

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO  
FACULTAD DE INGENIERÍA  
ESCUELA DE INGENIERÍA INFORMÁTICA

ESTUDIO DE APLICACIÓN DE REDES NEURONALES EN LA  
EVALUACIÓN DE RIESGO CREDITICIO

VALERIA ELIZABTEH MÉNDEZ ARAYA

INFORME FINAL DEL PROYECTO  
PARA OPTAR AL TÍTULO PROFESIONAL DE  
INGENIERO CIVIL EN INFORMÁTICA

DICIEMBRE - 2009

## **Dedicatoria y Agradecimientos**

Dedicado con mucho amor....

A Dios por la existencia.

A mi madre: base, sustento y fuerza motivadora en mi vida día a día. Gracias por el infinito amor y por hacer todo lo posible para que mis metas se fuesen logrando. Porque sin ti, nada de esto hubiese sido posible. Gracias por creer en mí.

A mis familiares y a todos aquellos que aportaron de una u otra forma para que mi caminar fuese más fácil.

No puede dejar de nombrar, a mi gran amiga, hermana elegida, apoyo y compañía en mi periodo universitario: Ina. Gracias por darme la oportunidad de conocer a una persona tan extraordinaria como tú; por hacerme participe de tu vida. Sabes que siempre estaré ahí para ti, así como tú lo estas para mí.

A mis profesores, por confiar en mí y tener la paciencia que requerí. Además agradezco su apoyo en momentos difíciles, por hacer la relación más cálida y humana. En especial a la profesora Silvana, quién con su calidez y su ternura hizo que esta etapa fuese más sencilla.

A la vida... por las experiencias y la madurez conseguida.

Valeria Méndez A.

## Tabla de Contenidos

1. Introducción.....	1
2. Análisis de Objetivos y Metodologías.....	3
2.1 Objetivo General.....	3
2.2 Objetivos Específicos.....	3
3. Análisis de problemática.....	4
3.1 Análisis de Crédito.....	4
3.1.1 Historia.....	4
3.1.2 Introducción al análisis de riesgo.....	5
3.1.3 Consideraciones.....	8
3.1.4 Riesgo Crediticio en Personas.....	9
3.2 Redes neuronales artificiales (RNA). Concepto y evolución histórica.....	9
3.2.1 El modelo biológico.....	10
3.2.2 Estructura de un sistema neuronal artificial.....	11
3.2.3 El modelo artificial .....	12
3.2.2.1 Entradas y salidas.....	12
3.2.2.2 Pesos sinápticos.....	13
3.2.2.3 Función de salida.....	13
3.3 Procesamiento en neurona Artificial.....	13
3.3.1 Ventajas de las Redes Neuronales.....	24
3.3.2 Modelos Neuronales.....	26
3.3.3 Fundamentos Biológicos de las Redes Neuronales.....	27
3.3.4 Modelos Computacionales.....	28
3.3.5 La Neurona Artificial.....	28
3.3.6 Estructura básica de la red.....	30
3.3.7 Aprendizaje.....	33
3.3.8 Funciones de Transferencia.....	36
3.3.9 Modelos de Redes Neuronales.....	39
3.3.9.1 Perceptron.....	39
3.3.9.2 Arquitectura del Perceptron.....	39
3.3.9.4 Aprendizaje del Perceptron.....	39
3.4 Generalización.....	50
4. Área Desarrollo.....	43
4.1 Codificación de datos.....	43
4.2 Acerca del Software utilizado.....	47
4.2.1 Proceso de Aplicación del modelo sobre los datos.....	48

4.2.2 Creación de la Red.....	48
4.2.3 Importación de los datos.....	51
4.2.4 Entrenamiento de la red.....	52
4.3 Resultados obtenidos.....	53
4.3.1 Análisis de Sensibilidad.....	54
5. Conclusiones.....	59
6. Referencias.....	59

## **Anexos**

Anexo A: Datos aplicación caso estudio.

Anexo B: Resultados obtenidos luego fase de entrenamiento.

Anexo C: Aplicación Ejemplo.

