

Las Redes Neuronales y la Gerencia de Riesgos

FERNANDO RICOTE GIL

UNIVERSIDAD COMPLUTENSE DE MADRID

La neurocomputación es la ciencia relacionada con los sistemas de procesamiento de la información que desarrolla capacidades operacionales como respuesta adaptativa a un ambiente de información, utilizando soluciones basadas en el funcionamiento del cerebro.

Esta ciencia resuelve problemas relacionados fundamentalmente con:

- El procesamiento por sensores
- El reconocimiento de patrones
- El análisis de datos
- El control

Se aplica fundamentalmente a:

- El reconocimiento de caracteres y de imágenes
- La optimización
- La predicción
- El pilotaje o conducción automática

Constituye el objeto del presente artículo el estudio de la aplicación de la ciencia neurocomputacional a la gerencia de los riesgos. Estos sistemas han recibido el nombre de Sistemas Neuronales Artificiales o simplemente Redes Neuronales (RN).

medios artificiales. La fascinación que la inteligencia como materia de estudio ha suscitado al género humano, puede verse reflejada en la aparición de una línea del estudio científico llamada «Inteligencia Artificial» que se corresponde con el estudio de la inteligencia.

Una de las múltiples líneas por las cuales se ha desarrollado esta investigación es el desarrollo de las llamadas «redes neuronales» encuadrada en una parte de la neurociencia que se denomina ciencia neuro-computacional.

Objeto de la Neurociencia computacional

El estudio de:

- Moléculas
- «Synapses»
- Neuronas
- Redes neuronales
- Proceso de la información

INTRODUCCIÓN

La Neurociencia es la ciencia que se ocupa del sistema nervioso o de cada uno de sus diversos aspectos y funciones especializadas.

El ser humano muestra un deseo profundo de poder reproducir la habilidad cognoscitiva por

Una red neuronal es el intento de poder realizar una simulación computacional del comportamiento de partes del cerebro humano mediante la réplica en pequeña escala de los patrones que éste desempeña para la formación de resultados a partir de los sucesos percibidos.

Concretamente, se trata de poder analizar y reproducir el mecanismo de aprendizaje y recono-

Cuadro 1. Breve historia

1943	Artículo Mc Culloch y Pitts donde RN simples podían calcular cualquier función aritmética o lógica Hubo Artículos y libros apoyando el tema
1948	Wiwner con su libro «Cibernetics»
1949	Donald Hebb con su libro «The organization of Behavior». El primero en aplicarlo a las sinapsis de las neuronas
1952	Ahsby con su libro «Design for a Braian: The Origin of Adpatative Behavior»
1957	Mark I Perceptron desarrollada por Rosenblatt Considerado el fundador de las RN Trabajó en reconocimiento de patrones Bernard Widrow. Creador del ADALINE. Regla Delta
1959	El libro de Nilson «Learning Machines» resumía el trabajo realizado en el período
1965	Minsky y Papert desacreditaron las RN y dividieron a los investigadores. Surge la IA
1972	Kokonen y Anderson con su el desarrollo de Associative memory neural net
hasta 1980	Pocos investigadores trabajan bajo el nombre de Procesamiento de señales adaptativo, reconocimiento de patrones y modelización biológica Ej: Grossberg y Kohonen
En los 80	Comienzos de los '80 Proyecto DARPA (Defense Adavanced Research Projects Agency) Parker (1985), Lecun (1986), Rumelhart, Hinton, and Williams (1986), McClelland and Rumelhart (1988): back-propagation algorithm
1983	Hopfield
1987	IEEE International conference on Neural Networks
1988	Journal de la INNS (International Neural Networks Society)
1989	LeCun: Learnig to recogniza handwriten characters Lang: Learnig to recognize spoken workds Cottrel: Face recogniti3n
1990-2006	IEEE Transaction on Neural Networks

ciendo de sucesos que poseen los seres más evolucionados. Se trata de poder analizar y reproducir el mecanismo de aprendizaje y reconociendo que posee el cerebro humano a la gestión del riesgo.

