas aplicaciones utilizando Matlab.

Se encontrará un bosquejo general sobre el desarrollo de la teoría de las redes neuronales artificiales desde su nacimiento hasta los últimos aportes que se han hecho; se encontrará también un resumen especializado acerca del funcionamiento del cerebro en general y de cómo ocurre el proceso neuronal, para terminar, haciendo la analogía que llevó a la transición del modelo biológico al modelo artificial.

Gran parte es dedicado a la explicación de los principales tipos de Redes Neuronales Artificiales, allí se profundiza en los aspectos fundamentales de cada tipo de red, sus antecedentes y evolución histórica, la estructura general de la red y el desarrollo matemático.

La parte central, muestra las principales relaciones entre la integración de las RNA, modelos de regresión y IBM Spss para el diseño, entrenamiento y simulación de redes neuronales. Se ilustra con ejemplos de las redes discutidas los referentes teóricos.

Es de aclarar que el estudio de esta teoría involucra además de tiempo y dedicación una buena fundamentación matemática y estadística. En la elaboración de los currículos de los cursos de estadística descriptiva, inferencial, probabilidades y Bioestadística, usando las nociones fundamentales del pensamiento complejo y la teoría de la complejidad, clave para estudiantes cuya formación tiene el enfoque de matemáticas y áreas afines, Este tema de tesis sea el primer paso para una cadena de proyectos de investigación que permitan a la facultad de ciencias exactas y naturales, al programa de matemática aplicada y a la maestría en estudios interdisciplinarios de

¹ Para ampliar esta información puede consultar, I. Acosta, M. Zuluaga C, Tutorial sobre redes neuronales aplicadas en ingeniería eléctrica. Universidad Tecnológica de Pereira, 2000

² Para ampliar esta información puede consultar, Vera Jasmidt, El modelo de Fitzhugh-Nagumo aplicado a una red neuronal, Tesis de maestría, universidad Eafit, Medellín, 2007

la complejidad y al grupo Dinusco ofrecer más temas de investigación propios y a los semilleros del mismo.

2. PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN

El pensamiento sistémico siempre ha estado presente en la historia de la humanidad. Para (O'CONNOR, 1998): "...entendido como la capacidad del ser humano para reconocer y aplicar los principios de la realimentación, las propiedades emergentes y el pensamiento circular tuvo su origen en el siglo III a.c. en Alejandría, cuando un griego llamado Ktesibios inventó un reloj de agua con un depósito autorregulable. Un siglo después, Herón de Alejandría, perfeccionó el trabajo de Ktesibios e inventó varias válvulas flotantes basándose en el mismo principio. Hasta el siglo XVII, cualquier sistema de realimentación mecánica se basaba en alguno de los prototipos inventados por Herón. En el campo de la medicina y la fisiología, el médico inglés William Harvey, en 1628 descubrió la circulación de la sangre, explicando cómo es bombeada por el corazón por todo el organismo y refutando la teoría de que el hígado era el órgano principal del sistema circulatorio. Desde entonces, la medicina ha descubierto muchos de los sistemas del cuerpo humano, los cuales son homeostáticos, es decir, se autorregulan y se acoplan para que el organismo funcione." Varios autores (O'CONNOR, 1998), (Montealegre, Londoño, & Polanía, 2002), (Briggs & Peat, 1994), señalan las etapas más destacadas en la historia del pensamiento sistémico. La Cibernética apareció en 1948, con Norbert Wiener, profesor de matemáticas del Instituto Tecnológico de Massachusetts (MIT), quien la definió como "la ciencia de la comunicación y el control de animales y máquinas".

La cibernética se centra en el funcionamiento de los sistemas, ya sean vivos, mecánicos o sociales. El principio básico de la cibernética fue la autorregulación de los sistemas mediante la realimentación, que Wiener definió como "método de controlar un sistema, que consiste en volver a introducirle los resultados de una ejecución anterior". En 1954 se publica *La Teoría General de Sistemas*, surgida del trabajo del biólogo Ludwig von Bertalanffy. En 1961, Jay Forrester aplicó los principios de la cibernética a los problemas de los sistemas económicos. Desde la década de 1960, se ha desarrollado una nueva rama de la cibernética que centra su estudio en la relación entre el observador y el sistema sometido a observación: La cibernética de segundo orden, que surge con el trabajo de Heinz von Foerster, explorando cómo las personas construyen modelos mentales de los sistemas con los que interactúan, sobre el principio de que no es posible separar a los observadores de los sistemas que están describiendo y, por tanto, debe haber realimentación entre observador y sistema observado, así como dentro del propio sistema en cuestión. A finales del siglo pasado, el matemático y físico Henri Poincaré cuestionó la perfección newtoniana en relación con las órbitas planetarias, lo que se conoce como el problema de los tres cuerpos. Planteaba una atracción gravitatoria múltiple, que hasta entonces se resolvía con las leyes de Newton y la suma de un pequeño valor que compensara la atracción del tercer elemento. Poincaré descubrió que, en situaciones críticas, ese tirón gravitatorio mínimo podía realimentarse hasta producir un efecto de resonancia que modificara la órbita o incluso lanzara el planeta fuera del sistema solar.

Los procesos de realimentación se corresponden en física con las ecuaciones iterativas, donde el resultado del proceso es utilizado nuevamente como punto de partida para el mismo proceso. De esta forma se constituyen los sistemas no lineales, que abarcan el 90% de los objetos existentes. El ideal clásico sólo contemplaba sistemas lineales, en los que efecto y causa se identifican plenamente; se sumaban las partes y se obtenía la totalidad. Poincaré introdujo el fantasma de la no linealidad, donde origen y resultado divergen y las fórmulas no sirven para resolver el sistema. Se había dado el primer paso hacia la Teoría del Caos. El término Caos se refiere a una interconexión subyacente que se manifiesta en acontecimientos aparentemente aleatorios. En la turbulencia de un arroyo es imposible predecir la trayectoria de una partícula de agua. Sin embargo, ese sistema es, a la vez, “*continuamente cambiante y siempre estable*”. Si tiramos una piedra al agua el sistema no se desestabilizará, cosa que sí ocurriría en un sistema no caótico. Esto es una metáfora de *nosotros mismos*: somos la misma persona que hace diez años, sin embargo, hace diez años estábamos formados por unos átomos diferentes y psicológicamente también somos diferentes. ¿Por qué un sistema caótico es tan cambiante? Porque todo está influido por todo. Todo está interconectado con todo. ¿Por qué un sistema caótico es, a la vez, tan estable? Por las interconexiones sutiles que se forman al estar todo influido por todo.

En 1960, el meteorólogo Edward Lorenz dio, sin proponérselo, el segundo paso hacia la Teoría del Caos. Entusiasta del tiempo, se dedicaba a estudiar las leyes atmosféricas y realizar simulaciones a partir de sus parámetros más elementales. Un día, para estudiar con más detenimiento una sucesión de datos, copió los números de la impresión anterior y los introdujo en la máquina. El resultado le conmocionó. Su tiempo, a escasa distancia del punto de partida, divergía algo del obtenido con anterioridad, pero al cabo de pocos meses -ficticios- las pautas perdían la semejanza por completo. Lorenz examinó sus números y descubrió que el problema se hallaba en los decimales; el ordenador guardaba seis, pero para ahorrar espacio él sólo introdujo tres, convencido de que el resultado apenas se resentiría.

Esta inocente actuación fijó el final de los pronósticos a largo plazo y puso de manifiesto la extremada sensibilidad de los sistemas no lineales: el llamado “efecto mariposa” o “dependencia sensible de las condiciones iniciales”. Se trata de la influencia que la más mínima perturbación en el estado inicial del sistema puede tener sobre el resultado final o, como recoge el escritor James Gleick (James, 1988), “si agita hoy, con su aleteo, el aire de Pekín, una mariposa puede modificar los sistemas climáticos de Nueva York el mes que viene”.

La iteración ofrece resultados estables hasta cierto punto, pero cuando éste se supera el sistema se derrumba en el caos. Los científicos (Briggs & Peat, 1994), aplican esta idea al ciclo vital humano: “Nuestro envejecimiento se puede abordar iteración constante de nuestras células al fin introduce un plegamiento y una como un proceso donde la divergencia que altera nuestras condiciones iniciales y lentamente nos desintegra”.

Este breve panorama constituye dos de las tres comprensiones de la complejidad del mundo y de la naturaleza (C. Maldonado, 2009) “la complejidad es un problema no una cosmovisión”: la complejidad como método o el pensamiento complejo de Edgar Morín, la complejidad como cosmovisión o enfoque sistémico que le adscribe un papel determinante al sujeto “observador” y la complejidad como ciencia o teoría (Ciencias) de la complejidad, la cual debido a dificultades principalmente lingüísticas, computacionales y matemáticas es la menos conocida por el gran público. No obstante, es igualmente, aquella que ha realizado las mayores contribuciones en dominios interdisciplinarios. Es precisamente desde esta última comprensión que se fundamenta teóricamente el programa de matemática aplicada de la universidad Surcolombiana: “El abordaje de la interdisciplinariedad desde el paradigma de los sistemas complejos,... se hacen con elementos heterogéneos en interacción, construyendo subsistemas que se mueven en diversas disciplinas y que se caracteriza por una serie de propiedades que incluyen: *Multiplicidad de perspectivas legítimas...*, *No linealidad...*, *Sistemas Emergentes...*, *Auto-organización...*, *Multiplicidad de escala...*, *Incertidumbre irreducible...*, *Modelizaciones*”.³

Para Carlos Maldonado (C. E. Maldonado & Gómez-Cruz, 2011), las ciencias de la complejidad estudian fenómenos, sistemas o comportamientos de complejidad *creciente*; esto es, fenómenos y sistemas que aprenden y se adaptan, y que, en el filo del caos o bien, lo que es equivalente, lejos del equilibrio, responden a la flecha del tiempo de la termodinámica del no-equilibrio (Nicholis & Prigogine, 1994). Por consiguiente, no se ocupan de todos y cada uno de los fenómenos y sistemas del mundo, puesto que no todas las cosas son complejas (C. E. Maldonado, 2014). Desde la perspectiva de las ciencias de la complejidad, los sistemas son simples o de complejidad creciente producto de la emergencia, no linealidad, aleatoriedad, interconexión, multiplicidad de escalas, entre otras características de los fenómenos estudiados.

El Paradigma de la Complejidad, según Clara Romero (Perez, 2003), aglutina a científicos de diversos campos de conocimiento que insisten en la conveniencia de adoptar nuevos modelos teóricos, metodológicos y, por ende, una nueva epistemología, que permita a la comunidad científica elaborar teorías más ajustadas de la realidad que posibilite, al mismo tiempo, diseñar y poner en práctica modelos de intervención –social, sanitaria, educativa, política, económica, ambiental, cultural, etc.– más eficaces que ayuden a pilotar y regular las acciones individuales y colectivas. Subyace en esta actitud reformista un firme intento de reformar la racionalidad sobre la que la ciencia y la tecnología se han venido apoyando. Esta actitud reformista afecta indistintamente tanto a las ciencias empírico-naturales como a las ciencias sociales y humanas y, en consecuencia, incide asimismo en las Ciencias de la Educación (Perez, 2003).

Las ciencias de la complejidad constituyen un conjunto o grupo de ciencias, disciplinas, teorías, lenguajes, métodos, metodologías, aproximaciones y lenguajes cuyo rasgo distintivo, según Maldonado (C. E. Maldonado, 2009), es la no-disciplinariedad o interdisciplinariedad.

³ Tomado de Condiciones de calidad para renovación de registro calificado. Documento maestro, 2015. Programa matemática aplicada, Universidad Surcolombiana

Exactamente en este sentido, las ciencias de la complejidad implican una revolución consistente en *indisciplinar* el conocimiento, las instituciones, el Estado y la sociedad. Esta indisciplinarización significa exactamente dejar de pensar en términos de disciplinas... Pues bien, la primera enseñanza al respecto consiste en el reconocimiento explícito de que la ciencia y las disciplinas no tienen ya, en absoluto, ningún objeto. Contra la tradición clásica que sostenía que la ciencia se caracteriza(ba) por tener un método propio, un lenguaje, una tradición, un ámbito y un modo propio de organización, las ciencias de la complejidad ponen abiertamente sobre la mesa, a plena luz del día, que la buena ciencia de punta hoy no se define por tener objetos sino, mejor y más radicalmente, por tener problemas. Al respecto, se plantea los siguientes interrogantes: ¿se enseña y se aprende en complejidad de la misma manera que se enseña y se aprende en la ciencia normal?, o ¿se aprende y se enseña de forma diferente?, ¿se investiga en complejidad de la misma manera como se investiga normalmente en la ciencia en la corriente principal de pensamiento?, ¿se investiga de otra forma?, ¿los puentes entre la ciencia y la sociedad pueden y deben ser los mismos que en el caso de la ciencia normal, o bien es preciso pensar en nuevos y diferentes canales y puentes?. Estos interrogantes se convierten por tanto en un marco de referencia para orientar la investigación que se pretende desarrollar en esta tesis.

En un contexto global, la comunidad académica de las ciencias de la complejidad integrada por profesores, investigadores y estudiosos e interesados por la complejidad en general es creciente. Sin embargo, según Maldonado (C. E. Maldonado & Gómez-Cruz, 2011) todos son complejólogos de primera generación. Se han formado en una disciplina, en una ciencia particular, y por diversas razones, han accedido a pensar, a trabajar y a vivir en términos de inter trans o multidisciplinariedad. No existen aún complejólogos de segunda generación. Que serán todos aquellos que se hayan formado-es decir, en, desde pregrado- en ciencias de la complejidad. Previsiblemente, en el futuro inmediato seguiremos siendo, todos, complejólogos de primera generación hasta tanto no se creen currículos, carreras y programas académicos en complejidad. A lo sumo, asistimos, en Colombia y en el mundo, a la presencia de materias (cursos, seminarios, etc.) de complejidad, con mayor o menor intensidad y transversalidad. Pero no aún asistimos a la creación de currículos de pregrado y postgrado enteramente dedicados al trabajo, formación e investigación en ciencias de la complejidad. En definitiva y en los términos de Maldonado, el mundo académico ha asimilado e incorporado en su quehacer las comprensiones de pensamiento complejo y la cibernética, pero aún no ha podido coptar a las ciencias de la complejidad.

Para el correcto desarrollo de esta tesis, cuyo propósito es diseñar una propuesta metodológica para el programa de matemática aplicada en la facultad de ciencias exactas y naturales de la universidad Surcolombiana que, desde una perspectiva interdisciplinaria y partiendo de un análisis tanto de la estructura curricular como de los requerimientos planteados en los contextos internacional, nacional e institucional, permita incorporar las redes neuronales en el plan de estudios, contribuyendo directamente con el perfil profesional: “El Matemático egresado de este programa curricular es un profesional con conocimientos bien fundamentados de la matemática universal y aplicaciones en: Biomatemáticas, Economía Matemática, Estadística y Sistemas

Dinámicos; es un docente e investigador con capacidad de utilizar la modelización matemática, los sistemas de computación simbólica y TIC en la resolución de problemas interdisciplinarios identificados a través de la gestión y ejecución de proyectos⁴”(Documento maestro, 2015), se hace necesario referenciar la situación actual del plan curricular para responder a las necesidades de formación en matemáticas a nivel de pregrado acorde con los requerimientos actuales. Interesa especialmente conocer cómo se fomenta el trabajo interdisciplinario enfocado a la solución de problemas de complejidad creciente. Al respecto, cabe decir que institucionalmente la universidad Surcolombiana ha adoptado el enfoque de la complejidad en su PEU⁵, donde se plantean planes de estudio basados en la identificación de problemas complejos para definir las competencias que debe desarrollar el estudiante orientando así el diseño curricular. De todas maneras, el currículo sigue siendo disciplinar y en este contexto deben formularse las propuestas que apunten al desarrollo del pensamiento complejo mediante estrategias tales como: vinculación a grupos y semilleros de investigación, proyectos de proyección social, desarrollo institucional, asignaturas de naturaleza interdisciplinaria, cursar asignaturas en otros programas entre otras.

Estas estrategias han sido sugeridas por (Badilla Saxe, 2009) “Desde la perspectiva de un nuevo paradigma emergente en educación, el diseño curricular debe evolucionar de una organización fragmentada y dividida en materias y disciplinas, hacia una concepción más orgánica, comprensiva y holista. La misión de todo sistema educativo que busque renovarse y atender la realidad social, científica, emocional, ecológica, artística y tecnológica actual debe ser tender hacia diseños curriculares holistas, orgánicos y complejos. Entre tanto, y mientras debemos atender planes y programas de estudio fragmentados y divididos, podemos echar mano de enfoques por proyectos o ejes transversales para procurar darles una integración...”

Teniendo en cuenta que este modelo curricular fragmentado en materias y disciplinas, como lo señalaba Maldonado (C. E. Maldonado & Gómez-Cruz, 2011), permanecerá por cierto tiempo hasta que se creen programas de pregrado para la formación de complejólogos de segunda generación, es en su contexto que se puntualiza un aspecto problemático que soporta en gran medida la necesidad de una propuesta para fomentar la interdisciplinariedad, mediante el desarrollo de proyectos para el estudio de sistemas de complejidad creciente y de esta manera, seguir formando complejólogos de primera generación.

He aquí la importancia y pertinencia de las instituciones de educación superior en la medida en que estas proponen y generan ambientes de aprendizaje para los estudiantes que por lo general están enmarcados en el modelo “transmisionista”. Precisamente, para este caso, se ha considerado pertinente revisar, atendiendo a la perspectiva interdisciplinar, el plan de estudios del programa de

⁴ Para ampliar esta información puede consultar Condiciones de calidad para renovación de registro calificado. Documento maestro, 2015. Programa matemática aplicada, Universidad Surcolombiana

⁵ Para ampliar esta información puede consultar ACUERDO NÚMERO 010 DE 2016 (11 de marzo) Por medio del cual se adopta el Proyecto Educativo Universitario P.E.U. de la Universidad Surcolombiana

matemática aplicada de la universidad Surcolombiana, lo cual, a su vez, nos lleva a profundizar en la caracterización de los programas de pregrado en matemáticas.

Estos surgieron ante la necesidad de formar el recurso humano para la investigación y para la docencia a nivel superior. A través de diferentes proyectos, se ha definido un marco común que orienta la formación de matemáticos a nivel mundial (Marco de fundamentación conceptual y especificaciones de la prueba ECAES. Asociación Colombiana de Facultades de Ciencias – ACOFACIEN- 2014). Este marco común incluye las siguientes áreas: cálculo en una y varias variables reales, álgebra lineal, ecuaciones diferenciales, variable compleja, probabilidad y estadística, métodos numéricos, geometría diferencial, estructuras algebraicas, matemáticas discretas, algoritmos y programación. Así mismo el marco define que las competencias para cualquier graduado en matemáticas son: la capacidad de idear demostraciones, la capacidad de modelar matemáticamente una situación y la capacidad de resolver problemas con técnicas matemáticas. Los programas deben tener la suficiente flexibilidad para incorporar las competencias que puedan emerger con el tiempo.

En un mundo globalizado la educación superior tiende a hacerlo para responder a estándares internacionales y nacionales. Este es el caso de los programas de pregrado en general, para los cuales se han fijado ciertos referentes como condiciones de calidad exigidas por el MEN. Para los propósitos de esta investigación, consideramos las siguientes⁶: (Informe de autoevaluación. Programa matemática aplicada. Universidad Surcolombiana, 2015):

- **Condición 3 contenido curricular:** Con base en la denominación del Programa, en los diagnósticos de la Justificación y en sus rasgos distintivos, debe establecer la pertinencia de la fundamentación teórica, de los propósitos formativos, de las competencias, de los perfiles profesional y ocupacional, del plan general de estudios con los respectivos contenidos de las actividades académicas y del componente de interdisciplinariedad.
- **Condición 4 organización de las actividades académicas:** Organización de las actividades académicas del programa (laboratorios, talleres, seminarios etc.), los cuales han de tener coherencia con la metodología, la estructura establecida por la institución, los lineamientos pedagógicos y didácticos, con el fin de lograr las metas de formación.
- **Condición 5 investigación:** El programa y la Institución tienen previstas las políticas y estrategias para lograr la formación investigativa de los estudiantes que les permitan desarrollar la actitud crítica y creativa, así como la incorporación del tic a su proceso formativo investigativo.

Reconocida la importancia del trabajo interdisciplinario, como estrategia para favorecer el desarrollo de competencias en los estudiantes y la consolidación de este en las prácticas

⁶ Tomado de Informe de autoevaluación. Programa matemática aplicada. Universidad Surcolombiana, 2015

investigativas de los programas de pregrado en matemáticas, surge entonces la necesidad de realizar una investigación, en la que se identifiquen las prácticas interdisciplinarias realizadas en el programa de matemática aplicada de la universidad Surcolombiana, que sin duda resulta complejo en cuanto que se encuentra inserto en unas dinámicas de trabajo por asignaturas bajo las cuales el estudiante, aun con los conceptos matemáticos adquiridos, no siempre encuentra su aplicación; sobre todo, por la dificultad de una visión holística (Documento maestro, 2015).

Si bien la esencia del pensamiento complejo se encuentra en las experiencias interdisciplinarias que permitan a los estudiantes conocer nuevos temas como las redes neuronales pasando del trabajo individual al colectivo, donde se obtienen conclusiones. Estas vivencias se encuentran en continua evolución y la búsqueda de nuevas posibilidades, es necesario que el futuro matemático tenga una amplia baraja de posibilidades para que sus búsquedas y hallazgos logren facilitarle la transición del pensamiento matemático elemental al pensamiento matemático avanzado.

En esta transición, es pertinente el planteamiento de Maldonado, referente a una nueva filosofía del conocimiento: “Ya sea desde la física y el estudio de fenómenos tales como los solitones, las funciones de onda, y los sistemas caóticos, o bien, igualmente, desde la simulación de problemas altísimamente más complicados como es el caso de los fenómenos meteorológicos, biológicos, sociales, económicos, políticos, o ecológicos, por ejemplo, lo cierto es que el rasgo más fuerte de los sistemas complejos se revela como el de su impredecibilidad. Este rasgo altera magníficamente el estatuto mismo de la ciencia y de la racionalidad humana en general. No es posible un estudio del conocimiento a la manera de la epistemología clásica, y ciertamente no cuando el conocimiento se relaciona con, se ocupa de, y está dirigido a, sistemas complejos. No es posible un conocimiento igualmente simple o igualmente complejo con independencia del objeto de que se ocupa. La complejidad del conocimiento *depende* exactamente *de* la dinámica del sistema con el que se ocupa. Pero lo propio del estudio de los sistemas complejos consiste en poner permanentemente de manifiesto que nos la vemos con sorpresas, que son sistemas esencialmente abiertos e indeterminados, que la incertidumbre es un rasgo central y no accesorio y que lo que el sujeto determine o quiera decidir no afecta en absoluto la libertad, la espontaneidad y la emergencia misma de los procesos de esta clase de sistemas. La epistemología de la complejidad no trabaja ya con leyes ni reglas, y el recurso a conceptos como patrones o pautas (patterns) apunta en realidad a la constitución y reconstitución continua de lógicas diversas y que sólo permanecen en dependencia de los microestados y las interacciones del sistema. Con todo, el problema fundamental para la filosofía del conocimiento cuando se ocupa de la complejidad permanece, a saber: ¿es posible hablar de patrones comunes universales para los diversos sistemas complejos, mejor, para un universo crecientemente complejo? La adecuada formulación y la eventual respuesta a esta pregunta producirán una transformación en el propio estatuto del acto de conocer, esto es, en la propia comprensión del conocer por parte de sí mismo.” (C. E. Maldonado, 2003)

Se hace apremiante entonces que desde este tipo de investigaciones se le dé la importancia que ameritan las prácticas interdisciplinarias y a la forma en que estas se llevan a cabo en el programa de matemática aplicada, puesto que son el producto de la dinámica del programa y posibilitan el

desarrollo integral de los estudiantes. Estos procesos de desarrollo integral, merecen un análisis profundo desde la perspectiva de la complejidad no solo por la necesidad de abrir otras miradas y nuevas perspectivas desde donde volver a preguntarse qué es propiamente lo interdisciplinar y cuáles son los nuevos caminos en el proceso de indisciplinar el conocimiento, sino también por la urgencia de construir una propuesta pedagógica que permita incluir temas de naturaleza interdisciplinar, no contemplados en el currículo, en el marco del currículo existente. En concreto, estudiar las redes neuronales integradas al diseño curricular de estadística y probabilidades, teniendo en cuenta la naturaleza estadística de las redes neuronales para contribuir así en la formación del pensamiento aleatorio de los futuros profesionales.

Por otro lado, las redes neuronales constituyen una herramienta de análisis, modelización y predicción que se encuentra integrada en diversos campos. De este modo adoptan diferentes connotaciones como herramientas de la ingeniería, réplicas del pensamiento racional o modelos de caja negra. En todos los casos, se trata de obtener modelos coherentes con la realidad observada, de tal modo que sean los datos los que determinen el comportamiento de la red, bien a través de la determinación de su estructura, bien de sus parámetros internos. Estas ideas acercan las redes neuronales a las ideas de los métodos no paramétricos de análisis de datos en particular y en general al ámbito de la estadística. (Mendez, 2009).

Así pues, abordar temáticas interdisciplinarias en un currículo asignaturista, es una alternativa válida y necesaria para el desarrollo de las competencias, que permitan al estudiante configurar su perfil profesional. Desde esta perspectiva, el aporte de esta investigación es ampliar el conocimiento sobre las redes neuronales como herramientas estadísticas para la modelación y simulación de sistemas complejos. Así, conocer las posibilidades de las redes neuronales como herramientas de clasificación y de predicción en particular al abordar la temática de la regresión permitirá comprender la interacción entre los contenidos curriculares de los campos de las redes neuronales, la estadística y las probabilidades.

Por otra parte, la propuesta permitirá construir referentes para la configuración de políticas institucionales orientadas a mejorar la implementación de estrategias que contribuyan al desarrollo del pensamiento complejo de los estudiantes. Teniendo en cuenta los anteriores derroteros, surgen entonces las siguientes preguntas de investigación:

- ¿Qué estrategias metodológicas desde la complejidad se pueden usar para integrar las redes neuronales al currículo de estadística y probabilidad con un enfoque interdisciplinario?
- ¿Qué es la regresión lineal y como se trabaja en el currículo tradicional?
- ¿Cómo definir las redes neuronales como sistemas dinámicos adaptativos complejos?
- ¿Qué son, ¿cómo se clasifican y cómo aprenden las redes neuronales artificiales?
- ¿Por qué usar las redes neuronales artificiales como modelos estadísticos?
- ¿Qué estrategia metodológica puede usarse para incorporar las redes neuronales artificiales en el currículo de estadística y probabilidad, para modelos de regresión lineal usando el software estadístico SPSS?

3. ANTECEDENTES Y JUSTIFICACIÓN

3.1. Antecedentes

3.1.1. Historia

Entre los años 1999 a 2004, el Ministerio de Educación Nacional llevó a cabo los proyectos: Incorporación de nuevas tecnologías al currículo de matemáticas de la educación básica secundaria y media de Colombia (Castiblanco, 2002) y Tecnología Informática: Innovación en el currículo de matemáticas de la educación básica secundaria y media (Castiblanco Paiba, Urquina Llanos, Camargo Uribe, & Moreno Armella, 2004), como resultado del interés por el mejoramiento de la calidad de la educación de un amplio grupo de educadores matemáticos y de un proceso de construcción participativa adelantado en el país para orientar la educación matemática colombiana...”. El proyecto tuvo como objetivo general: “Mejorar la calidad de la enseñanza de las matemáticas y la capacidad de aprendizaje mediante los recursos expresivos que la tecnología pone al alcance de las instituciones educativas”. El marco teórico del proyecto tuvo como referentes generales: la educación matemática, diseños didácticos incorporando tecnologías computacionales, la cultura informática, el currículo de matemáticas mediado por herramientas computacionales.

Teniendo en cuenta la capacidad de las nuevas herramientas para ofrecer medios alternativos de expresión matemática, formas innovadoras de manipulación de los objetos matemáticos y estrategias variadas de acercamiento al conocimiento matemático se definieron los siguientes campos conceptuales del proyecto: mediación instrumental, representaciones ejecutables, la naturaleza situada del conocimiento, exploración y sistematización, fluidez algorítmica y computacional. Una componente central del proyecto lo constituyó la investigación pedagógica mecanismo de autorregulación del proceso seguido (enfoques interpretativo y transformativo) y como reconocimiento de los avances en el trabajo de aula con tecnología. Para el caso particular del pensamiento estocástico componente esencial de este proyecto se adopta la propuesta curricular que enfatiza en el análisis exploratorio de datos y la modelización, Pensamiento estadístico y tecnologías computacionales) conceptualizando aspectos como: la lectura crítica de datos, el uso de diferentes representaciones, el establecimiento de regularidades (similitudes) y las variaciones. Para el currículo esto implica que a partir de situaciones problemáticas se posibilite el tránsito desde una lectura “literal” de los datos a una lectura en la que se integren relaciones como la comparación, la clasificación y la asociación entre las variables representadas en los datos y, por último, el que se puedan generar predicciones e inferencias a partir de los datos, que no se establecen directamente de las representaciones, sino que se requiere un mayor grado de elaboración conceptual.

3.1.2. Matemáticas para la creatividad

Es una propuesta curricular que surge de diferentes proyectos realizados por el grupo de investigación Dinusco, en particular la línea de investigación sobre el enfoque sistémico de las matemáticas escolares, así como también publicaciones (Montealegre Cárdenas, Londoño Betancourth, & Polanía Quiza, 2004) y que ha orientado el trabajo en programas de pregrado y postgrado, así como la capacitación para docentes de la educación básica y media.

La propuesta tiene las siguientes características: integra los valores característicos de los diferentes dominios (geométrico, numérico, estocástico, medida, variacional), vertientes del enfoque sistémico (Teoría cognitiva de Piaget, aprendizaje social de Vygotsky, unidad compleja de la Gestalt, aprendizaje significativo y la confrontación teoría-práctica explicada por el caos creativo), inclusión en el aprendizaje de aspectos lúdicos mediante juegos estructurados, enfoque de solución de problemas, modelación matemática y proyectos interdisciplinarios, apoyo computacional, enfoque de competencias y estándares del MEN. Fue precisamente esta propuesta la que orientó la creación y puesta en marcha del programa de matemática aplicada.

3.1.3. Modelización estadística con redes neuronales.

En esta tesis son consideradas las redes neuronales como potentes herramientas estadísticas de análisis de datos y se buscó como objetivo: clarificar la naturaleza de las redes neuronales, analizando sus fundamentos, poniendo de manifiesto su naturaleza estadística y sus conexiones con diversos métodos estadísticos ampliamente conocidos mostrando con ejemplos concretos la bondad de las redes neuronales como herramientas estadísticas de modelización y su capacidad para responder a las necesidades que presentan los problemas observados en el mundo real, mostrando cómo las redes neuronales constituyen modelos capaces de dar el salto del ámbito teórico a la realidad, aportando soluciones a problemas reales, al tiempo que se mantiene su rigor y esencia. Metodológicamente, se aborda: la aplicación de redes neuronales a la resolución de problemas de diversa índole en ámbitos medioambientales e industriales comparando modelos de redes neuronales y series de tiempo a la hora de modelar un proceso real de naturaleza continua, la capacidad de predecir variables binarias, esto es, en problemas de clasificación y aplicaciones al control de procesos reales (Castellano, 2009). La contribución al proyecto es pues de tipo teórico.

3.1.4. Redes Neuronales Artificiales aplicadas al análisis de datos.

Esta tesis describe tres líneas de investigación desarrolladas en torno a la aplicación de las redes neuronales artificiales en el ámbito del análisis de datos (Montaño, 2002). Los campos de aplicación tratados son: el análisis de datos aplicado a conductas adictivas, el análisis de supervivencia, y el estudio del efecto de las variables de entrada en una red neuronal. Los objetivos planteados: creación de una red neuronal capaz de discriminar entre sujetos consumidores y no consumidores de éxtasis a partir de las respuestas dadas a un cuestionario en la población de

jóvenes europeos, realizar una comparación en cuanto a capacidad de predicción entre dos modelos de RNA y un modelo de regresión, diseño del programa *Sensitivity Neural Network 1.0*, para simular el comportamiento de una red MLP asociada a la regla de aprendizaje *backpropagation*. Al igual que la tesis anterior la contribución al proyecto es de tipo teórico.

3.1.4.1. *El proceso de autoevaluación del programa de matemática aplicada.*

En este documento se presenta la evaluación de las condiciones de calidad en las que se puede evidenciar el estado actual del programa de pregrado Matemática Aplicada de la Universidad Surcolombiana, el cual está adscrito a la Facultad de Ciencias Naturales y Exactas y creado en 2009-1 iniciando actividades con la cohorte del año 2009-2. Su organización curricular ha sido revisada y modificada a través de dos procesos de autoevaluación (Informe autoevaluación, 2015), el primero de febrero 2009 al 2012-A y el segundo en Febrero 2012B- al 2014B, dando cumplimiento a las políticas de autoevaluación continúa establecidas por la Universidad Surcolombiana y la Facultad de Ciencias Exactas.

Las evaluaciones de las 15 condiciones presentadas aportan información necesaria que permite considerar que este documento cumple con las condiciones de calidad para solicitar la renovación de su registro calificado. En este segundo proceso se evaluaron los avances académicos, investigativos, administrativos que ha logrado superar el programa en la región Surcolombiana a partir del plan de mejoramiento que resultó de la primera Autoevaluación desarrollada en el año 2011. También, es pertinente señalar que este segundo ejercicio de Autoevaluación es fundamental para trazar un Plan de Acción que posibilite en los próximos años solucionar las dificultades que el programa presenta actualmente.

Presentamos a continuación los aspectos relevantes para este proyecto que corresponden a las condiciones 3 (contenido curricular), 4 (organización de las actividades académicas) y 5 (investigación).

3.1.4.2. *Condición 3: Contenido curricular*

3.1.4.2.1. *Fortalezas*

- Se realizó la actualización de los microdiseños curriculares.
- Se han fortaleciendo los siguientes énfasis del programa: Biomatemáticas, Economía Matemática y Sistemas Dinámicos.
- Se cuenta con una primera sala de computo, denominada “La Venada”, con 25 terminales, con licencias de Matlab y SPSS. Parcialmente se ha incorporado la Matemática Computacional, especialmente usando software Libre, como el Geogebra.
- Dado que el enfoque del programa es el de la Teoría de la Complejidad, se ha fortalecido en todas las asignaturas las interacciones teoría practica matemáticas.

3.1.4.2.2. *Debilidades*

- En el proceso anterior se tenía como debilidad la falta de implementación de la matemática computacional en todas las asignaturas. Para comenzar a trabajar en esta debilidad incorporado parcialmente la Matemática Computacional, especialmente usando software Libre, el Matlab y SPSS, en los seis primeros semestres de la carrera.
- Así mismo, en el proceso anterior se tenía como debilidad la necesidad de culminar con la revisión de los microdiseños curriculares. En este segundo ejercicio se realizó la revisión de los microdiseños curriculares en sus contenidos y metodologías, pero también ajustándolos al formato institucional.
- De este estudio surge la necesidad de ofrecer las nuevas asignaturas de Algebra Lineal II y de didáctica de las matemáticas.
- Falta legalizar por parte del consejo de Facultad de Ciencias Exactas y Naturales el reglamento de práctica profesional.

3.1.4.2.3. *Estrategias de mejoramiento*

- Trabajar en la integración de los componentes curriculares teóricos con los prácticos en las actividades registradas en cada uno de los microdiseños, este es factible en la Teoría del Pensamiento Complejo, la cual orienta este programa.
- Incluir como cursos obligatorios: Sucesiones y Series, curso de Teoría de Probabilidades los cuales estaban como cursos electivos del programa; se crean como cursos nuevos del Algebra Lineal II y Didáctica de las matemáticas.
- Presentación para aprobación en el Consejo de Facultad de un proyecto de reglamentación de la práctica profesional.

3.1.4.3. *Condición 4: Organización de las actividades académicas*

3.1.4.3.1. *Fortalezas*

- Las actividades académicas se encuentran organizados y se han actualizado de acuerdo a las necesidades del programa.
- Uso de software especializado en matemáticas para el desarrollo de las actividades académicas.
- La organización de actividades académicas a través del trabajo presencial y el trabajo independiente.

3.1.4.3.2. *Debilidades*

- En este segundo proceso se identificó la necesidad de fortalecer las actividades académicas y estrategias que el programa aplica para el desarrollo del trabajo independiente.

3.1.4.3.3. *Estrategias de mejoramiento*

- Continuar fomentando las distintas actividades académicas y estrategias que utiliza el programa para el acompañamiento del trabajo independiente del estudiante.

3.1.4.4. *Condición 5: Investigación*

3.1.4.4.1. *Fortalezas*

- Siendo un programa nuevo en la región la investigación es uno de los aspectos relevantes en el Programa y los profesores de planta continúan desarrollando la investigación y publican sus resultados.
- Los cursos avanzados continúan desarrollándose en torno a las líneas de investigación del programa.
- Se continúa fortaleciendo los semilleros de investigación con participación de estudiantes y docentes, con posters y ponencias en congresos nacionales.