## Índice Figuras

Figura 1 Modelo biológico de Red Neuronal.....	10
Figura 2 Modelo de Neurona Artificial.....	12
Figura 3 Neurona de McCulloch-Pitts.....	16
Figura 4 El Perceptrón de Rosenblatt.....	17
Figura 5 Perceptrones solucionando la función OR y la función XOR.....	18
Figura 6 Arquitectura de un Perceptrón multicapa .....	21
Figura 7 Esquema de un Modelo Neuronal.....	26
Figura 8 Modelo de Neurona Artificial.....	29
Figura 9 Unidad Típica de proceso de Red Neuronal.....	30
Figura 10 Equivalencia entre redes al utilizar funciones de activación lineales.....	31
Figura 11 Ejemplo de Red Neuronal con Capas Ocultas.....	32
Figura 12 Aprendizaje Supervisado.....	34
Figura 13 Aprendizaje No Supervisado.....	35
Figura 14 Arquitectura Red Perceptrón.....	40
Figura 16 Generalización. Situación ideal.....	42
Figura 17 Generalización. Situación real.....	42
Figura 18: Estructura de la red.....	47
Figura 19: Diseño Capa de entrada.....	48
Figura 20: Diseño Capa Intermedia .....	50
Figura 21: Diseño Capa de Salida.....	51
Figura 22: Caso de entrenamiento.....	52
Figura 23: Muestra error criterio de convergencia.....	53

## **Tablas**

Tabla 1: Factores de Riesgo Crediticio Individual .....	9
Tabla 2: Clasificación de Redes según Aprendizaje.....	24
Tabla 3: Relaciones de Transferencia.....	38
Tabla 4: Características demográficas de los sujetos consumidores y no consumidores de éxtasis.....	55

## 1. Introducción

La mente humana surge como modelo para máquinas inteligentes ya que en principio es una obvia idea el imitar su comportamiento. Una simulación en computador del funcionamiento del cerebro ha sido un objeto de estudio de la IA (Inteligencia Artificial) desde los años cuarenta del siglo pasado. Las poderosas cualidades de la mente en lo que respecta a pensamiento, recordación y solución de problemas, ha inspirado a los científicos el intentar el modelamiento computarizado de su operación.

Actualmente esta ciencia está comprendida por varios subcampos que van desde áreas de propósito general, como el aprendizaje y la percepción, a otras más específicas como la demostración de teoremas matemáticos, robótica, tratamiento de lenguajes naturales y diagnóstico de enfermedades, entre muchas otras.

Uno de los modelos que ha surgido para emular el proceso de aprendizaje es la red neuronal artificial. Las redes neuronales son modelos que intentan reproducir el comportamiento del cerebro. Una red neuronal consiste en un conjunto de elementos de procesamiento, llamados neuronas, los cuales se conectan entre sí. La organización y disposición de las neuronas dentro de una red neuronal se denomina topología, y viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad, y el tipo de conexión entre neuronas. Una vez determinada la topología de la red neuronal es necesario entrenarla. En la etapa de entrenamiento la red es capaz de aprender relaciones complejas entre entradas y salidas mediante el ajuste de los pesos de las conexiones entre neuronas. Las redes neuronales artificiales han sido aplicadas con éxito en gran cantidad de problemas como por ejemplo reconocimiento de patrones, clasificación, visión, control y predicción.

Las decisiones estratégicas son las que definen el futuro de las empresas y organizaciones. Típicamente estas decisiones requieren comprometer cantidades considerables de recursos y los resultados a obtener son inciertos y se conocerán años después de que se tomó la decisión. Actualmente, este tipo de decisiones pasan a través de un largo proceso de análisis, tanto del macro como del micro entorno en el que se encuentra inserta la organización, con el fin de poder aumentar de manera considerable las posibilidades de obtener resultados exitosos.

Hoy, más que nunca, es necesario poder controlar aquellas variables económicas que permitan decidir si la entidad que es objeto de análisis tiene o no las condiciones adecuadas para ser responsable de sus obligaciones adquiridas. Actualmente las entidades bancarias que se basan en redes neuronales para éstos análisis no divulgadas abiertamente dicha información y no contribuyen a la investigación. Esto se debe a que se debería tener acceso a información confidencial la cual no están dispuestos a entregar, de no ser estrictamente necesario.

Basado en esto, es que se propone como tema de proyecto de Titulación el Estudio de Aplicación de Redes Neuronales en la Evaluación de Riesgo Crediticio.

Buena parte de la información de crédito se obtiene de fuentes externas. Todos los aspectos o vías de información deben ser analizados, sin subestimar nada, ni dejar pasar ninguna oportunidad que permita conocer algo más sobre el cliente. Es muy importante que se obtenga información con respecto a la compañía para la cual el cliente en cuestión trabaja. Esa combinación de habilidades: buena, mala o indiferente, da lugar al éxito o fracaso del negocio. Esta información no puede ser descartada, y tiene que considerarse como un factor clave de tremendo impacto en la decisión de crédito.