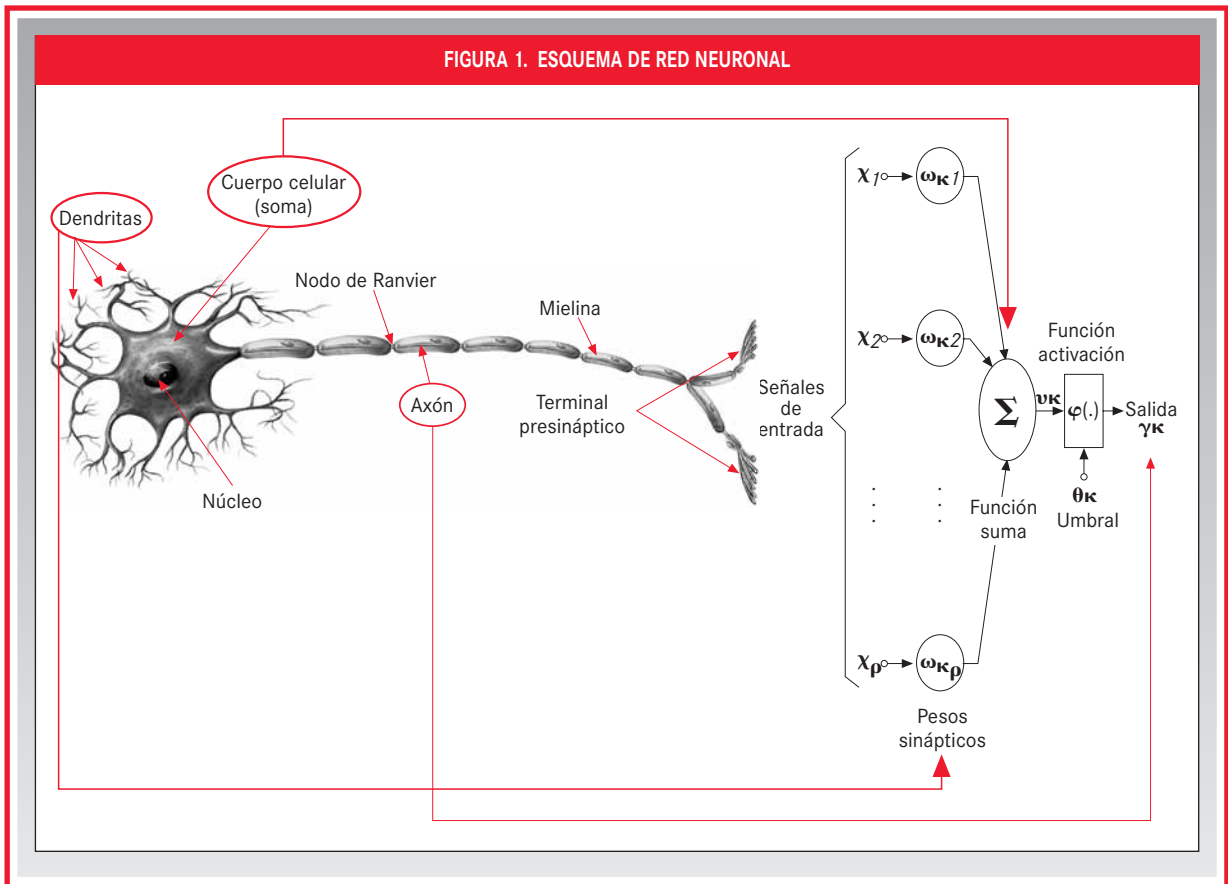
CONCEPTO DE RED NEURONAL

Las redes neuronales artificiales (RN) son sistemas paralelos para el procesamiento de la infor-

mación, inspirados en el modo en el que las redes de neuronas biológicas del cerebro procesan información.

Se ha podido comprobar que el cerebro humano es superior a los ordenadores en muchas tareas. Por ejemplo, en el procesamiento de información visual: un niño de 1 año reconoce objetos, caras,... mejor y más rápidamente que el mejor sistema de Inteligencia Artificial diseñado para dicha tarea, incluso ejecutándose en un avanzado ordenador. Es en las tareas basadas principalmente en aritmética sencilla, donde los ordenadores sobrepasan el cerebro humano.

FIGURA 1. ESQUEMA DE RED NEURONAL



Fuente: Mc) 2002 SNU CSE Biointelligence Lab and Center for Bioinformation Technology (CBIT).

Las RN deben tener así unas características similares a las del cerebro:

- Robustas y tolerantes a fallos. En el cerebro mueren todos los días gran cantidad de neuronas sin afectar sensiblemente a su funcionamiento.
- Flexibles. El cerebro se adapta a nuevas circunstancias mediante el aprendizaje.
- Podrán trabajar con información borrosa, incompleta, probabilística, con ruido o inconsistente.
- Serán altamente paralelas. El cerebro está formado por muchas neuronas interconectadas entre sí y es precisamente el comportamiento colectivo de todas ellas lo que caracteriza su forma de procesar la información.