3.1.4.4.2. *Debilidades*

- El laboratorio de computo denominado “La Venada”, aún es insuficiente para las practicas con software especializado.
- Los recursos para organizar eventos especializados en el área de Matemáticas es muy escaso.
- Falta implementar un programa de estímulos académicos para los mejores estudiantes, aunque la dependencia de bienestar Universitario ofrece alimentos a la gran mayoría de los estudiantes.
- En el proceso anterior se identificó como debilidad la falta de una mayor participación de los estudiantes en semilleros de investigación, se evidencia un aumento en la participación de estudiantes en semilleros.
- La Universidad aún no ha implementado el programa estímulos académicos a los estudiantes vinculados a procesos investigativos.
- Falta categorizar por Colciencias algunos grupos de investigación existentes y nuevos.

3.1.4.4.3. *Estrategia de mejoramiento*

- Constituir a corto plazo nuevos grupos de investigación con líneas de investigación específicas.
- Nombramiento de nuevos docentes investigadores de tiempo completo.
- Poner en marcha una revista especializada en la Facultad de Ciencias Exactas y Naturales para la publicación de resultados de investigación.

- Incentivar el programa de jóvenes investigadores y docentes para este campo del conocimiento.
- Promover un programa de postgrado centrado en las Matemáticas y con enfoque interdisciplinario.
- Participar más activamente en la actividad de ACOFACIEN y buscar nuevos convenios nacionales e internacionales.

3.2. Justificación

La Universidad Surcolombiana, creada por la ley 13 de 1976, ha definido como propósitos Institucionales, la participación en la construcción de una sociedad libre, justa, pluralista, desarrollando una cultura académica apoyada en el accionar del pensamiento complejo para construir y ejecutar proyectos, de investigación o proyección social, apoyados en un sólida cultura científica; así promueve el desarrollo integral de la región mediante la resolución de los problemas interdisciplinarios que afectan la vida de las personas y comunidades (Documento maestro. Programa matemática aplicada. Universidad Surcolombiana, 2015).

Particularmente La Universidad Surcolombiana, a través de la Facultad de Ciencias Exactas y Naturales y su Departamento de Matemáticas y Estadística, ha diseñado la propuesta curricular en el área de Matemáticas y sus aplicaciones en procura de una formación integral del estudiante y responder a las expectativas del desarrollo científico y tecnológico exigidas por la sociedad contemporánea. La propuesta ha sido elaborada al tenor de los parámetros consignados en el decreto 1295 del 20 de abril de 2010, *por el cual se reglamenta el registro calificado de que trata la Ley 1188 de 2008*, y en la normatividad expedida por **la Universidad Surcolombiana**, en particular el Acuerdo 026 del 17 de octubre de 2014 mediante el cual la Universidad Surcolombiana adopto, recientemente, su proyecto Educativo Universitario.

El equipo de docentes y estudiantes del programa, previa y detenida reflexión sobre la expectativas de formación que exige la región y el país a los futuros profesionales de las matemáticas, después de dos procesos de autoevaluación realizados entre los años 2011 y 2014, teniendo en cuenta los respectivos planes de mejoramiento y consultando los estándares definidos para Los Programas de Matemáticas por la Asociación Colombiana de Facultades de Ciencias Exactas y Naturales, ACOFACIEN, elaboró el documento maestro que contiene conclusiones de dos procesos de evaluación y que se materializa en un renovado plan curricular que consta de: un componente profesional básico, la fundamentación en la matemáticas universales; las matemáticas aplicadas que responden a las exigencias derivadas de los más recientes desarrollos científicos en esta área y a este nivel; así como un componente complementario, para responder a los intereses de los estudiantes.

El programa se ofrece en el nivel de pregrado y está orientado a la formación de un Profesional de la Matemática Aplicada, por ello con exigente formación en la matemática básica y con énfasis

en Matemáticas Aplicadas, en particular en: Biomatemáticas, Economía Matemática, Estadística y Sistemas Dinámicos. Este programa curricular conduce al título de Matemático, tiene una duración de nueve semestres y está inmerso dentro de los estándares de calidad exigidos por el Ministerio de Educación Nacional para estos propósitos.

Se busca que la propuesta metodológica aquí diseñada pueda integrarse a la dinámica del proceso de autoevaluación permanente que el programa viene adelantando hacia la obtención por parte del Ministerio de Educación Nacional de la acreditación de alta calidad. Una revisión cuidadosa de los microdiseños curriculares permite observar que el tema de las redes neuronales, fundamental para las ciencias de la complejidad, no aparece en la estructura temática de ninguno de ellos.

En este sentido, se hace necesario incorporar las redes neuronales al currículo del programa de matemática aplicada, dada la naturaleza interdisciplinaria de sus contenidos y específicamente su capacidad como herramienta estadística. Se responde así a las exigencias para los currículos de matemáticas a nivel internacional. Se espera que, como resultado del proceso de autoevaluación, que el comité curricular del programa apruebe la incorporación de estos ajustes en el plan curricular (documento maestro). Así mismo, se espera que en el futuro se consolide una línea de investigación en redes neuronales hacia el desarrollo de la metodología de modelación y simulación, básica para las ciencias de la complejidad. Esta propuesta también servirá de modelo para la incorporación de otros temas de la complejidad al currículo.

4. MARCO TEÓRICO

4.1. CARACTERÍSTICAS PRINCIPALES DE LAS REDES NEURONALES

4.1.1. INTRODUCCIÓN A LAS REDES NEURONALES

Resulta irónico pensar que máquinas de cómputo capaces de realizar 100 millones de operaciones en coma flotante por segundo, no sean capaces de entender el significado de las formas visuales o de distinguir entre distintas clases de objetos. Los sistemas de computación secuencial, son exitosos en la resolución de problemas matemáticos o científicos, en la creación, manipulación y mantenimiento de bases de datos, en comunicaciones electrónicas, en el procesamiento de textos, gráficos y auto edición, incluso en funciones de control de electrodomésticos, haciéndolos más eficientes y fáciles de usar, pero definitivamente tienen una gran incapacidad para interpretar el mundo.

Esta dificultad de los sistemas de cómputo que trabajan bajo la filosofía de los sistemas secuenciales, desarrollados por Von Neumann, ha hecho que un gran número de investigadores centre su atención en el desarrollo de nuevos sistemas de tratamiento de la información, que permitan solucionar problemas cotidianos, tal como lo hace el cerebro humano; este órgano biológico cuenta con varias características deseables para cualquier sistema de procesamiento digital, tales como:

- Es robusto y tolerante a fallas, diariamente mueren neuronas sin afectar su desempeño.
- Es flexible, se ajusta a nuevos ambientes por medio de un proceso de aprendizaje, no hay que programarlo.
- Puede manejar información difusa, con ruido o inconsistente.
- Es altamente paralelo
- Es pequeño, compacto y consume poca energía.

El cerebro humano constituye una computadora muy notable, es capaz de interpretar información imprecisa suministrada por los sentidos a un ritmo increíblemente veloz. Logra discernir un susurro en una sala ruidosa, un rostro en un callejón mal iluminado y leer entre líneas un discurso; lo más impresionante de todo, es que el cerebro aprende sin instrucciones explícitas de ninguna clase, a crear las representaciones internas que hacen posibles estas habilidades.

Basados en la eficiencia de los procesos llevados a cabo por el cerebro, e inspirados en su funcionamiento, varios investigadores han desarrollado desde hace más de 30 años la teoría de las Redes Neuronales Artificiales (RNA), las cuales emulan el comportamiento de las redes neuronales biológicas, y que se han utilizado para aprender estrategias de solución basadas en ejemplos de

comportamiento típico de patrones; estos sistemas no requieren que la tarea a ejecutar se programe, ellos generalizan y aprenden de la experiencia⁷.

La teoría de las RNA ha brindado una alternativa a la computación clásica, para aquellos problemas, en los cuales los métodos tradicionales no han entregado resultados muy convincentes, o poco convenientes. Las aplicaciones más exitosas de las RNA son:

- Procesamiento de imágenes y de voz
- Reconocimiento de patrones
- Planeamiento
- Interfaces adaptivas para sistemas Hombre/máquina
- Predicción
- Control y optimización
- Filtrado de señales

Los sistemas de cómputo tradicional procesan la información en forma secuencial; un computador serial consiste por lo general de un solo procesador que puede manipular instrucciones y datos que se localizan en la memoria, el procesador lee y ejecuta una a una las instrucciones en la memoria; este sistema serial es secuencial, todo sucede en una sola secuencia determinística de operaciones. Las RNA no ejecutan instrucciones, responden en paralelo a las entradas que se les presenta; el resultado no se almacena en una posición de memoria, este es el estado de la red para el cual se logra equilibrio. El conocimiento de una red neuronal no se almacena en instrucciones, el poder de la red está en su topología y en los valores de las conexiones (pesos) entre neuronas.

Las RNA son una teoría que aún está en proceso de desarrollo, su verdadera potencialidad no se ha alcanzado todavía; aunque los investigadores han desarrollado potentes algoritmos de aprendizaje de gran valor práctico, las representaciones y procedimientos de que se sirve el cerebro, son aún desconocidas. Tarde o temprano los estudios computacionales del aprendizaje con RNA acabarán por converger a los métodos descubiertos por evolución, cuando eso suceda, muchos datos empíricos concernientes al cerebro comenzarán súbitamente a adquirir sentido y se tornarán factibles muchas aplicaciones desconocidas de las redes neuronales.

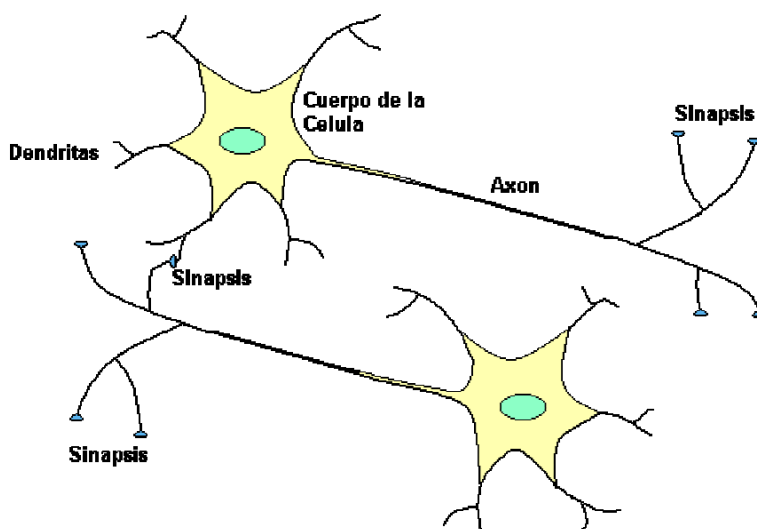
4.1.2. FUNCIONAMIENTO DE UNA NEURONA BIOLÓGICA

El cerebro consta de un gran número (aproximadamente 10^{11}) de elementos altamente interconectados (aproximadamente 10^4 conexiones por elemento), llamados neuronas. Estas neuronas tienen tres componentes principales, las dendritas, el cuerpo de la célula o soma y el

⁷ Consultar para mayor ampliación, Tutorial sobre aplicaciones de las redes neuronales a problemas de Ingeniería Eléctrica y su implementación en un sitio web, universidad tecnológica de Pereira Recuperado de <http://www.medicinaycomplejidad.org/pdf/redes/Contenido.pdf> , 2018.

axón. Las dendritas, son el árbol receptor de la red, son como fibras nerviosas que cargan de señales eléctricas el cuerpo de la célula. El cuerpo de la célula, realiza la suma de esas señales de entrada. El axón es una fibra larga que lleva la señal desde el cuerpo de la célula hacia otras neuronas. El punto de contacto entre un axón de una célula y una dendrita de otra célula es llamado sinapsis, la longitud de la sinapsis es determinada por la complejidad del proceso químico que estabiliza la función de la red neuronal. Un esquema simplificado de la interconexión de dos neuronas biológicas se observa en la (Figura 1)

Figura 1 Neuronas Biológicas



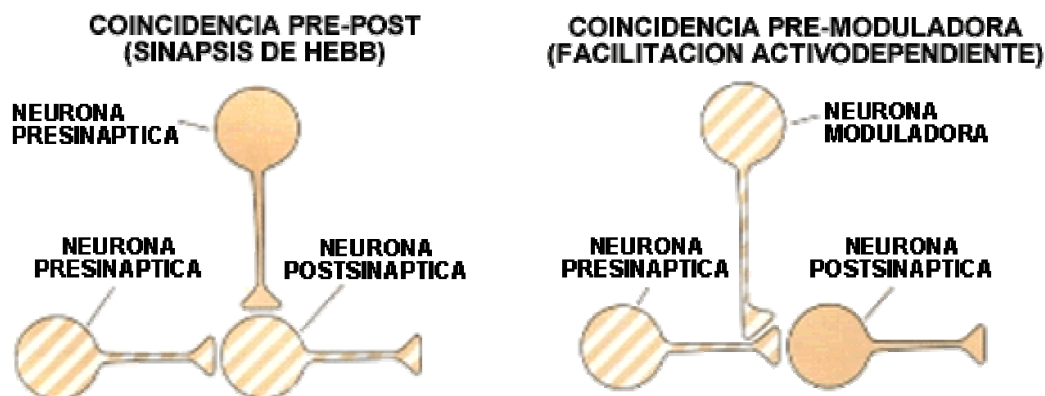
Fuente. (Buitrago & Muñoz, 2000)

Algunas de las estructuras neuronales son determinadas en el nacimiento, otra parte es desarrollada a través del aprendizaje, proceso en que nuevas conexiones neuronales son realizadas y otras se pierden por completo. El desarrollo neurológico se hace crítico durante los primeros años de vida, por ejemplo, está demostrado que si a un cachorro de gato, se le impide usar uno de sus ojos durante un período corto de tiempo, el nunca desarrollara una visión normal en ese ojo. Las estructuras neuronales continúan cambiando durante toda la vida, estos cambios consisten en el refuerzo o debilitamiento de las uniones sinápticas, por ejemplo, se cree que nuevas memorias son formadas por la modificación de esta intensidad entre sinapsis, así el proceso de recordar el rostro de un nuevo amigo, consiste en alterar varias sinapsis.

Como consecuencia de los primeros estudios sobre la base neural de los sistemas mnémicos (relacionados con la memoria), se creía que el almacenamiento de la memoria asociativa, requería de un circuito neuronal muy complejo. Entre quienes comenzaron a oponerse a este enfoque se hallaba Donald O. Hebb, profesor de la Universidad de Milner; Hebb sugirió que el aprendizaje asociativo podría ser producido por un mecanismo celular sencillo y propuso que las asociaciones

podrían formarse por una actividad neuronal coincidente: “Cuando un axón de la célula A excita la célula B y participa en su activación, se produce algún proceso de desarrollo o cambio metabólico en una o en ambas células, de suerte que la eficacia de A, como célula excitadora de B, se intensifica”. Según la regla Hebbiana de aprendizaje, el que coincida la actividad de las neuronas presinápticas (suministran el impulso de entrada) con la de las postsinápticas (reciben el impulso) es muy importante para que se refuerce la conexión entre ellas, este mecanismo es llamado pre-postasociativo, del cual puede observarse un ejemplo en la (Figura 2)

Figura 2 Cambios asociativos de las fuerzas sinápticas durante el aprendizaje



Fuente (Buitrago & Muñoz, 2000)

Todas las neuronas conducen la información de forma similar, ésta viaja a lo largo de axones en breves impulsos eléctricos, denominados potenciales de acción; los potenciales de acción que alcanzan una amplitud máxima de unos 100 mV y duran 1 ms, son resultado del desplazamiento a través de la membrana celular de iones de sodio dotados de carga positiva, que pasan desde el fluido extracelular hasta el citoplasma intracelular; la concentración extracelular de sodio supera enormemente la concentración intracelular⁸.

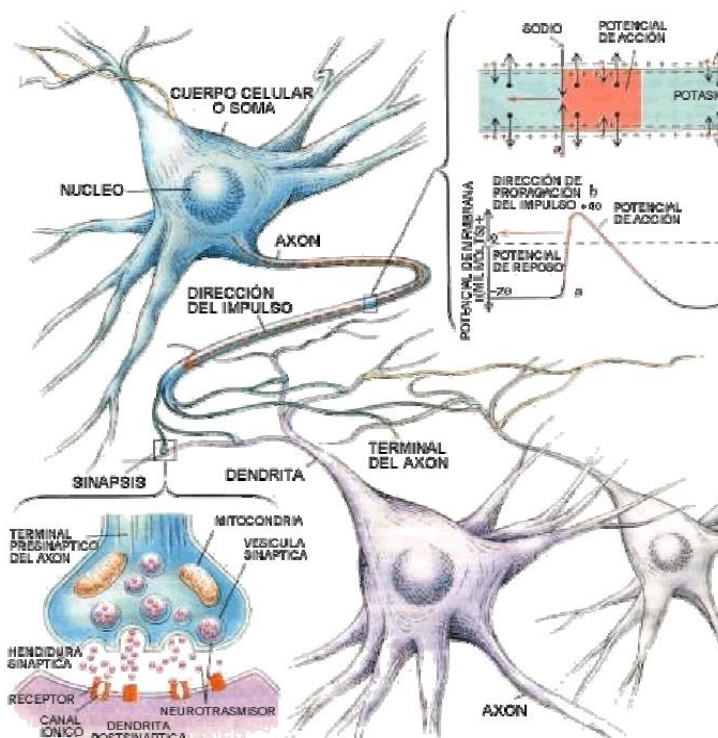
La membrana en reposo mantiene un gradiente de potencial eléctrico de -70mv , el signo negativo se debe a que el citoplasma intracelular está cargado negativamente con respecto al exterior; los iones de sodio no atraviesan con facilidad la membrana en reposo, los estímulos físicos o químicos que reducen el gradiente de potencial, o que despolaricen la membrana, aumentan su permeabilidad al sodio y el flujo de este ión hacia el exterior acentúa la despolarización de la membrana, con lo que la permeabilidad al sodio se incrementa más aún. Alcanzado un potencial crítico denominado “umbral”, la realimentación positiva produce un efecto regenerativo que obliga al potencial de membrana a cambiar de signo. Es decir, el interior de la célula se torna positivo con respecto al exterior, al cabo de 1 ms, la permeabilidad del sodio decae y el potencial de membrana retorna a -70mv , su valor de reposo. Tras cada explosión de actividad iónica, el

⁸ Para mayor información consultar, Biofísica (Neuronas), Recuperado de <http://biofísica.galeon.com/biof13.htm>, 2018

mecanismo de permeabilidad del sodio se mantiene refractario durante algunos milisegundos; la tasa de generación de potenciales de acción queda así limitada a unos 200 impulsos por segundo, o menos.

Aunque los axones puedan parecer hilos conductores aislados, no conducen los impulsos eléctricos de igual forma, como hilos eléctricos no serían muy valiosos, pues su resistencia a lo largo del eje es demasiado grande y la resistencia de la membrana demasiado baja; la carga positiva inyectada en el axón durante el potencial de acción queda disipada uno o dos milímetros más adelante, para que la señal recorra varios centímetros es preciso regenerar frecuentemente el potencial de acción a lo largo del camino, la necesidad de reforzar repetidamente esta corriente eléctrica limita a unos 100 metros por segundo la velocidad máxima de viaje de los impulsos, tal velocidad es inferior a la millonésima de la velocidad de una señal eléctrica por un hilo de cobre. Los potenciales de acción, son señales de baja frecuencia conducidas en forma muy lenta, éstos no pueden saltar de una célula a otra, la comunicación entre neuronas viene siempre mediada por transmisores químicos que son liberados en las sinapsis. de Pereira Un ejemplo de comunicación entre neuronas y del proceso químico de la liberación de neurotransmisores se ilustra en la (Figura 3)

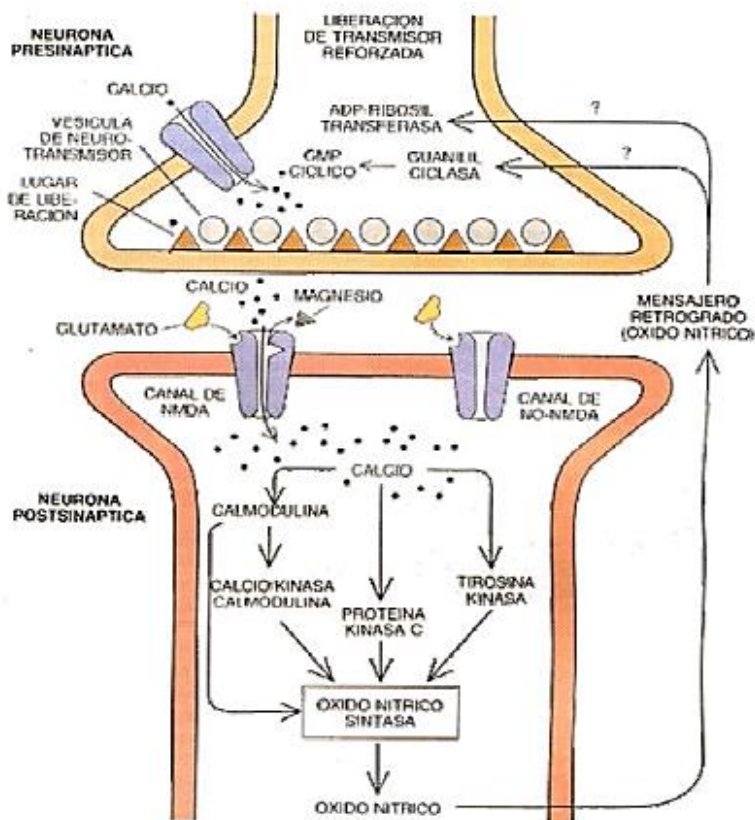
Figura 3 Comunicación entre neuronas



Fuente. (Buitrago & Muñoz, 2000)

Las sinapsis se clasifican según su posición en la superficie de la neurona receptora en tres tipos: axo-somática, axo-dendrítica, axo-axónica. Los fenómenos que ocurren en la sinapsis son de naturaleza química, pero tienen efectos eléctricos laterales que se pueden medir. En la (Figura 4) se visualiza el proceso químico de una sinapsis y los diferentes elementos que hacen parte del proceso tanto en la neurona presináptica, como en la postsináptica.

Figura 4. Proceso químico de una sinapsis



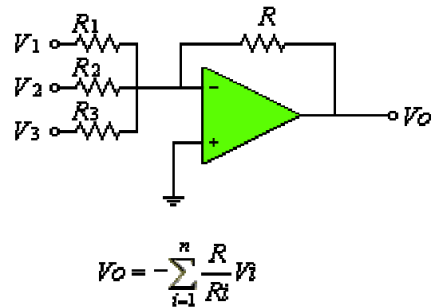
Fuente. (Análisis entre neuronas biológicas y artificiales.2012)

Las RNA no alcanzan la complejidad del cerebro, sin embargo, hay dos aspectos similares entre redes biológicas y artificiales, primero los bloques de construcción de ambas redes son sencillos elementos computacionales (aunque las RNA son mucho más simples que las biológicas) altamente interconectados; segundo, las conexiones entre neuronas determinan la función de la red.

4.2. CARACTERÍSTICAS DE UNA RED NEURONAL ARTIFICIAL

El modelo de una neurona artificial es una imitación del proceso de una neurona biológica, puede también asemejarse a un sumador hecho con un amplificador operacional tal como se ve en la (Figura 5)

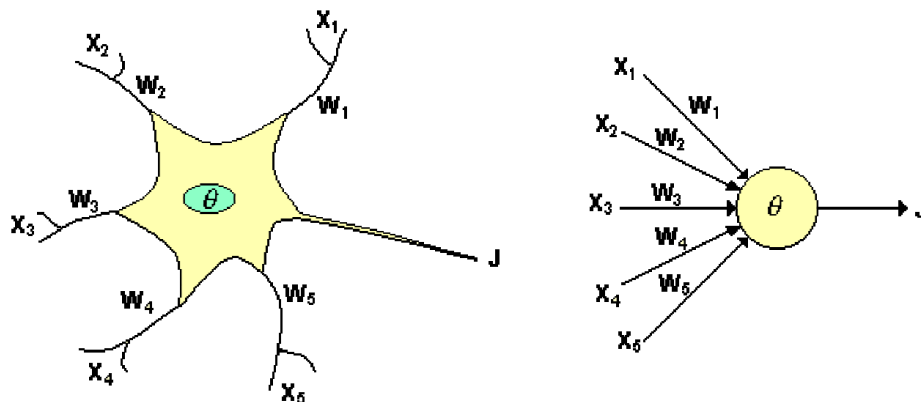
Figura 5 Neurona Artificial



Fuente. (Buitrago & Muñoz, 2000)

Existen varias formas de nombrar una neurona artificial, es conocida como nodo, neuronodo, celda, unidad o elemento de procesamiento (PE); En la (Figura 6) se observa un PE en forma general y su similitud con una neurona biológica

Figura 6. De la neurona biológica a la neurona artificial



Fuente. (Buitrago & Muñoz, 2000)

De la observación detallada del proceso biológico se han hallado los siguientes análogos con el sistema artificial:

- Las entradas X_i representan las señales que provienen de otras neuronas y que son capturadas por las dendritas.
- Los pesos W_i son la intensidad de la sinapsis que conecta dos neuronas; tanto X_i como W_i son valores reales.
- θ es la función umbral que la neurona debe sobrepasar para activarse; este proceso ocurre biológicamente en el cuerpo de la célula.

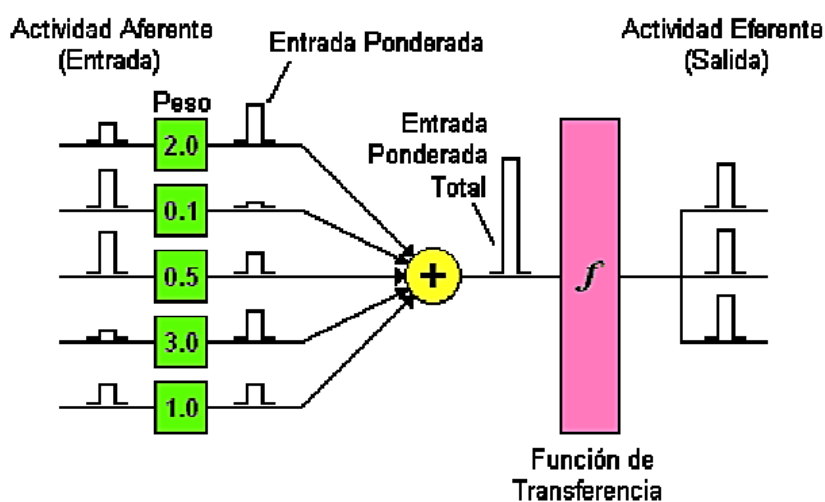
Las señales de entrada a una neurona artificial X_1, X_2, \dots, X_n son variables continuas en lugar de pulsos discretos, como se presentan en una neurona biológica. Cada señal de entrada pasa a través

de una ganancia o peso, llamado peso sináptico o fortaleza de la conexión cuya función es análoga a la de la función sináptica de la neurona biológica. Los pesos pueden ser positivos (excitatorios), o negativos (inhibitorios), el nodo sumatorio acumula todas las señales de entradas multiplicadas por los pesos o ponderadas y las pasa a la salida a través de una función umbral o función de transferencia. La entrada neta a cada unidad puede escribirse de la siguiente manera

$$neta_i = \sum_{i=1}^n W_i X_i = \bar{X}\bar{Y}$$

Una idea clara de este proceso se muestra en la (Figura 7), en donde puede observarse el recorrido de un conjunto de señales que entran a la red.

Figura 7. Proceso de una red neuronal



Fuente.(Buitrago & Muñoz, 2000)

Una vez que se ha calculado la activación del nodo, el valor de salida equivale a

$$x_i = f_i(neta_i)$$

Donde f_i representa la función de activación para esa unidad, que corresponde a la función escogida para transformar la entrada $neta_i$ en el valor de salida x_i , y que depende de las características específicas de cada red.

4.2.1. Notación.

Una notación matemática estándar no ha sido aún establecida para las redes neuronales, ya que sus aplicaciones son útiles en muchos campos, Ingeniería, Física, Psicología y Matemáticas. En este trabajo se adoptó la siguiente convención para identificar las variables, de manera que fuera compatible con las diferentes áreas, siendo lo más sencilla posible:

- Valores escalares: se representarán por medio de letra minúscula itálica

- Vectores: se representarán con letra itálica minúscula en negrilla.
- Matrices: se representarán con letra mayúscula itálica en negrilla. Para redes multicapa, los parámetros adoptarán la siguiente forma:

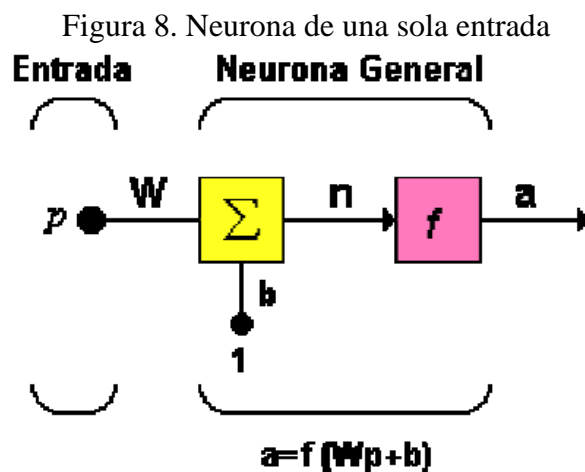
$$W_{s^c, s^c}^c$$

Donde c , es el número de la capa a la que corresponde dicho peso, y s representa las neuronas que participan en proceso. Así $W_{1,1}^2$ representa el peso de la segunda capa que comunica la primera neurona de dicha capa con la primera neurona de la primera capa. De igual manera el peso que representa la conexión desde la última neurona de la capa dos a la última neurona de la capa uno estará representada por: W_{s^2, s^1}^2

Esta convención es adoptada para todos los parámetros de la red.

4.2.2. Funciones de Transferencia

Un modelo más académico que facilita el estudio de una neurona, puede visualizarse en la (Figura 8)



Fuente. (Buitrago & Muñoz, 2000)

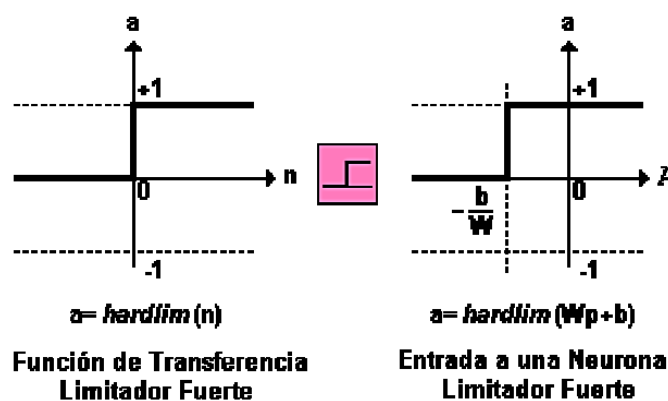
Las entradas a la red serán ahora presentadas en el vector \mathbf{p} , que para el caso de una sola neurona contiene sólo un elemento, w sigue representando los pesos y la nueva entrada b es una ganancia que refuerza la salida del sumador n , la cual es la salida neta de la red; la salida total está determinada por la función de transferencia f , la cual puede ser una función lineal o no lineal de n , y que es escogida dependiendo de las especificaciones del problema que la neurona tenga que resolver; aunque las RNA se inspiren en modelos biológicos no existe ninguna limitación para realizar modificaciones en las funciones de salida, así que se encontrarán modelos artificiales que nada tienen que ver con las características del sistema biológico.

4.2.2.1. Limitador fuerte (Hardlim).

La (), muestra como esta función de transferencia acerca la salida de la red a cero, si el argumento de la función es menor que cero y la lleva a uno si este argumento es mayor que uno. Esta función crea neuronas que clasifican las entradas en dos categorías diferentes, característica que le permite ser empleada en la red tipo Perceptrón:

$$a = \begin{cases} 1 & \text{si } n \geq 0 \\ 0 & \text{si } n < 0 \end{cases}$$

Figura 9. Función de transferencia Hardlim

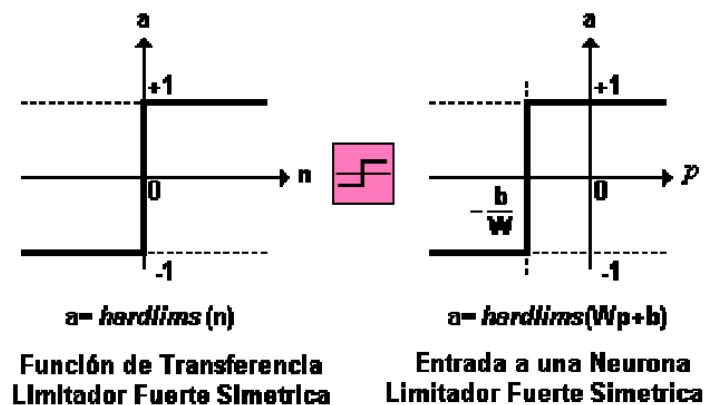


Fuente: (Buitrago & Muñoz, 2000)

El icono para la función Hardlim reemplazará a la letra f en la expresión general, cuando se utilice la función Hardlim.

Una modificación de esta función puede verse en la figura 1.3.6, la que representa la función de transferencia Hardlims que restringe el espacio de salida a valores entre 1 y -1.

Figura 10. Función de transferencia Hardlims

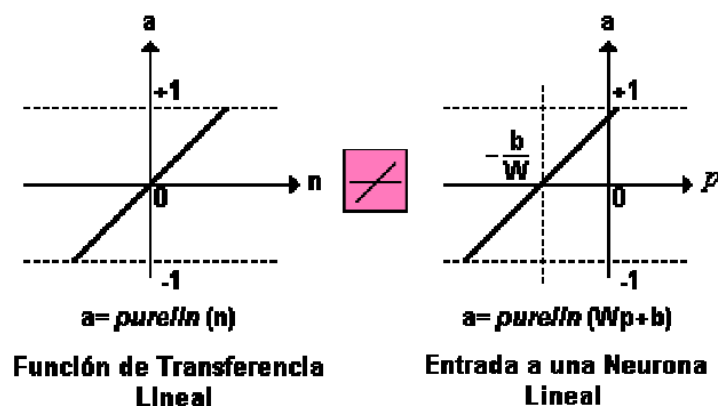


Fuente: (Buitrago & Muñoz, 2000)

4.2.2.2. Función de transferencia lineal (purelin):

La salida de una función de transferencia lineal es igual a su entrada, $a=n$

Figura 11. Función de transferencia lineal



Fuente. (Buitrago & Muñoz, 2000)

En la gráfica del lado derecho de la (Figura 11), puede verse la relación entre la salida a de la red, y la entrada p , teniendo en cuenta el valor de ganancia b ; neuronas que emplean esta función de transferencia son utilizadas en la red tipo Adaline.

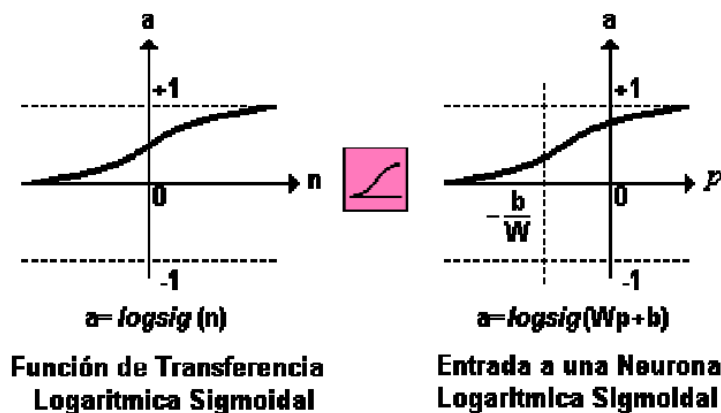
4.2.2.3. Función de transferencia sigmoidal (logsig):

Esta función toma los valores de entrada, los cuales pueden oscilar entre más y menos infinito, y restringe la salida a valores entre cero y uno, de acuerdo a la expresión

$$a = \frac{1}{1 + e^{-n}}$$

Esta función es comúnmente usada en redes multicapa, como la Backpropagation, en parte porque la función logsig es diferenciable.