En la sección 2, se presentan los objetivos, tanto generales como específicos. En la sección 3 se presenta la base teórica de conocimientos Económicos de Riesgo Crediticio y la base de conocimientos de Redes Neuronales que se requiere para el desarrollo del tema. Además se presenta el modelo y resolución de una problemática abordada con redes neuronales. En la sección 4 se presenta un ejercicio de prueba realizado como entrenamiento en trabajo con Redes Neuronales. A continuación en la sección 5 se muestra la implementación del proyecto. En la sección 6 y final, se presentan las conclusiones.



## **2. Análisis de Objetivos y Metodología**

### **2.1 Objetivo General**

- Implementación de red neuronal aplicado a problemática de riesgo crediticio en personas.

### **2.2 Objetivos Específicos**

- Estudio preliminar de redes neuronales y riesgo crediticio.
- Implementación y análisis mediante herramienta Neuronal Network Matlab.
- Análisis de datos y resultados.

### **3. Análisis de la problemática**

#### **3.1. Análisis de Crédito**

En términos muy simples existe riesgo en cualquier situación en que no se sabe con exactitud lo que ocurrirá al futuro. También “riesgo” es sinónimo de incertidumbre, la dificultad de poder predecir lo que ocurrirá. En general es importante conocer al riesgo en el ámbito económico, ya que para decisiones financieras tales como otorgamiento de créditos, análisis de inversión, es la base para predecir el futuro y si no se da según se había previsto, seguramente se habrá tomado una mala decisión [1].

Hay quienes se rehúsan a aceptar riesgos y quienes no son tan reacios, en todos los casos la idea es asumir el menor riesgo posible en la medida de las posibilidades. Pero no siempre el riesgo es malo se puede convivir con el a través de un incentivo. Es decir se acepta más riesgo en la medida que haya recompensa. Es por ello que existe una relación muy estrecha entre riesgo y rentabilidad.

##### **3.1.1. Historia**

A lo largo de toda la evolución del riesgo crediticio, y desde sus inicios, el concepto de análisis y criterios utilizados han sido los siguientes: desde principios de 1930 la herramienta clave de análisis ha sido el balance. A principios de 1952, se cambiaron al análisis de los estados de resultados, lo que más importaban eran las utilidades de la empresa. Desde 1952 hasta nuestros tiempos el criterio utilizado ha sido el flujo de caja. Se otorga un crédito si un cliente genera suficiente caja para pagarlo, ya que los créditos no se pagan con utilidad, ni con inventarios ni menos con buenas intenciones, se pagan con caja [2].

El análisis de créditos se considera un arte ya que no hay esquemas rígidos y que por el contrario es dinámico y exige creatividad por parte del oficial de crédito o de negocios. Sin embargo es importante dominar las diferentes técnicas de análisis de créditos y complementarla con una buena cuota de experiencia y buen criterio, así mismo es necesario contar con la información necesaria y suficiente que permita minimizar el número de incógnitas para poder tomar la decisión correcta.

### 3.1.2 Introducción al análisis de riesgos

Un aspecto de extraordinaria importancia en la gestión de los riesgos crediticios, es el relativo al análisis y evaluación del riesgo, así como la clasificación de los clientes. Estos procesos de análisis de riesgos precisan de fuentes de información, tanto internas como externas, y de sistemas específicos.

La gestión de riesgos puede afrontarse de una forma masiva o adoptando criterios cualitativos. La gestión masiva es posible cuando existen miles de millones de clientes, encontrándose el cliente individualmente considerado falto o carente de excesivo valor, esto es, se acomete la gestión desde un proceso de insensibilidad y distanciamiento, de forma objetiva y aplicando la ley de los grandes números, el número está por encima del individuo [3]. Por el contrario, la gestión cualitativa, que tiene un carácter más personal e individualizado, tomando en consideración al individuo frente al número, es factible en empresas con menor número de clientes.