Las aplicaciones principales estarán centradas en campos donde la inteligencia humana no pueda ser emulada de forma satisfactoria por algoritmos aritméticos que pueden ser implementados en ordenadores. El punto clave de las RN es la nueva estructura de estos sistemas para el procesamiento de la información.

Las RN están compuestas, al igual que el cerebro, por un número muy elevado de elementos básicos (las neuronas), altamente interconectadas entre ellos y con modelo de respuesta para cada elemento en función de su entorno muy parecido al comportamiento de las neuronas biológicas.

Estos modelos son simulados en ordenadores y es el comportamiento colectivo de todos los elementos lo que le confiere esas características tan

ESTUDIO

peculiares para la resolución de problemas complejos. Las RN, como las personas, aprenden a partir de ejemplos.

Aprender en sistemas biológicos involucra la modificación de la ínter conectividad entre las neuronas y esto es también ocurre con las RN.

El aprendizaje en las RN se efectúa mediante la asignación de ponderaciones a los nodos.

$$Y = w_0 + w_1x_1 + \dots + w_nx_n$$

$$Y_i = w_{i0} + \sum_{j=1}^n w_{ij}x_j$$

Con las funciones suma y transferencia se obtienen las salidas correspondientes de cada neurona

Función Suma Σ	Función de transferencia $f(x)$
--------------------------	------------------------------------

$$f(w_{ij}x_j) = Y_i$$

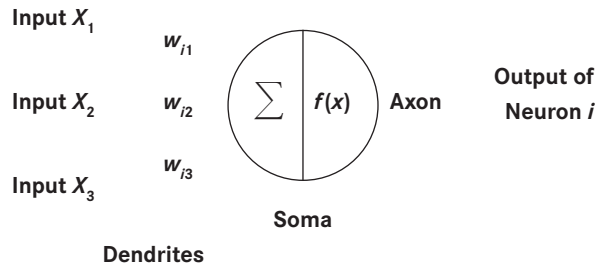
Siendo:

Y_i : Salida del neuron i

$f(x)$: Función de transferencia

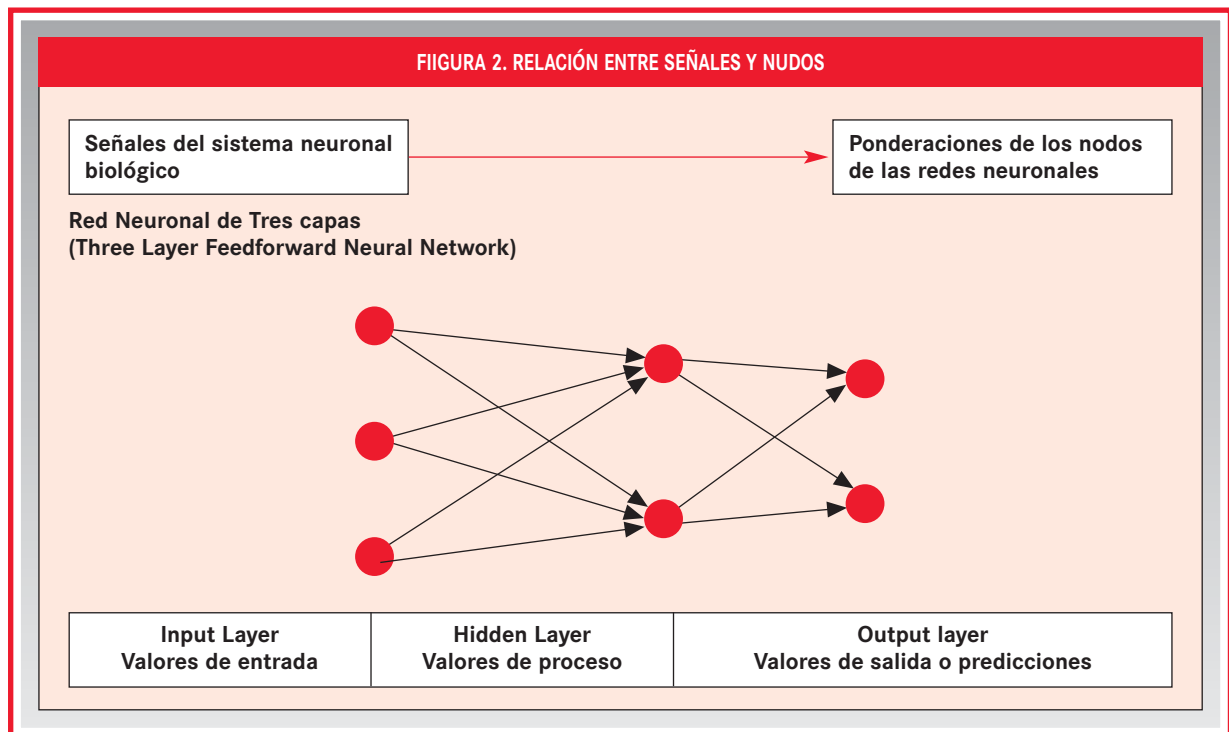
w_{ij} : Ponderación del input j del neuron i