Figura 12. Función de transferencia sigmoidal



Fuente. (Buitrago & Muñoz, 2000)

La Tabla 1, hace una relación de las principales funciones de transferencia empleadas en redes neuronales

Tabla 1. Funciones de Transferencia

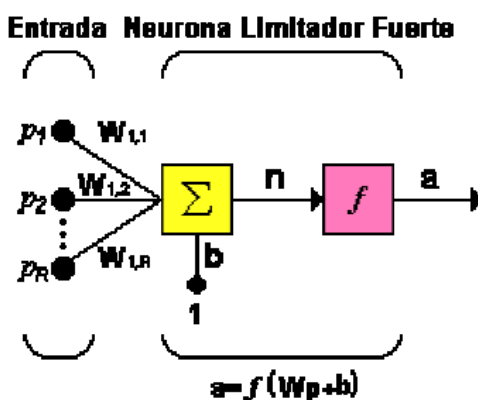
Nombre	Relación Entrada /Salida	Icono	Función
Limitador Fuerte	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$		<i>hardlim</i>
Limitador Fuerte Simétrico	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$		<i>hardlims</i>
Lineal Positiva	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$		<i>poslin</i>
Lineal	$a = n$		<i>purelin</i>
Lineal Saturado	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$		<i>satlin</i>
Lineal Saturado Simétrico	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = +1 \quad n > 1$		<i>satlins</i>
Sigmoidal Logarítmico	$a = \frac{1}{1 + e^{-n}}$		<i>logsig</i>
Tangente Sigmoidal Hiperbólica	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		<i>tansig</i>
Competitiva	$a = 1 \quad \text{Neurona con } n \text{ max}$ $a = 0 \quad \text{El resto de neuronas}$		<i>compet</i>

Fuente: (Buitrago & Muñoz, 2000)

4.2.3. Topología de una Red.

Típicamente una neurona tiene más de una entrada; en la (Figura 13) se observa una neurona con R entradas; las entradas individuales p_1, p_2, \dots, p_R son multiplicadas por los pesos correspondientes $w_{1,1}, w_{1,2}, \dots, w_{1,R}$ pertenecientes a la matriz de pesos W

Figura 13. Neurona con múltiples entradas



Fuente. (Buitrago & Muñoz, 2000)

La neurona tiene una ganancia b , la cual llega al mismo sumador al que llegan las entradas multiplicadas por los pesos, para formar la salida n ,

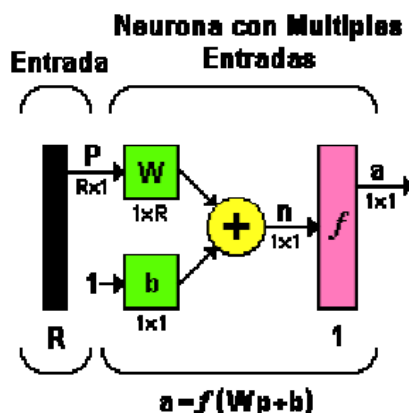
$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b$$

Esta expresión puede ser escrita en forma matricial

$$n = w_p + b$$

Los subíndices de la matriz de pesos representan los términos involucrados en la conexión, el primer subíndice representa la neurona destino y el segundo representa la fuente de la señal que alimenta a la neurona. Por ejemplo, los índices de $w_{1,2}$ indican que este peso es la conexión desde el segundo elemento del vector de entrada a la primera neurona. Esta convención se hace más útil cuando hay más de una neurona, o cuando se tiene una red con demasiados parámetros, en este caso la notación de la (Figura 13) puede resultar inapropiada y se prefiere emplear la notación compacta de la (Figura 14)

Figura 14. Neurona con múltiples entradas, notación abreviada



Fuente. (Buitrago & Muñoz, 2000)

El vector de entrada p es representado por la barra sólida vertical a la izquierda, las dimensiones de p son mostradas en la parte inferior de la variable como $R \times 1$ indicando que el vector de entrada es un vector fila de R elementos. Las entradas van a la matriz de pesos W , la cual tiene R columnas y sólo una fila para el caso de una sola neurona. Una constante 1 entra a la neurona multiplicada por la ganancia escalar b . La salida de la red a es en este caso un escalar, si la red tuviera más de una neurona a , sería un vector.

Dentro de una red neuronal, los elementos de procesamiento se encuentran agrupados por capas, una capa es una colección de neuronas; de acuerdo a la ubicación de la capa en la RNA, esta recibe diferentes nombres.

4.2.3.1. Capa de entrada

Recibe las señales de la entrada de la red, algunos autores no consideran el vector de entrada como una capa pues allí no se lleva a cabo ningún proceso, en este trabajo se seguirá este criterio.

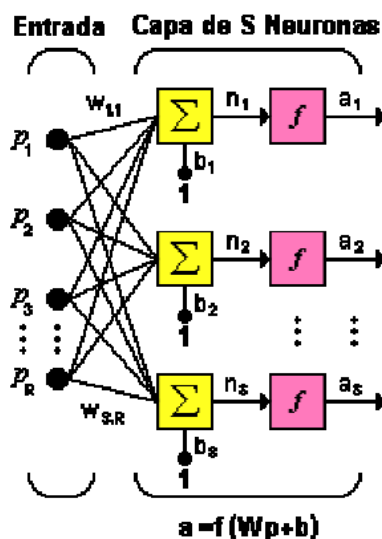
4.2.3.2. Capas ocultas

Estas capas son aquellas que no tienen contacto con el medio exterior, sus elementos pueden tener diferentes conexiones y son éstas las que determinan las diferentes topologías de la red.

4.2.3.3. Capa de salida

Recibe la información de la capa oculta y transmite la respuesta al medio externo. Una red de una sola capa con un número S de neuronas, se observa en la (Figura 15) en la cual, cada una de las R entradas es conectada a cada una de las neuronas, la matriz de pesos tiene ahora S filas.

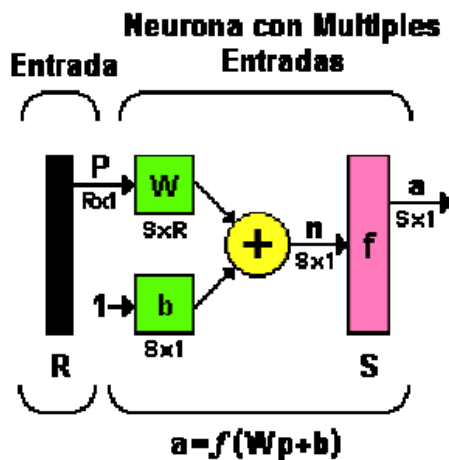
Figura 15. Capa de S neuronas



Fuente. (Buitrago & Muñoz, 2000)

La capa incluye la matriz de pesos, los sumadores, el vector de ganancias, la función de transferencia y el vector de salida. Esta misma capa se observa en notación abreviada en la (Figura 16)

Figura 16. Capa de S neuronas con notación abreviada

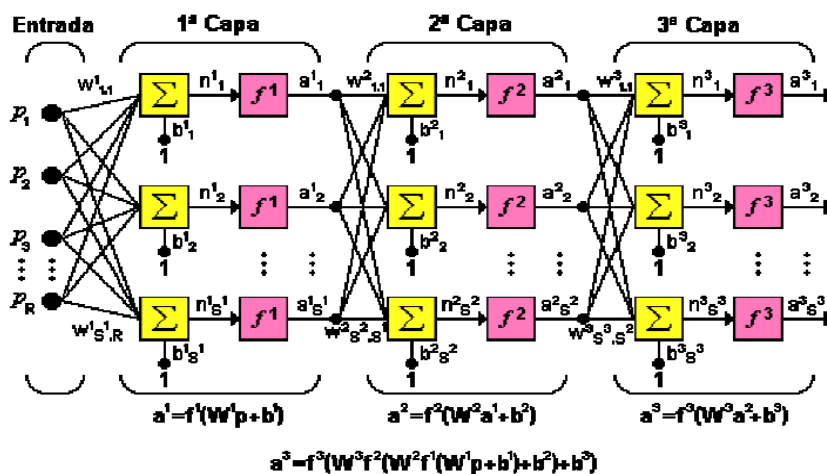


Fuente. (Buitrago & Muñoz, 2000)

En la (Figura 16) se han dispuesto los símbolos de las variables de tal manera que describan las características de cada una de ellas, por ejemplo la entrada a la red es el vector \mathbf{p} cuya longitud R aparece en su parte inferior, \mathbf{W} es la matriz de pesos con dimensiones $S \times R$ expresadas debajo del símbolo que la representa dentro de la red, \mathbf{a} y \mathbf{b} son vectores de longitud S el cual, como se ha dicho anteriormente representa el número de neuronas de la red.

Ahora, si se considera una red con varias capas, o red multicapa, cada capa tendrá su propia matriz de peso \mathbf{W} , su propio vector de ganancias \mathbf{b} , un vector de entradas netas \mathbf{n} , y un vector de salida \mathbf{a} . La versión completa y la versión en notación abreviada de una red de tres capas, pueden ser visualizada en la (Figura 17)

Figura 17. Red de tres capas

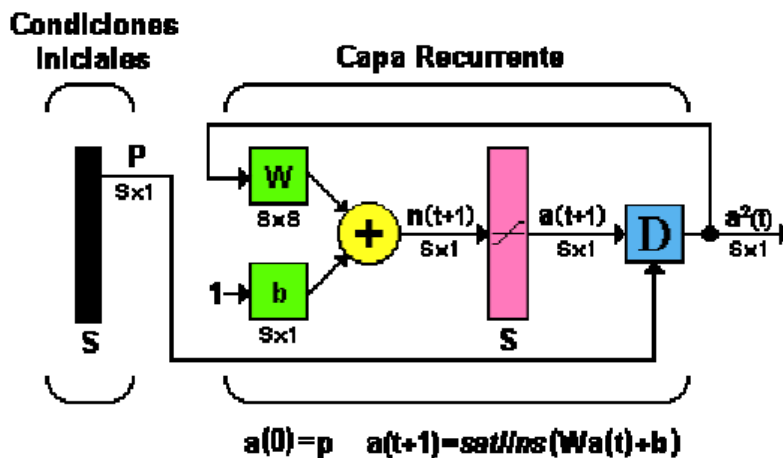


Fuente. (Buitrago & Muñoz, 2000)

Las redes multicapa son más poderosas que las redes de una sola capa, por ejemplo, una red de dos capas que tenga una función sigmoideal en la primera capa y una función lineal en la segunda, puede ser entrenada para aproximar muchas funciones de forma aceptable.

Un tipo de redes, un poco diferente a las que se han estudiado hasta el momento, son las redes recurrentes, estas contienen una realimentación hacia atrás o retroalimentación, es decir algunas de sus salidas son conectadas a sus entradas. Un tipo de red recurrente de tiempo discreto es mostrado en la (Figura 18)

Figura 18. Redes Recurrentes



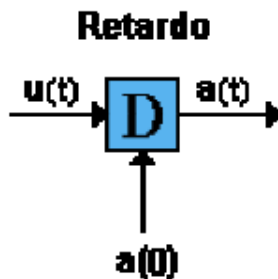
Fuente. (Buitrago & Muñoz, 2000)

La red alcanzará su estado estable cuando la salida para un instante de tiempo sea la misma salida del instante de tiempo anterior.

Las redes recurrentes son potencialmente más poderosas que las redes con realimentación hacia delante. En este tipo de redes se introducen también dos nuevos conceptos, el bloque de retardo de la (Figura 19) y el bloque integrador de la (Figura 20)

- Retardo

Figura 19. Bloque de retardo

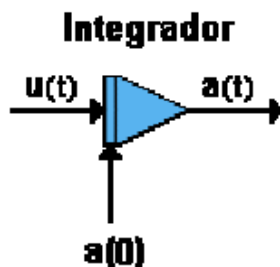


Fuente. (Buitrago & Muñoz, 2000; Freeman & Skapura, 1993)

La salida del bloque de retardo es el valor de entrada retrasado en un paso de tiempo, este bloque requiere que la salida sea inicializada con el valor $a(0)$ para el tiempo $t=0$; $a(0)$ se convierte en la salida de la red para el instante de tiempo inicial.

- Integrador

Figura 20. Bloque integrador



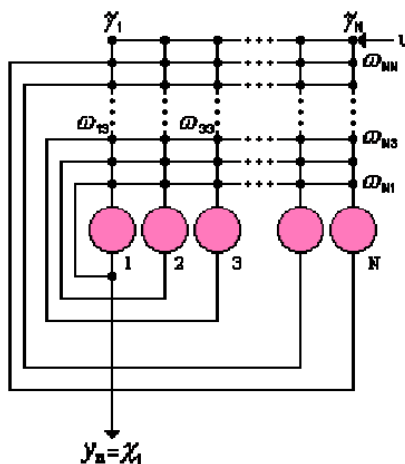
Fuente. (Freeman & Skapura, 1993)

La salida del integrador es calculada de acuerdo a la expresión

$$a(t) = \int_0^t u(t)dt + a(0)$$

Una red recurrente cuya implementación necesita un bloque integrador se ilustra en la (Figura 21)

Figura 21. Red neuronal dinámica recurrente

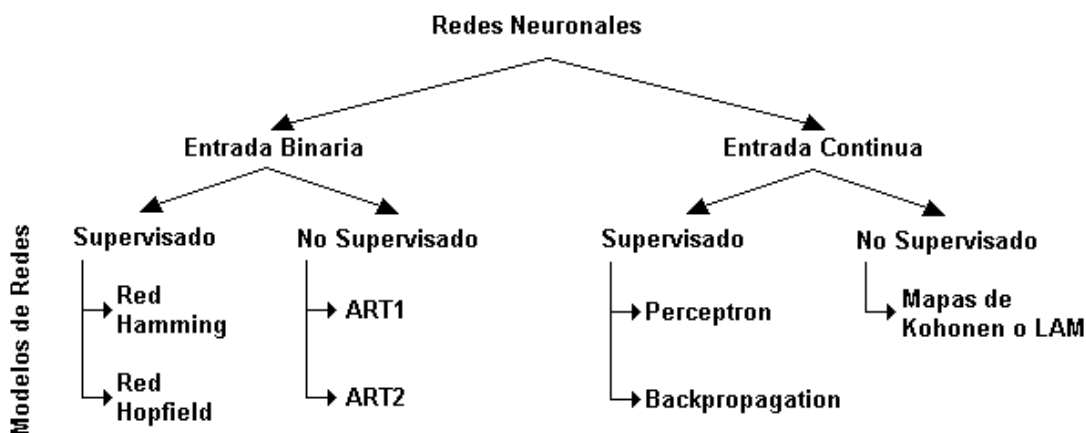


Fuente.(Freeman & Skapura, 1993)

Para garantizar la estabilidad de las redes dinámicas recurrentes en el proceso de identificación de sistemas no lineales, (Delgado, 1998) formuló condiciones estrictas para los pesos la red y su desarrollo se basa en la función de Lyapunov.

En general las redes neuronales se pueden clasificar de diversas maneras, según su topología, forma de aprendizaje (supervisado o no supervisado), tipos de funciones de activación, valores de entrada (binarios o continuos); un resumen de esta clasificación se observa en la (Figura 22)

Figura 22. Clasificación de las Redes Neuronales



Fuente. (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2006)

4.3. Teoría de la Complejidad (interdisciplinariedad)

En un sentido no exhaustivo, cuestiones relativas al desorden, el caos, la no-linealidad, el no-equilibrio, la indecibilidad, la incertidumbre, la contradicción, el azar, la temporalidad, la emergencia, la auto-organización (Morin & Pakman, 1994). La complejidad puede entenderse, por lo tanto, como un paradigma científico emergente que involucra un nuevo modo de hacer y entender la ciencia, extendiendo los límites y criterios de científicidad, más allá de las fronteras de la ciencia moderna, ancladas sobre los principios rectores del mecanicismo, el reduccionismo y el determinismo. Por otro lado, para (Rodríguez Zoya & Leónidas Aguirre, 2011) la complejidad se ubica en una zona marginal del saber científico contemporáneo, aunque sin duda sus grados de penetración, y por consiguiente de marginalidad y desconocimiento, varían de una ciencia y/o disciplina a otra. El hecho relativo a que estas ciencias hayan sido más receptivas del estudio científico de la complejidad obedece también a otro hecho no menos significativo, el cual refiere al lenguaje propio en el que se expresa el pensamiento de las ciencias físico-naturales, más proclive o sensible al empleo del lenguaje formal y matemático.

Así, el estudio y dominio de la complejidad en el terreno científico ha estado estrechamente ligado a la invención y el desarrollo de la computación moderna. De este modo, la teoría matemática de la complejidad y las ciencias de la computación constituyen el andamiaje necesario, pero no exclusivo ni exhaustivo, para el abordaje de una mirada científica de la complejidad en el campo de las ciencias de la vida y de la materia en sentido amplio. Lo que hoy suele llamarse ‘teoría de la complejidad’ (Morin, 2008) -en singular-, o en su denominación más

pluralista, ‘teorías de la complejidad’ -en plural-, es en realidad el nombre de un campo con límites borrosos que abarca, en su formulación científica, a las teorías de los sistemas complejos en sentido amplio, la teoría del caos y los fractales.

las teorías contemporáneas de la complejidad, del otro, son dos mundos con escasos puntos de conexión y con pocos puentes articuladores.⁹ Un académico destacado, se pronuncia en estos términos:

“Existen dos grandes comprensiones de complejidad, usualmente indiferentes entre sí, distantes incluso, y quizás radicalmente distintas. De un lado, la complejidad como ciencia, y de la otra, la complejidad como método. Resulta más apropiado referirnos a la primera como las ciencias de la complejidad o también, más prudentemente, como el estudio de los sistemas complejos adaptativos. En cuanto a la segunda concepción, es conocida genéricamente como el pensamiento complejo. Mientras que la primera hace referencia a diversos, incluso numerosos, autores y líneas de trabajo e investigación, en el segundo caso se trata prioritariamente de la obra de un solo autor, aunque sean numerosos sus seguidores y epígonos” (Maldonado 2007:19).

Si el solo criterio cuantitativo no se es suficiente para precisar el grado de complejidad de un sistema, menos aún lo es para clarificar las tensiones y los antagonismos entre estos dos modos de abordaje: ciencias vs. pensamiento de la complejidad, qué intereses rigen la producción de conocimiento en ciencias de la complejidad, y a quiénes benefician los saberes por ellas elaborados, apuesta a constituirse como la estrategia de conocimiento de un sujeto quien construye y redefine su estrategia de conocimiento conforme se desarrolla el proceso cognoscitivo o simplemente, toda teoría científica puede ser concebida como una sub-clase del discurso científico. El pensamiento complejo intenta vertebrar un método no clásico para el estudio de la complejidad, éste método atribuye de modo ineludible un rol central al sujeto de conocimiento en la elaboración de su estrategia cognitiva, es un sujeto abierto a la complejidad humana

La metodología de las ciencias de la complejidad, específicamente la modelización y simulación de sistemas complejos, constituye una tercera vía de hacer ciencia, distinta y complementaria a la deducción y la inducción (C. E. Maldonado & Gómez-Cruz, 2010), La oposición real en el plano metodológico entre el pensamiento complejo y las ciencias de la complejidad, concierne al lugar del sujeto en la metodología científica, las ciencias de la complejidad borran al sujeto de su formulación metodológica y epistemológica, con esta propuesta, el desafío más fundamental es estimular el desarrollo de las ciencias de la complejidad guiadas por un pensamiento complejo. Proceso por el cual se crea una abstracción de un sistema u objeto real, con la finalidad de desarrollar procesos de inferencia sobre los mismos. Una nueva civilización requiere, efectivamente, de una nueva ciencia e incluir la potencialidad metodológica

⁹ De destacar el trabajo de Carlos Eduardo Maldonado, quien, desde una óptica de filosofía de la ciencia, se ha atrevido a plantear el problema de una teoría general de la complejidad (Maldonado 2007:101-132)

de las ciencias de la complejidad en un marco epistémico ampliado a la ética, la política la educación como propone el pensamiento complejo (Rodríguez Zoya & Leónidas Aguirre, 2011).

4.4. Estadística y Probabilidad

4.4.1. Ideas estocásticas fundamentales

Es importante identificar estos conceptos, que son los que debemos enseñar en los niveles no universitarios. Cualquier tema puede ser enseñado efectivamente en alguna forma correcta a cualquier niño en cualquier estado de desarrollo, lo que implica que las ideas fundamentales (Batanero, 2001) son una guía necesaria desde la escuela primaria a la universidad para garantizar una cierta continuidad. La transición a un nivel cognitivo superior se facilita si el tema subyacente ha sido preparado en una representación conveniente en etapas cognitivas anteriores.

Por ejemplo, si un niño, al lanzar dos dados concede más probabilidad al 7, porque hay más sumas con éste valor, tiene un modelo apropiado, que puede evolucionar al más complejo de aplicación de la regla de Laplace, e incluso al de variable aleatoria y moda de su distribución de probabilidad. Por el contrario, si un niño al lanzar dos monedas explica la mayor proporción de casos «mixtos» argumentando que tras «cara» es más probable «cruz», usa un modelo de explicación que le satisface, pero no permite continuación a un estadio más elaborado. En una edad temprana ayudan al niño a entender su entorno por sus propios medios, mucho antes de que sea capaz de comprender la complejidad del modelo matemático subyacente.

El gran número de paradojas estocásticas puede confundir incluso a los expertos. Por ello, es más importante construir intuiciones correctas en este campo que en ningún otro. Como señala (Feller, 2008), la intuición estadística puede ser entrenada incluso en los adultos, aunque si un niño adquiere «intuiciones erróneas» cuando es muy pequeño, esto le puede impedir más tarde la adquisición de un conocimiento adecuado. Por ello parece necesario ofrecer a los niños actividades estocásticas, en forma de juegos y experimentos en edad temprana. Sin embargo, las actividades a desarrollar no deben escogerse al azar.

A continuación, analizamos los 10 grupos de ideas fundamentales en estocástica (Heitele, 1975) las cuales son:

- La probabilidad como normalización de nuestras creencias.

La primera idea fundamental es asignar números a los sucesos aleatorios, de forma que estos números reflejen nuestro grado de creencia en su verificación. En el lenguaje ordinario se usan expresiones del tipo de «casi cierto», «más probable que», etc., para comparar los sucesos aleatorios, pero estas expresiones son poco precisas, porque diferentes personas les conceden diferente valor. De este modo ponemos en correspondencia la multidimensionalidad del complejo mundo a nuestro alrededor con el intervalo, y se hace accesible a los dispositivos matemáticos. Podemos comparar sucesos muy dispares, en base a su mayor o menor probabilidad.

En la idea de probabilidad, tan sencilla, pero tan potente, encontramos un ejemplo de cómo el hombre idea modelos matemáticos para comprender y predecir la realidad.

Si en la idea de probabilidad incluimos todos los procesos de cálculo que llevan a asignar estos números, habremos incluido toda la teoría de la probabilidad en esta primera idea. No obstante, en el estudio que hemos hecho con anterioridad, vemos que, a pesar de su aparente simplicidad, esta idea no está libre de controversias.

- El espacio muestral como conjunto de todas las posibilidades

No menos fundamental es la idea debida a Kolmogorov de asignar un espacio muestral de sucesos observables a cada experimento aleatorio y representar los sucesos observables como subconjuntos del espacio muestral, dando una interpretación probabilística a las operaciones con sucesos. Esta idea permitió axiomatizar la probabilidad, como medida normada aditiva, sobre el álgebra de conjuntos, puesto que las operaciones en esta álgebra de conjuntos permitían definir operaciones sobre la misma probabilidad. Puesto que todo suceso elemental forma parte del conjunto de referencia, se dota de sentido al muestreo, ya que, al observar repetidamente una serie de repeticiones del experimento, siempre observaremos elementos del espacio muestral. Esta propiedad permite, desde la escuela elemental realizar actividades de probabilidad comparativa, incluso sin cualificación.

El inventariar todos los posibles resultados del experimento tiene, sin embargo, a veces dificultades para los niños que no han alcanzado un nivel de razonamiento combinatorio suficiente. La dificultad está en que hay que considerar no sólo el suceso que ha ocurrido realmente o incluso el suceso de interés sino todos los sucesos que podrían ocurrir.

- Regla de adición de probabilidades

Una regla general en matemáticas es construir modelos complejos a partir de otros más simples. Aunque las probabilidades pueden asignarse fácilmente en espacios muestrales sencillos, la regla de Laplace es difícil de aplicar directamente en casos más complejos, como obtener tres números iguales al lanzar 3 dados. La regla de adición permite obtener este tipo de probabilidades. Consiste en calcular la probabilidad de un suceso compuesto calculando por separado la probabilidad de cada uno de los sucesos simples que la componen y luego sumando estas probabilidades.

Otra aplicación de esta regla sería calcular la probabilidad del suceso contrario a uno dado. Vemos por tanto que esta idea puede aprenderse a diversos niveles de profundidad y que el aprendizaje temprano para casos finitos prepara la comprensión del caso general.

- Independencia y regla del producto

El producto cartesiano permite construir el espacio muestral del experimento compuesto. De nuevo vemos como a partir de la idea de experimento simple pasamos a la de experimento compuesto y posteriormente a la idea de proceso estocástico en tiempo discreto. Por un lado, mientras que la definición de independencia de sucesos de un mismo experimento se entiende con facilidad cuando

se usa la regla de producto, es más complejo de explicar cuando se trata de independencia de sucesos de distintos experimentos, porque en este caso, tenemos que identificar con claridad que entendemos por la intersección de los sucesos intervinientes en la fórmula del producto. Por otro lado, la idea de que un experimento aleatorio se puede repetir en las mismas condiciones y que los resultados de cada experimento son independientes es un buen ejemplo de la diferencia que hay entre entender un modelo teórico y saber aplicarlo en una situación concreta.

Por otro, ocurre que en la vida cotidiana mucha gente, incluso científicos, no son capaces de aplicar esta idea consecuentemente en las situaciones prácticas.

- Equidistribución y simetría

Así cuatro primeras ideas fundamentales no dan reglas prácticas de cómo calcular las probabilidades. Es una idea estratégica descubrir y usar las simetrías físicas o de otro tipo en las situaciones problemáticas, para decidir que ninguno de los resultados posibles tiene mayor ventaja que el resto y que, por lo tanto, podemos asignarles la misma probabilidad. Una vez que se acepta esta conclusión, a partir de los axiomas se llega con facilidad al cálculo de las probabilidades de los sucesos elementales en los espacios muestrales finitos. Por ejemplo, al lanzar un dado, la simetría supone que ninguna cara se distingue de las demás.

Hay que recalcar que la equidistribución de los sucesos elementales de un experimento no puede ser separada de la simetría estadística, es decir, la simetría confirmada por los registros estadísticos de resultados del experimento. El que un dado u otro dispositivo generador de resultados aleatorios cumpla las condiciones de simetría no es un hecho que pueda deducirse de la teoría matemática, sino de la experiencia. Parece ser que la idea de simetría es difícil de enseñar a los niños, por éste motivo y porque los niños tienen creencias sobre que algunos resultados son más fáciles que otros, a pesar de la simetría física. Sólo con el trabajo repetido de ejemplos de diversos materiales simétricos y no simétricos se irá desarrollando esta idea.

- Combinatoria

El utilizar técnicas de recuento para calcular el número de elementos favorables y desfavorables a un suceso y usar estos números para calcular las probabilidades es otra idea fundamental, sobre todo en el cálculo de probabilidades complejas. Sin embargo, es demasiado simple considerar la combinatoria tan sólo como auxiliar de la probabilidad, como puede parecer a la vista de su estructura matemática. La extracción al azar de una urna de tres objetos entre cuatro posibles es un experimento aleatorio en tres fases, que puede ser interpretado significativamente en el espacio muestral de las variaciones $V_{4,3}$. Este punto de vista es soportado por los resultados de psicología del desarrollo, en particular por los trabajos de Piaget, quien sostiene que el camino de la comprensión de los fenómenos de azar pasa por el de las operaciones combinatorias básicas.

- Modelos de urnas y simulación

La palabra clave «simulación» en estadística significa algo parecido al isomorfismo en otras ramas de las matemáticas. Consiste en poner en correspondencia dos experimentos aleatorios diferentes. Un ejemplo sería «simular» el experimento aleatorio consistente en observar el sexo de un recién nacido por el experimento aleatorio consistente en lanzar una moneda al aire. Lo importante de la simulación es que podemos operar y observar resultados del segundo experimento y utilizarlos para obtener información del primero.

Por ejemplo, si queremos saber cuál es la probabilidad que entre 100 recién nacidos hay más de un 60% de varones, podemos lanzar, por ejemplo 1000 veces 100 monedas al aire, estudiar en cada uno de los 1000 experimentos si hubo o no más de un 60% de nacimientos y obtener una estimación para la probabilidad pedida. La ventaja de la simulación es obvia, incluso en este ejemplo tan sencillo, pues permite condensar el experimento en un tiempo y espacio concreto. Otro hecho importante es que, en principio es posible asignar un modelo de urnas a la mayor parte de experimentos aleatorios, al menos a aquellos con espacio muestral numerable. Además, no sólo el experimento original puede componerse para formar nuevos experimentos, sino también los modelos de urnas asociados.

Esto permite, como ha mostrado (Polya & Zugazagoitia, 1965), simular procesos aleatorios complejos mediante una secuencia de extracciones de bolas en urnas.

- La variable aleatoria

La idea de variable aleatoria ha sido responsable de las múltiples aplicaciones actuales del cálculo de probabilidades, puesto que el cálculo de probabilidades pasó de ocuparse del estudio de la probabilidad de sucesos aislados al estudio de las distribuciones de probabilidad. La variable aleatoria y su distribución, así como el estudio de las familias de distribuciones y sus propiedades son una herramienta muy potente, porque permite trabajar con el aparato del análisis matemático. Hay serias razones que justifican por qué los matemáticos pasados, que no conocían esta idea, tuvieron serios problemas con diversas paradojas matemáticas y, por ejemplo, Bernoulli necesitara 20 años para descubrir y probar su ley débil de los grandes números, temas que hoy se explica en unas pocas líneas. Desde el punto de vista elemental, manejamos intuitivamente la idea de variable aleatoria cuando nos encontramos con juegos y experimentos en los que usamos dados, monedas, etc. Así mismo tenemos experiencias cotidianas con variables aleatorias continuas, como el tiempo de espera del autobús, o el necesario para llegar de nuestra casa al trabajo.

Mientras que la idea de media es muy intuitiva, lo es menos la idea de distribución, especialmente cuando unos valores son más probables que otros. Sea cual sea la respuesta, no debe negarse el hecho de que la distribución normal tenga una parte fundamental para explicar el mundo que nos rodea, ya que la encontramos en muchos fenómenos físicos, psicológicos o de otro tipo. Un modelo matemático convincente para explicar la presencia de la distribución normal en tantos fenómenos es el teorema central del límite que, por otra parte, no es accesible en su forma

deductiva en ningún nivel por debajo de la universidad. Esto no excluye que sea interesante con los alumnos de secundaria considerar alguna aproximación intuitiva y experimental-inductiva, por ejemplo, usando el aparato de Galton u otro dispositivo de simulación.

- Las leyes de los grandes números

Para analizar la dificultad de comprensión de la convergencia, hay que distinguir entre las leyes empíricas de los grandes números y las correspondientes leyes matemáticas deducidas en forma de teoremas por diferentes probabilistas y que pueden ser demostradas formalmente. La convergencia empírica es observable en la realidad, por ejemplo, al observar las gotas de lluvia sobre un pavimento, o la proporción de recién nacidos varones en un hospital a lo largo del año. Esta convergencia empírica hace que las correspondientes leyes matemáticas de los grandes números se justifiquen como un buen modelo para los fenómenos aleatorios, aunque no contesta la pregunta de si es posible que los alumnos sean capaces de diferenciar entre el modelo y realidad, ya que de hecho vemos que con frecuencia se espera una convergencia empírica demasiado rápida o demasiado exacta. Las oportunidades didácticas de experiencias empíricas sobre la convergencia serían deseables desde la escuela, para preparar la comprensión posterior de los teoremas matemáticos.

- Muestreo

La última idea fundamental es la de muestra que nos introduce en la inferencia y establece otro nuevo puente entre estadística y probabilidad. El conocimiento científico se adquiere a partir de las experiencias empíricas y estas son siempre limitadas, por lo que las conclusiones deben ser más amplias de los datos que obtenemos en las observaciones. La representatividad nos indica que la muestra se parece, en cierto modo a la población y debe ser una característica importante, ya que un prejuicio es sólo juicio basado en una muestra no representativa.

Por otro lado, la variabilidad indica que una muestra puede ser diferente de otra, por lo que al enjuiciar, pensar e inferir sólo es posible a base de muestras, la gente debería ser cauta y crítica al argumentar.

4.4.2. Análisis exploratorio de datos

Hasta los comienzos del siglo XX la estadística se restringía a la estadística descriptiva, que, a pesar de sus limitaciones, hizo grandes aportaciones al desarrollo de las ciencias experimentales. A partir de esa época, comenzaría la inferencia estadística clásica, con los trabajos de Fisher, Pearson y sus colaboradores y progresivamente se incorporaría la aportación de la escuela bayesiana. Cabría señalar que los avances del cálculo de probabilidades llevaron a la creación de la estadística teórica, que, en cierto modo, se alejó de las ideas estadísticas primitivas centradas en el análisis y recogida de datos. De este modo, en los años 60, la mayor parte de los libros de texto se ocupaban especialmente de los modelos inferenciales clásicos o bayesianos con respecto a

conjunto simple de datos y hubo una tendencia a la matematización, junto con un descuido en la enseñanza de los aspectos prácticos del análisis de datos.

Puesto que era fácil analizar grandes muestras ya no había por qué limitarse a los métodos estadísticos basados en distribuciones conocidas, cuya principal aplicación eran las pequeñas muestras. Tampoco había por qué restringirse a analizar una o unas pocas variables, porque el tiempo de cálculo se había eliminado y era preferible aprovechar toda la información disponible. Como consecuencia, durante las últimas décadas se han desarrollado una serie de tipos de análisis de datos que se sitúan entre la estadística descriptiva y la inferencia o estadística teórica. Entre estos tipos se encuentran el análisis exploratorio de datos desarrollado por (Tukey, 1977) y el análisis multivariante.

En primer lugar, se disminuía la importancia visual de la representación de los datos, dándose la exclusividad a los cálculos y en segundo se pensaba que para obtener conclusiones de los datos era preciso recurrir a la inferencia. Partiendo de esta hipótesis, que es previa a la recogida de datos, se calculan los estadísticos para aceptar o no la hipótesis. En realidad, los datos se recogen con el único propósito de poner tal hipótesis a prueba. Al contemplar solamente dos alternativas, los datos no se exploraban para extraer cualquier otra información que pueda deducirse de los mismos.

Las diferencias de los datos con respecto a esta estructura, representan las desviaciones o residuos de los datos, que usualmente no tienen por qué presentar una estructura determinada. Tradicionalmente el estudio se ha concentrado en la búsqueda de un modelo que exprese la regularidad de las observaciones. En lugar de imponer, en hipótesis, un modelo a las observaciones, se genera dicho modelo desde las mismas.

4.4.2.1. *Características educativas del análisis exploratorio de datos*

Como hemos indicado, nos encontramos ante una nueva filosofía en la aplicación de los métodos de análisis de datos, aunque unida a ella se han desarrollado también algunas técnicas concretas para su aplicación, como el gráfico de la caja o el de tallo y hojas. Esta filosofía consiste en el estudio de los datos desde todas las perspectivas y con todas las herramientas posibles, incluso las ya existentes. Al considerar la conveniencia o no de incluir un tema como objeto de enseñanza hemos de tener en cuenta su utilidad y que este tema se halle al alcance de los alumnos. Estos conjuntos de datos pueden ser obtenidos por los mismos estudiantes, mediante la realización de una encuesta a sus compañeros sobre temas diversos, como características físicas, aficiones, empleo del tiempo libre, etc., o a partir de datos obtenidos en anuarios, publicaciones estadísticas o Internet.

Por ejemplo, pasar de lista de números a representaciones como la del «tallo y hojas», puede facilitar la exploración de la estructura total, así como construyendo gráficos, como el de la «caja» que hace posible la comparación de varias muestras.

4.4.2.2. *Asociación*

4.4.2.2.1. *Importancia en estadística*

De la extensión de la idea de correlación a variables cualesquiera, incluso no numéricas, surge el concepto general de asociación.

- Regresión

Esta relación se representa por un modelo matemático, dado por la ecuación de regresión, a la que se añaden un conjunto de hipótesis básicas. En particular, si la ecuación de regresión es lineal respecto a estos parámetros hablamos de regresión lineal. Entre los tipos de regresión no lineal podemos citar la regresión logística y la regresión polinómica.

Se quiere predecir la variable dependiente, a partir de los valores de las variables independientes, lo cual es muy útil si la variable dependiente es costosa o difícil de medir.

Los problemas estadísticos implicados en el análisis de regresión son:

- Obtener un buen estimador de los parámetros del modelo;
- Contrastar hipótesis sobre dichos parámetros;
- Determinar la bondad del modelo para los datos particulares
- Comprobar que se cumplen las hipótesis exigidas.

4.4.2.2.2. *Asociación y Correlación*

La medida de la intensidad de la relación entre dos variables, mediante un coeficiente adecuado, constituye el problema de correlación y asociación, en el caso general.

4.4.2.2.3. *Papel de la asociación y regresión en los métodos estadísticos*

En los casos mencionados, las ecuaciones obtenidas expresan la variable dependiente como combinación lineal de las variables independientes. Supongamos que no es necesario distinguir entre variables dependientes e independientes, sino que lo que se trata es de re expresar cada variable en función del resto. Ello puede interpretarse como obtener un cambio de base en un espacio vectorial en el que los ejes primitivos de coordenadas son las variables observadas y los nuevos ejes combinaciones lineales de las mismas. Si se combina esta idea con la de que el nuevo sistema de ejes esté formado por un conjunto de variables incorrelacionadas y que, además, se obtengan en un orden decreciente en cuanto a la varianza explicada, se obtiene el modelo de componentes principales.

En el caso de sólo dos variables se obtiene la recta de regresión lineal, por lo que el análisis de componentes principales puede considerarse una extensión de la regresión. De ahí surge primitivamente el análisis factorial y posteriormente las técnicas derivadas de él, hasta llegar al

análisis de correspondencias, que extiende el análisis factorial al estudio de las tablas de contingencia de variables cualitativas dobles o múltiples. Otra posibilidad de extensión es la de suponer que no existe una diferenciación estricta entre variables dependientes e independientes, sino que las diversas variables están ligadas entre sí formando redes causales donde unas variables dependen de otras formando un sistema de interrelaciones múltiples.

4.4.2.2.4. *Asociación y causalidad*

Por ejemplo, Kendall y Lazarsfeld, Campbell y Stanley, o Blalock, consideran que si una variable A es «causa» de otra variable B entonces A está correlacionada positivamente con B. Otros autores tales como Asher, consideran que si A es «causa» de B entonces el coeficiente de correlación puede servir confirmar el grado de la influencia «causal» de A en B. Esto justifica la práctica de rechazar las hipótesis «causales» cuando A no está correlacionado positivamente con B. No entraremos aquí en un análisis completo de la noción de causa a la que se ha atribuido diferentes significados en filosofía y metodología de investigación. Lo que haremos es precisar el punto de vista «estadístico», base de los modelos causales-estructurales. Estos modelos estructurales extienden las ideas de correlación y regresión en el sentido de admitir en el estudio redes complejas de variables interconectadas, en lugar de efectuar una separación estricta de las variables en dependientes e independientes.