En la gestión de riesgos pueden adoptarse mecanismos preventivos o defensivos y mecanismos curativos. Dentro de los procesos preventivos, que tienen el objetivo de evitar la asunción de riesgos por encima de la política de la empresa, se encuadran el análisis previo del cliente, la vigilia de riesgos, los informes comerciales, los sistemas de análisis financiero, el control de la deuda del cliente, los sistemas de control de límites, el scoring (especialmente aplicable a sistemas masivos según reglas estadísticas), etc. En los mecanismos curativos, tendentes a asegurar la recuperación de la deuda en las condiciones pactadas con el cliente, pueden incluirse los sistemas de cobertura, las fianzas, avales y garantías, los seguros de crédito, los sistemas de aviso de vencimiento, la gestión proactiva de incidencias, los sistemas de recobro y las acciones jurídicas. De igual modo, el gestor de riesgos ha de estar atento a todas las fases del ciclo de la operación: vigilar la definición de límites de riesgo en el momento de las altas de clientes, controlar las condiciones de entrega y facturación del pedido, la gestión de comprobantes, las incidencias en la facturación y el control de las condiciones y formas de pago al vencimiento [2].

En una fase de análisis previo se debe medir y calificar el riesgo, esto es, analizar y valorar las contingencias, cuantificando cuál se va a asumir con el cliente y qué valoración tiene el mismo, asignándose límites de riesgos. Para ello se aplicarán sistemas de gestión y modelos de análisis de riesgos, que van alcanzando cada vez mayor grado de automatización. En este proceso de análisis de la solvencia, el “**credit manager**” debe estar en permanente contacto, no sólo con el departamento financiero, sino también con el departamento comercial, pues debe tenerse presente que una venta no se da por concluida hasta el momento de su cobro, lo que implica una coordinación entre ambos departamentos que procure acuerdos con los clientes, coberturas adecuadas, cumplimiento de los límites de riesgo asignados, autorizaciones de excedidos, etc.

La gestión del riesgo precisa también de información externa que se obtendrá tanto de los registros oficiales como de empresas especializadas y “**bureaus**” de crédito.

Para llegar a fijar un límite de riesgo, que no es más que el resultado final del proceso de análisis, han de tomarse en consideración aspectos tales como la clasificación del cliente, su implantación en el sector, su volumen de negocio, su relevancia comercial respecto a la empresa, la rentabilidad final que genera y su solvencia técnica, cifrada en su comportamiento histórico de pagos, en su comportamiento externo, y el resultado de un análisis económico-financiero de los ratios, el balance, la cuenta de resultados, y la información externa de empresas especializadas.

Una de las técnicas más utilizadas en la valoración del riesgo para asignación de límites es el scoring, basado en la aplicación de técnicas estadísticas de análisis multivariable, con el objetivo de determinar las leyes cuantitativas que rigen la vida económica de la empresa. De este modo se determina el comportamiento en función de las variables con mayor potencia predictiva dada su correlación con el resultado, y con el peso de cada una de las variables dentro de cada riesgo considerado, se asigna una puntuación.

También se utiliza el rating, aplicando técnicas estadísticas de análisis cuantitativo y la opinión de expertos, otorgándose la puntuación en función de las variables consideradas como relevantes en el análisis y evaluación del riesgo y ponderando cada una de ellas, para finalmente agrupar los riesgos puntuados en clases homogéneas, segmentando la población global en grupos de similar valoración [4].

Una vez debidamente valoradas y ponderadas estas variables según el modelo de análisis tomado, y efectuadas las oportunas correcciones en coordinación con el departamento comercial, se fijará el límite de riesgo del cliente que vendrá dado por la máxima pérdida o quebranto económico que puede ocasionar en la empresa.

Las nuevas tecnologías aplicadas a este campo facilitan enormemente la actividad del gestor, permitiendo la automatización de procesos repetitivos y posibilitando la asignación de límites de riesgo por cliente de una forma fiable.

En términos muy simples existe riesgo en cualquier situación en que no se sabe con exactitud lo que ocurrirá al futuro [4]. En otros lugares riesgo es sinónimo de incertidumbre, es la dificultad de poder predecir lo que sucederá. En general es importante conocer al riesgo en el ámbito crediticio, ya que la mayoría de las decisiones financieras de importancia son en base a predecir el futuro y no se da en base a lo que se había previsto.

Hay quienes se rehúsan a aceptar riesgos y quienes no son tan reacios, en todos los casos la idea es asumir el menor riesgo posible en la medida de las posibilidades.

Pero no siempre el riesgo es malo, se puede convivir con él a través de un incentivo. Es decir se acepta más Riesgo en la medida que haya recompensa.

Es por ello que existe una relación muy estrecha entre riesgo y rentabilidad.

Existen varios tipos de riesgos crediticios, dependiendo del contexto en que este se dé. En este caso en particular es de nuestro interés el análisis del riesgo individual de crédito.