x_j : El valor de entrada de j



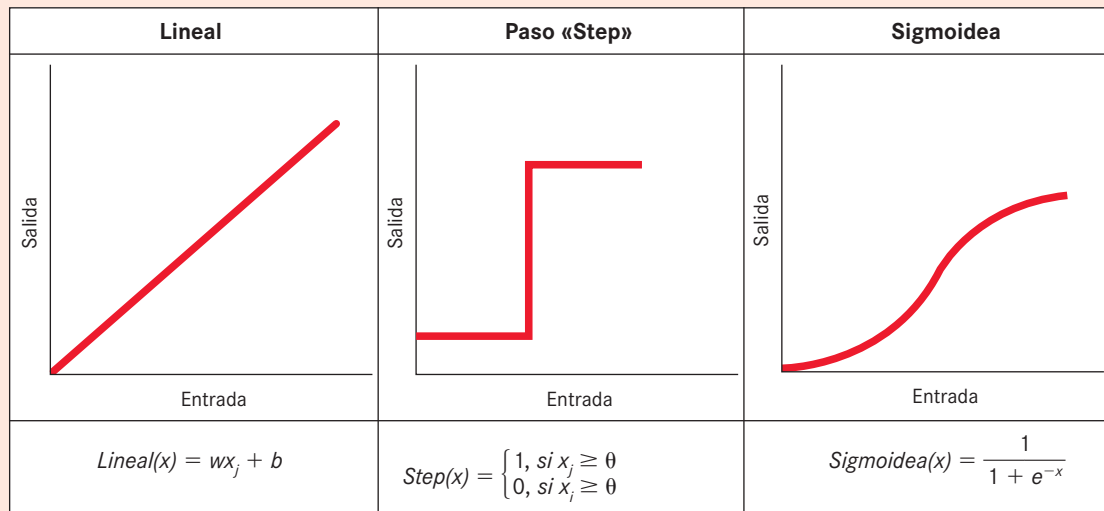
Tipos de función de transferencia o activación

La función de activación que se aplica a la suma de ponderaciones puede ser de diferentes tipos aunque suelen utilizarse fundamentalmente funciones de tipo lineal, paso o sigmoideas, como se refleja en la figura 2.



Fuente: Neural Networks Demystified by Louise Francis.

FIGURA 3. TIPOS DE FUNCIÓN DE TRANSFERENCIA



La función de activación que se aplica a la suma de ponderaciones es típicamente una función sigmoidea.

La más común de las funciones es la función logística:

$$f(Y) = h_i = \frac{1}{1 + e^{-Y}}$$

Otra función sigmoidea utilizada es la función tangente hiperbólica (transformada no lineal):

$$f(Y) = h_i = \tanh(v_i) = \frac{e^{-Y} - e^Y}{e^{-Y} + e^Y}$$

PROCESO DE APRENDIZAJE DE LAS REDES NEURONALES

El proceso de estimar las mejores ponderaciones para la red neuronal es el llamado entrenamiento o aprendizaje.

El acercamiento usado por el software comercial para estimar las ponderaciones o pesos es el «backpropagation».

En cada ciclo de red con los datos de entrenamiento, se produce un valor predicho para la variable designada. Este valor se compara con el valor real para la variable designada y se calcula el error para cada observación.

Los errores son realimentados a través de la red y se calculan las nuevas ponderaciones para reducir el error global.

El proceso de entrenamiento es realmente un procedimiento de optimización estadístico.