4.4.3. Modelos de regresión

En esta sesión nos dedicaremos al estudio de una parte de la estadística bivariada, cuando se tienen dos variables de estudio, se puede configurar una de las siguientes situaciones:

- Dos variables categóricas: Tablas de contingencia, pruebas de asociación, estadístico χ^2 , pruebas de concordancia, Análisis de correspondencia simples, modelos para tablas de contingencias.
- Dos variables numéricas: Gráficos de dispersión, correlación, histogramas bivariados, regresión simple.
- Una variable categórica y una numérica: Tablas de resumen, comparaciones por grupos, boxplot por grupos, ajuste de modelos de diseños de experimentos. Modelos de regresión logística.

Nuestro análisis de regresión se centra principalmente en la situación donde se tiene 2 variables de tipo numérico (regresión simple) y luego puede ser extendida a un grupo más amplio de variables (regresión múltiple).

4.4.3.1. *Modelo de regresión lineal simple*

Cuando se tienen dos variables numéricas, el análisis de regresión supone la existencia de:

- Una variable explicativa o dependiente Y.

- Una variable explicativa o independiente o regresora X.

Las cuales pueden ser relacionadas a través de un modelo de la forma

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, i = 1, 2, 3 \dots n$$

sobre un conjunto de n individuos.

- β_0 representa el intercepto del modelo e regresión, es decir, el valor estimado para Y cuando X=0.
- β_1 representa el incremento marginal esperado en Y, cuando X aumenta una unidad.
- ϵ_i se asocia a los residuales del modelo, ya que es normal que los puntos no se ajusten completamente a la línea recta, los residuales son entonces las diferencias que hay los valores observados y los proyectados según la recta de regresión. Los residuales del modelo no deben interpretarse como debilidades o errores en la construcción del modelo, deben interpretarse como componentes no explicados por el modelo.

El problema se centra entonces en conocer (estimar) los parámetros del modelo β_0, β_1 .

4.4.3.1.1. Métodos de Mínimos Cuadrados

El método de mínimos cuadrados es un procedimiento matemático que busca minimizar la suma de cuadrados de los errores usando:

$$\sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2$$

La solución del proceso se conoce como los estimadores de mínimos cuadrados para el modelo de regresión simple es el siguiente:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{Cov(x, y)}{S_x^2} = \frac{S_{xy}}{S_x^2} = r \frac{S_y}{S_x}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

4.4.3.1.2. Valores estimados y residuales del modelo

Una vez se tiene las estimaciones para los parámetros del modelo, es posible construir estimaciones para Y en cada uno de los valores observados de X y los residuales, que son las diferencias entre los valores observados y los ajustados/estimados/pronosticados por el modelo. Los valores estimados se calculan como:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Mientras que los residuales toman la forma:

$$\hat{r}_i = y_i - \hat{y}_i$$

Estos son los valores que caen sobre la recta ajustada para los diferentes valores de X

4.4.3.1.3. *Supuestos del modelo de regresión*

Para realizar los análisis posteriores, es necesario asumir unos supuestos sobre los residuales del modelo:

- Los residuales provienen de una distribución normal.
- Los residuales son independientes entre sí.
- Los residuales tienen varianza constante σ^2

Consideraciones a tener en cuenta:

- Los supuestos sobre los residuales no son necesarios para encontrar las estimaciones del parámetro.
- Los supuestos son necesarios para hacer pruebas de hipótesis sobre los parámetros del modelo.

4.4.3.1.4. *Coefficiente de determinación*

Como una medida de la calidad del ajuste del modelo a los datos, se construye el coeficiente de determinación R^2 bajo la siguiente expresión:

$$R^2 = 1 - \frac{\sum r_i^2}{\sum (y_i - \bar{y})^2}, \quad \text{con } 0 < R^2 < 1$$

Aunque no existe una frontera para clasificar con base en R^2 los modelos en buenos y malos, puede decirse que este modelo no sería del todo confiable en la predicción de Y

Por esta razón cuando un modelo de regresión simple (una sola variable predictiva x), el coeficiente de determinación no es muy alto, debe explorarse la situación para vincular más variables al modelo con el propósito de explicar mayor porcentaje de la variación de y.

4.4.4. **Otros modelos de regresión**

Cuando se tienen dos variables X, Y no siempre la relación entre las dos variables es de tipo lineal, también puede ser de otro tipo como los siguientes:

- Lineal: $Y = \beta_0 + \beta_1 X$

- Logarítmico: $Y = \beta_0 + \beta_1 \log(X)$
- Inverso: $Y = \beta_0 + \beta_1 \frac{1}{X}$
- Cuadrático: $Y = \beta_0 + \beta_1 X + \beta_2 X^2$
- Cubica: $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$
- Potencia: $Y = \beta_0 (X^{\beta_1})$
- Compuesto: $Y = \beta_0 (\beta_1^X)$
- Exponencial: $Y = \beta_0 \exp^{\beta_1 X}$
- Crecimiento: $Y = \exp^{\beta_0 + \beta_1 X}$

4.5. NRA¹⁰ y modelos estadísticos

Uno de los aspectos que han fomentado la idea errónea acerca de las diferencias entre RNA y estadística versa sobre la terminología utilizada en la literatura de ambos campos. Recordemos que el campo de las RNA surge como una rama de la IA con una fuerte inspiración neurobiológica se presenta la Tabla 2. Equivalencia en la terminología usada en estadística y de redes neuronales. Tabla 2 en la que se pone de manifiesto que las RNA y la estadística utilizan términos diferentes para nombrar el mismo objeto (Sarle, 1994; Vicino, 1998)

Tabla 2. Equivalencia en la terminología usada en estadística y de redes neuronales.

Terminología estadística	Terminología de redes neuronales
Observación	Patrón
Muestra	Datos de entrenamiento
Muestra de validación	Datos de validación, test
Variabes explicativas	Variabes de entrada
Variable de respuesta	Variable de salida
Modelo	Arquitectura
Residual	Error
Error aleatorio	Ruido
Estimación	Entrenamiento, aprendizaje
Interpolación	Generalización
Interacción	Conexión funcional
Coefficientes	Pesos de conexión
Constante	Peso umbral
Regresión y análisis discriminante	Aprendizaje supervisado o heteroasociación
Reducción de datos	Aprendizaje no supervisado o autoasociación
Análisis de cluster	Aprendizaje competitivo

Fuente. (Moreno, 2004)

¹⁰ Redes Neuronales Artificiales vs modelos estadísticos

Se puede establecer una similitud entre modelos estadísticos y modelos de redes neuronales ver Tabla 3 (Moreno, 2004; Sarle & FI, 1994)

Tabla 3. Equivalencia entre modelos estadísticos y modelos de red neuronal

Modelo estadístico	Modelo de red neuronal
Regresión lineal múltiple	Perceptrón simple con función lineal
Regresión logística	Perceptrón simple con función logística
Función discriminante lineal	Perceptrón simple con función umbral
Regresión no lineal múltiple	Perceptrón multicapa con función lineal en la salida
Función discriminante no lineal	Perceptrón multicapa con función logística en la salida
Análisis de componentes principales	Regla de Oja Perceptrón multicapa autoasociativo
Análisis de clusters	Mapas autoorganizados de Kohonen
K vecinos más cercanos	Learning Vector Quantization (LVQ)
Regresión kernel	Funciones de Base Radial (RBF)

Fuente (Moreno, 2004)

La mayoría de redes neuronales aplicadas al análisis de datos son similares y, en algunos casos, equivalentes a modelos estadísticos conocidos. A continuación, se describirán algunas de las relaciones que se han establecido a nivel teórico entre ambas perspectivas.

Para (Sarle, 2002) un Perceptrón simple puede ser considerado como un Modelo Lineal Generalizado (MLG), debido a la equivalencia entre el concepto de función de enlace en un MLG y la función de activación de la neurona de salida en un Perceptrón (Moreno, 2004):

$$Y \equiv f(X, W)$$

El valor de la variable de respuesta Y se obtiene aplicando una función de enlace sobre una combinación lineal de coeficientes W y variables explicativas X . La función de enlace en un MLG no suele estar acotada y, en la mayoría de casos, es necesario que sea monótona como las funciones identidad, recíproca y exponencial.

El concepto de discrepancia en un MLG y el concepto de función de error en un Perceptrón también son equivalentes (Biganzoli, Boracchi, Mariani, & Marubini, 1998). En el caso del Perceptrón la función que en general se intenta minimizar es la suma del error cuadrático:

$$E = \sum_{p=1}^P \frac{1}{2} \sum_{k=1}^M (d_{pk} - y_{pk})^2$$

Donde P es el número de patrones, M el número de neuronas de salida, d_{pk} es la salida deseada para la neurona de salida k para el patrón p y y_{pk} es la salida obtenida por la red para la neurona de salida k para el patrón p.

Una diferencia radica en que el Perceptrón normalmente estima los parámetros del modelo mediante el criterio de mínimos cuadrados, es decir, intentando minimizar la función E, el MLG ajusta el modelo mediante el método de máxima verosimilitud para una variedad de distribuciones de la clase exponencial, la función de error que se intenta minimizar se denomina cross entropy (Bishop & Bishop, 1995) que viene dada por:

$$E = - \sum_{p=1}^P \sum_{k=1}^M [d_{pk} \log y_{pk} + (1 - d_{pk}) \log(1 - y_{pk})]$$

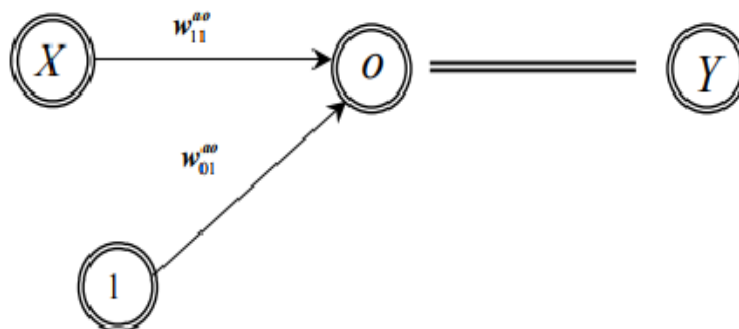
Estableciendo analogías entre RNA y modelos concretos pertenecientes a MLG, un Perceptrón simple con función de activación lineal en la neurona de salida y utilizando la suma del error cuadrático equivale a un modelo de regresión lineal (Liestol, Andersen y Andersen, 1994; Michie, Spiegelhalter y Taylor, 1994; Sarle, 1994; Kemp, McAulay y Palcic, 1997) ver (Figura 23. Perceptrón simple con función lineal = Modelo de regresión lineal Figura 23)

$$Y = aX + b + \varepsilon$$

Con ε una variable aleatoria de medida cero, y varianza infinita. Si se desea predecir el valor de la variable Y para $X=x$, se emplea el valor esperado de la distribución continuada:

$$\hat{Y} = E[Y/X = x] = ax + b$$

Figura 23. Perceptrón simple con función lineal = Modelo de regresión lineal



Fuente:(Castellano, 2009)

La (Figura 23) muestra la estructura de la red que describiría una regresión lineal simple. La salida de la red será:

$$O = W_{11}^{a0} x + W_{01}^{a0}$$

Que coincide con la estructura de la predicción que proporciona un modelo de regresión lineal. Se precisa pues de un perceptrón sin capa oculta, con función de activación, la identidad, el modelo más sencillo de perceptrón, para recrear la regresión lineal simple.

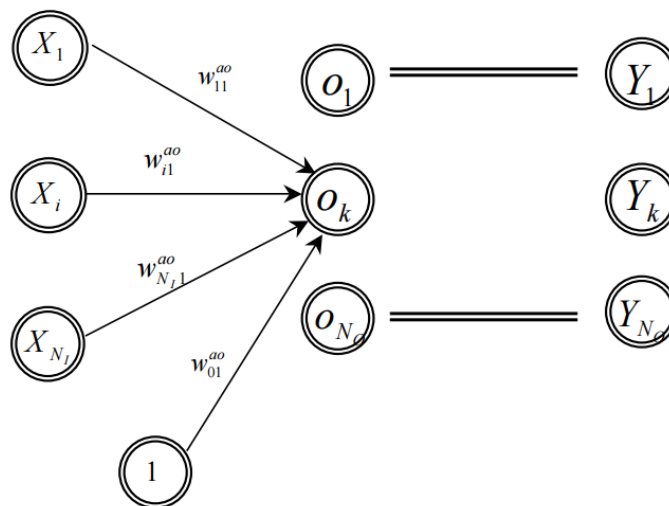
La generalización del caso anterior (Cachero, 1996; Montgomery et al., 2005)) consiste en considerar las variables explicativa y dependiente como multidimensionales (Figura 24), obteniéndose el modelo de regresión:

$$Y_k = \sum_{i=1}^{N_I} W_{ik} X_i + W_{0k} + \varepsilon, \text{ para } 1 \leq k \leq N_0$$

La predicción será

$$\hat{Y}_k = \sum_{i=1}^{N_I} W_{ik}^{a0} x_i + W_{0k}^{a0}, \text{ para } 1 \leq k \leq N_0$$

Figura 24. Red Neuronal para la Regresión Lineal Múltiple Multidimensional

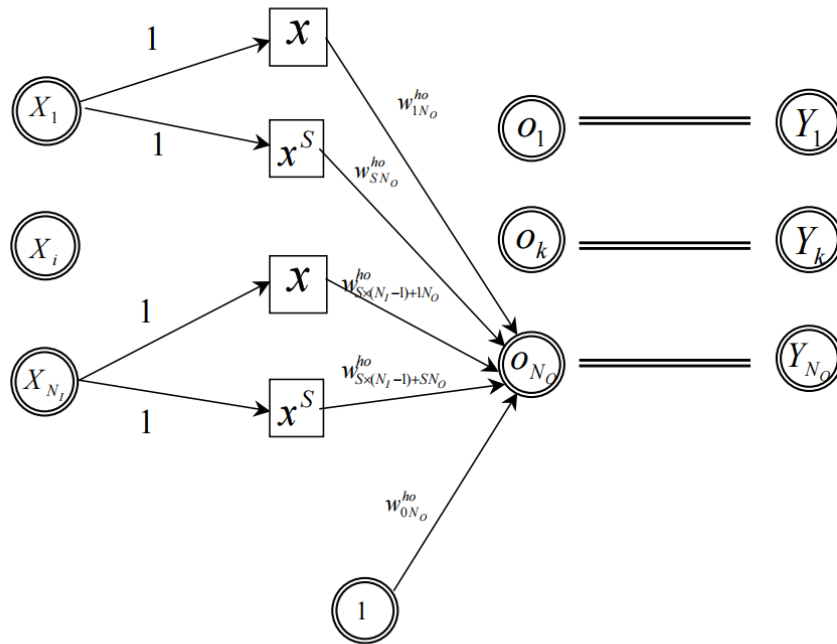


Fuente:(Castellano, 2009)

La llamada regresión polinómica, que tratará de reescribir la función () que relaciona la variable regresora con la variable respuesta que se desea predecir o estimar. La predicción asociada a este el modelo de la variable dependiente k-ésima será:

$$\hat{Y}_k = \sum_{i=1}^{N_I} \sum_{j=1}^S W_{ijk} (x_i)^j + W_{0k}, \quad \text{para } 1 \leq k \leq N_O$$

Figura 25. Red Neuronal para la Regresion Polinmica



Fuente:(Castellano, 2009)

5. OBJETIVOS

5.1. Objetivo general

Diseñar una propuesta metodológica desde la complejidad para integrar las redes neurales al currículo de estadística y probabilidad con un enfoque interdisciplinario.

5.2. Objetivos Específicos

- Estudiar los conceptos básicos de redes neuronales biológicas y artificiales.
- Estudiar el modelo de regresión en el marco del currículo de estadística y probabilidad.
- Analizar la relación entre las redes neuronales artificiales y los modelos estadísticos.
- Modelar y simular sistemas complejos usando IBM Spss Statistics
- Diseñar un microdiseño curricular sobre redes neuronales artificiales

6. METODOLOGÍA

6.1. Tipo y enfoque de la investigación

El enfoque del estudio es cualitativo y el tipo resulta ser una combinación de teoría fundamentada e investigación acción.

Según Hernández (2010), este enfoque se emplea cuando el tema del estudio ha sido poco explorado, o no se ha hecho investigación al respecto en algún grupo social específico. Los planteamientos cualitativos están enfocados en profundizar en los fenómenos, explorándolos desde la perspectiva de los participantes. Este es el caso para abordar desde la complejidad la integración de las redes neuronales con el currículo de estadística y probabilidades con un enfoque interdisciplinario. Este es un proceso cualitativo el cual no es lineal, sino iterativo o recurrente, las supuestas etapas en realidad son acciones para adentrarnos más en el problema de investigación y la tarea de recolectar y analizar datos es permanente.

El tipo o diseño de teoría fundamentada (Hernández Sampieri et al., 2006) utiliza un procedimiento sistemático cualitativo para generar una teoría que explique en un nivel conceptual una acción, una interacción o un área específica. En este caso como estudiar las redes neuronales desde su naturaleza estadística. Esta teoría es denominada sustantiva y es de naturaleza “local” (se relacionan con una situación y un contexto particular). Sus explicaciones se circunscriben a un ámbito determinado, pero poseen riqueza interpretativa y aportan nuevas visiones de un fenómeno.

La teoría fundamentada puede representar toda la complejidad descubierta en el proceso. Asimismo, la teoría fundamentada va más allá de los estudios previos y los marcos conceptuales preconcebidos, en búsqueda de nuevas formas de entender los procesos sociales que tienen lugar en ambientes naturales.

Este diseño se ha definido para este proceso ya que está inmerso en otro más general que es el proceso de autoevaluación del programa de matemática aplicada, con su propia dinámica: “El modelo se fundamenta en los principios de participación y continuidad establecidos por el Consejo Nacional de Acreditación para la acreditación de programas y de instituciones, y en consecuencia debe garantizar que su desarrollo involucra a todos los actores del programa, así como que su carácter de proceso permanente debe garantizar que hay continuidad en el tiempo entre las acciones y actividades, al margen de los resultados esperados en cada etapa” (Modelo de autoevaluación de la universidad Surcolombiana, 2015 Dirección general de currículo)

Esta dinámica posee características propias del diseño de la investigación-acción, cuyo propósito es resolver problemas cotidianos e inmediatos y mejorar prácticas concretas. Se centra en aportar información que guíe la toma de decisiones para programas, procesos y reformas estructurales (Hernández Sampieri et al., 2006). El diseño participativo implica que las personas

interesadas en resolver el problema ayudan a desarrollar todo el proceso de la investigación: de la idea a la presentación de resultados. Las etapas o ciclos para efectuar una investigación-acción son: detectar el problema de investigación, formular un plan o programa para resolver el problema o introducir el cambio, implementar el plan y evaluar resultados, además de generar retroalimentación, la cual conduce a un nuevo diagnóstico y a una nueva espiral de reflexión y acción.

6.2. Universo de estudio, población y muestra

El universo de estudio lo constituye el programa de matemática aplicada de la Universidad Surcolombiana. La población y muestra está constituida por los miembros del comité de currículo y del equipo de autoevaluación junto con todos los documentos por ellos producidos: Informes de autoevaluación, documento maestro del programa, plan de estudios y microdiseños curriculares.

Hernández (2010) engloba estos términos en uno llamado contexto: para responder a las preguntas de investigación es necesario elegir un contexto o ambiente donde se lleve a cabo el estudio; asimismo, es preciso ubicar el planteamiento en espacio y tiempo. Ya que se ha elegido un ambiente o lugar apropiado, comienza la tarea de responder a las preguntas de investigación. El ambiente puede ser tan variado como el planteamiento del problema. Tal ambiente puede variar, ampliarse o reducirse y es explorado para ver si es el apropiado. Dos dimensiones resultan esenciales con respecto a la selección del ambiente: conveniencia y accesibilidad.

La inmersión total implica observar eventos, establecer vínculos con los participantes, comenzar a adquirir su punto de vista; recabar datos sobre sus conceptos, lenguaje y maneras de expresión, historias y relaciones; detectar procesos sociales fundamentales. Tomar notas y empezar a generar datos en forma de apuntes, mapas, esquemas, cuadros, diagramas y fotografías, así como recabar objetos y artefactos; elaborar descripciones del ambiente. Estar consciente del propio papel como investigador y de las alteraciones que se provocan; así como reflexionar acerca de las vivencias.

Los tipos de muestras que suelen utilizarse en las investigaciones son las *no probabilísticas* o *dirigidas*, cuya finalidad no es la generalización en términos de probabilidad. También se les conoce como “guiadas por uno o varios propósitos”, pues la elección de los elementos depende de razones relacionadas con las características de la investigación.

6.3. Estrategias Metodológicas

El proceso se basó fundamentalmente en el análisis documental, identificación de categorías conceptuales: Redes neuronales, procesos estocásticos, apoyo computacional, aplicaciones. En un

análisis de segundo nivel se configuró la estructura temática que produce la teoría sobre las redes neuronales como herramientas de clasificación y predicción con el apoyo del software SPSS¹¹.

Para Hernández (2010), existe una estrecha vinculación entre la conformación de la muestra, la recolección de los datos y su análisis. Asimismo, se revisa el papel del investigador en dichas tareas. Los principales métodos para recabar datos cualitativos son la observación, la entrevista, los grupos de enfoque, la recolección de documentos y materiales, y las historias de vida.

El análisis cualitativo implica organizar los datos recogidos, transcribirlos a texto cuando resulta necesario y codificarlos. La codificación tiene dos planos o niveles. Del primero, se generan unidades de significado y categorías. Del segundo, emergen temas y relaciones entre conceptos. Al final se produce teoría enraizada en los datos. El análisis cualitativo es iterativo y recurrente.

En el análisis de datos cualitativos el proceso esencial consiste en que recibimos datos no estructurados y los estructuramos e interpretamos. Para efectuar un análisis cualitativo los datos se organizan. Al revisar el material, las unidades de análisis emergen de los datos. El investigador analiza cada unidad y extrae su significado. De las unidades surgen las categorías, por el método de comparación constante (similitudes y diferencias entre las unidades de significado). Así se efectúa la codificación en un primer plano. La codificación en un segundo plano, implica comparar categorías y agruparlas en temas (también mediante la comparación constante). Las categorías y temas son relacionados para obtener clasificaciones, hipótesis y teoría.

6.4. Técnicas e instrumento de Investigación

El instrumento de investigación es el investigador y las técnicas: el análisis de documentos y entrevistas abiertas para complementar la información documental.

Muestreo, recolección y análisis resultan actividades casi paralelas (Hernández Sampieri et al., 2006). La recolección de datos ocurre en los ambientes naturales y cotidianos de los participantes o unidades de análisis. El instrumento de recolección de los datos en el proceso cualitativo es el investigador. Las unidades de análisis pueden ser personas, casos, significados, prácticas, episodios, encuentros, papeles desempeñados, relaciones, grupos, organizaciones, comunidades, subculturas, estilos de vida, etcétera. Los datos se recolectan por medio de diversas técnicas o métodos y que también pueden cambiar en el transcurso del estudio: observaciones, entrevistas, análisis de documentos y registros, etcétera.

¹¹ Copyright IBM Corp. _2018. Reservados todos los derechos

7. RESULTADOS Y DISCUSIÓN

7.1. Análisis de la evaluación de las condiciones 3, 4, 5

En este documento se presenta el análisis de la evaluación de las condiciones 3, 4, 5 de calidad en las que se puede evidenciar el estado actual del programa de pregrado Matemática Aplicada de la Universidad Surcolombiana, el cual está adscrito a la Facultad de Ciencias Naturales y Exactas y creado en 2009-1 iniciando actividades con la cohorte del año 2009-2

Su organización curricular ha sido revisada y modificada a través de dos procesos de autoevaluación, el primero de febrero 2009 al 2012-A y el segundo en Febrero 2014A- al 2015A, dando cumplimiento a las políticas de autoevaluación continua establecidas por la Universidad Surcolombiana y la Facultad de Ciencias Exactas, en nuestro caso después de estos procesos de autoevaluación se analizó y encontró en las fortalezas, debilidades y estrategias de mejoramiento de las condiciones 3, 4, 5 ver Tabla 4 con el objeto de hacer un diagnóstico de como la integración de los componentes curriculares teóricos con los prácticos en las actividades registradas en cada uno de los microdiseños, este es factible en la Teoría del Pensamiento Complejo, la cual orienta este programa.

Se encontró que La Universidad Surcolombiana en su Misión también adoptó el enfoque del Pensamiento Complejo, que es referente epistemológico de este programa curricular.

Se encontró que la Teoría de la Complejidad, como fundamento epistemológico del quehacer científico en todas las áreas del conocimiento, en particular de las matemáticas, hoy ha sido oficialmente adoptado en el Proyecto Educativo de La Universidad Surcolombiana, Acuerdo 026 del 2014 emitido por el Consejo Superior Universitario, lo nos permite interactuar mejor en el contexto interno y externo de nuestra Universidad, en particular para articular los componentes flexibles del programa con otros programas de la Facultad de Ciencias Exactas y Naturales, así como con los otros programas de pregrado y postgrado¹².

Tabla 4. Análisis de las condiciones encontradas en el proceso de autoevaluación

Fortalezas	Debilidades	Estrategias de Mejoramiento
<i>Condición 3. Contenido curricular</i>		
<ul style="list-style-type: none"> Se realizó la actualización de los microdiseños curriculares. 	<ul style="list-style-type: none"> En el proceso anterior se tenía como debilidad la falta de implementación de la 	<ul style="list-style-type: none"> Trabajar en la integración de los componentes curriculares

¹² Para ampliar esta información se puede consultar segundo proceso informe autoevaluación 2014-a-2015-a programa de matemática aplicada

<ul style="list-style-type: none"> • Se han fortaleciendo los siguientes énfasis del programa: Bio-matemáticas, Economía Matemática y Sistemas Dinámicos. • Se cuenta con una primera sala de computo, denominada “La Venada”, con 25 terminales, con licencias de Matlab y SPSS. Parcialmente se ha incorporado la Matemática Computacional, especialmente usando software Libre, como el Geogebra. • Dado que el enfoque del programa es el de la Teoría de la Complejidad, se ha fortalecido en todas las asignaturas las interacciones teoría practica matemáticas. 	<p>matemática computacional en todas las asignaturas. Para comenzar a trabajar en esta debilidad incorporado parcialmente la Matemática Computacional, especialmente usando software Libre, el Matlab y SPSS, en los seis primeros semestres de la carrera.</p> <ul style="list-style-type: none"> • Así mismo, en el proceso anterior se tenía como debilidad la necesidad de culminar con la revisión de los microdiseños curriculares. En este segundo ejercicio se realizó la revisión de los microdiseños curriculares en sus contenidos y metodologías, pero también ajustándolos al formato institucional. • De este estudio surge la necesidad de ofrecer las nuevas asignaturas de Algebra Lineal II y de didáctica de las matemáticas. • Falta legalizar por parte del consejo de Facultad de Ciencias Exactas y Naturales el reglamento de práctica profesional. 	<p>teóricos con los prácticos en las actividades registradas en cada uno de los microdiseños, este es factible en la Teoría del Pensamiento Complejo, la cual orienta este programa.</p> <ul style="list-style-type: none"> • Incluir como cursos obligatorios: Sucesiones y Series, curso de Teoría de Probabilidades los cuales estaban como cursos electivos del programa; se crean como cursos nuevos del Algebra Lineal II y Didáctica de las matemáticas. • Presentación para aprobación en el Consejo de Facultad de un proyecto de reglamentación de la práctica profesional.
--	--	---

Condición 4 Organización de las actividades académicas

<ul style="list-style-type: none"> • Las actividades académicas se encuentran organizados y se han actualizado de acuerdo a las necesidades del programa. • Uso de software especializado en matemáticas 	<ul style="list-style-type: none"> • En este segundo proceso se identificó la necesidad de fortalecer las actividades académicas y estrategias que el programa aplica para el desarrollo del trabajo independiente. 	<ul style="list-style-type: none"> • Continuar fomentando las distintas actividades académicas y estrategias que utiliza el programa para el acompañamiento del trabajo independiente del estudiante.
--	--	--

<p>para el desarrollo de las actividades académicas.</p> <ul style="list-style-type: none"> • La organización de actividades académicas a través del trabajo presencial y el trabajo independiente. 		
Condición 5 Investigación		
<ul style="list-style-type: none"> • Siendo un programa nuevo en la región la investigación es uno de los aspectos relevantes en el Programa y los profesores de planta continúan desarrollando la investigación y publican sus resultados. • Los cursos avanzados continúan desarrollándose en torno a las líneas de investigación del programa. • Se continúa fortaleciendo los semilleros de investigación con participación de estudiantes y docentes, con posters y ponencias en congresos nacionales. 	<ul style="list-style-type: none"> • El laboratorio de computo denominado “La Venada”, aún es insuficiente para las practicas con software especializado. • Los recursos para organizar eventos especializados en el área de Matemáticas es muy escaso. • Falta implementar un programa de estímulos académicos para los mejores estudiantes, aunque la dependencia de bienestar Universitario ofrece alimentos a la gran mayoría de los estudiantes. • En el proceso anterior se identificó como debilidad la falta de una mayor participación de los estudiantes en semilleros de investigación, se evidencia un aumento en la participación de estudiantes en semilleros. • La Universidad aún no ha implementado el programa estímulos académicos a los estudiantes vinculados a procesos investigativos. 	<ul style="list-style-type: none"> • Constituir a corto plazo nuevos grupos de investigación con líneas de investigación específicas. • Nombramiento de nuevos docentes investigadores de tiempo completo. • Poner en marcha una revista especializada en la Facultad de Ciencias Exactas y Naturales para la publicación de resultados de investigación. • Incentivar el programa de jóvenes investigadores y docentes para este campo del conocimiento. • Promover un programa de postgrado centrado en las Matemáticas y con enfoque interdisciplinario. • Participar más activamente en la actividad de ACOFACIEN y buscar nuevos convenios nacionales e internacionales.

	<ul style="list-style-type: none"> Falta categorizar por Colciencias algunos grupos de investigación existentes y nuevos. 	
--	--	--

Fuente: Elaboración Propia

7.2. Análisis del plan de estudios

Se realizó un análisis en la descripción del plan de estudios del documento maestro donde se evidencia que las redes neuronales se trabajan en los proyectos de semilleros o en proyectos de investigación no están insertados en el currículo ver Tabla 5, en los cursos de estadística y probabilidades no aparecen las RNA como herramientas en el manejo de datos.

Tabla 5. Análisis de los cursos del componente básico del programa aplicada de naturaleza matemático

Curso	Resumen	Estructura Temática
<i>I Semestre</i>		
Fundamentos de Matemáticas	El curso de fundamentos de matemáticas, como el nombre lo indica, pretende ser posibilitador de los procesos de reflexión, discusión, análisis, construcción de conceptos, exploración de métodos en lo referente al razonamiento matemático, al pensamiento aleatorio, el conteo y la medida, los sistemas numéricos, el pensamiento variacional junto con los procesos funcionales (sus inversas) y sus invariantes	<ul style="list-style-type: none"> -El Razonamiento Matemático. -El Pensamiento Aleatorio. -El Conteo y la Medida. -Sistema de los números reales y sistema de los números complejos. -El Pensamiento Variacional.
Geometría Euclidiana	Este curso se inicia con la discusión de los cinco postulados de Euclides y la apertura heurística a otros modelos de geometrías; tratamos la geometría desde la óptica vectorial en el plano y el espacio para estudiar la geometría dinámica al partir del grupo básico de los movimientos rígidos, sus composiciones y reconocimiento de invariantes de la formas regulares en el	<ul style="list-style-type: none"> -Axiomas de Euclides, Teorema de Pitágoras y Tales de Mileto. Aplicaciones. -El Grupo de la Geometría Dinámica: Simetrías e Isometrías en el plano.

	plano y en el espacio; y se termina con un estudio vectorial de la Geometría analítica y sus aplicaciones.	<ul style="list-style-type: none"> -Introducción a la Geometría Analítica Vectorial en el plano. -Descripción de las Cónicas y sus aplicaciones.
Métodos de resolución de problemas en álgebra y geometría	Los contenidos de este curso fortalecerá el conocimiento de los estudiantes de matemática aplicada en las áreas básicas de aritmética, álgebra y geometría, dándoles los instrumentos necesarios para que sean competentes en la solución de problemas matemáticos.	<ul style="list-style-type: none"> -Teorema del binomio de Newton. -Progresiones y otras sucesiones. -Sistemas y modelos lineales. -Modelización trigonométrica. -Curvas en la naturaleza. -Optimización: método gráfico. -Método de exhaustión: áreas y volumen. -Sistemas dinámicos.
<i>II semestre</i>		
Calculo Diferencial	En este curso se interpretan, se formalizan y se aplican dos de las nociones básicas de Calculo (análisis) como son las aproximaciones tanto como en ambientes discretos como en ambientes continuos, y las variaciones, teniendo esta su soporte en el concepto de Limite.	<ul style="list-style-type: none"> -Convergencia Discreta (Límites de sucesiones). -Convergencia continua y continuidad (Límites de funciones reales y continuidad). -Cálculo de Variaciones (Derivación). -Fundamentos de optimización (Aplicaciones de la derivada).

<p>Algebra Lineal I</p>	<p>Este curso contribuye a formar un profesional integral competente en el desempeño en el área de matemática aplicada, específicamente, en la modelización de algunas situaciones reales usando uno de los conceptos básicos de la matemática como lo es la de función lineal, junto con su estructura básica, el espacio vectorial. También, a través del curso se adquieren las habilidades necesarias para la confrontación de métodos cualitativos y cuantitativos para las formalizaciones en ecuaciones diferenciales ordinarias, específicamente, en los sistemas lineales con coeficientes constantes.</p>	<ul style="list-style-type: none"> -Álgebra Matricial y sus aplicaciones -Sistemas de ecuaciones lineales y aplicaciones -Inversas y determinante de una matriz -Vectores y espacios vectoriales -Transformaciones lineales y matrices
<p>Estadística Descriptiva</p>	<p>En este curso se busca proporcionar los conocimientos básicos en Estadística descriptiva y relacional necesarios para resolver problemas que se presentan frecuentemente en el rol profesional del Matemático asociado a las áreas del conocimiento.</p>	<ul style="list-style-type: none"> -Conceptos Fundamentales de Estadística. -Métodos para la organización de datos: Distribuciones de frecuencias y gráficos estadísticos. -Métodos Estadísticos de Tendencia central. -Métodos Estadísticos de Variabilidad. -Estadística Relacional Bivariada.
	<p>Este curso pretende dar herramientas informáticas a los estudiantes en lo relacionado con el lenguaje LATEX y la preparación de todo tipo de documento científico, como también realizar una introducción a la programación de algoritmos en Matlab u otro software de matemáticas. Allí se</p>	<ul style="list-style-type: none"> -Procesamiento de textos científicos. -Introducción a Matlab. -Introducción a la programación de códigos.

Sistemas de Computación Simbólica y TIC.	implementarán la construcción de diagramas de flujo y pseudocódigos, ya que son la base para programar diferentes tipos de códigos.	-Entornos interactivos de látex y aplicaciones en Matlab.
<i>III semestre</i>		
Calculo Integral	Aplicar los conceptos básicos del cálculo integral como herramienta analítica, en la modelación y solución de situaciones problema, en contextos específicos de la ciencia y la tecnología. En este curso se interpretan, se formaliza y se aplican los conceptos básicos del cálculo integral. La idea de integral según Riemann, teorema fundamental del cálculo, convergencia de integrales, y aproximación funcional.	<ul style="list-style-type: none"> -Concepto de Integral. -Integral de Riemann. -Teorema Fundamental del Cálculo. -Convergencia de Integrales. -Aproximación Funcional.
Algebra Lineal II	El álgebra lineal es la rama de la matemática que estudia aquellas estructuras llamadas espacios vectoriales y las relaciones entre ellas. Utilizando para este estudio lo más característico que es la clasificación. Entre los espacios vectoriales existen diversas relaciones; entre las de mayor frecuencia están las funciones entre ellas que respetan las estructuras de espacio vectorial, y las operaciones que permiten construir nuevos espacios vectoriales a partir de otros ya conocidos, entre ellos las funciones lineales, los subespacios y los cocientes. Cuando los espacios vectoriales son de dimensión finita, existe el isomorfismo entre el espacio vectorial de las funciones o transformaciones lineales y el espacio de las matrices. También tienen asociadas matrices, las formas bilineales y cuadráticas.	<ul style="list-style-type: none"> -Espacio vectorial de dimensión finita y el isomorfismo $\mathcal{L}(k^n, k^m) \simeq M_{n,m}(k)$ -Espacio vectorial con producto interno y la adjunta. -Algunos operadores lineales en un espacio vectorial con producto interno, y la pseudo-inversa. -Formas cuadráticas. -Determinantes. -Polinomio característico y los espacios vectoriales complejos.