Según muchos autores, entre los que se encuentran Ricardo Pascale, José Miguel Bulnes y Lawrence J. Gitman, se recomienda seguir el siguiente orden para análisis de créditos [4].

a) Generalidades

- Todo crédito debe pasar por una etapa de evaluación por más simple y rápida que ésta sea - Todo crédito tiene riesgo por fácil y bueno y bien garantizado que parezca - El análisis de crédito no pretende acabar con el 100% de la incertidumbre del futuro, sino que solo disminuya, por eso en materia de crédito no hay que pretender ser infalible, incluso quien lleva muchos años en análisis de créditos y nunca se ha equivocado, podría no ser un analista por su excesivo conservatismo, dado que en las decisiones de crédito no se está actuando con variables exactas, debe jugar un rol muy importante el buen criterio y sentido común.

b) Antecedentes generales anteriores a crédito: en este punto es necesario saber cual es el destino que se le dará al crédito y/o préstamo solicitado, por ello existen cuatro motivos por los cuales es necesario conocer el destino del crédito:

- Para comprobar la coherencia con las políticas de créditos de la institución.
- Para poder evaluar correctamente el crédito.
- Para poder fijar condiciones acordes con las necesidades.
- Para poder ejercer un control al deudor.

Por ello y para efectos de sugerencias es necesario saber que las causas más comunes que pueden producir una solicitud de crédito son:

- Incremento de las ventas permanentes (aumento de cuentas por cobrar, inventarios, etc.)
- Disminución velocidad de cobranza y/o índice de cobranza (aumento de cuentas por cobrar)
- Aumento de plazo de ciclo operacional
- Compra de materia prima (comercio exterior)
- Todo lo anterior a través de las líneas de créditos rotativas
- Reemplazo y/o aumento de activos fijos y/o mantenimiento y reparación
- Inversiones en otras empresas
- Problemas de caja por pagos inesperados
- Financiar gastos operacionales en negocios poco rentables

- Prepago para obtener mejores condiciones de créditos
- Pago de pasivos a otra institución financiera (compra de deuda)

c) Fuentes de pago y estructura: aquí surge la pregunta al igual que en el anterior ¿con qué recursos se pagará el crédito?, es decir cuál es la fuentes de donde se obtendrán los recursos para cancelar el crédito.

### **3.1.3 Consideraciones**

Se deben considerar las variables macroeconómicas que afectan a un país, tales como políticas de incentivo a importaciones o exportaciones, políticas tributarias, costo del dinero, movimiento de capital de entes capitalistas, política monetaria, precios internacionales, conflictos internacionales, inflación, crecimiento económico, mediterraneidad de un país, pobreza y subdesarrollo, dependencia de otros países, desarrollo social de un país, huelgas sindicales o problemas sociales, etc.

- Otras de las variables de mucha importancia es el análisis del sector de la empresa, dentro de las que se deben considerar variables como vulnerabilidad del sector, desarrollo, F.O.D.A., dependencia de otros sectores, estancamiento por diferentes razones, poco incentivo del gobierno, poco interés por parte de inversionista, fuerte inversión inicial, etc.

- Tipo de empresa solicitante del crédito, es decir si corresponde a una SRL (sociedad de responsabilidad limitada), empresas Unipersonales, S.A., Cooperativas, Asociaciones, empresas familiares, etc. Este dato es muy importante ya que permitirá conocer la calidad de la administración de la empresa y como operan, su organigrama, entre otros [2].

- Considerar el sector al que pertenece el solicitante, ya que esto permitirá saber cual es su forma de actuar y su ciclo operativo para tener una mejor herramienta y juicio de análisis - De preferencia se deben analizar balances de las tres últimas gestiones - Balance con antigüedad no mayor a 6 meses - Calificación de la auditoria, hay que tener en cuenta que no todos los auditores califican.

- Los comentarios deben ser de fondo y no de forma, deben permitir identificar las causas y dar respuestas sobre el rubro - Los comentarios del balance deben responder a los ¿por qué?

- Deben analizarse los balances consolidados en caso de grupos económicos manteniendo cuidado de consolidar cada una de las cuentas.

### 3.1.4 Riesgo Crediticio en Personas

Dependiendo del contexto en el que se centre el análisis, son los factores que inciden en la toma de decisiones. En este estudio en particular, se centra en el crédito en personas.

La tabla 1 presenta un esquema de los factores más importantes que se deben considerar al momento de analizar el riesgo individual [3]. Todos estos son factores que son esenciales y deben ser considerados en el momento de decidir si se otorga o no el crédito a una persona o entidad.