Estudio de dependencias lineales

Los valores de las unidades internas (neuronas) son linealmente combinados para obtener una predicción por la red neuronal

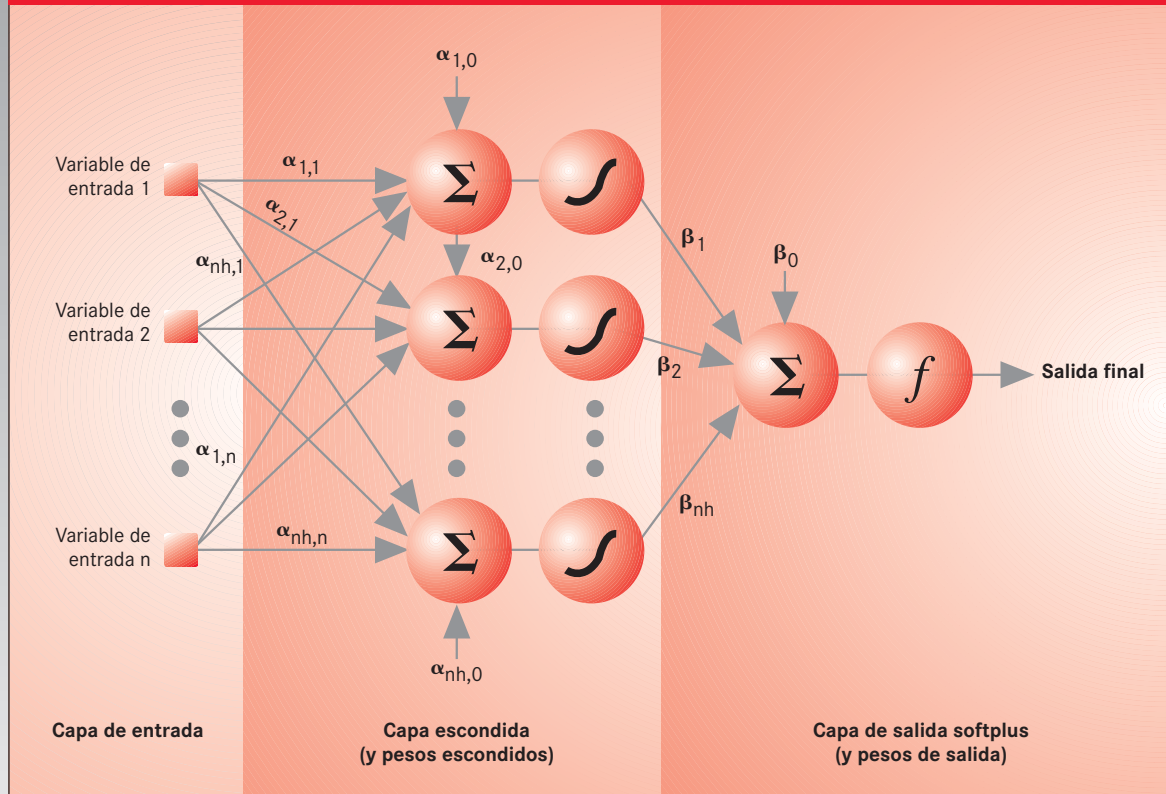
$$p(\vec{x}) = \beta_0 + \sum_{i=1}^{n_h} \beta_i x_i$$

$p(\vec{x})$ Salida de la red neuronal

n_h número de unidades internas de la red

β_0, β_i coeficientes de la combinación lineal

FIGURA 4. ESQUEMA DE PROCESO DE RELACIÓN



$$p(\vec{x}) = \beta_0 + \sum_{i=1}^{n_h} \beta_i \tanh(\alpha_{i0} + \sum_{j=1}^n \alpha_{ij}x_j)$$

El error de predicción vendrá dado por la expresión:

$$e = p(\vec{x}) - y$$

El procedimiento minimiza la suma de los residuos cuadrados:

$$\begin{aligned} \text{Min}(\sum Y - \hat{Y})^2 \\ q = e^2 = (p(\vec{x}) - y)^2 \end{aligned}$$

Elección del número de unidades internas «hidden units»

Es preciso tener en cuenta que:

- A mayor número de unidades el error de proceso disminuye

- Existe un punto a partir de que el incremento del número de unidades incrementa el error del proceso.