Sucesiones y Series.	<p>En este curso se pretende formalizar y a la vez aplicar las nociones básicas del análisis, tal es el caso de las ideas de aproximación y variación, los resultados con ellas relacionados desde las perspectivas analítica, numérica y cualitativa en tiempo discreto (sucesiones y series).</p>	<ul style="list-style-type: none"> -Sucesiones de Números Reales -Series de Números Reales -Sucesiones y Series de Funciones -Ecuaciones en Diferencias y Sistemas Dinámicos Discretos -Series de Fourier
----------------------	---	--

IV Semestre

Ecuaciones Diferenciales	<p>Con este curso se pretende expresar en el lenguaje de las ecuaciones diferenciales aquellos principios que gobiernan muchos fenómenos, ya sean físicos, químicos, biológicos, económicos y ambientales desde los tres enfoques: cualitativo (o geométrico), analítico y numérico.</p>	<ul style="list-style-type: none"> -Ecuaciones Diferenciales Ordinarias de primer Orden. -Sistemas de Ecuaciones diferenciales de primer Orden. -Sistemas Lineales EDO'S -Sistemas de Ecuaciones diferenciales no Lineales. -Transformadas de Laplace.
--------------------------	--	---

Cálculo Vectorial	<p>Se comienza con las ecuaciones de planos, rectas, superficies cilíndricas y superficies en 3D.</p> <p>A partir del concepto de vector se definen campos escalares y vectoriales</p> <p>Se tratan los principales temas del cálculo en varias variables como son límites derivadas. e integrales. Todo el curso está orientado para estudiar los teoremas fundamentales del cálculo vectorial. El teorema de Green, el teorema fundamental para integrales de línea y el teorema de Stokes. Como aplicaciones están Optimización,</p>	<ul style="list-style-type: none"> -Vectores y funciones vectoriales. -Funciones de varias variables. Límites. -Derivadas parciales.
-------------------	---	---

	momentos de primer y segundo orden, planos tangentes, vectoriales, potencial escalar, gradiente, rotacional y divergencia, entre otras de las aplicaciones.	-Integrales múltiples. -Análisis vectorial.
Teoría de Números.	El curso consiste en un desarrollo de tipo axiomático de los temas centrales de los números enteros como son: Divisibilidad, Máximo Común Divisor, Mínimo Común Múltiplo, Números Primos, Teorema Fundamental de la Aritmética, Ecuaciones Diofánticas, Congruencias y Criterios de Divisibilidad	-Números naturales. -Números enteros. -Congruencias y ecuaciones diofánticas. -Criterios de divisibilidad. -Fracciones continuas.
Lógica y Teoría de Conjuntos.	En este curso se pretende que los estudiantes adquieran los elementos necesarios para aprender a pensar matemáticamente, construir hipótesis y probarlas y comprender el alcance de los diferentes tipos de razonamiento en la construcción de conocimiento matemático; además, teniendo en cuenta que modernamente la lógica se aplica significativamente en varios campos del saber, se hace necesario proporcionar una base que le permita al estudiante emprender sus propios caminos de investigación.	-Conceptos básicos, proposiciones, simbolización, conectores lógicos. -Inferencia lógica. -Esquemas de razonamiento. -Teoría de conjuntos. -Algebra Booleana.

V Semestre

Geometría Diferencial	El desarrollo de la mayoría de las actividades programadas en el curso de Geometría Diferencial le debe proporcionar al estudiante los elementos necesarios para: Identificar y estudiar propiedades Invariantes de curvas y superficies en una vecindad de cada uno de sus puntos. Contrastar las propiedades locales de una figura geométrica con sus propiedades globales. Identificar la	-Curvas. -Superficies. -Geodésicas y teorema de Gauss – Bonnet.
-----------------------	---	---

	naturaleza de la geometría diferencial a través del conocimiento de la curva solamente en una vecindad del punto e involucrando un proceso de límite.	
Teoría de Probabilidades.	En este curso se busca proporcionar los conocimientos básicos en Estadística inferencial o Probabilidad necesarios para resolver problemas que se presentan frecuentemente en la función profesional del Matemático circunscrito en las ciencias exactas y naturales particularmente.	<ul style="list-style-type: none"> -Probabilidad y principales distribuciones de probabilidad. -Diseño de muestras. -Estimación de parámetros. -Prueba de hipótesis. -Diseño experimental /Análisis de varianza.
Seminario de Investigación.	El seminario de investigación se ubica en quinto semestre del programa de Matemáticas aplicadas, tiene el énfasis en el abordaje a través de la teoría de la Complejidad de los aspectos matemáticos que contribuyen a resolver problemas de diversos contextos: ciencias naturales, ciencias sociales, vida cotidiana, tecnología e innovación y sus interrelaciones. Supone una buena fundamentación matemática, de los fundamentos teóricos de para la investigación, apertura hacia otras disciplinas, y creciente uso de las tecnologías.	<ul style="list-style-type: none"> -Enfoques Interdisciplinarios de la investigación -Investigación en Matemática Aplicada -Formulación de un Proyecto de Investigación
Análisis Numérico	En este curso se analizará la teoría de los métodos de aproximación numérica, la convergencia y el error de problemas como: Solución de Ecuaciones no lineales, Interpolación polinomial, Integración y diferenciación, sistemas lineales, Ecuaciones diferenciales y Sistemas no lineales.	<ul style="list-style-type: none"> -Sistemas dinámicos discretos. -Análisis lineal. -Solución de ecuaciones. -Interpolación y aproximación. -Integración y diferenciación numérica.

		<ul style="list-style-type: none"> -Ecuaciones diferenciales. -Teoría de la aproximación. -Sistemas no lineales.
Teoría de Grupos.	<p>En álgebra abstracta, la teoría de grupos estudia las estructuras algebraicas conocidas como grupos. Sus objetivos son, entre otros, la clasificación de los grupos, sus propiedades y sus aplicaciones tanto dentro como fuera de las matemáticas.</p> <p>En este curso se estudian algunas estructuras algebraicas, tal es el caso de semigrupo, monoide y grupo. Dentro de los grupos los finitos e infinitos, grupos cíclicos, grupos de permutaciones, grupos alternante, grupo simple, así como también los homomorfismos de grupos y los teoremas de Sylow para los grupos finitos.</p>	<ul style="list-style-type: none"> -Nociones preliminares. -Teoría Grupos. -Homomorfismos de grupos. -Grupos Finitos.
<i>VI Semestre</i>		
Análisis Real	<p>En este curso se pretende formalizar y a la vez generalizar las nociones básicas del análisis, tal es el caso de las ideas de convergencia y continuidad, los resultados con ellas relacionados y adquiridas en los cursos de cálculo, tanto en tiempo discreto como en tiempo continuo.</p>	<ul style="list-style-type: none"> -Conjuntos finitos e infinitos. -El cuerpo ordenado completo de los números reales. -Sucesiones y series de números reales. -Topología de la recta. -Continuidad de funciones reales. -Sucesiones y series de funciones.
		<ul style="list-style-type: none"> -Teoría de muestreo.

Estadística Aplicada.	Como ya se tiene las herramientas de la estadísticas descriptivas, ahora en este curso se aplicaran métodos de correlación, análisis de varianzas en dos; tres y cuatro variables; lo mismo que algunas series de tiempo y estadística no paramétrica tendremos como soporte de ayuda el stat - graphic, minitab y el spss.	<ul style="list-style-type: none"> -Regresión y correlación – series de tiempo. -Análisis de varianza. -Bloques aleatorizados, cuadrados latinos y diseños relacionados. -Diseño factorial 2k.
Matemática Computacional.	Este curso tiene varios componentes indispensables para adquirir habilidades informáticas para desarrollar códigos en un Lenguaje de programación como es el Matlab. Se realiza un estudio de métodos numéricos que puedan ser usados en otros cursos de la carrera. Básicamente se trata de dar herramientas para solucionar Ecuaciones ordinarias y parciales como también algunas ecuaciones diferenciales estocásticas.	<ul style="list-style-type: none"> -Introducción al lenguaje de programación matlab y latex. -Métodos de integración y sistemas lineales. -Métodos de interpolación y ecuaciones diferenciales ordinarias. -Métodos para solucionar numéricamente ecuaciones diferenciales parciales.
Teoría de Anillos y Módulos.	Se estudian nociones fundamentales sobre anillos, dominios de integridad, módulos y álgebras sobre un cuerpo conmutativo, como también las relaciones entre anillos y cuerpos como es el homomorfismo.	<ul style="list-style-type: none"> -Preliminares y nociones fundamentales sobre anillos. -Dominios de integridad. -Álgebras sobre anillos conmutativos. -Módulos.
Modelización Matemática.	Se estudian y analizan métodos, técnicas para modelizar distintos tipos de sistemas y el apoyo computacional para su simulación.	<ul style="list-style-type: none"> -Teoría General de Sistemas y complejidad. -Modelos Analíticos: Determinístico y Aleatorios. -Optimización y Modelos numéricos.

		-Estrategias para analizar y usar modelos de sistemas ambientales (Stella)
<i>VII Semestre</i>		
Análisis en Varias Variables	En este curso se pretende generalizar al espacio euclidiano n-dimensional, las nociones básicas del análisis, tal es el caso de las ideas de convergencia, continuidad y diferenciabilidad; los resultados con ellos relacionados y formalizados para sistema de los números reales, tanto en tiempo discreto como en tiempo continuo.	<ul style="list-style-type: none"> -Elementos de Topología en el espacio euclidiano n-dimensional. -Límites y Continuidad para funciones de \mathbb{R}^n en \mathbb{R}^m -Aplicaciones Diferenciables de \mathbb{R}^n en \mathbb{R}^m -Teoremas de la Función Inversa y la Función Implícita.
Economía Matemática	La Economía y la Matemáticas se relacionan interdisciplinariamente en el contexto de las ciencias sociales, la cual comprende los contextos donde el ser humano se desenvuelve, por ello el profesional de las matemáticas de conocer y aplicar los modelos básicos que la Ciencia Económica ha ido construyendo a los largo de la historia para resolver los problemas de la cotidianidad o relacionados con la previsión del futuro en la vida individual y de las comunidades, estos modelos se presentan con un lenguaje matemático para adquirir la categoría de universalidad.	<ul style="list-style-type: none"> -Matemáticas financieras e inventarios. -Análisis convexo y optimización. -Teoría de juegos en economía. -Series temporales en economía. -Proyecto empresarial.
	Con el presente curso damos un panorama general de las Biomatemática enfatizando en los modelos matemáticos para entender ciertos fenómenos en biología y disciplinas afines con apoyo computacional. Se parte de que la Biomatemática es una disciplina que combina el uso de ambas. Así, los	-Modelos Poblacionales.

Biomatemáticas	modelos matemáticos son herramientas importantes en la investigación científica en muchas áreas de la Biología, que incluye la fisiología, la Ecología, la evolución, toxicología, inmunología, mantenimiento de fuentes naturales, dinámica de la población, nutrición etc. Por consiguiente, es importante que el alumno conozca y analice la aplicación directa de ciertos modelos matemáticos en su área.	<ul style="list-style-type: none"> -Dinámica Poblacional. -Ondas de Población. -Modelos Sobre Epidemiología.
Sistemas Dinámicos.	Con esta asignatura el alumno aprenderá los conceptos básicos de los Sistemas Dinámicos. Será capaz de implementar algoritmos que permitan visualizar el comportamiento de sistemas dinámicos tanto reales como complejos, lineales y no lineales, determinísticos y estocásticos, discretos y continuos, detectando la existencia de caos, vía bifurcaciones.	<ul style="list-style-type: none"> -Historia y filosofía de los sistemas dinámicos -Sistemas dinámicos determinísticos, aleatorios y conceptos básicos. -Sistemas dinámicos discretos, continuos lineales y no lineales. -Linealización - puntos hiperbólicos en el plano, y el teorema de Hartman-Grobman. -Puntos no hiperbólicos en el plano y bifurcaciones.
VIII Semestre		
Variable Compleja.	En este curso se presenta los recursos teóricos necesarios con la finalidad de que los estudiantes conozcan los fundamentos y las relaciones que configuran esta rama de la matemática, conocida como la Variable Compleja, y además las aplique eficientemente mediante el seguimiento de ejercicios resueltos, así como también en la aplicabilidad del Teorema del residuo en el cálculo de integrales reales definida así como también integrales	<ul style="list-style-type: none"> -Funciones Analíticas u holomorfas -Series infinitas -Integración Compleja -Mapeos conformes -Problemas con valores en la frontera y valores iniciales

	reales impropias; así como también aplicaciones físicas en flujos de calor y electrostática y la transformada de Fourier.	
Topología	Mirando la topología como el estudio de las propiedades de las figuras geométricas que son invariantes, así esas figuras se sometan a deformaciones críticas que las hagan perder todas sus propiedades métricas y proyectivas. Este curso pretende establecer una línea de aprendizaje y de investigación que se centre en la construcción de estructuras topológicas, en el estudio de la continuidad y convergencia, así como también en la demostración, significado y aplicación de teoremas interesantes como la fórmula de Euler para los poliedros, el teorema de la curva de Jordan, el problema de los cuatros colores entre otros.	<ul style="list-style-type: none"> -Preliminares sobre conjuntos, relaciones y funciones -La topología del espacio Euclidiano -Espacios métricos -Espacios topológicos
Análisis funcional y EDP'S.	Este curso ofrece la formación necesaria para representar los modelos matemáticos en cuanto a: un contexto, espacio topológico o de HILBERT o de BANACH; Operadores con propiedades específicas, según el problema que se esté representado, y el tipo de soluciones que se esperan.	<ul style="list-style-type: none"> -Estudio comparado entre espacios de HILBERT y Espacios de BANACH. -El teorema del punto fijo. -Ecuación del Calor, Ecuación de Ondas, de Transporte. -El problema de Auto-valores. -Fundamentos de TEORÍA ESPECTRAL.
<i>IX Semestre</i>		
	El curso de Física Matemática pone de manifiesto muchas de las majestuosas aplicaciones que tiene la matemática para explicar fenómenos del mundo	-Algunos Tópicos del Análisis Vectorial y de Fourier.

Física Matemática	natural, macroscópico y microscópico, en particular aquellos que tienen relación con la física clásica y la física moderna. Vale la pena mencionar que para la formulación y explicación de tales fenómenos se tiene como base el análisis vectorial, las series de números complejos y por supuesto las ecuaciones diferenciales ordinarias (EDO) y parciales (EDP).	<ul style="list-style-type: none"> -Ecuaciones fundamentales de la Física Matemática. -Teoría de Sturm-Liouville y Funciones de Green. -Funciones especiales de la Física Matemática.
Teoría de Bifurcaciones	<p>En este curso se interpreta, se formaliza y se aplica las nociones básicas del fenómeno de la bifurcación, tales como:</p> <ul style="list-style-type: none"> - Puntos fijos, puntos periódicos y puntos hiperbólicos en una dimensión y en dos dimensiones. - Tipos de bifurcación, especialmente, la bifurcación de Hopf Supercrítica y Subcrítica. 	<ul style="list-style-type: none"> -Generalidades sobre Sistemas Dinámicos -Bifurcaciones en sistemas de una dimensión $\frac{dx}{dt} = f_{\alpha}(x)$ -Formas Normales -Puntos hiperbólicos en el plano -Bifurcaciones en Sistemas de dos dimensiones -Bifurcación y pérdida de hiperbolicidad

Fuente: Elaboración propia

7.3. Propuesta Curricular

Se proponen incorporar a los contenidos de los cursos del plan de estudios las redes neuronales y RNA ver Tabla 6 con sus propuestas metodológicas. Así como también se elabora la propuesta de microdiseño curricular para el componente complementario flexible llamado “fundamentos de las redes neuronales artificiales” ver (anexo A)

Tabla 6. Contenidos de redes neuronales a incorporar en cursos del plan de estudios y sugerencias metodológicas

Contenido propuestos	Cursos del plan de estudios	Sugerencias metodológicas
1. Neurona biológica 1.1. Estructura de la neurona 1.2. Excitabilidad de la neurona 1.3. Modelo del potencial de acción y su dinámica	Componente complementario flexible Biología, seminario de investigación modelización matemática ecuaciones diferenciales sistemas dinámicos, análisis numérico	Estudiar el modelo de FitzHug-Nagumo para el potencial de acción o impulso nervioso de la membrana celular
2. Redes neuronales artificiales (RNA). 2.1. Concepto y evolución histórica. 2.2. Terminología y notación 2.3. Clasificación: por conexiones, por topología y por tipo de aprendizaje 2.4. Redes probabilísticas 2.5. Modelación y simulación con redes neuronales 2.6. Los mapas autoorganizados de Kohonen	Componente complementario flexible seminario investigación modelización estadística y probabilidad sistemas dinámicos análisis numérico matemática computacional biomatemática	Estudiar una red neuronal para la predicción de variables consideradas a lo largo del tiempo. Estudio de caso: Modelación de la cuenca hidrográfica de un río. Estudiar un problema de aplicación a las ciencias medioambientales: predicción de los niveles de riesgo de polen de betula en el aire.

<p>3. Estadística y redes neuronales</p> <p>3.1. Predicción con redes neuronales</p> <p>3.1.1. Regresión lineal, polinómica, logística</p> <p>3.1.2. Otros modelos de regresión</p> <p>3.2. Clasificación con redes neuronales</p> <p>3.2.1. Métodos clásicos</p> <p>3.2.2. Métodos no paramétricos</p> <p>3.3. Otros modelos de análisis de datos</p> <p>3.4. Aproximadores universales</p> <p>3.5. Redes neuronales para datos con respuesta binario y función de error para variables objetivo binario</p>	<p>Componente complementario flexible</p> <p>seminario de investigación</p> <p>modelización matemática</p> <p>matemática computacional</p> <p>sistemas dinámicos</p> <p>teoría del caos y la complejidad</p> <p>matemática computacional</p> <p>estadística y probabilidad</p> <p>economía matemática</p>	<p>Estudiar aplicaciones de las redes neuronales en psicología: conductas adictivas y análisis de supervivencia.</p> <p>Para diferentes escenarios de simulación comparar el modelo lineal generalizado con las redes neuronales.</p> <p>Estudiar aplicaciones de las redes neuronales a problemas de control. Casos: producción y reciclado</p>
---	---	--

Fuente: Elaboración propia

7.4. Estrategias curriculares

Se estructura una propuesta metodología en el cual se usa el software IBM Spss Statistics se anexa un tutorial de redes neuronales con IBM Spss ver (anexo B), donde relacionan interdisciplinariamente, las redes neuronales, regresión lineal y IBM Spss Statistics, el cual busca apoyar los cursos de estadística y probabilidades en la inclusión de este tipo de temas usando los siguientes ejemplos.

7.4.1. Ejemplo 1: Perceptrón Multicapa Usando IBM Spss Statistics v22

El procedimiento de perceptrón multicapa (MLP) produce un modelo predictivo para una o más variables dependientes (objetivo) basadas en los valores de las variables de predicción.

7.4.1.1. Usar un Perceptrón multicapa para evaluar el riesgo de crédito

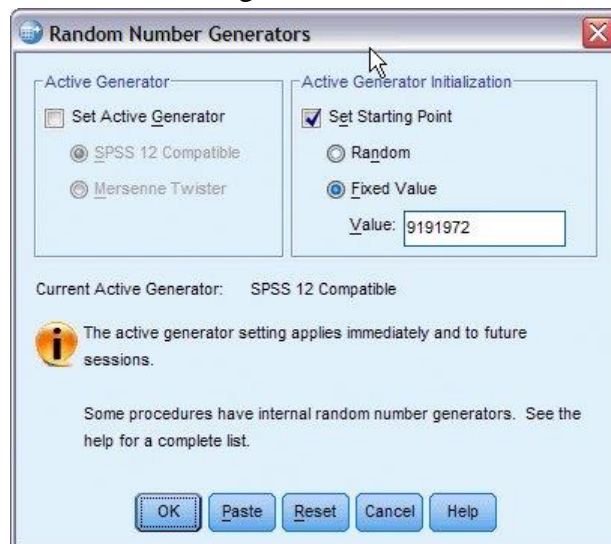
Un asesor de crédito en un banco debe poder identificar las características que son indicativas de personas que probablemente no cumplirán con los préstamos y usar esas características para identificar los riesgos de crédito buenos y malos.

Supongamos que los datos de 850 clientes anteriores y posibles clientes se encuentran en el archivo *bankloan.sav*. Si desea obtener más información, consulte Archivos muestrales en Apéndice del Spss. Los primeros 700 casos son clientes a los que anteriormente se les ha concedido un préstamo. Utilice una muestra aleatoria de estos 700 clientes para crear un perceptrón multicapa, reservando los clientes restantes para validar el análisis. A continuación, utilice el modelo para clasificar los 150 posibles clientes según los riesgos de crédito que presentan. Además, el encargado de préstamos ha analizado anteriormente los datos mediante regresión logística (en la opción Regresión) y quiere evaluar el perceptrón multicapa como herramienta de clasificación en el cual crearemos la red neuronal artificial cuyo modelo será el siguiente.

7.4.1.2. Preparación de los datos para su análisis

Establecer la semilla aleatoria, puede replicar el análisis de manera exacta. Para establecer la semilla aleatoria, elija en los menús: Transformar Generadores de números aleatorios (Figura 26).

Figura 26 Cuadro de diálogo Generadores de números aleatorios



Fuente. Copyright IBM Corp. _2018

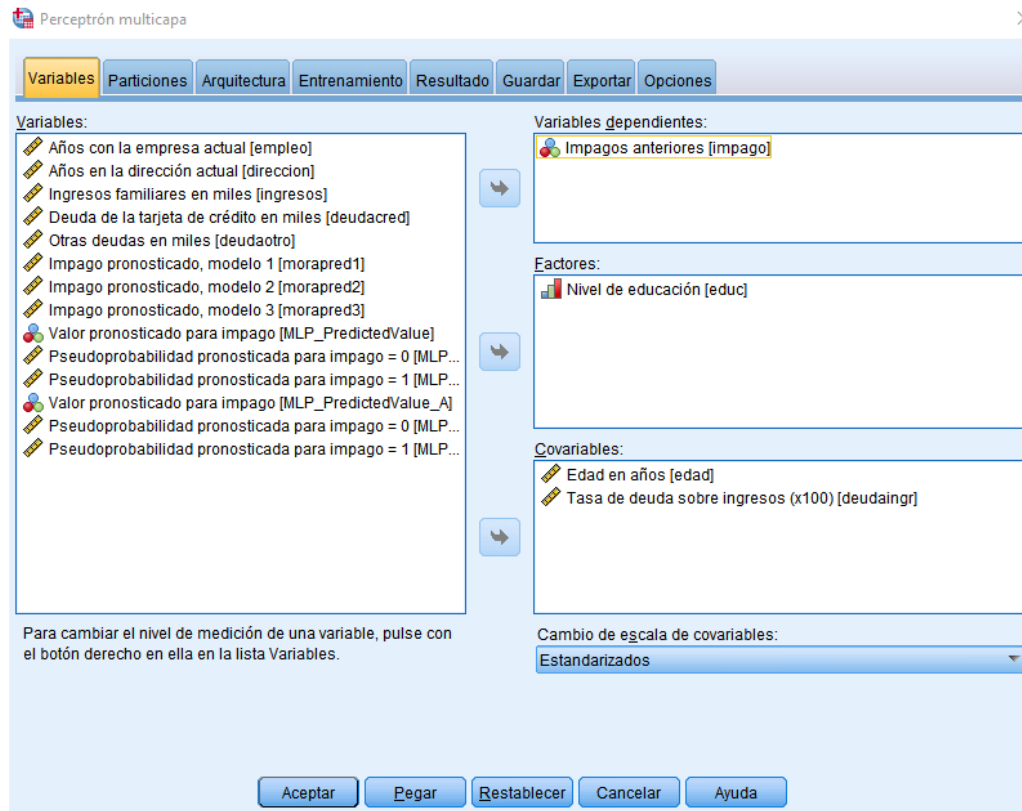
1. Seleccione Establecer punto de inicio.
2. Seleccione Valor fijo, y escriba 9191972 como el valor.
3. Click en aceptar

Luego en la ruta analizar/ redes neuronales / perceptrón multicapa como lo muestra la (Figura 27).

- Seleccione Impagos anteriores [impago] como la variable dependiente.
- Seleccione Nivel educativo [educ] como un factor.

- Seleccione Edad en años [edad] y Tasa de deuda sobre ingresos [deudaingr] como covariables.

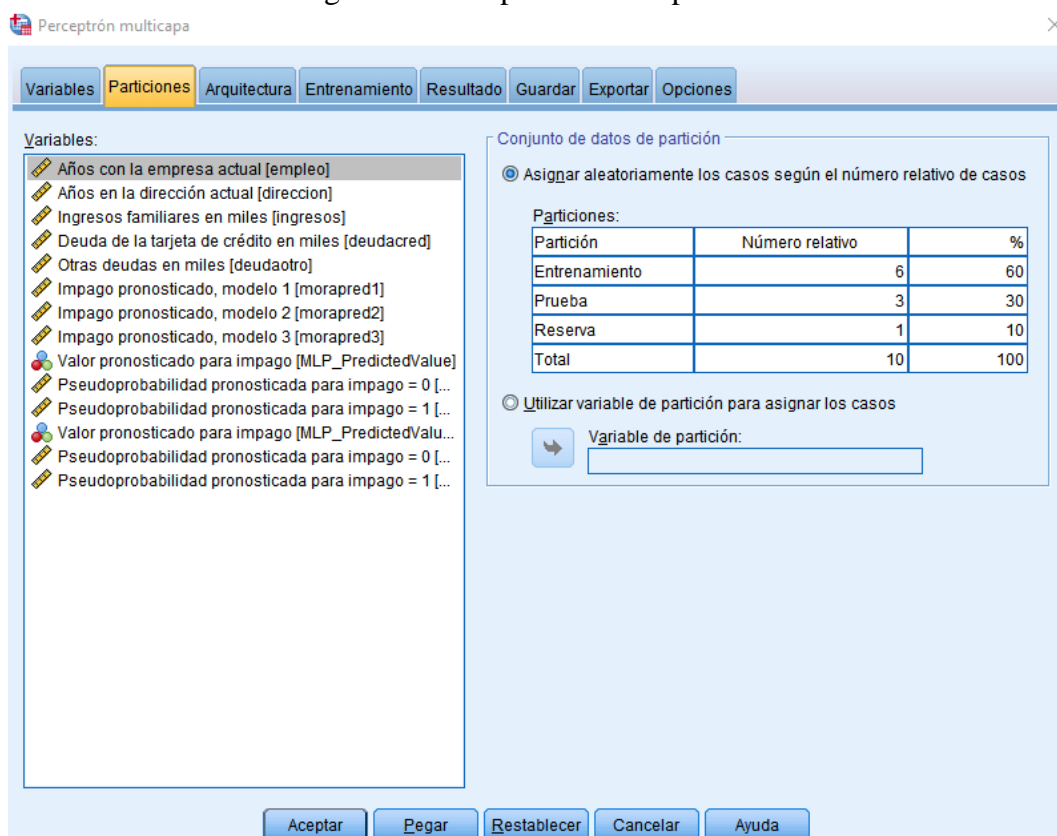
Figura 27 Perceptrón multicapa: Pestaña Variables fuente



Fuente. Copyright IBM Corp. _2018

- Pulse en la pestaña Particiones.
- Asignar en entrenamiento 6, prueba 3 y reserva 1 como lo muestra la figura 12.
- Dejar las demás pestañas por defecto y dar click en aceptar (Figura 28)

Figura 28 Perceptrón multicapa: Pestaña Particiones



Fuente. Copyright IBM Corp. _2018

Los resultados encontrados serán los visto en la siguiente Tabla 7 :

Tabla 7 Resumen de procesamiento de casos

		N	Porcentaje
Muestra	Entrenamiento	403	57,6%
	Pruebas	208	29,7%
	Reserva	89	12,7%
Válido		700	100,0%
Excluido		150	
Total		850	

Fuente. Elaboración Propia

Los primeros 700 casos son clientes a los que anteriormente se les ha concedido un préstamo el 100%. Con un 57.6% de entrenamiento, 29,7% en pruebas y de reserva un 12.7%, los cuales son valores muy cercanos a los asignados y representar una muy buena aproximación.

En la Tabla 8, la información sobre la red me indica que el número de variables de entrada son 7 y 2 capas ocultas, la función de activación de la capa oculta es Tangente hiperbólica y la de salida es Softmax.

Tabla 8 Información de red

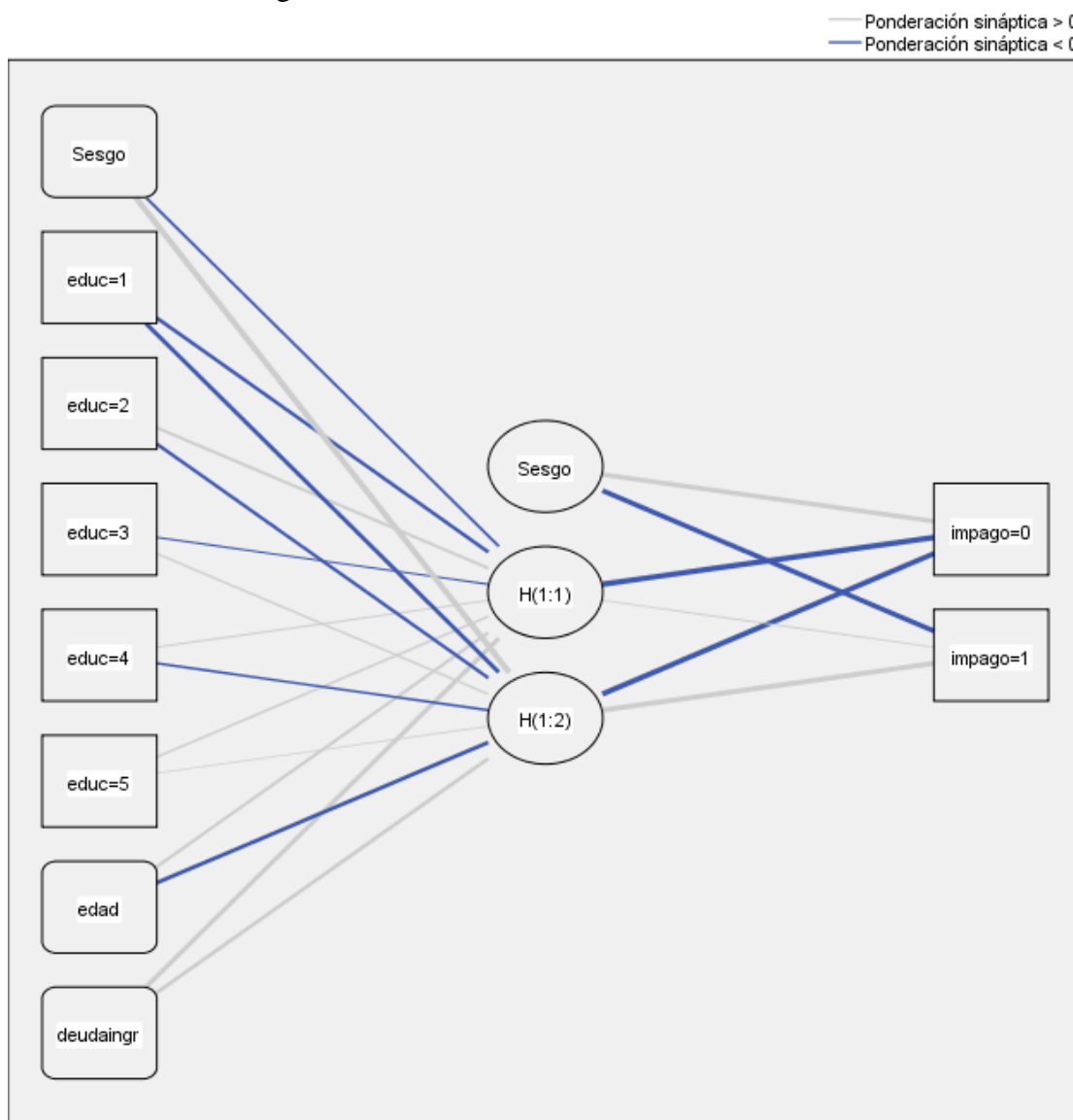
Capa de entrada	Factores	1	Nivel de educación
	Covariables	1	Edad en años
		2	Tasa de deuda sobre ingresos (x100)
	Número de unidades ^a		7
	Método de cambio de escala para las covariables		Estandarizados
Capas ocultas	Número de capas ocultas		1
	Número de unidades en la capa oculta 1 ^a		2
	Función de activación		Tangente hiperbólica
Capa de salida	Variables dependientes	1	Impagos anteriores
	Número de unidades		2
	Función de activación		Softmax
	Función de error		Entropía cruzada

a. Se excluye la unidad de sesgo

Fuente. Elaboración Propia

El modelo de red neuronal artificial que genera el IBM Spss, se observan que las variables están organizadas por categorías que son las de entrada según la (Figura 29).

Figura 29 Modelo simulación de entrenamiento



Función de activación de capa oculta: Tangente hiperbólica

Función de activación de capa de salida: Softmax

Fuente. Copyright IBM Corp. _2018

Nuevamente nos muestra un resumen del modelo en la Tabla 9 donde los Porcentaje de pronósticos incorrectos tienen muy buenos valores en el entrenamiento, prueba y reserva y los cálculos de error se basan en la muestra de comprobación.

Tabla 9 Resumen del modelo

Entrenamiento	Error de entropía cruzada	204,041
	Porcentaje de pronósticos incorrectos	26,1%
	Regla de parada utilizada	1 paso(s) consecutivo(s) sin disminución del error ^a
	Tiempo de entretamiento	0:00:00,17
Pruebas	Error de entropía cruzada	88,972
	Porcentaje de pronósticos incorrectos	17,8%
Reserva	Porcentaje de pronósticos incorrectos	20,2%

Variable dependiente: Impagos anteriores

a. Los cálculos de error se basan en la muestra de comprobación.

Fuente. Elaboración Propia

En la siguiente Tabla 10 se analizan los pronósticos y se observan que los porcentajes de comprobación de la muestra están muy cercanos entre sí para el No con un 89.5% muestra de entrenamiento, un 94.3% muestra de pruebas y un 95.5% para la muestra de reserva, un asertividad global muy buena para nuestro modelo como se observa en la tabla.

Tabla 10 Clasificación

Muestra	Observado	Pronosticado		
		No	Sí	Porcentaje correcto
Entrenamiento	No	263	31	89,5%

	Sí	74	35	32,1%
	Porcentaje global	83,6%	16,4%	73,9%
Pruebas	No	148	9	94,3%
	Sí	28	23	45,1%
	Porcentaje global	84,6%	15,4%	82,2%
Reserva	No	63	3	95,5%
	Sí	15	8	34,8%
	Porcentaje global	87,6%	12,4%	79,8%

Variable dependiente: Impagos anteriores

Fuente. Elaboración Propia

Se pueden plantear las preguntas de cambiar los porcentajes asignados en las particiones y observar el comportamiento del modelo, así como también asignarle otras variables dependientes que pueden ser más de una y hacer el análisis pertinente al comportamiento del modelo.

7.4.2. Ejemplo 2: Función de base radial con IBM Spss Statistics

El procedimiento Función de base radial (RBF) genera un modelo predictivo para una o más variables dependientes (de destino) basado en los valores de las variables predictoras.