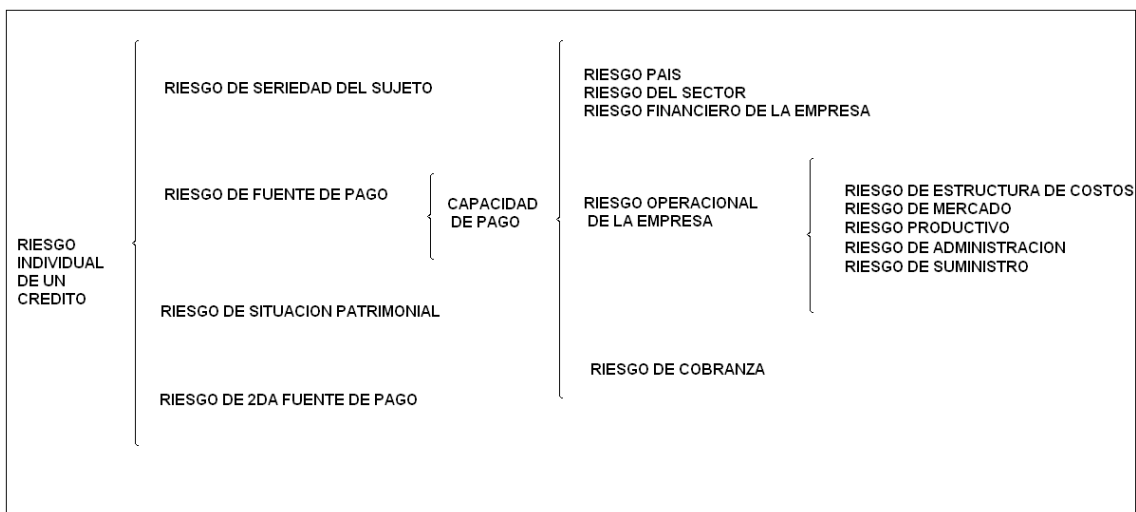


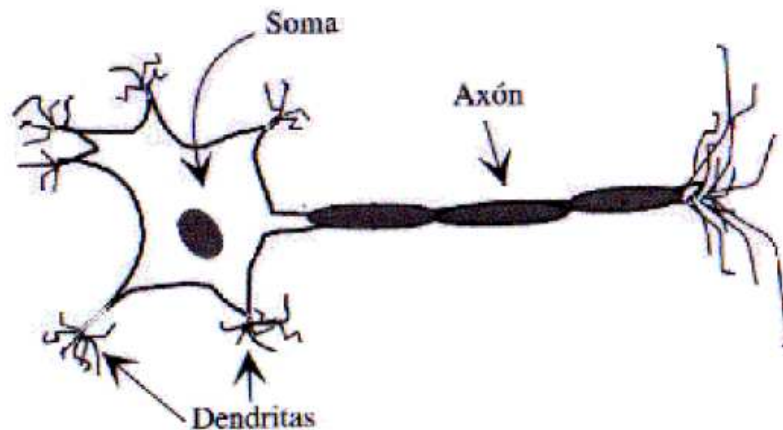
Tabla 1: Factores de Riesgo Crediticio Individual

### 3. 2 Redes neuronales artificiales (RNA): Concepto y evolución histórica

Las redes neuronales artificiales (RNA) son modelos matemáticos que intentan reproducir el funcionamiento del sistema nervioso. Como todo modelo, realizan una simplificación del sistema real que simulan y toman las características principales del mismo para la resolución de una tarea determinada. Estos sistemas conexionistas consisten en un conjunto de elementos simples de procesamiento llamados nodos o neuronas conectadas entre sí por conexiones que tienen un valor numérico modificable llamado peso.

### 3.2.1 El modelo biológico

El cerebro es el elemento principal del sistema nervioso humano y está compuesto por un tipo especial de célula llamada neurona. Una neurona es una célula viva y como tal posee todos los elementos comunes de las células biológicas. A su vez, las neuronas tienen características propias que le permiten comunicarse entre ellas, lo que las diferencia del resto de las células biológicas. La figura 1 muestra la estructura típica de una neurona biológica.



**Figura 1: Modelo Biológico de Red Neuronal.**

De la figura se observa que la neurona biológica está compuesta por un *cuerpo celular* o *soma*, del cual se desprende el árbol de ramificaciones llamado *árbol dendrítico*, compuesto por las *dendritas*. Del soma también parte una fibra tubular, llamada *axón*, el cual suele ramificarse cerca de su extremo. Las dendritas actúan como un canal de entrada de señales provenientes desde el exterior hacia la neurona, mientras que el axón actúa como un canal de salida.