Estudio de Dependencias no lineales

La función de activación que se aplica a la suma de ponderaciones:

$$h = \tanh(\alpha_0 + \alpha_1x_1 + \alpha_2x_2)$$

Esta función se puede desarrollar por Taylor, de la forma siguiente:

$$\begin{aligned} \tanh(\alpha_0 + \alpha_1x_1 + \alpha_2x_2) = \\ = \beta + (1 - \beta^2)(\alpha_1x_1 + \alpha_2x_2) + \\ + (-\beta + \beta^3)(\alpha_1x_1 + \alpha_2x_2)^2 + \end{aligned}$$

$$+ \left(\frac{2\beta}{3} + \frac{5\beta^3}{3} + \beta^5 \right) (\alpha_1 x_1 + \alpha_2 x_2)^4 + O(\alpha_1 x_1 + \alpha_2 x_2)^5$$

VENTAJAS E INCONVENIENTES DE LAS REDES NEURONALES

Debido a su constitución y a sus fundamentos, las redes neuronales artificiales presentan un gran número de características semejantes a las del cerebro.

La ventaja más importante de las redes neuronales es que pueden utilizarse para la solución de problemas que son demasiado complejos para las técnicas convencionales: problemas que no tienen un algoritmo específico para su solución, o cuyo algoritmo es demasiado complejo para ser encontrado. Por lo tanto, son de especial utilidad cuando existen grandes cantidades de datos y mucha incertidumbre en cuanto a la manera de como estos son producidos.

Entre las ventajas se incluyen:

- Aprendizaje Adaptativo. Capacidad de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial.
- Tolerancia a fallos. La destrucción parcial de una red conduce a una degradación de su estructura; sin embargo, algunas capacidades de la red se pueden retener, incluso sufriendo un gran daño.
- Operación en tiempo real. Los cálculos neuronales pueden ser realizados en paralelo; para esto se diseñan y fabrican máquinas con hardware especial para obtener esta capacidad.

- Fácil inserción dentro de la tecnología existente. Se pueden obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilitará la integración modular en los sistemas existentes.

En el Cuadro 2 se expone un resumen de las ventajas e inconvenientes de los modelos neuronales.

LAS APLICACIONES EMPRESARIALES DE LAS REDES NEURONALES

Es un sistema compuesto por un gran número de elementos básicos, agrupados en capas y que se encuentran altamente interconectados; Esta estructura posee varias entradas y salidas, las cuales serán entrenadas para reaccionar (valores OUTPUT), de una manera deseada, a los estímulos de entrada (valores INPUT).

Las redes neuronales son de utilidad para sistemas que son sólo parcialmente conocidos, que producen una cantidad inmensa de datos; estos datos con frecuencia contienen valiosa información.

Las soluciones de software de predicción están basadas en una combinación de redes reglas bases y nerviosas.

Si se construyen las redes neuronales de los datos históricos; la red es entonces «especializada» para que reconozca qué entradas tienen un efecto significativo en la predicción.

La red nerviosa descubre automáticamente no sólo las relaciones lineales aditivas entre los datos sino también las relaciones no lineales.

De los datos históricos, la red aprende a combinar las variables independientes para producir el resultado deseado. La red empieza entrenando haciendo las predicciones con los pesos al azar

ESTUDIO

Cuadro 2. Ventajas e Inconvenientes

Ventajas	Inconvenientes
Universalidad de la función de aproximación. Capacidad de representar cualquier dependencia funcional	Es preciso conocer bien el problema que se desea modelar
Trabajan con gran cantidad de variables y datos	El efecto caja negra «black box» Los datos entran en la caja negra y se obtienen las predicciones, pero no se revela normalmente la naturaleza de las relaciones entre las variables independientes y dependientes. En algunos casos las redes neuronales no dan explicaciones, al igual que en otros planteamientos tradicionales
Disponen de gran poder de predicción	Los procesos son muy largos
Autoorganización. Una red neuronal puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje	
Son explicativos	Requieren la definición de muchos parámetros antes de poder aplicar la metodología
Pueden ser no lineales y por lo tanto son capaces de representar mejor cualquier comportamiento que los modelos lineales	
Aprendizaje adaptativo: disponen de capacidad de aprendizaje (automática), que elimina la necesidad de adaptación de los sistemas expertos	
Las entradas de datos se incorporan directamente al software de predicción, sin interpretación o modificación. (no interviene un experto que realiza previamente un modelo mental)	
Resistencia a una cierta falta de fiabilidad de los datos	
Posibilidades de la variable de predicción, clasificación	
Menos trabajo personal que en la aplicación de estadística clásica	
Aplicación a la segmentación, determinando por si mismas cuantos cluster encierran cada clase	
Comportamiento aceptable cuando no se dispone de una fiable base de datos. Posibilidad de tratar de completar los datos mediante redes	
Más fácil que al análisis multivariante	
Facilidad de análisis y modelización de las relaciones espaciales, geográficas, etc.	
El modelo obtenido es continuo y derivable por lo que se presta bien a la visualización, con indudables aplicaciones prácticas	
Se puede demostrar que la regresión lineal y logística son casos particulares de las redes neuronales	