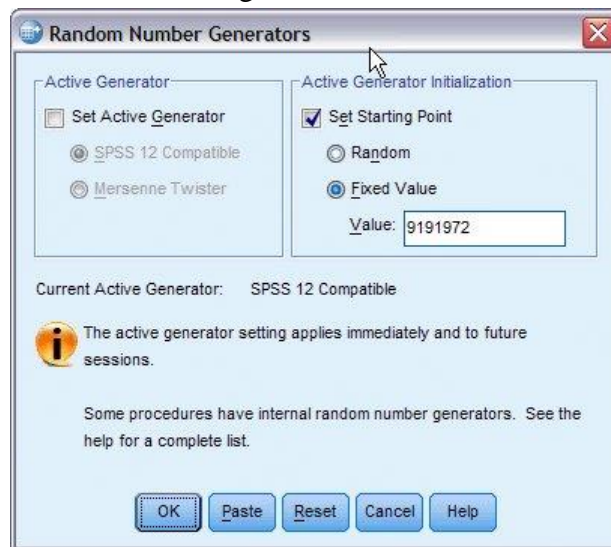
7.4.2.1. *Uso de la función de base radial para la clasificación de clientes de telecomunicaciones*

Un proveedor de telecomunicaciones ha segmentado su base de clientes por patrones de uso de servicio, categorizando los clientes en cuatro grupos. Si los datos demográficos se pueden utilizar para predecir la pertenencia a un grupo, se pueden personalizar las ofertas para cada uno de los posibles clientes. Supongamos que la información sobre los clientes actuales se encuentra en *telco.sav*. Si desea obtener más información, consulte Archivos muestrales en Apéndice de IBM Spss. Uso de la función de base radial para la clasificación de clientes de telecomunicaciones

Establecer la semilla aleatoria, puede replicar el análisis de manera exacta. Para establecer la semilla aleatoria, elija en los menús: Transformar Generadores de números aleatorios (Figura 30)

7.4.2.2. Preparación de los datos para su análisis

Figura 30 Cuadro de diálogo Generadores de números aleatorios



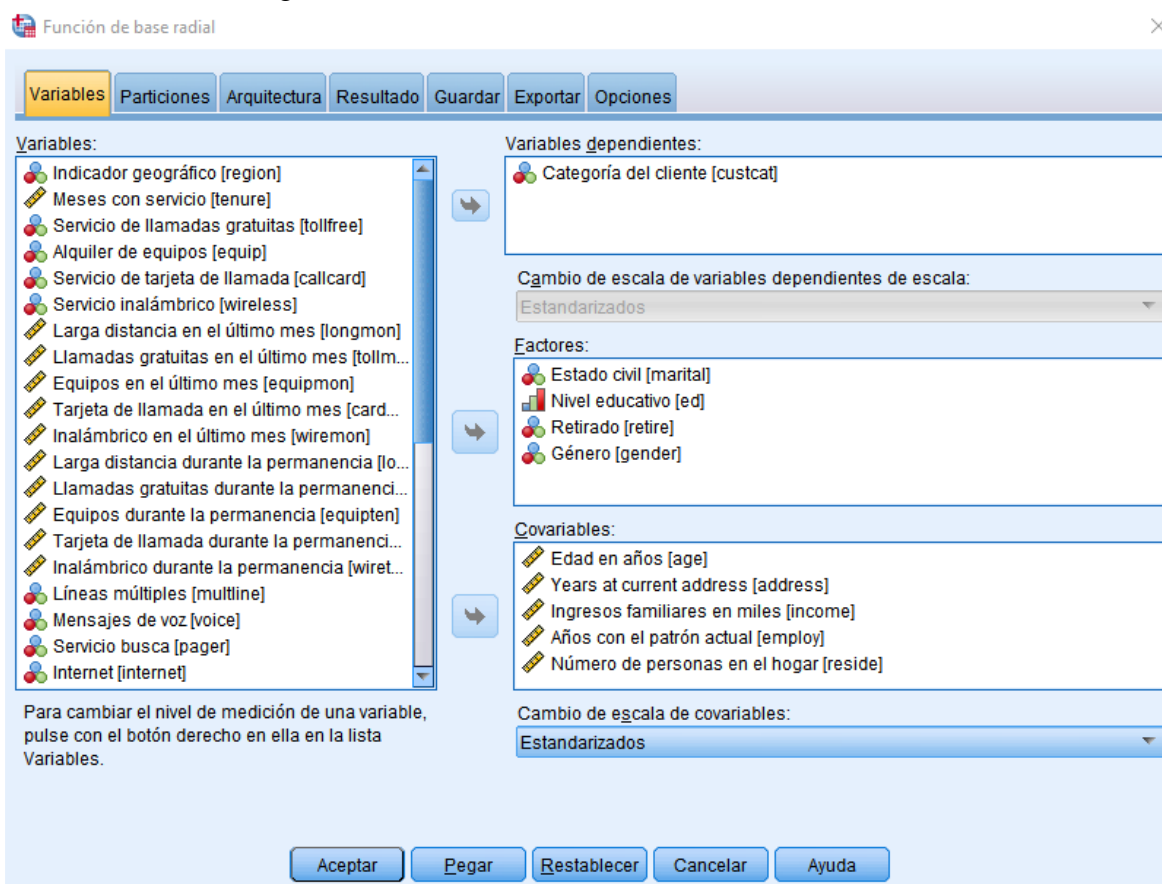
Fuente. Copyright IBM Corp. _2018

1. Seleccione Establecer punto de inicio.
2. Seleccione Valor fijo, y escriba 9191972 como el valor.
3. Click en aceptar

Para ejecutar un análisis de función de base radial, elija en los menús: Analizar Redes neuronales Función de base radial (Figura 31)

- Seleccione Categoría del cliente [custcat] como variable dependiente.
- Seleccione Estado civil [marital], Nivel educativo [ed], Retirado [retire] y Género [gender] como factores.
- Seleccione desde Edad en años [age] hasta Número de personas en el hogar [reside] como covariables.
- Seleccione Corregida normalizada como método para cambiar la escala de las covariables.
- Pulse en la pestaña Particiones y escriba 6 como número relativo de la muestra de entrenamiento. Escriba 2 como número relativo de la muestra de prueba. Escriba 1 como número relativo de la muestra reservada

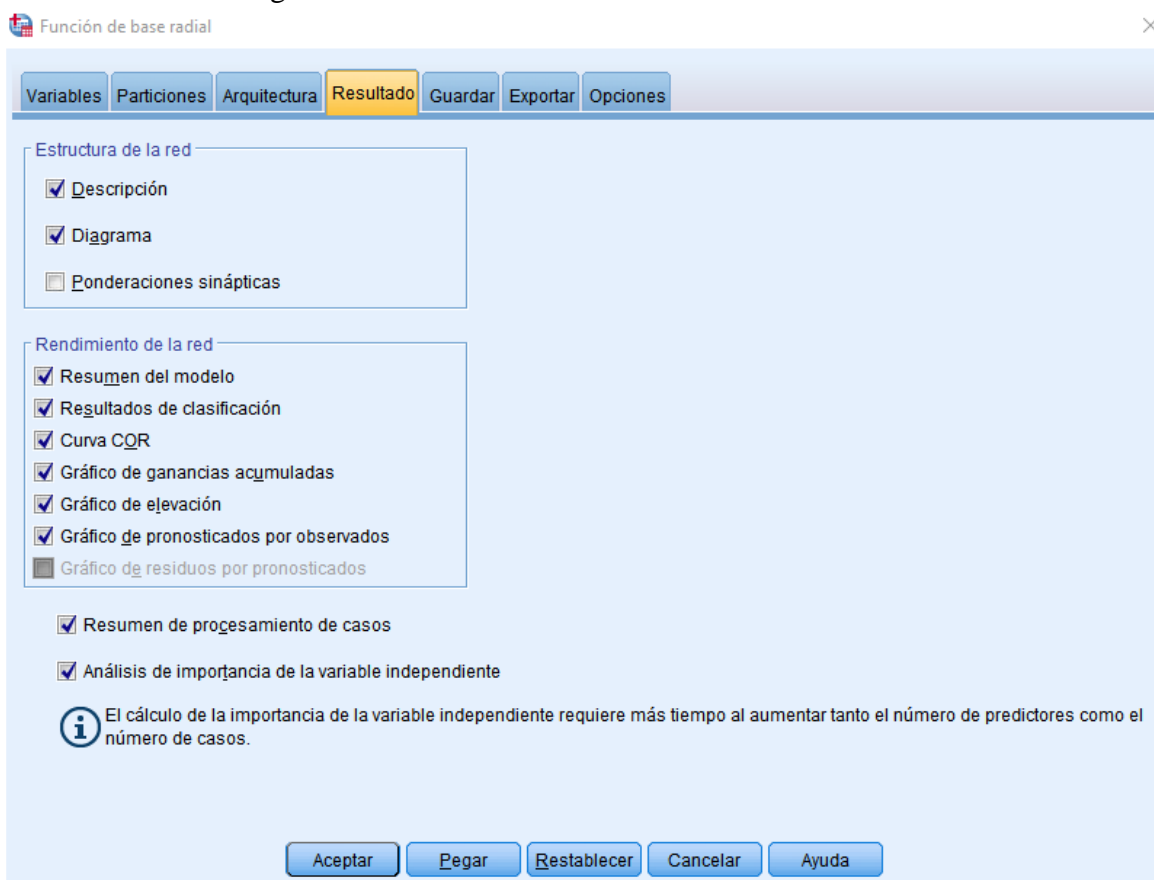
Figura 31 Función de base radial: Pestaña Variables



Fuente. Copyright IBM Corp. _2018

- Anule la selección en Diagrama en el grupo de estructura de red.
- Seleccione Curva COR, Gráfico de ganancias acumuladas, Gráfico de elevación (índice) y Gráfico de pronosticados por observados en el grupo de rendimiento de red.
- Pulse en la pestaña Guardar y dejar por defecto las demás pestañas y aceptar (Figura 32)

Figura 32 Función de base radial: Pestaña Resultados



Fuente. Copyright IBM Corp. _2018

El resumen según la Tabla 11 del procesamiento de los casos muestra que se asignaron 665 casos a la muestra de entrenamiento, 224 a la muestra de prueba y 111 a la muestra reservada. No se excluyó ningún caso del análisis.

Tabla 11. Resumen de procesamiento de casos

		N	Porcentaje
Muestra	Entrenamiento	665	66,5%
	Pruebas	224	22,4%
	Reserva	111	11,1%
Válido		1000	100,0%
Excluido		0	
Total		1000	

Fuente. Elaboración Propia

La Tabla 12 de información de red muestra información sobre la red neuronal y resulta útil para garantizar que las especificaciones son correctas. Aquí se ve en particular que:

- El número de unidades en la capa de entrada es el número de covariables más el número total de niveles de factor.
- Se crea una unidad independiente para cada categoría de Estado civil, Nivel educativo, Retirado y Género y ninguna de las categorías se considera como una unidad “redundante”, como es habitual en muchos procedimientos de modelado. ,,
- De igual manera, se crea una unidad de resultado independiente para cada categoría de Categoría del cliente, para un total de 4 unidades en la capa de salida. ,,
- Se cambia la escala de las covariables utilizando el método normalizado corregido. ,,
- La selección de arquitectura automática ha elegido 9 unidades en la capa oculta. ,, El resto de la información de red se toma por defecto para el procedimiento.

Tabla 12. Información de red

Capa de entrada	Factores	1	Estado civil
		2	Nivel educativo
		3	Retirado
		4	Género
	Covariables	1	Edad en años
		2	Years at current address
		3	Ingresos familiares en miles
		4	Años con el patrón actual
		5	Número de personas en el hogar
	Número de unidades		16
Método de cambio de escala para las covariables		Normalizada	
Capa oculta	Número de unidades		8 ^a
	Función de activación		Softmax
Capa de salida	Variables dependientes	1	Categoría del cliente
	Número de unidades		4
	Función de activación		Identidad

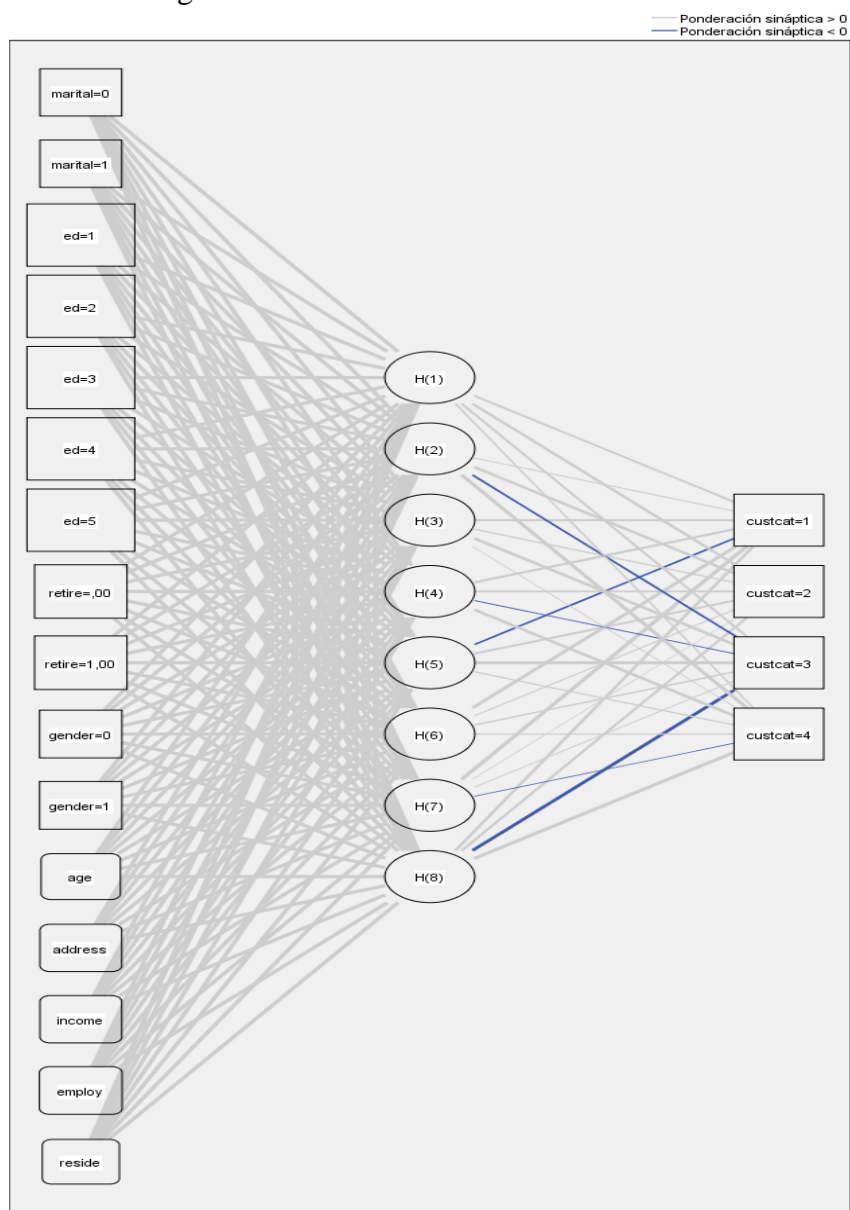
	Función de error	Suma de cuadrados
--	------------------	-------------------

a. Determinado por el criterio de datos de prueba: El "mejor" número de unidades ocultas es el que produce el error menor en los datos de prueba.

Fuente. Elaboración Propia

El modelo de red neuronal artificial es:

Figura 33. modelo de red neuronal artificial



Función de activación de capa oculta: Softmax

Función de activación de capa de salida: Identidad

Fuente. Copyright IBM Corp. _2018

El resumen del modelo de la Tabla 13, muestra información sobre los resultados del entrenamiento, la prueba y la aplicación de la red final a la muestra reservada.

- Se muestra la suma de errores cuadráticos, ya que siempre se utiliza para redes RBF. Ésta es la función de error que la red intenta minimizar durante el entrenamiento y la prueba. ,,
- El porcentaje de pronósticos incorrectos se toma de la tabla de clasificación, y se discutirá más adelante en ese tema.

Tabla 13. Resumen del modelo

Entrenamiento	Error de suma de cuadrados	232,605
	Porcentaje de pronósticos incorrectos	59,5%
	Tiempo de entretamiento	0:00:01,88
Pruebas	Error de suma de cuadrados	80,783 ^a
	Porcentaje de pronósticos incorrectos	67,0%
Reserva	Porcentaje de pronósticos incorrectos	62,2%

Variable dependiente: Categoría del cliente Fuente: Spss v 22

a. El número de unidades ocultas se determina por el criterio de los datos de prueba: El "mejor" número de unidades ocultas es la que produce el menor error en los datos de prueba.

Fuente. Elaboración Propia

La Tabla 14 de clasificación muestra los resultados prácticos de la utilización de la red. Para cada caso, la respuesta pronosticada es la categoría con la mayor pseudoprobabilidad pronosticada por el modelo.

- Las casillas de la diagonal son los pronósticos correctos.
- Las casillas fuera de la diagonal son los pronósticos incorrectos

Dados los datos observados, el modelo “nulo” (es decir, el que no incluye ningún predictor) clasificaría a todos los clientes en el grupo modal, Servicio plus. Por tanto, el modelo nulo sería correcto $281/1000 = 28,1\%$ de las veces. La red RBF consigue un 10,1% más, es decir, un 40.5% de los clientes. En concreto, el modelo es particularmente bueno para identificar los clientes de Servicio Plus y Servicio total. Sin embargo, funciona excepcionalmente mal para clasificar los

clientes de Servicio electrónico. Tal vez necesite encontrar otro predictor para separar estos clientes. Otra posibilidad, dado que estos clientes se suelen clasificar erróneamente como clientes Servicio plus y Servicio total, es que la compañía intentase sencillamente vender un producto superior a los posibles clientes que normalmente figurarían en la categoría E-servicio. Las clasificaciones basadas en los casos utilizados para crear el modelo tienden a ser demasiado “optimistas” en el sentido de que aumentan en exceso la tasa de clasificación. La muestra reservada ayuda a validar el modelo. En este caso, el modelo clasificó correctamente el 37.8% de casos. Aunque la muestra reservada es bastante pequeña, eso sugiere que el modelo es correcto aproximadamente dos de cada cinco veces.

Tabla 14. Clasificación

Muestra	Observado	Pronosticado				Porcentaje correcto
		Servicio básico	E-Servicio	Servicio Plus	Servicio Total	
Entrenamiento	Servicio básico	55	12	64	44	31,4%
	E-Servicio	22	17	40	62	12,1%
	Servicio Plus	36	13	105	31	56,8%
	Servicio Total	26	14	32	92	56,1%
	Porcentaje global	20,9%	8,4%	36,2%	34,4%	40,5%
Pruebas	Servicio básico	16	5	21	17	27,1%
	E-Servicio	11	2	13	27	3,8%
	Servicio Plus	13	6	32	14	49,2%
	Servicio Total	5	4	14	24	51,1%
	Porcentaje global	20,1%	7,6%	35,7%	36,6%	33,0%
Reserva	Servicio básico	6	1	15	10	18,8%
	E-Servicio	5	4	5	9	17,4%
	Servicio Plus	8	1	18	4	58,1%
	Servicio Total	6	1	4	14	56,0%
	Porcentaje global	22,5%	6,3%	37,8%	33,3%	37,8%

Variable dependiente: Categoría del cliente, Fuente: Elaboración propia usando Spss v22

Fuente. Elaboración Propia.

8. CONCLUSIONES

- Que el comité curricular del programa apruebe la incorporación de estos ajustes en el plan curricular. Así mismo, se espera que en el futuro se consolide una línea de investigación en redes neuronales hacia el desarrollo de la metodología de modelación y simulación, básica para las ciencias de la complejidad.
- Se evidencio que los microdiseños del plan curricular, Registro Calificado 02853 del 16 de febrero de 2016 no estaban presentes los temas de las redes neuronales, se sugiere y diseña la propuesta del microdiseño RNA para que se incluya en el plan de acción del programa de matemática aplicada.
- La comparación teórica entre RNA y modelos estadísticos es que no se trata de metodologías contrapuestas (White, 1989). Se ha puesto de manifiesto que existe un solapamiento entre ambos campos. Las RNA incluyen diversos modelos, como el MLP, que son de gran utilidad en las aplicaciones de análisis de datos y estadística. La metodología estadística es directamente aplicable a las RNA de diversas formas, incluyendo criterios de estimación y algoritmos de optimización.
- Las RNA no se ven afectadas por el cumplimiento del supuesto de proporcionalidad, tampoco es necesario introducir de forma explícita términos de interacción entre predictores ni funciones concretas entre predictores y variable de respuesta, debido a que son aprendidos de forma automática en el proceso de entrenamiento del modelo.
- Se evidencia la relación estrecha entre la terminología estadística y las redes neuronales usados para la predicción, se anexa un tutorial sobre RNA en Spss tomado de IMB Spss v22, ajustado y con ejemplos propuestos a manera de estrategia metodológica.
- La herramienta Red neuronal de IMB Spss Statistics, es igualmente versátil, ya que permite variar parámetros como el número de patrones para el entrenamiento, las funciones de activación, el número de neuronas en la capa oculta, diferentes variaciones del algoritmo de Back Propagation para el entrenamiento, o la construcción de un algoritmo propio, así como aplicar los conceptos ya revisados para la implementación de otro tipo de redes neuronales, en diferentes clases de usos.
- Los algoritmos matemáticos en los que han sido desarrollados, para todos los tipos de redes son modelos sencillos, que, aunque exigen cierto grado de conocimientos de cálculo diferencial, regresión y probabilidades, pueden desarrollados en cualquier lenguaje de programación.

9. RECOMENDACIONES

No solo hay que fortalecer las líneas de investigación de la facultad de ciencias exactas y naturales, frente a los diferentes grupos de investigación de la universidad Surcolombiana, si no extender a las demás Instituciones de educación superior del Huila, es necesario aunar esfuerzos para generar nuevas redes de cooperación entre las instituciones en temas como RNA, para generar nuevo conocimiento y proyectos de investigación.

10. BIBLIOGRAFÍA

- Análisis entre neuronas biológicas y artificiales. (2012). Retrieved from <http://redesneuronales-uba.blogspot.com/>
- Badilla Saxe, E. (2009). Diseño curricular: De la integración a la complejidad. *Revista Electrónica" Actualidades Investigativas En Educación"*, 9(2)
- Batanero, C. (2001). Didáctica de la estadística. *Granada: Universidad De Granada*,
- Biganzoli, E., Boracchi, P., Mariani, L., & Marubini, E. (1998). Feed forward neural networks for the analysis of censored survival data: A partial logistic regression approach. *Statistics in Medicine*, 17(10), 1169-1186.
- Bishop, C., & Bishop, C. M. (1995). *Neural networks for pattern recognition* Oxford university press.
- Briggs, J., & Peat, F. (1994). Espejo y reflejo: Del orden al caos. *Barcelona: Gedisa*,
- Buitrago, M. I. A., & Muñoz, C. Z. (2000). Tutorial sobre redes neuronales aplicadas en ingeniería eléctrica y su implementación en un sitio web. *Tutorial Sobre Redes Neuronales Aplicadas En Ingeniería Eléctrica Y Su Implementación En Un Sitio Web*,
- Castellano, M. (2009). Modelización estadística con redes neuronales. aplicados a la hidrología, aerobiología y modelización de procesos. *Modelizacion Estadistica Con Redes Neuronales*,
- Castiblanco Paiba, A. C., Urquina Llanos, H., Camargo Uribe, L., & Moreno Armella, L. E. (2004). Tecnología informática: Innovación en el currículo de matemáticas de la educación básica secundaria y media.
- Castiblanco, A. C. (2002). El proyecto de “incorporación de nuevas tecnologías al currículo de matemáticas de la educación media” y sus avances.

- Delgado, A. (1998). Propiedades matemáticas y aplicaciones de las redes neuronales dinámicas recurrentes. *Universidad Nacional De Colombia. Santafé De Bogotá, Colombia,*
- Feller, W. (2008). *An introduction to probability theory and its applications* John Wiley & Sons.
- Freeman, J. A., & Skapura, D. M. (1993). Redes neuronales. *Algoritmos, Aplicaciones Y,*
- Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics, 6(2), 187-205.*
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2006). *Metodología de la investigación* México: McGraw-Hill.
- James, G. (1988). Chaos: Making a new science. *Viking Penguin Inc,*
- Maldonado, C. (2009). La complejidad es un problema, no una cosmovisión. *UCM Revista De Investigación, 13, 42-54.*
- Maldonado, C. E. (2003). El problema de la filosofía del conocimiento y el estudio de los sistemas complejos. *Práxis Filosófica, (17), 6.*
- Maldonado, C. E. (2009). Complejidad de los sistemas sociales: Un reto para las ciencias sociales. *Cinta De Moebio, (36), 146-157.*
- Maldonado, C. E. (2014). ¿ Qué es un sistema complejo? *Revista Colombiana De Filosofía De La Ciencia, 14(29), 71-93.*
- Maldonado, C. E., & Gómez-Cruz, N. A. (2010). Modelamiento y simulación de sistemas complejos. *Borradores De Investigación: Serie Documentos Administración, ISSN 0124-8219, no.66 (Febrero De 2010),*
- Maldonado, C. E., & Gómez-Cruz, N. A. (2011). El mundo de las ciencias de la complejidad. *Borradores De Investigación: Serie Documentos Administración, ISSN 0124-8219, no.76 (Septiembre De 2010),*

- Mendez, M. C. (2009). Modelización estadística con redes neuronales. aplicaciones a la hidrología, aerobiología y modelización de procesos. *Modelización Estadística Con Redes Neuronales. Aplicaciones a La Hidrología, Aerobiología Y Modelización De Procesos*,
- Montaño, J. J. (2002). Redes neuronales artificiales aplicadas al análisis de datos [artificial neural networks applied to data analysis]. *Unpublished Doctoral Dissertation, Balearic Islands University, Spain*,
- Montealegre Cárdenas, M., Londoño Betancourth, G., & Polanía Quiza, L. A. (2004). *Enfoque sistémico de las matemáticas escolares: Educación básica y media* (v. 1 ed.). Colombia: Guadalupe Ltda.
- Montealegre, M., Londoño, G., & Polanía, L. (2002). Fundamentos de los sistemas dinámicos. *Universidad Surcolombiana, 1*
- Moreno, J. J. M. (2004). *Redes neuronales artificiales aplicadas al análisis de datos*
- Morin, E. (2008). Complejidad restringida, complejidad general biblioteca virtual participativa de la complejidad, *Complejidad Restringida, Complejidad General Biblioteca Virtual Participativa De La Complejidad, 2004a [Cited 19-11 2008]*,
- Morin, E., & Pakman, M. (1994). *Introducción al pensamiento complejo* Gedisa Barcelona.
- Nicholis, G., & Prigogine, I. (1994). La estructura de lo complejo. en el camino hacia una nueva comprensión de las ciencias. *Editorial Alianza SA Madrid, España*,
- O'CONNOR, J. (1998). McDermott, Ian. *Introducción Al Pensamiento Sistémico. Barcelona: Urano*,

Perez, C. R. (2003). Paradigma de la complejidad, modelos científicos y conocimiento educativo.

Ágora Digital, (6)

Polya, G., & Zugazagoitia, J. (1965). Cómo plantear y resolver problemas. *Cómo Plantear Y*

Resolver Problemas,

Rodríguez Zoya, L. G., & Leónidas Aguirre, J. (2011). Teorías de la complejidad y ciencias sociales. nuevas estrategias epistemológicas y metodológicas. *Nómadas*, (30)

Sarle, W. S. (1994). Neural networks and statistical models.

Sarle, W. S., & Fl, R. (1994). Neural network implementation in SAS® software.

Sarle, W.S. (Ed.) (2002). *Neural network FAQ*. Recuperado 05/07/18, desde <ftp://ftp.sas.com/pub/neural/FAQ.html>.

Tukey, J. W. (1977). *Exploratory data analysis* Reading, Mass.

Vicino, F. (1998). Some reflections on artificial neural networks and statistics: Two ways of obtaining solutions by working with data. *Substance use & Misuse*, 33(2), 221-231.

11. ANEXOS

Anexo A: Microdiseño curricular fundamentos de las redes neuronales artificiales

FORMATO OFICIAL DE MICRODISEÑO CURRICULAR

FACULTAD: Ciencias Exactas y Naturales

PROGRAMA: Matemática Aplicada

1. IDENTIFICACIÓN DEL CURSO

NOMBRE DEL CURSO: Componente complementario flexible: fundamentos de las redes neuronales artificiales

CÓDIGO: BEEEXMA04 **No. DE CRÉDITOS ACADÉMICOS:** 4 **HORAS SEMANALES:** 4

REQUISITOS: Teoría de Probabilidades

ÁREA DEL CONOCIMIENTO: Matemática aplicada

UNIDAD ACADÉMICA RESPONSABLE DEL DISEÑO CURRICULAR:

Comité de Currículo Departamento de Matemática y estadística.

COMPONENTE BÁSICO **COMPONENTE FLEXIBLE**

TIEMPO (en horas) DEL TRABAJO ACADÉMICO DEL ESTUDIANTE

Actividad Académica Del Estudiante	Trabajo Presencial	Trabajo Independiente	Total (Horas)
Horas	64	128	192
TOTAL	64	128	192

2. PRESENTACION RESUMEN DEL CURSO

Se estudian los conceptos básicos sobre redes neuronales como modelos matemáticos para procesos que impliquen aprendizaje: naturaleza y evolución histórica, clasificación y aplicaciones como herramientas estadísticas en tareas de predicción y clasificación.

3. JUSTIFICACIÓN.

Las redes neuronales artificiales constituyen un campo de investigación donde interactúan conocimientos provenientes de diversas áreas. Esta naturaleza interdisciplinaria las hace un campo de aplicación para el matemático contemporáneo, dadas sus características, se constituyen en herramientas para la modelación y simulación de sistemas dinámicos complejos en el abordaje de problemas reales de complejidad creciente dándoles también un carácter transdisciplinar.

4. COMPETENCIAS GENERALES

COMPETENCIAS GENERALES		
SABER	INTERPRETATIVA	El alumno debe demostrar dominio de los conceptos básicos sobre las redes neuronales como herramientas para la modelación y simulación de sistemas complejos
	ARGUMENTATIVA	El alumno debe buscar y dar a conocer el porqué de un concepto, el porqué de una definición, el porqué de una propiedad, el porqué de un proceso y en general, el porqué de determinada situación problema.
	PROPOSITIVA	El alumno debe construir el por qué y el cómo de un resultado (teorema, algoritmo)
HACER	El alumno debe tener la capacidad de formular el problema, y de modelarlo, encontrarle soluciones, de simularlo y de buscar, encontrar otras alternativas de solución en otros contextos.	
SER	Se Quiere ante todo, un egresado de Matemáticas con una amplia formación en ética y valores, en lo social, en lo epistemológico, en lo estético y en lo ontológico	

5. DEFINICION DE UNIDADES TEMATICAS Y ASIGNACIÓN DE TIEMPO DE TRABAJO PRESENCIAL E INDEPENDIENTE DEL ESTUDIANTE POR CADA EJE TEMATICO

No.	NOMBRE DE LAS UNIDADES TEMÁTICAS	DEDICACIÓN DEL ESTUDIANTE (horas)		HORAS TOTALES (a + b)
		a) Trabajo Presencial	b) Trabajo Independiente	
1	Redes neuronales artificiales (RNA)	20	40	60
2	Predicción con RNA	20	40	60
3	Clasificación con RNA	24	48	72
TOTAL		64	128	192

6. PROGRAMACION SEMANAL DEL CURSO

Unidad Temática	No. Semanas	CONTENIDOS TEMÁTICOS	ACTIVIDADES Y ESTRATEGIAS PEDAGOGICAS	H. T. P.		H.T.I.	
				Clases	Laboratorio y/o practica	Trabajo dirigido	Trabajo independiente
1	1	Modelo biológico de la neurona Naturaleza y evolución histórica de las redes neuronales artificiales	Exposición del profesor. Modelación con ecuaciones diferenciales				
	2	Terminología y notación para las redes neuronales artificiales Clasificación: por las conexiones entre las capas	Exposición del profesor, utilización de software.				
2	3	Clasificación por la topología: Perceptrones	Exposición del profesor, análisis de algoritmos				
	4	Funciones de base radial	Exposición del profesor, solución de talleres				
	5	Clasificación por tipo de aprendizaje:	Exposición del profesor, solución de talleres y aplicaciones .				

		Redes on-line y off-line					
	6	Redes con aprendizaje supervisado	Uso de software y modelos de aplicación.				
3	7	Redes con aprendizaje no supervisado	Exposición del profesor, solución de talleres.				
	8	Modelación y simulación con redes neuronales	Exposición del profesor, Estudio de casos				
	9	Estadística y redes neuronales Predicción con redes neuronales Regresión lineal	Exposición del profesor, Aplicaciones mediante software estadístico y de redes neuronales. Comparación de resultados.				
4	10	Regresión polinómica	Exposición del profesor, Comparación de metodologías.				
	11	Regresión logística	Exposición del profesor, Comparación de metodologías.				
	12	Clasificación con redes neuronales Métodos clásicos	Exposición del profesor, Comparación de metodologías.				
	13	Métodos no paramétricos	Exposición del profesor, Comparación de metodologías.				

5	14	Otros modelos de análisis de datos	Exposición del profesor, Comparación de metodologías.				
	15	Aproximadores universales	Exposición del profesor, Comparación de metodologías.				
	16	Redes probabilísticas	Exposición del profesor, Comparación de metodologías.				

H. T. P. = Horas De trabajo presencial

H. T. I. = Horas de trabajo independiente

7. EVALUACIÓN DEL APRENDIZAJE

UNIDAD TEMÁTICA	ESTRATEGIA DE EVALUACION	PORCENTAJE (%)
Redes neuronales artificiales	Prueba escrita + informes sobre resultados relevantes	30%
Predicción con RNA	Prueba escrita + solución de talleres en clase	30%
Clasificación con RNA	Prueba escrita + aplicación de modelos tipo estudiados en clase y solución de ejercicios	40%

8. BIBLIOGRAFÍA

8.1. Bibliografía Básica:

8.2. Bibliografía Complementaria:

REVISADO POR **Diego Mauricio Echeverri Suaza**

FECHA DE DILIGENCIAMIENTO: **Junio 2018**

Anexo B Tutorial de Introducción a las Redes Neuronales con IBM Spss Statistics

1. Tutorial de Introducción a las Redes Neuronales con IBM Spss Statistics¹³

Las redes neuronales son la herramienta preferida para muchas aplicaciones de minería de datos predictiva por su potencia, flexibilidad y facilidad de uso. Las redes neuronales predictivas son especialmente útiles en las aplicaciones cuyo proceso subyacente sea complejo; por ejemplo:

- Prever la demanda de los consumidores para racionalizar los costes de producción y entrega.
- Predecir la probabilidad de respuesta al marketing mediante correo directo para determinar a qué unidades familiares de una lista de correo debe enviarse una oferta.
- Puntuar a un solicitante para determinar el riesgo que supone concederle un crédito.
- Detectar transacciones fraudulentas en una base de datos de reclamaciones de seguros.

Las redes neuronales utilizadas en las aplicaciones predictivas, como las redes de **perceptrones multicapa (MLP)** y las de **función de base radial (RBF)**, se supervisan en el sentido de que los resultados pronosticados por el modelo se pueden comparar con los valores conocidos de las variables de destino. La opción de redes neuronales le permite ajustar las redes MLP y RBF y guardar los modelos resultantes para la puntuación.

1.1 ¿Qué es una red neuronal?

El término **red neuronal** se aplica a una familia de modelos relacionada de manera aproximada que se caracteriza por un gran espacio de parámetro y una estructura flexible y que proviene de los estudios sobre el funcionamiento del cerebro. Conforme fue creciendo la familia, se diseñó la mayoría de los nuevos modelos para aplicaciones no biológicas, aunque gran parte de la terminología asociada refleja su origen.

Las definiciones específicas de redes neuronales son tan variadas como los campos en que se utilizan. Aunque no hay una sola definición que abarque a toda la familia de modelos, tenga en cuenta la siguiente descripción por ahora¹⁴:

Una red neuronal es un procesador distribuido en paralelo de forma masiva con una propensión natural a almacenar conocimiento experimental y convertirlo en disponible para su uso. A semeja al cerebro en dos aspectos:

¹³ Copyright IBM Corp. _2018. Reservados todos los derechos.

¹⁴ Haykin, S. 198. Neural Networks: A Comprehensive Foundation, 2nd ed. New York: Macmillan Colege Publishing.

- El conocimiento se adquiere por la red mediante un proceso de aprendizaje.
- Las fuerzas de conexión interneuronal, conocidas como ponderaciones sinápticas, se utilizan para almacenar el conocimiento.

Si desea ver información detallada acerca de por qué esta definición puede ser demasiado restrictiva, consulte¹⁵.

Para poder diferenciar las redes neuronales de los métodos estadísticos tradicionales que usan esta definición, lo que *no* se dice es tan significativo como el texto real de la definición. Por ejemplo, el modelo de regresión lineal tradicional puede adquirir conocimiento a través del método de mínimos cuadrados y almacenar ese conocimiento en los coeficientes de regresión. En este sentido, es una red neuronal. De hecho, se puede argumentar que la regresión lineal es un caso especial de determinadas redes neuronales. Sin embargo, la regresión lineal tiene una estructura de modelo rígida y un conjunto de supuestos que se imponen antes de aprender a partir de los datos.

Por lo contrario, las definiciones anteriores plantean exigencias mínimas sobre la estructura y los supuestos del modelo. Por tanto, una red neuronal puede aproximar una amplia gama de modelos estadísticos sin que tenga que hipotetizar de antemano determinadas relaciones entre las variables dependientes e independientes. En lugar de eso, la forma de las relaciones se determina durante el proceso de aprendizaje. Si una relación lineal entre las variables dependientes e independientes es adecuada, los resultados de la red neuronal deben aproximarse lo máximo posible a los del modelo de regresión lineal. Si una relación no lineal es más adecuada, la red neuronal se aproximará automáticamente a la estructura del modelo "correcta".

El inconveniente que supone esta flexibilidad radica en que las ponderaciones sinápticas de una red neuronal no pueden interpretarse fácilmente. Por tanto, si intenta explicar un proceso subyacente que produce las relaciones entre las variables dependientes e independientes, sería mejor usar un modelo estadístico más tradicional. Sin embargo, si la interpretación del modelo no es importante, puede obtener rápidamente modelos con buenos resultados mediante

1.2 Estructura de red neuronal en IBM Spss Statistics

Aunque las redes neuronales plantean exigencias mínimas sobre los supuestos y la estructura del modelo, resulta útil comprender la **arquitectura** general de la red. La red de perceptrones multicapa (MLP) o de función de base radial (RBF) es una función de predictores (denominados también entradas o variables independientes) que minimiza el error de predicción de las variables de destino (también denominadas salidas).

Tomemos como ejemplo el conjunto de datos *bankloan.sav* incluido con el producto, en el que desea poder identificar a los posibles morosos entre un grupo de solicitantes de préstamos. Una

¹⁵ Ripley, B. D. 196. Pattern Recognition and Neural Networks. Cambridge: Cambridge University Press.

red MLP o RBF aplicada a este problema es una función de las mediciones que minimiza el error al pronosticar la posibilidad de mora. La figura siguiente resulta muy útil para ilustrar la forma de esta función.