El espacio entre dos neuronas vecinas se denomina *sinapsis*. En el córtex cerebral se observa una organización horizontal en capas, así como también una organización vertical en columnas de neuronas.

La intensidad de una sinapsis no es fija, sino que puede ser modificada en base a la información proveniente del medio. De esta manera la estructura del cerebro no permanece fija sino que se va modificando por la formación de nuevas conexiones, ya sean excitadoras o inhibitorias, la destrucción de conexiones, la modificación de la intensidad de la sinapsis, o incluso por muerte neuronal.

Desde un punto de vista funcional, las neuronas conforman un procesador de información sencillo. Constan de un subsistema de entrada (dendritas), un subsistema de procesamiento (el soma) y un subsistema de salida (axón). Como se mencionó antes, una de las características principales de las neuronas, y que la distinguen del resto de las células, es su capacidad de comunicarse. Las señales nerviosas pueden ser eléctricas o químicas. La transmisión química se da principalmente en la



comunicación entre neuronas, mientras que la eléctrica se produce dentro de una neurona [5]. En general, una neurona recibe información de cientos de neuronas vecinas y la transmite a otras tantas neuronas. La comunicación entre neuronas se lleva a cabo de la siguiente manera: en el soma de las neuronas transmisoras o presinápticas se genera un pulso eléctrico llamado potencial de acción. El pulso eléctrico se propaga a través del axón en dirección a las sinapsis. La información se transmite a las neuronas vecinas utilizando un proceso químico, mediante la liberación de neurotransmisores. Estos neurotransmisores se transmiten a través de la sinapsis hacia la neurona receptora. La neurona receptora o postsináptica toma la señal enviada por cientos de neuronas a través de las dendritas y la transmite al cuerpo celular. Estas señales pueden ser excitadoras (positivas) o inhibitoras (negativas) [6]. El soma es el encargado de integrar la información proveniente de las distintas neuronas. Si la señal resultante supera un determinado umbral (umbral de disparo) el soma emite un pulso que se transmite a lo largo del axón dando lugar a la transmisión eléctrica a lo largo de la neurona. Al llegar la señal al extremo del axón se liberan neurotransmisores que permiten transmitir la señal a las neuronas vecinas [7].

**3.2.1.1 Estructura de un sistema neuronal artificial:** Como se dijo anteriormente, las redes neuronales son modelos matemáticos que intentan reproducir el comportamiento del cerebro humano. El principal objetivo de este modelo es la construcción de sistemas capaces de presentar un cierto comportamiento inteligente. Esto implica la capacidad de aprender a realizar una determinada tarea.

Las características principales que reproducen las redes neuronales artificiales se pueden reducir a los siguientes tres conceptos: procesamiento paralelo, distribuido y adaptativo [8]. El verdadero poder de este modelo radica en el procesamiento paralelo realizado por las neuronas artificiales. La neurona artificial es un elemento de procesamiento simple y constituye el elemento principal de un sistema neuronal artificial.

Una red neuronal artificial está compuesta por un conjunto de capas. De esta manera, la información se encuentra distribuida a lo largo de las sinapsis de la red, dándole a este sistema cierta tolerancia a fallos. A su vez, las redes neuronales artificiales son capaces de adaptar su funcionamiento a distintos entornos modificando sus conexiones entre neuronas. De esta manera pueden aprender de la experiencia y generalizar conceptos.

Por último, un conjunto de redes neuronales, junto con las interfaces de entrada y salida, y los módulos lógicos adicionales conforman un sistema neuronal artificial.

### 3.2.2 Modelo artificial

La neurona artificial es un elemento de procesamiento simple que a partir de un vector de entradas produce una única salida. En general se pueden encontrar tres tipos de neuronas artificiales, donde cada una de las cuales tiene su contraparte en el sistema nervioso:

1. Las que reciben información directamente desde el exterior, a las cuales se las denomina neuronas de entrada.

2. Las que reciben información desde otras neuronas artificiales, a las cuales se las denomina neuronas ocultas. Es en estas neuronas, en particular en sus sinapsis, donde se realiza la representación de la información almacenada.

3. Las que reciben la información procesada y las devuelven al exterior. A estas neuronas se las denomina neuronas de salida.

La figura 2 muestra de manera gráfica como es un modelo de neurona artificial.