ESTUDIO

Gestión de riesgo en la empresa	Evaluación de campañas publicitarias
Modelos para análisis de riesgos	Análisis de la fidelidad de clientes (churning)
Predicción: De mercados (bolsa, acciones, fondos de inversión, bonos, renta fija y futuros) y de insolvencias	Detección y prevención de fraude
Optimización de procesos	Análisis de la demanda de servicios
Control de procesos	Tarificación y Solvencia

ajustados. Compara sus predicciones entonces con los resultados conocidos y ajusta cada ponderación para que cause menos error. Después de repasar miles de ejemplos en innumerables situaciones de tiempo, la red aprende los modelos y tendencias que le permiten que haga las predicciones exactas.

En general se trata de la sustitución de la estadística clásica y de modelos lineales por las RN en diversos procesos empresariales.

Las RN pueden ser combinadas con otras herramientas de la Inteligencia Artificial (IA), tal como la lógica difusa (lógica fuzzy), los algoritmos genéticos y los sistemas expertos.

Cuadro 3. Resumen de aplicaciones empresariales de las redes neuronales

Seguros	Riesgo operacional Credibilidad Análisis de supervivencia	Banca	Credit scoring: Riesgo en función de un historial Predicción del comportamiento de nuevos clientes Selección de buenas clases de riesgo Riesgo operacional
Ciencia e Ingeniería	Análisis de datos y clasificación Gestión de costes en proyectos de construcción Ingeniería Química Ingeniería Eléctrica Climatología Conducción automática de vehículos Reacción química a productos Sistemas dinámicos Control Interpretación de señales	Medicina y salud	Ayuda al diagnóstico Diagnóstico clínico Análisis de Imágenes e interpretación Análisis de señales e interpretación Desarrollo de medicamentos Desarrollo de drogas (anticancerígenos y otras) Distribución de recursos
Finanzas	Predicciones Medida de la volatilidad Precio de opciones Gestión de acciones Formación y Gestión de carteras Clasificación-Agencias de Rating Predicción de riesgos Gestión de Fondos de Pensiones y Fondos de Inversión	Marketing	Análisis sociodemográficos de la clientela Ciblage de mailing y ciblage de marketing Campanas de venta Predicción de fidelización Venta cruzada Minería de datos Investigaciones de mercado con campos en blanco

Cuadro 3. Resumen de aplicaciones empresariales de las redes neuronales (continuación)

Administración Pública	Predicción de la criminalidad Análisis de las informaciones fiscales. Detección del fraude fiscal y de otro tipo. Predicciones económicas Proyección de prestaciones de los sistemas de salud Proyección de prestaciones sociales Análisis de los riesgos Predecir la demanda de servicios públicos Gestión de recursos de agua	Transportes y Comunicaciones	Optimización de rutas. Optimización en la distribución de recursos Sistema Inteligente de gestión de stocks Predicción de stocks
Energía	Predicción consumo eléctrico Distribución recursos hidráulicos para la producción eléctrica Predicción de consumo de gas, etc.	Alimentación	Análisis de olor y aroma Perfiles de clientes en función de su compra Desarrollo de productos Control de Calidad
Meteorología	Predicción del tiempo Modelos de predicción	Tratamiento de textos y proceso de formas	Reconocimiento de caracteres impresos mecánicamente Reconocimiento de gráficos Verificación de firmas Reconocimiento de caracteres escritos a mano Reconocimiento de escritura manual cursiva OCR
Industria manufacturera	Predicción de la demanda de un producto. Planificación de la producción Planificar los avisos y alarmas de las máquinas y reactores químicos Control de procesos Control de calidad Control de robots	Industria de servicios	Predicción de la demanda de un servicio

LAS APLICACIONES RELACIONADAS CON EL RIESGO Y LA SEGURIDAD

La aplicación general de las RN es la gestión eficaz y la rentabilidad de la dirección de riesgo.

Con procesos informáticos de gestión de RN, muy elaborados y después de largos procesos, como ocurre en el cerebro humano, se puede predecir el resultado de una situación, más rápidamente y con más precisión que con cualquier otro método actualmente en el uso, extrayendo el valor de una fuente de datos que es por otra parte demasiado pesada para ser de utilidad a través de otras metodologías.