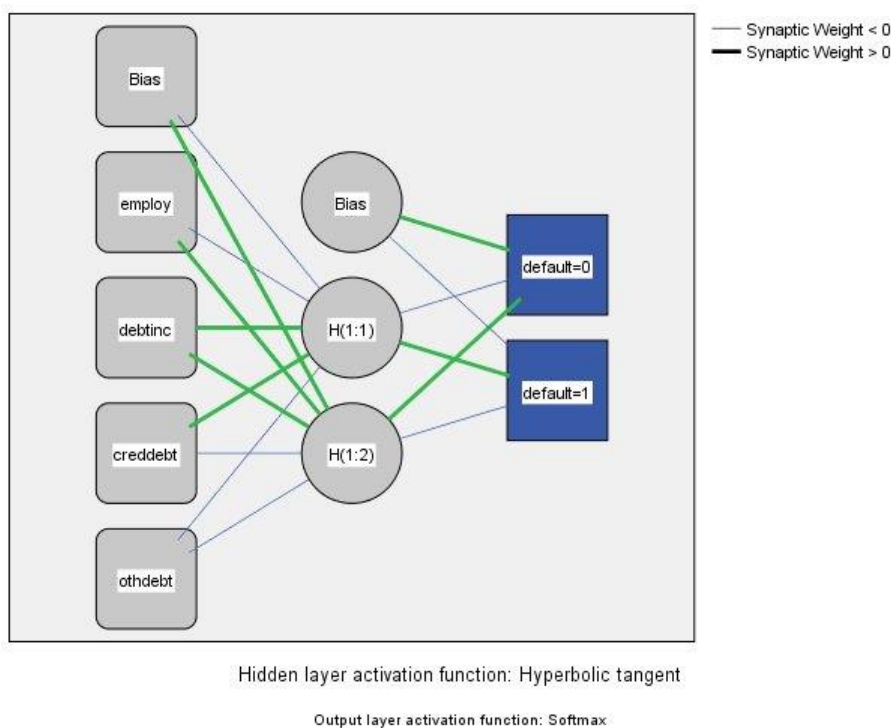


Figura 34 Arquitectura Feedforward con una capa oculta

Fuente. Copyright IBM Corp. _2018

Esta estructura se denomina **arquitectura feedforward** porque las conexiones de la red fluyen unidimensionalmente desde la capa de entrada hasta la capa de salida sin ciclos de retroalimentación. En esta figura:

- La **capa de entrada** contiene los predictores.
- La **capa oculta** contiene nodos (o unidades) no observables. El valor de cada unidad oculta es una función de los predictores; la forma exacta de la función depende, por un lado, del tipo de red y, por otro lado, de especificaciones controlables por el usuario.
- La **capa de salida** contiene las respuestas. Como el historial de moras es una variable categórica con dos categorías, se recodifica como dos variables de indicador. Cada unidad de salida es una función de las entradas ocultas. Nuevamente, la forma exacta de la función depende, por un lado, del tipo de red y, por otro lado, de especificaciones controlables por el usuario.

La red MLP permite una segunda capa oculta; en ese caso, cada unidad de la segunda capa oculta es una función de las unidades de la primera capa oculta, y cada respuesta es una función de las unidades de la segunda capa oculta.

1.3 Perceptrón multicapa

El procedimiento Perceptrón multicapa (MLP) genera un modelo predictivo para una o más variables dependientes (de destino) basada en los valores de las variables predictoras.

Ejemplos. A continuación, se describen dos situaciones en las que se utiliza el procedimiento MLP:

El encargado de préstamos de un banco desea poder identificar características que sean indicativas de personas que puedan causar mora en los créditos y utilizar dichas características para identificar riesgos de crédito positivos y negativos. Utilizando una muestra de clientes anteriores, puede entrenar un perceptrón multicapa, validar el análisis utilizando una muestra reservada de clientes anteriores y posteriormente utilizar la red para clasificar a los posibles clientes según los riesgos de crédito que presenten.

Un sistema hospitalario está interesado en realizar un seguimiento de los costes y las duraciones de las estancias de los pacientes ingresados para tratamientos de infartos de miocardio (IM o "ataque al corazón"). La obtención de estimaciones precisas de estas medidas permite a la administración gestionar adecuadamente las camas disponibles a medida que se trata a los pacientes. Utilizando los registros de tratamiento de una muestra de pacientes que recibieron tratamiento a consecuencia de un infarto de miocardio, el administrador puede entrenar una red que pronostique tanto el coste como la duración de la estancia.

Consideraciones de los datos

1.4 Variables dependientes.

Según la Figura 2. Las variables dependientes pueden ser:

- *Nominal.* Una variable puede ser tratada como nominal cuando sus valores representan categorías que no obedecen a una clasificación intrínseca. Por ejemplo, el departamento de la compañía en el que trabaja un empleado. Algunos ejemplos de variables nominales son: región, código postal o confesión religiosa.
- *Ordinal.* Una variable puede ser tratada como ordinal cuando sus valores representan categorías con alguna clasificación intrínseca. Por ejemplo, los niveles de satisfacción con un servicio, que abarquen desde muy insatisfecho hasta muy satisfecho. Entre los ejemplos de variables ordinales se incluyen escalas de actitud que representan el grado de satisfacción o confianza y las puntuaciones de evaluación de las preferencias.

- *Escalas.* Una variable puede tratarse como escala (continua) cuando sus valores representan categorías ordenadas con una métrica con significado, por lo que son adecuadas las comparaciones de distancia entre valores. Son ejemplos de variables de escala: la edad en años y los ingresos en dólares.

El procedimiento supone que se ha asignado el nivel de medición adecuado a todas las variables dependientes. No obstante, puede cambiar temporalmente el nivel de medición para una variable pulsando con el botón derecho en la variable en la lista de variables de origen y seleccionar un nivel de medición en el menú emergente. Para modificar permanentemente el nivel de medición de una variable, consulte Nivel de medición de variable.

Un icono situado junto a cada variable de la lista de variables identifica el nivel de medición y el tipo de datos.

Tabla 15 Iconos de nivel de medición

	Numérico	Cadena	Fecha	Hora
Escala (Continuo)		n/a		
Ordinal				
Nominal				

Fuente. Copyright IBM Corp. _2018

1.5 Variables predictoras.

Los predictores se pueden especificar como factores (categóricos) o covariables (de escala).

1.6 Codificación de la variable categórica.

El procedimiento recodifica temporalmente predictores categóricos y variables dependientes utilizando la codificación "una de c " para todo el procedimiento. Si hay c categorías de una variable, la variable se almacena como vectores c , con la primera categoría denotada $(1,0,\dots,0)$, la siguiente categoría $(0,1,0,\dots,0)$, ..., y la última categoría $(0,0,\dots,0,1)$.

Este esquema de codificación incrementa el número de ponderaciones sinápticas y puede dar como resultado un entrenamiento más lento. No obstante, los métodos de codificación más "compactos" normalmente conducen a redes neuronales con un peor ajuste. Si el entrenamiento de la red avanza muy lentamente, pruebe a reducir el número de categorías en los predictores categóricos combinando categorías similares o eliminando los casos que tengan categorías extremadamente raras. Si desea obtener más información sobre la recodificación de variables, consulte Recodificar en las mismas variables o Recodificar en distintas variables.

Toda codificación "Una de c " se basa en los datos de entrenamiento, incluso si se define una muestra reservada o de comprobación (consulte Particiones (Perceptrón multicapa)). De este modo, si las muestras reservadas o de comprobación contienen casos con categorías de predictores que no están presentes en los datos de entrenamiento, esos casos no son utilizados por el procedimiento o en la puntuación. De este modo, si las muestras reservadas o de comprobación contienen casos con categorías de variables dependientes que no están presentes en los datos de entrenamiento, esos casos no son utilizados por el procedimiento o en la puntuación.

1.7 Cambio de escala.

De forma predeterminada, se cambia la escala de las covariables y las variables dependientes de escala para mejorar el entrenamiento de la red. Todo cambio de escala se realiza en base a los datos de entrenamiento, incluso si se define una muestra reservada o de comprobación (consulte Particiones (Perceptrón multicapa)). Es decir, dependiendo del tipo de cambio de escala, la media, la desviación estándar, el valor mínimo o el valor máximo de una covariable o variable dependiente se calculan utilizando sólo los datos de entrenamiento. Si especifica una variable para definir particiones, es importante que estas covariables o variables dependientes tengan distribuciones similares en todas las muestras reservadas, de entrenamiento o comprobación. Utilice por ejemplo, Explorar para examinar las distribuciones en las particiones.

Ponderaciones de frecuencia. Este procedimiento ignora las ponderaciones de frecuencia.

Replicación de los resultados. Si desea replicar exactamente los resultados, utilice el mismo valor de inicialización para el generador de números aleatorios, el mismo orden de datos y el mismo orden de variables, además de utilizar la misma configuración del procedimiento. A continuación, se incluyen más detalles sobre este tema:

- **Generación de números aleatorios.** El procedimiento utiliza la generación de números aleatorios durante la asignación aleatoria de particiones, el submuestreo aleatorio para la inicialización de las ponderaciones sinápticas, el submuestreo aleatorio para la selección automática de arquitectura y el algoritmo de recocido simulado que se utiliza en la inicialización de las ponderaciones y en la selección automática de arquitectura. Para reproducir los mismos resultados aleatorios en el futuro, utilice el mismo valor de

inicialización para el generador de números aleatorios antes de cada ejecución del procedimiento Perceptrón multicapa. Para obtener más información, consulte el tema Generadores de números aleatorios.

- **Orden de casos.** Los métodos de entrenamiento en línea y por mini lotes (Entrenamiento (Perceptrón multicapa)) dependen explícitamente del orden de casos; sin embargo, incluso el entrenamiento por lotes depende del orden de casos ya que la inicialización de las ponderaciones sinápticas implica el submuestreo del conjunto de datos.

Para minimizar los efectos del orden, ordene los casos aleatoriamente. Para comprobar la estabilidad de una solución dada, puede obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios. En situaciones con tamaños de archivo extremadamente grandes, se pueden llevar a cabo varias ejecuciones con una muestra de casos ordenados con distintos órdenes aleatorios.

- **Orden de las variables.** Los resultados pueden verse afectados por el orden de las variables en las listas de factores y covariables, debido al diferente patrón de los valores iniciales asignados al cambiar el orden de las variables. Como con los efectos del orden de casos, puede probar diferentes órdenes de variables (sencillamente arrastrándolas y soltándolas en las listas de factores y covariables) para evaluar la estabilidad de una solución determinada.

Creación de una red de perceptrones multicapa

Esta característica requiere la opción Redes neuronales.

En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

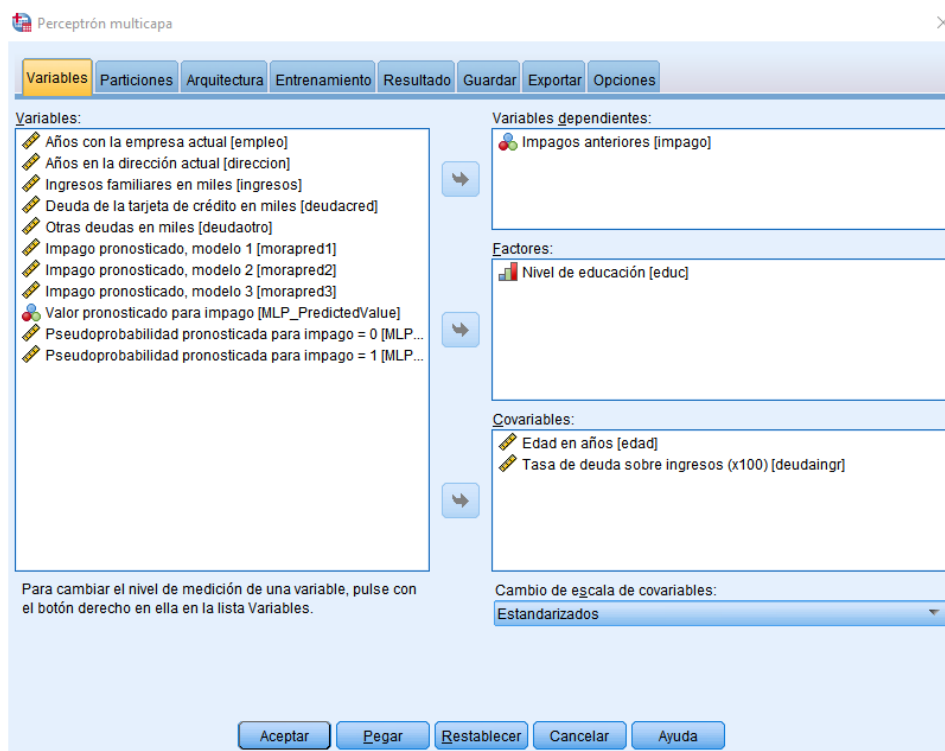


Figura 35 Perceptrón Multicapa En Spss ruta: Analizar/redes Neuronales

Fuente. Copyright IBM Corp. _2018

1. Seleccione al menos una variable dependiente.
2. Seleccione al menos un factor o covariable.

Si lo desea, en la pestaña Variables puede cambiar el método para cambiar la escala de las covariables. Las diferentes opciones son:

- **Tipificados.** Se resta la media y se divide por la desviación estándar, $(x - \text{media})/s$.
- **Normalizados.** Se resta el mínimo y se divide por el rango, $(x - \text{mín})/(\text{máx} - \text{mín})$. Los valores normalizados quedan comprendidos entre 0 y 1.
- **Normalizados corregidos.** Versión ajustada de la resta del mínimo y división por el rango, $[2 * (x - \text{mín})/(\text{máx} - \text{mín})] - 1$. Los valores normalizados ajustados quedan comprendidos entre -1 y 1.
- **Ninguno.** Sin cambio de escala de covariables.

Campos con un nivel de medición desconocido: La alerta de nivel de medición se muestra si el nivel de medición de una o más variables (campos) del conjunto de datos es desconocido. Como el nivel de medición afecta al cálculo de los resultados de este procedimiento, todas las variables deben tener un nivel de medición definido.

Explorar datos. Lee los datos del conjunto de datos activo y asigna el nivel de medición predefinido en cualquier campo con un nivel de medición desconocido. Si el conjunto de datos es grande, puede llevar algún tiempo.

Asignar manualmente. Abre un cuadro de diálogo que contiene todos los campos con un nivel de medición desconocido. Puede utilizar este cuadro de diálogo para asignar el nivel de medición a esos campos. También puede asignar un nivel de medición en la Vista de variables del Editor de datos.

Como el nivel de medición es importante para este procedimiento, no puede acceder al cuadro de diálogo para ejecutar este procedimiento hasta que se hayan definido todos los campos en el nivel de medición.

Este procedimiento pega la sintaxis de comando de MLP.

1.8 Particiones (Perceptrón multicapa)

Conjunto de datos de partición. Este grupo especifica el método de crear particiones en el conjunto de datos activo correspondientes a las muestras de entrenamiento, prueba y reserva. La **muestra de entrenamiento** comprende los registros de datos utilizados para entrenar la red neuronal; cierto porcentaje de casos del conjunto de datos debe asignarse a la muestra de entrenamiento para poder obtener un modelo. La **muestra de prueba** es un conjunto independiente de registros de datos que se utiliza para realizar un seguimiento de los errores durante el entrenamiento, con el fin de evitar un exceso de entrenamiento. Es muy recomendable crear una muestra de entrenamiento y el entrenamiento de la red generalmente será más eficiente si la muestra de prueba es más pequeña que la muestra de entrenamiento. La **muestra reservada** es otro conjunto independiente de registros de datos que se utiliza para evaluar la red neuronal final; el error de la muestra reservada ofrece una estimación "sincera" de la capacidad predictora del modelo, ya que los casos reservados no se utilizan para crear el modelo.

- **Asigne aleatoriamente los casos en función del número relativo de casos.** Especifique el número relativo (proporción) de casos asignados aleatoriamente a cada muestra (entrenamiento, prueba y reserva). Las columnas de % indican el porcentaje de casos que se asignarán a cada muestra según los números relativos que se han especificados.

Por ejemplo, especificar 7, 3, 0 como números relativos de las muestras de entrenamiento, prueba y reserva, equivale a especificar 70%, 30% y 0%. Especificar 2, 1, 1 como números relativos equivale a 50%, 25% y 25%; 1, 1, 1 equivale a dividir el conjunto de datos en tres partes iguales para entrenamiento, prueba y reserva.

- **Utilizar variable de partición para asignar los casos.** Especifique una variable numérica que asigne cada caso del conjunto de datos activo a la muestra de entrenamiento, prueba o reserva. Los casos con un valor positivo de la variable se asignarán a la muestra de entrenamiento, los casos con un valor 0 a la muestra de prueba y los casos con un valor negativo se asignarán a la muestra reservada. Los casos con un valor perdido del sistema se excluirán del análisis. Todos los valores perdidos de usuario de la variable de partición se tratarán siempre como válidos.

Nota: utilizar una variable de partición no garantiza resultados idénticos en posteriores ejecuciones de este procedimiento. Consulte "Replicación de los resultados" en el tema principal Perceptrón multicapa.

Para configurar las particiones del perceptrón multicapa

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

2. En el cuadro de diálogo Perceptrón multicapa, pulse en la pestaña **Particiones**.

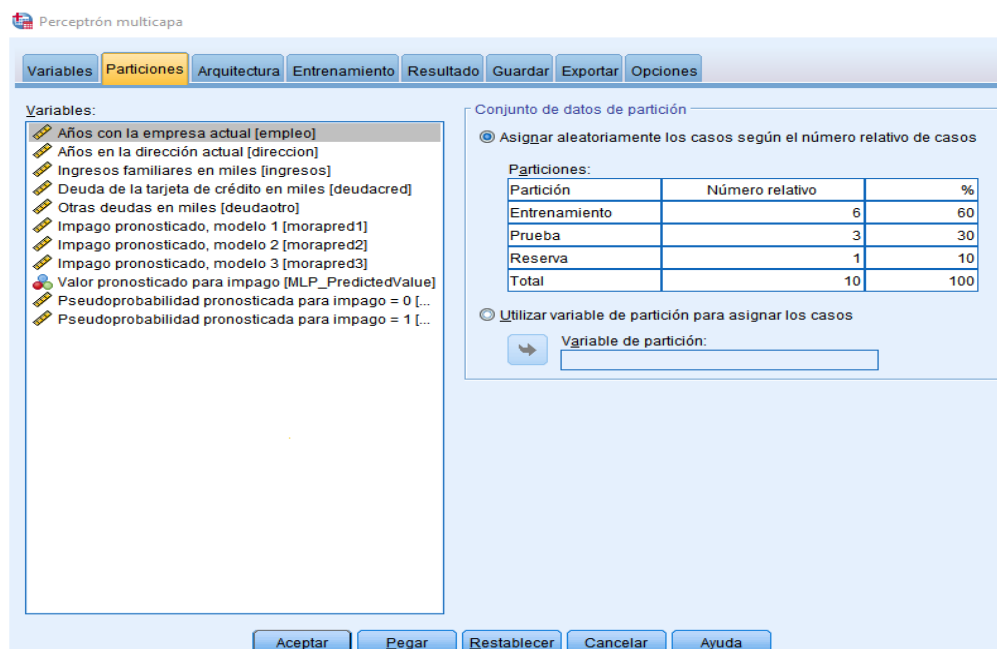


Figura 36 pestaña Particiones

Fuente. Copyright IBM Corp. _2018

2.1 Arquitectura (Perceptrón multicapa)

La pestaña Arquitectura se utiliza para especificar la estructura de la red. El procedimiento puede seleccionar automáticamente la "mejor" arquitectura o el usuario puede especificar una arquitectura personalizada.

La selección automática de arquitectura genera una red con una capa oculta. Especifique el número máximo y el mínimo de unidades permitido en la capa oculta, y la selección automática de arquitectura calculará cuál es el "mejor" número de unidades de la capa oculta. La selección automática de arquitectura utiliza las funciones de activación predeterminadas para las capas ocultas y de salida.

La selección de arquitectura personalizada le da un control de experto sobre las capas ocultas y de salida, y puede ser muy útil cuando se sabe de antemano qué arquitectura se desea o cuando necesita refinar los resultados de la selección automática de arquitectura.

2.1.1 Capas ocultas

La capa oculta contiene nodos de red no observables (unidades). Cada unidad oculta es una función de la suma ponderada de las entradas. La función es la función de activación y los valores de las ponderaciones se determinan mediante el algoritmo de estimación. Si la red contiene una segunda capa oculta, cada unidad oculta de la segunda capa es una función de la suma ponderada de las unidades de la primera capa oculta. La misma función de activación se utiliza en ambas capas.

Número de capas ocultas. Una percepción multicapa puede tener una o dos capas ocultas.

2.1.2 Función de activación.

La función de activación "relaciona" la suma ponderada de unidades de una capa, con los valores de unidades en la capa correcta.

- **Tangente hiperbólica.** Esta función tiene la forma: $\gamma(c) = \tanh(c) = (e^c - e^{-c}) / (e^c + e^{-c})$. Toma argumentos de valor real y los transforma al rango (-1, 1). Cuando se utiliza la selección automática de arquitectura, ésta es la función de activación para todas las unidades de las capas ocultas.
- **Sigmoide.** Esta función tiene la forma: $\gamma(c) = 1 / (1 + e^{-c})$. Toma argumentos de valor real y los transforma al rango (0, 1).

Número de unidades. El número de unidades en cada capa oculta se puede especificar explícitamente o el algoritmo de estimación puede determinarlo automáticamente.

2.1.3 Capa de salida

La capa de salida contiene las variables (dependientes) de destino.

Función de activación. La función de activación "relaciona" la suma ponderada de unidades de una capa, con los valores de unidades en la capa correcta.

- **Identidad.** Esta función tiene la fórmula: $\gamma(c) = c$. Toma argumentos de valor real y los devuelve sin modificar. Cuando se utiliza la selección automática de arquitectura, ésta es la función de activación para las unidades de la capa de salida si hay variables dependientes de escala.
- **Softmax.** Esta función tiene la forma: $\gamma(c_k) = \exp(c_k) / \sum_j \exp(c_j)$. Toma un vector de argumentos de valor real y lo transforma en un vector cuyos elementos quedan comprendidos en el rango (0, 1) y suman 1. Softmax está sólo disponible si todas las variables dependientes son categóricas. Cuando se utiliza la selección automática de arquitectura, ésta es la función de activación para las unidades de la capa de salida si todas las variables dependientes son categóricas.
- **Tangente hiperbólica.** Esta función tiene la forma: $\gamma(c) = \tanh(c) = (e^{-c} - e^c) / (e^{-c} + e^c)$. Toma argumentos de valor real y los transforma al rango (-1, 1).
- **Sigmoide.** Esta función tiene la forma: $\gamma(c) = 1 / (1 + e^{-c})$. Toma argumentos de valor real y los transforma al rango (0, 1).

2.1.4 Cambio de escala de variables dependientes de escala.

Estos controles sólo se encuentran disponibles si se ha seleccionado al menos una variable dependiente de escala.

- **Tipificados.** Se resta la media y se divide por la desviación estándar, $(x - \text{media}) / s$.
- **Normalizados.** Se resta el mínimo y se divide por el rango, $(x - \text{mín}) / (\text{máx} - \text{mín})$. Los valores normalizados quedan comprendidos entre 0 y 1. Éste es el método de cambio de escala necesario para las variables dependientes de escala si la capa de salida utiliza la función de activación sigmoide. La opción de corrección especifica un pequeño número que se aplica como corrección a la fórmula de cambio de escala; esta corrección garantiza que los valores de todas las variables dependientes reescaladas quedan comprendidos dentro del rango de la función de activación. En concreto, los valores 0 y 1, que se producen en la fórmula no corregida cuando x toma los valores mínimo o máximo, definen los límites del rango de la función sigmoide, pero no se encuentran dentro de dicho rango. La fórmula correcta es $[x - (\text{mín} - \epsilon)] / [(\text{máx} + \epsilon) - (\text{mín} - \epsilon)]$. Especifique un número mayor o igual que 0.

- **Normalizados corregidos.** Versión ajustada de la resta del mínimo y división por el rango, $[2*(x-\text{mín})/(\text{máx}-\text{mín})]-1$. Los valores ajustados normalizados quedan comprendidos entre -1 y 1 . Éste es el método de cambio de escala necesario para las variables dependientes de escala si la capa de salida utiliza la función de activación de tangente hiperbólica. La opción de corrección especifica un pequeño número que se aplica como corrección a la fórmula de cambio de escala; esta corrección garantiza que los valores de todas las variables dependientes reescaladas quedan comprendidos dentro del rango de la función de activación. En concreto, los valores -1 y 1 , que se producen en la fórmula no corregida cuando x toma los valores mínimo o máximo, definen los límites del rango de la función tangente hiperbólica, pero no se encuentran dentro de dicho rango. La fórmula corregida es $\{2*[(x-(\text{mín}-\epsilon))/((\text{máx}+\epsilon)-(\text{mín}-\epsilon))]\}-1$. Especifique un número mayor o igual que 0 .
- **Ninguna.** Sin cambio de escala de las variables dependientes de escala.

Para especificar la arquitectura de un perceptrón multicapa

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

2. En el cuadro de diálogo Perceptrón multicapa, pulse en la pestaña **Arquitectura**.

Perceptrón multicapa

Variables Particiones **Arquitectura** Entrenamiento Resultado Guardar Exportar Opciones

Selección automática de arquitectura
 Número mínimo de unidades de la capa oculta:
 Número máximo de unidades de la capa oculta:

Arquitectura personalizada

Capas ocultas

Número de capas ocultas:
 Ung
 Dos

Función de activación:
 Tangente hiperbólica
 Sigmoide

Número de unidades:
 Calcular automáticamente
 Personalizado
 Capa oculta 1:
 Capa oculta 2:

Capa de salida

Función de activación:
 Identidad
 Softmax
 Tangente hiperbólica
 Sigmoide

La función de activación elegida para la capa de salida determina los métodos de cambio de escala que están disponibles.

Cambio de escala de variables dependientes de escala:
 Estandarizados
 Normalizados
 Corrección:
 Corregidos normalizados
 Corrección:
 Ninguno

Aceptar Pegar Restablecer Cancelar Ayuda

Figura 37 Pestaña Arquitectura

Fuente. Copyright IBM Corp. _2018

2.2 Entrenamiento (Perceptrón multicapa)

La pestaña Entrenamiento se utiliza para especificar el modo en que debe entrenarse la red. El tipo de entrenamiento y el algoritmo de optimización determinan qué opciones de entrenamiento se encuentran disponibles.

2.2.1 Tipo de entrenamiento.

El tipo de entrenamiento determina el modo en que la red procesa los registros. Seleccione uno de los siguientes tipos de entrenamiento:

- **Lote.** Actualiza las ponderaciones sinápticas sólo tras pasar todos los registros de datos de entrenamiento; es decir, el entrenamiento por lotes utiliza la información de todos los registros del conjunto de datos de entrenamiento. El entrenamiento por lotes se suele preferir porque minimiza directamente el error total; sin embargo, el entrenamiento por

lotes puede obligar a actualizar muchas veces las ponderaciones hasta que se cumpla alguna de las reglas de parada y por tanto pueden ser necesarias muchas lecturas de datos. Resulta muy útil para conjuntos de datos "más pequeños".

- **En línea.** Actualiza las ponderaciones sinápticas después de cada registro de datos de entrenamiento; es decir, el entrenamiento en línea utiliza la información de un solo registro por vez. El entrenamiento en línea obtiene continuamente un registro y actualiza las ponderaciones hasta que se cumple una regla de parada. Si se han utilizado todos los registros una vez y no se cumple ninguna regla de parada, el proceso continúa reciclando los registros de datos. El entrenamiento en línea es superior al entrenamiento por lotes para los grupos de datos "más grandes" con predictores asociados; es decir, si hay muchos registros y muchas entradas, y sus valores no son independientes entre sí, el entrenamiento en línea puede obtener una respuesta razonable con mayor rapidez que el entrenamiento por lotes.
- **Por mini lotes.** Divide los registros de datos de entrenamiento en grupos de tamaño parecido y actualiza las ponderaciones sinápticas tras pasar un grupo; es decir, el entrenamiento por mini lotes utiliza la información de un grupo de registros. A continuación, el proceso recicla el grupo de datos si es necesario. El entrenamiento por mini lotes ofrece una solución intermedia entre el entrenamiento por lotes y en línea, y puede ser el idóneo para conjuntos de datos de "tamaño medio". El procedimiento puede determinar automáticamente el número de registros de entrenamiento por mini lote, o bien puede especificar un entero mayor que 1 y menor o igual que el número máximo de casos para almacenar en memoria. Puede establecer el número máximo de casos a almacenar en memoria en la pestaña Opciones.

2.2.2 Algoritmo de optimización.

Éste es el método utilizado para estimar las ponderaciones sinápticas.

- **Gradiente conjugado escalado.** Los supuestos que justifican el uso de métodos de gradiente conjugado se aplican únicamente a los tipos de entrenamiento por lotes, de modo que este método no se encuentra disponible para el entrenamiento en línea o por mini lotes.
- **Pendiente de gradiente.** Este método debe utilizarse con el entrenamiento en línea o por mini lotes; también puede utilizarse con el entrenamiento por lotes.

2.2.3 Opciones de entrenamiento.

Las opciones de entrenamiento le permiten ajustar el algoritmo de optimización. Generalmente no tendrá que cambiar estos ajustes a menos que la red experimente problemas con la estimación.

Entre las opciones de entrenamiento del algoritmo de gradiente conjugado escalado se incluyen:

- **Lambda inicial.** El valor inicial del parámetro lambda para el algoritmo de gradiente conjugado escalado. Especifique un número mayor que 0 y menor que 0,000001.
- **Sigma inicial.** El valor inicial del parámetro sigma para el algoritmo de gradiente conjugado escalado. Especifique un número mayor que 0 y menor que 0,0001.
- **Centro de intervalo y desplazamiento de intervalo.** El centro del intervalo (a_0) y el desplazamiento de intervalo (a) definen el intervalo $[a_0 - a, a_0 + a]$, en el que se generan aleatoriamente vectores de ponderación cuando se utiliza el recocido simulado. Simulated annealing is used to break out of a local minimum, with the goal of finding the global minimum, during application of the optimization algorithm. Este enfoque se utiliza en la inicialización de ponderaciones y la selección automática de arquitectura. Especifique un número para el centro de intervalo y un número mayor que 0 para el desplazamiento de intervalo.

Entre las opciones de entrenamiento del algoritmo de pendiente de gradiente se incluyen:

- **Tasa de aprendizaje inicial.** El valor inicial de la tasa de aprendizaje para el algoritmo de pendiente de gradiente. Una tasa de aprendizaje superior significa que la red se entrenará más rápidamente, con el inconveniente de que puede volverse inestable. Especifique un número mayor que 0.
- **Límite inferior de la tasa de aprendizaje.** El límite inferior de la tasa de aprendizaje para el algoritmo de pendiente de gradiente. Este ajuste se aplica sólo al entrenamiento en línea y por mini lotes. Especifique un número mayor que 0 y menor que la tasa de aprendizaje inicial.
- **Impulso.** El parámetro de impulso inicial para el algoritmo de pendiente de gradiente. El término de impulso ayuda a evitar las inestabilidades provocadas por una tasa de aprendizaje demasiado elevada. Especifique un número mayor que 0.
- **Reducción de la tasa de aprendizaje, en épocas.** El número de épocas (p), o lecturas de datos de la muestra de entrenamiento, para reducir la tasa de aprendizaje inicial al límite inferior de la tasa de aprendizaje cuando la pendiente de gradiente se utiliza con el entrenamiento en línea o por mini lotes. Esto le da control sobre el factor de decaimiento de la tasa de aprendizaje $\beta = (1/p K) * \ln(\eta_0/\eta_{low})$, donde η_0 es la tasa de aprendizaje inicial, η_{low} es el límite inferior de la tasa de aprendizaje y K es el número total de mini lotes (o el número de registros de entrenamiento, para el entrenamiento en línea) en el conjunto de datos de entrenamiento. Introduzca un entero mayor que 0.

Para configurar los criterios de entrenamiento del perceptrón multicapa

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

2. En el cuadro de diálogo Perceptrón multicapa, pulse en la pestaña **Entrenamiento**.

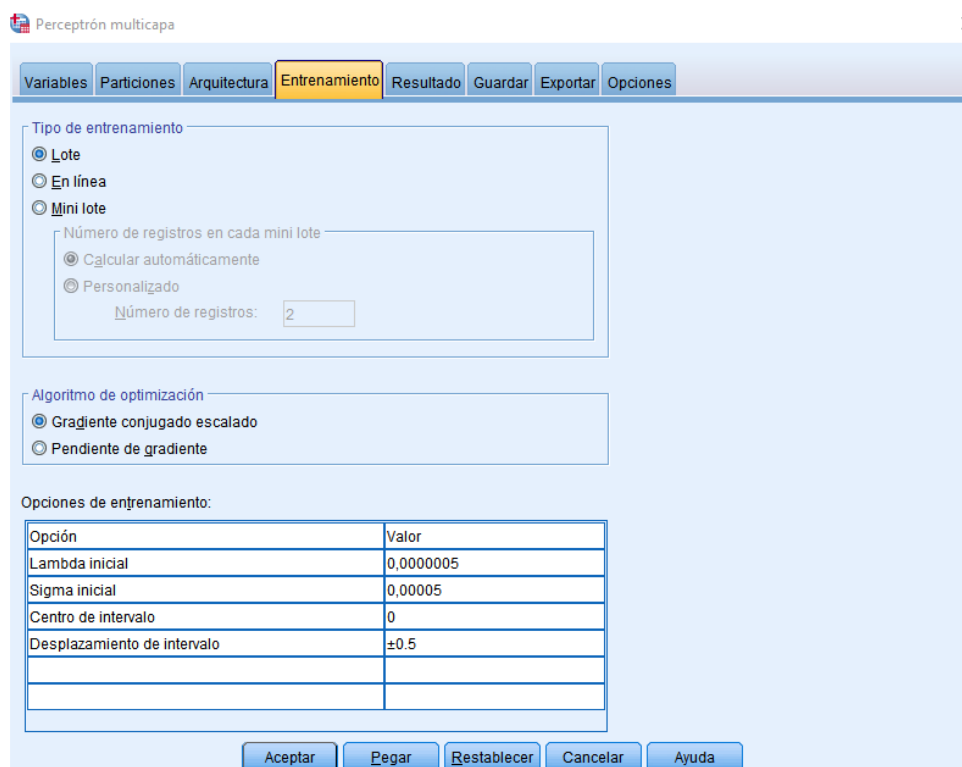


Figura 38 Pestaña Entrenamiento

Fuente. Copyright IBM Corp. _2018

2.3 Salida (Perceptrón multicapa)

2.3.1 Estructura de red.

Muestra información resumida sobre la red neuronal.

- **Descripción.** Muestra información sobre la red neuronal, incluyendo las variables dependientes, el número de unidades de entrada y de salida, el número de unidades y capas ocultas, y las funciones de activación.
- **Diagrama.** Muestra el diagrama de red como un gráfico que se puede editar. Tenga en cuenta que a medida que aumenta el número de covariables y niveles del factor, el diagrama se hace más difícil de interpretar.

- **Ponderaciones sinápticas.** Muestras las estimaciones de coeficiente que muestran la relación entre las unidades de una capa determinada con las unidades de la capa siguiente. Las ponderaciones sinápticas se basan en la muestra de entrenamiento incluso si el conjunto de datos activo se divide en datos de entrenamiento, comprobación y reservados. Tenga en cuenta que el número de ponderaciones sinápticas puede llegar a ser bastante elevado, y estas ponderaciones generalmente no se utilizan para interpretar los resultados de red.

2.3.2 Rendimiento de red.

Muestra los resultados utilizados para determinar si el modelo es "bueno". *Nota:* los gráficos de este grupo se basan en las muestras de entrenamiento y comprobación combinadas, o sólo la muestra de entrenamiento si no existe muestra de comprobación.

- **Resumen de modelo.** Muestra un resumen de los resultados de la red neuronal por partición y global, incluyendo el error, el error relativo o el porcentaje de predicciones incorrectas, la regla de parada utilizada para detener el entrenamiento y el tiempo de entrenamiento.

El error es el error de sumas cuadráticas cuando se aplica la función de activación de tangente hiperbólica, identidad o sigmoide a la capa de salida. Éste es el error de entropía cruzada cuando se aplica la función de activación softmax a la capa de salida.

Los porcentajes o errores relativos de predicciones incorrectas se muestran dependiendo de los niveles de medición de la variable dependiente. Si alguna variable dependiente tiene un nivel de medición de escala, se muestra el error relativo global promedio (relativo al modelo de la media). Si todas las variables dependientes son categóricas, se muestra el porcentaje promedio de predicciones incorrectas. También se muestran los errores relativos o porcentajes de predicciones incorrectas para las variables dependientes individuales.

- **Resultados de clasificación.** Muestra una tabla de clasificación para cada variable dependiente categórica por partición y global. Cada tabla proporciona el número de casos clasificados correcta e incorrectamente para cada categoría de variable dependiente. También se informa del porcentaje de casos totales que se clasificaron correctamente.
- **Curva COR.** Muestra una curva COR (del inglés Receiver Operating Characteristic, características operativas del receptor) para cada variable dependiente categórica. También muestra una tabla que proporciona el área bajo cada curva. Para una variable dependiente dada, el gráfico COR muestra una curva para cada categoría. Si la variable dependiente tiene dos categorías, cada curva trata la categoría en cuestión como el estado positivo frente a la otra categoría. Si la variable dependiente tiene más de dos categorías, cada curva trata la categoría en cuestión como el estado positivo frente a la agregación de las demás categorías.

- **Gráfico de ganancias acumuladas.** Muestra un gráfico de ganancias acumuladas para cada variable dependiente categórica. La presentación de una curva para cada categoría de variable dependiente es la misma que para las curvas COR.
- **Gráfico de elevación.** Muestra un gráfico de elevación para cada variable dependiente categórica. La presentación de una curva para cada categoría de variable dependiente es la misma que para las curvas COR.
- **Gráfico de predichos por observación.** Muestra un gráfico de valores pronosticados por observados para cada variable dependiente. Para las variables dependientes categóricas, se muestran diagramas de caja agrupados de pseudoprobabilidades pronosticadas para cada categoría de respuesta, con la categoría de respuesta observada como la variable de clúster. Para las variables dependientes de escala se muestra un diagrama de dispersión.
- **Gráfico de residuos por pronosticados.** Muestra un gráfico de residuos por valores pronosticados para cada variable dependiente de escala. No debería haber patrones visibles entre los residuos y los valores pronosticados. Este gráfico sólo se genera para las variables dependientes de escala.

2.3.3 Resumen de procesamiento de casos.

Muestra la tabla de resumen de procesamiento de casos, que resume el número de casos incluidos y excluidos en el análisis, en total y por muestras de entrenamiento, comprobación y reservadas.

2.3.4 Análisis de importancia de la variable independiente.

Realiza un análisis de sensibilidad, que calcula la importancia de cada predictor en la determinación de la red neuronal. El análisis se basa en las muestras de entrenamiento y comprobación combinadas, o sólo la muestra de entrenamiento si no existe muestra de comprobación. Así se crea una tabla y un gráfico que muestra la importancia y la importancia normalizada de cada predictor. Tenga en cuenta que el análisis de sensibilidad supone un proceso de cálculo muy extenso que puede tardar bastante si el número de predictores o de casos es elevado.

Para seleccionar los resultados del perceptrón multicapa

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

2. En el cuadro de diálogo Perceptrón multicapa, pulse en la pestaña **Resultado**.

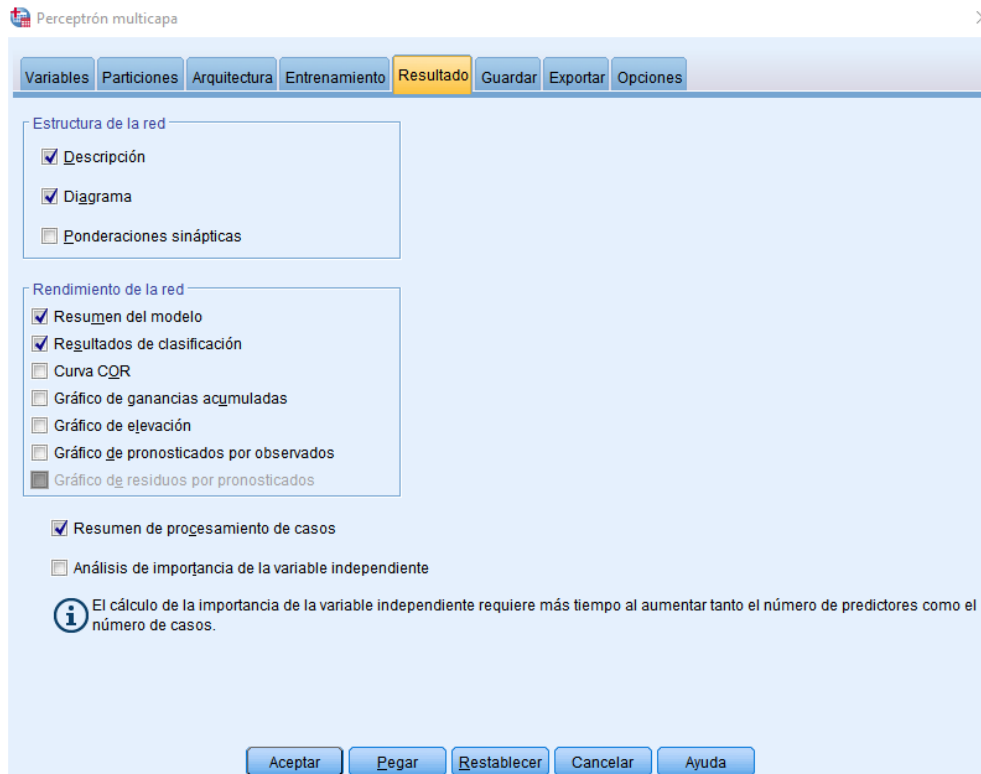


Figura 39 Pestaña Resultado

Fuente. Copyright IBM Corp. _2018

2.4 Guardar (Perceptrón multicapa)

La pestaña Guardar se utiliza para guardar predicciones como variables en el conjunto de datos.

- **Guardar categoría o valor predicho para cada variable dependiente.** Esta opción guarda el valor predicho para variables dependientes de escala y la categoría pronosticada para variables dependientes categóricas.
- **Guardar pseudoprobabilidad pronosticada o categoría para cada variable dependiente.** Esta opción guarda las pseudoprobabilidades pronosticadas para variables dependientes categóricas. Para cada una de las primeras n categorías se guarda una variable diferente, donde n se especifica en la columna **Categorías para guardar**.

2.4.1 Nombres de las variables guardadas.

La generación automática de nombres garantiza que conserva todo su trabajo. Los nombres personalizados le permiten descartar/reemplazar los resultados de las ejecuciones anteriores sin eliminar antes las variables guardadas en el Editor de datos.

Probabilidades y pseudoprobabilidades

Las variables dependientes categóricas con activación softmax y error de entropía cruzada tendrán un valor predicho para cada categoría, donde cada valor predicho es la probabilidad de que el caso pertenezca a la categoría.

Las variables dependientes categóricas con error de sumas cuadráticas tendrá un valor predicho para cada categoría, pero los valores pronosticados no se pueden interpretar como probabilidades. El procedimiento guarda estas pseudoprobabilidades pronosticadas incluso si son menores que 0 o mayores que 1, o la suma para una variable dependiente dada no es 1.

Los gráficos de curva COR, ganancias acumuladas y elevación (consulte Salida (Perceptrón multicapa)) se crean en base a pseudoprobabilidades. En el caso de que alguna de las pseudoprobabilidades sea menor que 0 o mayor que 1, o la suma para una variable dada no sea 1, primero se reescalan para que estén entre 0 y 1 y para que sumen 1. Las pseudoprobabilidades se reescalan dividiéndolas por su suma. Por ejemplo, si un caso tiene pseudoprobabilidades pronosticadas de 0,50, 0,60 y 0,40 para una variable dependiente de tres categorías, cada pseudoprobabilidad se divide por la suma 1,50 para obtener 0,33, 0,40 y 0,27.

Si alguna de las pseudoprobabilidades es negativa, se suma el valor absoluto de la menor a todas las pseudoprobabilidades antes del cambio de escala anterior. Por ejemplo, si las pseudoprobabilidades son -0,30, 0,50 y 1,30, primero se añade 0,30 a cada valor para obtener 0,00, 0,80 y 1,60. A continuación, se divide cada nuevo valor por la suma 2,40 para obtener 0,00, 0,33 y 0,67.

Para guardar las variables en el conjunto de datos activo

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

2. En el cuadro de diálogo Perceptrón multicapa, pulse en la pestaña **Guardar**.

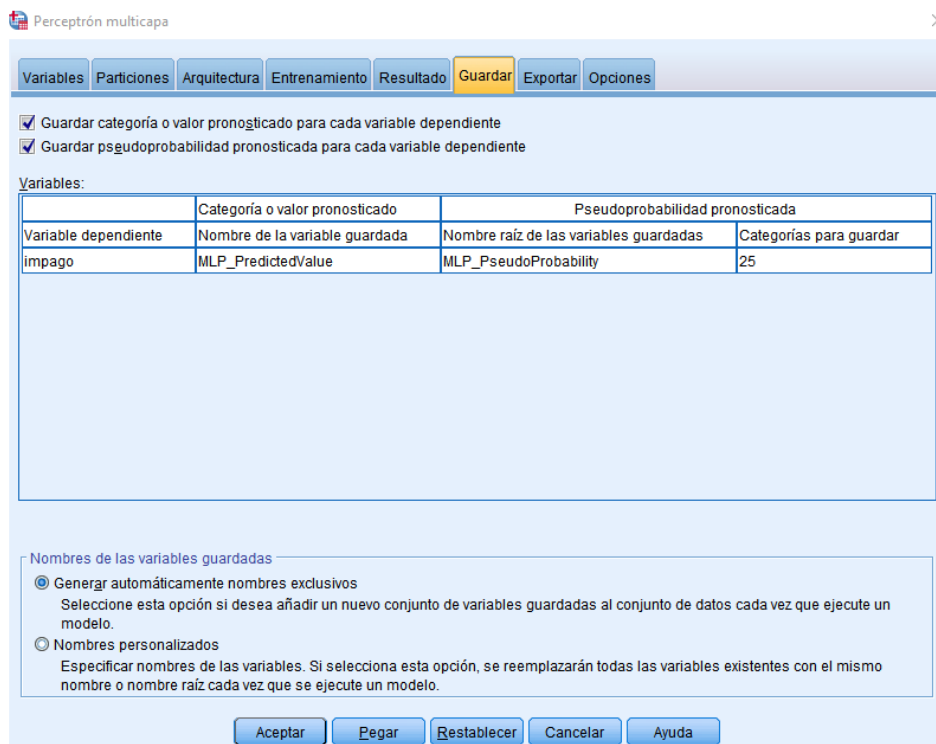


Figura 40 Pestaña Guardar

Fuente. Copyright IBM Corp. _2018

2.5 Exportar (Perceptrón multicapa)

La pestaña Exportar se utiliza para guardar las estimaciones de ponderaciones sinápticas para cada variable dependiente en un archivo XML (PMML). Puede utilizar este archivo de modelo para aplicar la información del modelo a otros archivos de datos para puntuarlo. Para obtener más información, consulte el tema Asistente para puntuación. Esta opción no se encuentra disponible si se han definido archivos segmentados.

Para exportar las ponderaciones sinápticas

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

2. En el cuadro de diálogo Perceptrón multicapa, pulse en la pestaña **Exportar**.

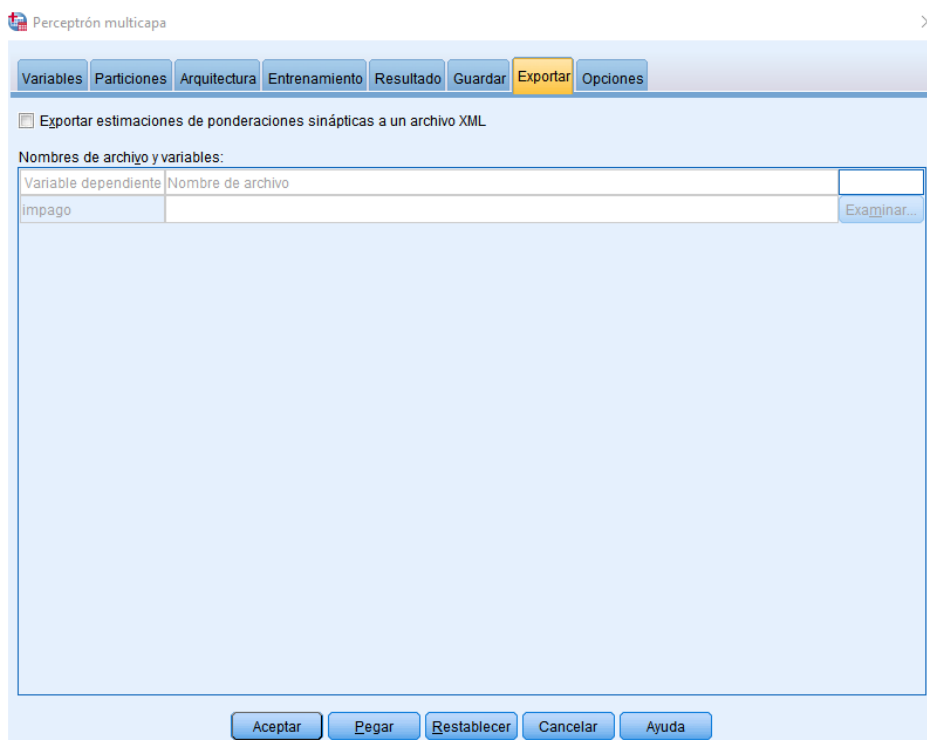


Figura 41 Pestaña Exportar.

Fuente. Copyright IBM Corp. _2018

2.6 Opciones (Perceptrón multicapa)

2.6.1 Valores perdidos del usuario.

Los factores deben tener valores válidos para el caso para que se incluyan en el análisis. Estos controles permiten decidir si los valores perdidos del usuario se deben tratar como válidos entre los factores y las variables dependientes categóricas.

2.6.2 Reglas de parada.

Son las reglas que determinan cuando detener el entrenamiento de la red neuronal. El entrenamiento pasa al menos por una lectura de los datos. El entrenamiento puede entonces detenerse de acuerdo con los siguientes criterios, que se comprueban en el orden indicado en la lista. En las siguientes definiciones de regla de parada, un paso corresponde a una lectura de datos para los métodos en línea y por mini lotes, una iteración para el método por lotes.

- **Número máximo de pasos sin una disminución del error.** El número de pasos que se permitirán antes de comprobar si existe una disminución del error. Si no existe una disminución del error tras el número especificado de pasos, se detiene el entrenamiento.

Especifique un número entero mayor que 0. También puede especificar la muestra de datos que se va a utilizar para calcular el error. **Elegir automáticamente** utiliza la muestra de comprobación si la hay y, si no, la muestra de entrenamiento. Tenga en cuenta que el entrenamiento por lotes garantiza una disminución de los errores de la muestra de entrenamiento tras cada lectura de los datos, por lo que esta opción se aplica sólo al entrenamiento por lotes si existe una muestra de comprobación. **Datos de entrenamiento y de prueba** comprueba el error para cada una de estas muestras; esta opción se aplica únicamente si existe una muestra de comprobación.

Nota: tras cada lectura completa de los datos, el entrenamiento en línea y por mini lotes requiere una lectura adicional de los datos para calcular el error de entrenamiento. Esta lectura adicional de los datos puede ralentizar considerablemente el entrenamiento, por lo que generalmente se recomienda suministrar una muestra de comprobación y seleccionar **Elegir automáticamente** en cualquier caso.

- **Tiempo máximo de entrenamiento.** Elija si desea especificar un número máximo de minutos para que se ejecute el algoritmo. Especifique un número mayor que 0.
- **Número máximo de épocas de entrenamiento.** El número máximo de épocas (lecturas de datos) permitido. Si se supera el número máximo de épocas, se detiene el entrenamiento. Introduzca un entero mayor que 0.
- **Cambio mínimo relativo del error de entrenamiento.** El entrenamiento se detiene si el cambio relativo del error de entrenamiento comparado con el del paso anterior es menor que el criterio. Especifique un número mayor que 0. Para el entrenamiento en línea y por mini lotes, este criterio se ignora si sólo se utilizan datos de comprobación para calcular el error.
- **Cambio mínimo relativo de la tasa de errores de entrenamiento.** El entrenamiento se detiene si el cociente del error de entrenamiento respecto al error del modelo nulo es menor que el valor del criterio. El modelo nulo pronostica el valor promedio para todas las variables dependientes. Especifique un número mayor que 0. Para el entrenamiento en línea y por mini lotes, este criterio se ignora si sólo se utilizan datos de comprobación para calcular el error.

2.6.3 Número máximo de casos que se almacenarán en memoria.

Esta opción controla los siguientes ajustes de los algoritmos del perceptrón multicapa. Especifique un entero mayor que 1.

- En la selección automática de la arquitectura, el tamaño de la muestra utilizada para determinar la arquitectura de la red es $\min(1000, \text{tamañoem})$, donde *tamañoem* es el número máximo de casos que se pueden almacenar en la memoria.

- En el entrenamiento por mini lotes con cálculo automático del número de mini lotes, el número de mini lotes es $\min(\max(M/10,2), \text{tamaño mem})$, donde M es el número de casos de la muestra de entrenamiento.

Para configurar las opciones del perceptrón multicapa

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Perceptrón multicapa...

2. En el cuadro de diálogo Perceptrón multicapa, pulse en la pestaña **Opciones**.

Figura 42 Pestaña Opciones.

Fuente. Copyright IBM Corp. _2018

2.7 Función de base radial

El procedimiento Función de base radial (RBF) genera un modelo predictivo para una o más variables dependientes (de destino) basado en los valores de las variables predictoras.

Ejemplo. Un proveedor de telecomunicaciones ha segmentado su base de clientes por patrones de uso del servicio, clasificando a los clientes en cuatro grupos. Una red RBF que utiliza los datos demográficos para pronosticar la pertenencia a los grupos permite a la compañía personalizar las ofertas para cada uno de los posibles clientes.

Consideraciones de los datos

2.8 Variables dependientes.

Las variables dependientes pueden ser:

- *Nominal.* Una variable puede ser tratada como nominal cuando sus valores representan categorías que no obedecen a una clasificación intrínseca. Por ejemplo, el departamento de la compañía en el que trabaja un empleado. Algunos ejemplos de variables nominales son: región, código postal o confesión religiosa.
- *Ordinal.* Una variable puede ser tratada como ordinal cuando sus valores representan categorías con alguna clasificación intrínseca. Por ejemplo, los niveles de satisfacción con un servicio, que abarquen desde muy insatisfecho hasta muy satisfecho. Entre los ejemplos de variables ordinales se incluyen escalas de actitud que representan el grado de satisfacción o confianza y las puntuaciones de evaluación de las preferencias.
- *Escalas.* Una variable puede tratarse como escala (continua) cuando sus valores representan categorías ordenadas con una métrica con significado, por lo que son adecuadas las comparaciones de distancia entre valores. Son ejemplos de variables de escala: la edad en años y los ingresos en dólares.

El procedimiento supone que el nivel de medición adecuado se ha asignado a todas las variables dependientes, aunque puede cambiar temporalmente el nivel de medición para una variable si pulsa el botón derecho la variable en la lista de variables de origen y selecciona un nivel de medición en el menú emergente. Para modificar permanentemente el nivel de medición de una variable, consulte Nivel de medición de variable.

Un icono situado junto a cada variable de la lista de variables identifica el nivel de medición y el tipo de datos como lo muestra la tabla 6.

Tabla 16 *Iconos de nivel de medición*

	Numérico	Cadena	Fecha	Hora
Escala (Continuo)		n/a		
Ordinal				
Nominal				

Fuente. Copyright IBM Corp. _2018

2.9 Variables predictoras.

Los predictores se pueden especificar como factores (categóricos) o covariables (de escala).

2.10 Codificación de la variable categórica.

El procedimiento recodifica temporalmente predictores categóricos y variables dependientes utilizando la codificación "una de c " para todo el procedimiento. Si hay c categorías de una variable, la variable se almacena como vectores c , con la primera categoría denotada $(1,0,\dots,0)$, la siguiente categoría $(0,1,0,\dots,0)$, ..., y la última categoría $(0,0,\dots,0,1)$.

Este esquema de codificación incrementa el número de ponderaciones sinápticas y puede dar como resultado un entrenamiento más lento, pero los métodos de codificación más "compactos" normalmente conducen a redes neuronales con un peor ajuste. Si el entrenamiento de la red avanza muy lentamente, pruebe a reducir el número de categorías en los predictores categóricos combinando categorías similares o eliminando los casos que tengan categorías extremadamente raras. Si desea obtener más información sobre la recodificación de variables, consulte Recodificar en las mismas variables o Recodificar en distintas variables.

Toda codificación "Una de c " se basa en los datos de entrenamiento, incluso si se define una muestra reservada o de comprobación (consulte Particiones (Función de base radial)). De este modo, si las muestras reservadas o de comprobación contienen casos con categorías de predictores que no están presentes en los datos de entrenamiento, esos casos no son utilizados por el procedimiento o en la puntuación. De este modo, si las muestras reservadas o de comprobación

contienen casos con categorías de variables dependientes que no están presentes en los datos de entrenamiento, esos casos no son utilizados por el procedimiento o en la puntuación.

2.11 Cambio de escala.

De forma predeterminada, se cambia la escala de las covariables y las variables dependientes de escala para mejorar el entrenamiento de la red. Todo cambio de escala se realiza en base a los datos de entrenamiento, incluso si se define una muestra reservada o de comprobación (consulte Particiones (Función de base radial)). Es decir, dependiendo del tipo de cambio de escala, la media, la desviación estándar, el valor mínimo o el valor máximo de una covariable o variable dependiente se calculan utilizando sólo los datos de entrenamiento. Si especifica una variable para definir particiones, es importante que estas covariables o variables dependientes tengan distribuciones similares en todas las muestras reservadas, de entrenamiento o comprobación. Utilice por ejemplo, Explorar para examinar las distribuciones en las particiones.

2.12 Ponderaciones de frecuencia.

Este procedimiento ignora las ponderaciones de frecuencia.

2.13 Replicación de los resultados.

Si desea replicar exactamente sus resultados, utilice el mismo valor de inicialización para el generador de números aleatorios y el mismo orden de datos, además de utilizar la misma configuración del procedimiento. A continuación se incluyen más detalles sobre este tema:

- **Generación de números aleatorios.** El procedimiento utiliza la generación de números aleatorios durante la asignación de particiones. Para reproducir los mismos resultados aleatorios en el futuro, utilice el mismo valor de inicialización para el generador de números aleatorios antes de cada ejecución del procedimiento Función de base radial. Para obtener más información, consulte el tema Generadores de números aleatorios.
- **Orden de casos.** Los resultados dependen también del orden de los datos porque se utiliza el algoritmo de agrupación en clústeres de dos fases para determinar las funciones de base radial. Para obtener más información, consulte el tema Análisis de clústeres en dos fases.

Para minimizar los efectos del orden, ordene los casos aleatoriamente. Para comprobar la estabilidad de una solución dada, puede obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios. En situaciones con tamaños de archivo extremadamente grandes, se pueden llevar a cabo varias ejecuciones con una muestra de casos ordenados con distintos órdenes aleatorios.

Esta característica requiere la opción Redes neuronales.

En los menús, elija:

Analizar > Redes neuronales > Función de base radial...

1. Seleccione al menos una variable dependiente.
2. Seleccione al menos un factor o covariable.

Si lo desea, en la pestaña Variables puede cambiar el método para cambiar la escala de las covariables. Las diferentes opciones son:

- **Tipificados.** Se resta la media y se divide por la desviación estándar, $(x-\text{media})/s$.
- **Normalizados.** Se resta el mínimo y se divide por el rango, $(x-\text{mín})/(\text{máx}-\text{mín})$. Los valores normalizados quedan comprendidos entre 0 y 1.
- **Normalizados corregidos.** Versión ajustada de la resta del mínimo y división por el rango, $[2*(x-\text{mín})/(\text{máx}-\text{mín})]-1$. Los valores normalizados ajustados quedan comprendidos entre -1 y 1.
- **Ninguno.** Sin cambio de escala de covariables.

2.13.1 Campos con un nivel de medición desconocido

La alerta de nivel de medición se muestra si el nivel de medición de una o más variables (campos) del conjunto de datos es desconocido. Como el nivel de medición afecta al cálculo de los resultados de este procedimiento, todas las variables deben tener un nivel de medición definido.

Explorar datos. Lee los datos del conjunto de datos activo y asigna el nivel de medición predefinido en cualquier campo con un nivel de medición desconocido. Si el conjunto de datos es grande, puede llevar algún tiempo.

Asignar manualmente. Abre un cuadro de diálogo que contiene todos los campos con un nivel de medición desconocido. Puede utilizar este cuadro de diálogo para asignar el nivel de medición a esos campos. También puede asignar un nivel de medición en la Vista de variables del Editor de datos.

Como el nivel de medición es importante para este procedimiento, no puede acceder al cuadro de diálogo para ejecutar este procedimiento hasta que se hayan definido todos los campos en el nivel de medición.

Este procedimiento pega la sintaxis de comando de RBF.

2.14 Particiones (Función de base radial)

Conjunto de datos de partición. Este grupo especifica el método de crear particiones en el conjunto de datos activo correspondientes a las muestras de entrenamiento, prueba y reserva. La **muestra de entrenamiento** comprende los registros de datos utilizados para entrenar la red neuronal; cierto porcentaje de casos del conjunto de datos debe asignarse a la muestra de entrenamiento para poder obtener un modelo. La **muestra de prueba** es un conjunto independiente de registros de datos que se utiliza para realizar un seguimiento de los errores durante el entrenamiento, con el fin de evitar un exceso de entrenamiento. Es muy recomendable crear una muestra de entrenamiento y el entrenamiento de la red generalmente será más eficiente si la muestra de prueba es más pequeña que la muestra de entrenamiento. La **muestra reservada** es otro conjunto independiente de registros de datos que se utiliza para evaluar la red neuronal final; el error de la muestra reservada ofrece una estimación "sincera" de la capacidad predictora del modelo, ya que los casos reservados no se utilizan para crear el modelo.

- **Asigne aleatoriamente los casos en función del número relativo de casos.** Especifique el número relativo (proporción) de casos asignados aleatoriamente a cada muestra (entrenamiento, prueba y reserva). Las columnas de % indican el porcentaje de casos que se asignarán a cada muestra según los números relativos que se han especificados.

Por ejemplo, especificar 7, 3, 0 como números relativos de las muestras de entrenamiento, prueba y reserva, equivale a especificar 70%, 30% y 0%. Especificar 2, 1, 1 como números relativos equivale a 50%, 25% y 25%; 1, 1, 1 equivale a dividir el conjunto de datos en tres partes iguales para entrenamiento, prueba y reserva.

- **Utilizar variable de partición para asignar los casos.** Especifique una variable numérica que asigne cada caso del conjunto de datos activo a la muestra de entrenamiento, prueba o reserva. Los casos con un valor positivo de la variable se asignarán a la muestra de entrenamiento, los casos con un valor 0 a la muestra de prueba y los casos con un valor negativo se asignarán a la muestra reservada. Los casos con un valor perdido del sistema se excluirán del análisis. Todos los valores perdidos de usuario de la variable de partición se tratarán siempre como válidos.

Para configurar las particiones de la función de base radial

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Función de base radial...

2. En el cuadro de diálogo Función de base radial, pulse en la pestaña **Particiones**.

2.15 Arquitectura (Función de base radial)

La pestaña Arquitectura se utiliza para especificar la estructura de la red. El procedimiento crea una red neuronal con una capa "función de base radial"; en general, no será necesario cambiar esta configuración.

2.15.1 Número de unidades en la capa oculta.

Hay tres formas de elegir el número de unidades ocultas.

1. **Busque el mejor número de unidades en un rango calculado automáticamente.** El procedimiento calcula automáticamente los valores mínimo y máximo del rango y localiza el mejor número de unidades ocultas en el rango.

Si se define una muestra de comprobación, el procedimiento utiliza el criterio de los datos de comprobación: el mejor número de unidades ocultas es el que produce el error menor en los datos de comprobación. Si no se define una muestra de comprobación, el procedimiento utiliza el criterio de información bayesiano (BIC): el mejor número de unidades ocultas es el que produce el menor BIC según los datos de entrenamiento.

2. **Busque el mejor número de unidades en un rango especificado.** Puede proporcionar su propio rango y el procedimiento encontrará el "mejor" número de unidades ocultas en ese rango. Como antes, el mejor número de unidades ocultas del rango se determina utilizando el criterio de datos de comprobación o el BIC.
3. **Utilizar un número especificado de unidades.** Puede anular el uso de un rango y especificar un número determinado de unidades directamente.

2.15.2 Función de activación para la capa oculta.

La función de activación para la capa oculta es la función de base radial, que "enlaza" las unidades de una capa con los valores de unidades de la capa siguiente. Para la capa de resultado, la función de activación es la función de identidad, por lo que las unidades de resultado son simplemente las sumas ponderadas de las unidades ocultas.

- **Función de base radial normalizada.** Utiliza la función de activación softmax de modo que las activaciones de todas las unidades ocultas se normalizan para sumar uno.
- **Función de base radial ordinaria.** Utiliza la función de activación exponencial de modo que la activación de la unidad oculta es una "curva" de Gauss como una función de las entradas.

2.15.3 Superposición entre unidades ocultas.

El factor de superposición es un multiplicador aplicado al ancho de las funciones de base radial. El valor calculado automáticamente del factor de superposición $1+0,1d$, donde d es el número de unidades de entrada (la suma del número de categorías de todos los factores y el número de covariables).

Para especificar la arquitectura de la función de base radial

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Función de base radial...

2. En el cuadro de diálogo Función de base radial, pulse en la pestaña **Arquitectura**.

2.16 Resultado (Función de base radial)

2.16.1 Estructura de red.

Muestra información resumida sobre la red neuronal.

- **Descripción.** Muestra información sobre la red neuronal, incluyendo las variables dependientes, el número de unidades de entrada y de salida, el número de unidades y capas ocultas, y las funciones de activación.
- **Diagrama.** Muestra el diagrama de red como un gráfico que se puede editar. Tenga en cuenta que a medida que aumenta el número de covariables y niveles del factor, el diagrama se hace más difícil de interpretar.
- **Ponderaciones sinápticas.** Muestras las estimaciones de coeficiente que muestran la relación entre las unidades de una capa determinada con las unidades de la capa siguiente. Las ponderaciones sinápticas se basan en la muestra de entrenamiento incluso si el conjunto de datos activo se divide en datos de entrenamiento, comprobación y reservados. Tenga en cuenta que el número de ponderaciones sinápticas puede llegar a ser bastante elevado, y estas ponderaciones generalmente no se utilizan para interpretar los resultados de red.

2.16.2 Rendimiento de red.

Muestra los resultados utilizados para determinar si el modelo es "bueno". *Nota:* los gráficos de este grupo se basan en las muestras de entrenamiento y comprobación combinadas, o sólo la muestra de entrenamiento si no existe muestra de comprobación.

- **Resumen de modelo.** Muestra un resumen de los resultados de la red neuronal por partición y global, incluyendo el error, el error relativo o el porcentaje de predicciones incorrectas y el tiempo de entrenamiento.

El error es el error de sumas cuadráticas. Además, se muestran los errores relativos o porcentajes de predicciones incorrectas dependiendo de los niveles de medición de la variable dependiente. Si alguna variable dependiente tiene un nivel de medición de escala, se muestra el error relativo global promedio (relativo al modelo de la media). Si todas las variables dependientes son categóricas, se muestra el porcentaje promedio de predicciones incorrectas. También se muestran los errores relativos o porcentajes de predicciones incorrectas para las variables dependientes individuales.

- **Resultados de clasificación.** Muestra una tabla de clasificación para cada variable dependiente categórica. Cada tabla proporciona el número de casos clasificados correcta e incorrectamente para cada categoría de variable dependiente. También se informa del porcentaje de casos totales que se clasificaron correctamente.
- **Curva COR.** Muestra una curva COR (del inglés Receiver Operating Characteristic, características operativas del receptor) para cada variable dependiente categórica. También muestra una tabla que proporciona el área bajo cada curva. Para una variable dependiente dada, el gráfico COR muestra una curva para cada categoría. Si la variable dependiente tiene dos categorías, cada curva trata la categoría en cuestión como el estado positivo frente a la otra categoría. Si la variable dependiente tiene más de dos categorías, cada curva trata la categoría en cuestión como el estado positivo frente a la agregación de las demás categorías.
- **Gráfico de ganancias acumuladas.** Muestra un gráfico de ganancias acumuladas para cada variable dependiente categórica. La presentación de una curva para cada categoría de variable dependiente es la misma que para las curvas COR.
- **Gráfico de elevación.** Muestra un gráfico de elevación para cada variable dependiente categórica. La presentación de una curva para cada categoría de variable dependiente es la misma que para las curvas COR.
- **Gráfico de predichos por observación.** Muestra un gráfico de valores pronosticados por observados para cada variable dependiente. Para las variables dependientes categóricas, se muestran diagramas de caja agrupados de pseudoprobabilidades pronosticadas para cada categoría de respuesta, con la categoría de respuesta observada como la variable de clúster. Para las variables dependientes de escala se muestra un diagrama de dispersión.
- **Gráfico de residuos por pronosticados.** Muestra un gráfico de residuos por valores pronosticados para cada variable dependiente de escala. No debería haber patrones visibles entre los residuos y los valores pronosticados. Este gráfico sólo se genera para las variables dependientes de escala.

2.16.3 Resumen de procesamiento de casos.

Muestra la tabla de resumen de procesamiento de casos, que resume el número de casos incluidos y excluidos en el análisis, en total y por muestras de entrenamiento, comprobación y reservadas.

2.16.4 Análisis de importancia de la variable independiente.

Realiza un análisis de sensibilidad, que calcula la importancia de cada predictor en la determinación de la red neuronal. El análisis se basa en las muestras de entrenamiento y comprobación combinadas, o sólo la muestra de entrenamiento si no existe muestra de comprobación. Así se crea una tabla y un gráfico que muestra la importancia y la importancia normalizada de cada predictor. Tenga en cuenta que el análisis de sensibilidad supone un proceso de cálculo muy extenso que puede tardar bastante si el número de predictores o de casos es elevado.

Para seleccionar los resultados de la función de base radial

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Función de base radial...

2. En el cuadro de diálogo Función de base radial, pulse en la pestaña **Resultado**.

2.17 Guardar (Función de base radial)

La pestaña Guardar se utiliza para guardar predicciones como variables en el conjunto de datos.

- **Guardar categoría o valor predicho para cada variable dependiente.** Esta opción guarda el valor predicho para variables dependientes de escala y la categoría pronosticada para variables dependientes categóricas.
- **Guardar pseudoprobabilidad predicha para cada variable dependiente.** Esta opción guarda las pseudoprobabilidades pronosticadas para variables dependientes categóricas. Para cada una de las primeras n categorías se guarda una variable diferente, donde n se especifica en la columna *Categorías para guardar*.

2.17.1 Nombres de las variables guardadas.

La generación automática de nombres garantiza que conserva todo su trabajo. Los nombres personalizados le permiten descartar/reemplazar los resultados de las ejecuciones anteriores sin eliminar antes las variables guardadas en el Editor de datos.

2.17.2 Probabilidades y pseudoprobabilidades

Las pseudoprobabilidades pronosticadas no se pueden interpretar como probabilidades porque el procedimiento Función de base radial utiliza el error de sumas cuadráticas y la función de activación de identidad para la capa de salida. El procedimiento guarda estas pseudoprobabilidades pronosticadas incluso si son menores que cero o mayores que uno, o la suma para una variable dependiente dada no es 1.

Los gráficos de curva COR, ganancias acumuladas y elevación (consulte Resultado (Función de base radial)) se crean en base a pseudoprobabilidades. En el caso de que alguna de las pseudoprobabilidades sea menor que 0 o mayor que 1, o la suma para una variable dada no sea 1, primero se reescalan para que estén entre 0 y 1 y para que sumen 1. Las pseudoprobabilidades se reescalan dividiéndolas por su suma. Por ejemplo, si un caso tiene pseudoprobabilidades pronosticadas de 0,50, 0,60 y 0,40 para una variable dependiente de tres categorías, cada pseudoprobabilidad se divide por la suma 1,50 para obtener 0,33, 0,40 y 0,27.

Si alguna de las pseudoprobabilidades es negativa, se suma el valor absoluto de la menor a todas las pseudoprobabilidades antes del cambio de escala anterior. Por ejemplo, si las pseudoprobabilidades son -0,30, 0,50 y 1,30, primero se añade 0,30 a cada valor para obtener 0,00, 0,80 y 1,60. A continuación, se divide cada nuevo valor por la suma 2,40 para obtener 0,00, 0,33 y 0,67.

Para guardar las variables en el conjunto de datos activo

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Función de base radial...

2. En el cuadro de diálogo Función de base radial, pulse en la pestaña **Guardar**.

2.18 Exportar (Función de base radial)

La pestaña Exportar se utiliza para guardar las estimaciones de ponderaciones sinápticas para cada variable dependiente en un archivo XML (PMML). Puede utilizar este archivo de modelo para aplicar la información del modelo a otros archivos de datos para puntuarlo. Para obtener más información, consulte el tema Asistente para puntuación. Esta opción no se encuentra disponible si se han definido archivos segmentados.

Para exportar las ponderaciones sinápticas

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Función de base radial...

2. En el cuadro de diálogo Función de base radial, pulse en la pestaña **Exportar**.

2.19 Opciones (Función de base radial)

2.19.1 Valores perdidos del usuario.

Los factores deben tener valores válidos para el caso para que se incluyan en el análisis. Estos controles permiten decidir si los valores perdidos del usuario se deben tratar como válidos entre los factores y las variables dependientes categóricas.

Para especificar las opciones

Esta característica requiere la opción Redes neuronales.

1. En los menús, elija:

Analizar > Redes neuronales > Función de base radial...

2. En el cuadro de diálogo Función de base radial, pulse en la pestaña **Opciones**.

Referencias

Bell, E. H. 1961. Social foundations of human behavior: Introduction to the study of sociology. Nueva York: Harper & Row.

Bishop, C. M. 1995. *Neural Networks for Pattern Recognition*, 3rd ed. Oxford: Oxford University Press.

Blake, C. L., y C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mlearn/MLRepository.html>.

Breiman, L., y J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580–598.

Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

Fine, T. L. 1999. *Feedforward Neural Network Methodology*, 3rd ed. Nueva York: Springer-Verlag.

Green, P. E., y V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

Green, P. E., y Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

Greenacre, M. J. 1984. *Theory and applications of correspondence analysis*. Londres: Academic Press.

Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, 469–506.

Hartigan, J. A. 1975. *Clustering algorithms*. Nueva York: John Wiley and Sons. Hastie, T., y R. Tibshirani. 1990. *Generalized additive models*. Londres: Chapman and Hall.

Haykin, S. 1998. *Neural Networks: A Comprehensive Foundation*, 2nd ed. Nueva York: Macmillan College Publishing.

Kennedy, R., C. Riquier, y B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, 56–70.

McCullagh, P., y J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. Londres: Chapman & Hall.

Menec, V., N. Roos, D. Nowicki, L. MacWilliam, G. Finlayson, y C. Black. 1999. *Seasonal Patterns of Winnipeg Hospital Use*. : Manitoba Centre for Health Policy. Price, R. H., y D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, 579–586.

Rickman, R., N. Mitchell, J. Dingman, y J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228, 54–58.

Ripley, B. D. 1996. *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press. 95 96 Bibliografía Rosenberg, S., y M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489–502.