Estos sistemas ofrecen una alternativa automatizada, más rápida y más flexible a los tratamientos tradicionales para analizar la información y

ESTUDIO

por tanto son de gran utilidad en la seguridad, por fiabilidad y rapidez.

Se detallan algunas aplicaciones particulares a la seguridad y gerencia de riesgos:

Cuadro 4. Aplicación a la seguridad y la gerencia de riesgos

Análisis de datos químicos	Predicción de insolvencias
Análisis de la demanda	Investigaciones de mercado con campos en blanco
Análisis espectral: permiten representar sistemas lineales y señales de aproximación de fenómenos físicos	Lectores de huellas digitales para control de acceso
Control de clientes	Oído electrónico
Control de combustión	Olfato electrónico (Descubridores de gases y contaminantes)
Control de emisiones para la gestión de riesgos medioambientales	Predicción de cumplimiento de contratos en tiempo
Control del riesgo operacional	Predicción del tiempo
Detección de fallos en general	Previsión de demanda de productos o servicios
Detección del fraude en la utilización de servicios, tarjetas, prestaciones, etc.	Previsión de necesidades energéticas y recursos
Detección y prevención de fallos en comunicaciones	Prueba de drogas
Detectores inteligentes de riesgos (fuego, emisiones, etc.)	Rating de clientes
Diagnóstico (técnico, médico u otros)	Retina artificial
Dirección del riesgo financiero	Sistemas de diagnóstico
Dirección eficaz de procesos	Sistemas de remuneración a empleados
Eficacia del negocio	Sonar. Para control de aproximación de objetos y reconocimiento
Eliminación de pruebas en humanos o animales, que pueden reemplazarse en algunos casos por modelos neuronales	Supervisión medioambiental
Evaluación de los procesos de cada prestación o servicio	Suscripción y tarificación de riesgos
Gestión de las prestaciones	Valoración de riesgos para la toma de decisiones de transferencia o retención
Gestión de los contratos de seguro	Verificación de firmas
Gusto electrónico (Comprobación farmacéutica y calidad de los alimentos)	Zonificación de los riesgos

LA PRÓXIMA GENERACIÓN DE REDES

El análisis de vectores de contexto *Context Vector Analysis* es considerado la próxima generación en las soluciones de software de predicción.

Permiten la interpretación de textos y notas como pueden ser un e-mail, una nota de reparación, una prescripción, etc, con un tratamiento automático de dichas fuentes de información.

Estas representaciones matemáticas, llamadas «vectores», pueden asociarse a bibliotecas de información matemáticamente representada, como pueden ser directivas de gestión de riesgos o cuidados, información de servicio de clientes, nuevas ofertas de productos, etc.

Los vectores del contexto pueden aplicarse a:

Automatización de comunicaciones
Interpretación de correspondencia e informes
Identificación automática de excepciones de tratamiento de riesgos
Diseño de nuevos productos y servicios, para los clientes, basadas en sus necesidades de información y modelos de la compra

Además de las redes Fuzzy, están apareciendo nuevas redes como:

- Complex neuron
- Mixture of expert
- ...

CONCLUSIÓN

En resumen, las soluciones de software de RN, software de predicción en combinación con otras técnicas permitirán a las empresas mejorar la gestión de riesgos, los procesos y su rentabilidad, efectuando una mejor prestación de servicios y productos adaptados.

La mejora creciente de la rentabilidad se obtiene a través de la reducción de riesgo, eficacia de los procesos, la capacidad de gestión y la mejora de los productos con diseños más eficaces.

BIBLIOGRAFÍA

- LAU. L.: «Neural Networks, Theoretical Foundations and Analysis», 1991, IEEE Press.
- KOSKO, B.: *Neural Networks and Fuzzy Systems: A Dynamical Approach to Machine Intelligence*, Prentice Hall, Englewood Cliffs, 1992.
- KUNG, S. Y.: «Digital Neural Networks», by PTR Prentice Hall, Inc. 1993.
- FAUSETT. L.: *Fundamentals of Neural Networks*, Prentice-Hall, 1994.
- GURNEY. K.: *An Introduction to Neural Networks*, UCL Press, 1997.
- MUTCH, J.: «Technology: Unlocking the Neural Network», January 1999.
- HAYKIN, S.: *Neural Networks*, 2nd Edition, Prentice Hall, 1999.
- Analytic Techniques for Intelligence & Security. A Fair Isaac. White Paper. May 2003.