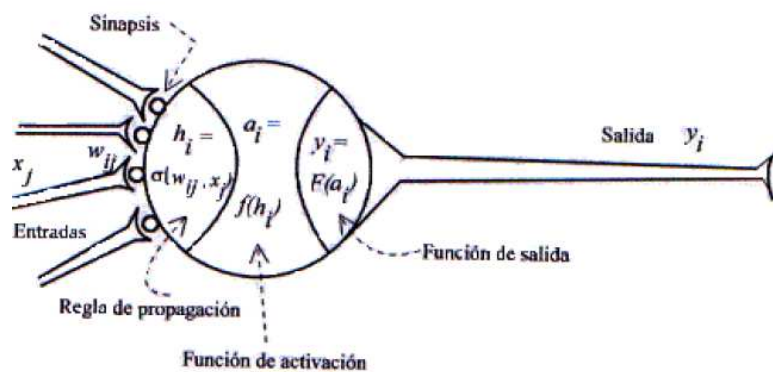


Figura 2: Modelo de Neurona Artificial

#### 3.2.2.1 Entradas y salidas

Las entradas y salidas de una neurona pueden ser clasificadas en dos grandes grupos: binarias o continuas. Las neuronas binarias (digitales) sólo admiten dos valores posibles; en general en este tipo de neurona se utilizan los siguientes dos alfabetos  $\{0,1\}$  o  $\{-1,1\}$ . Por su parte, las neuronas continuas (analógicas) admiten valores dentro de un determinado rango, que en general suele definirse como  $[-1, 1]$ .

La selección del tipo de neurona a utilizar depende de la aplicación y del modelo a construir.

### 3.2.2.2 Pesos sinápticos

El peso sináptico  $w_{ij}$  define la fuerza de una conexión sináptica entre dos neuronas, la neurona presináptica  $i$  y la neurona postsináptica  $j$ . Los pesos sinápticos pueden tomar valores positivos, negativos o cero. En caso de una entrada, un peso positivo actúa como excitador, mientras que un peso negativo actúa como inhibidor. En caso de que el peso sea cero, no existe comunicación entre el par de neuronas. Mediante el ajuste de los pesos sinápticos la red es capaz de adaptarse a cualquier entorno y realizar una determinada tarea.

### 3.2.2.3 Función de salida

La función de salida proporciona el valor de salida de la neurona, en base al estado de activación de la neurona. En general se utiliza la función identidad, es decir:

$$y(t) = F_i(a_i(t)) = a_i(t)$$

## 3.3 Procesamiento en neurona artificial

La actividad que una unidad de procesamiento o neurona artificial realiza en un sistema de este tipo es simple. Normalmente, consiste en sumar los valores de las entradas (*inputs*) que recibe de otras unidades conectadas a ella, comparar esta cantidad con el valor umbral y, si lo iguala o supera, enviar activación o salida (*output*) a las unidades que esté conectada. Tanto las entradas que la unidad recibe como las salidas que envía dependen a su vez del peso o fuerza de las conexiones por las cuales se realizan dichas operaciones [9].

La arquitectura de procesamiento de la información de los sistemas de RNA se distingue de la arquitectura convencional Von Neumann (fundamento de la mayor parte de los computadores existentes) en una serie de aspectos fundamentales.

En primer lugar, el procesamiento de la información de un modelo Von Neumann es secuencial, esto es, una unidad o procesador central se encarga de realizar una tras otra determinadas transformaciones de expresiones binarias almacenadas en la memoria del computador. Estas transformaciones son realizadas de acuerdo con una serie de instrucciones (algoritmo, programa), también almacenadas en la memoria. La operación básica de un sistema de este tipo sería: localización de una expresión en la memoria, traslado de dicha expresión a la unidad de procesamiento, transformación de la expresión y colocación de la nueva expresión en otro compartimento de la memoria [10]. Por su parte, el procesamiento en un sistema conexionista no es secuencial sino paralelo, esto es, muchas unidades de procesamiento pueden estar funcionando simultáneamente.

En segundo lugar, un rasgo fundamental de una arquitectura Von Neumann es el carácter discreto de su memoria, que está compuesta por un gran número de ubicaciones físicas o compartimentos independientes donde se almacenan en código digital tanto las instrucciones (operaciones a realizar) como los datos o números que el computador va a utilizar en sus operaciones. En redes neuronales, en cambio, la información que posee un sistema no está localizada o almacenada en compartimentos discretos, sino que está distribuida a lo largo de los parámetros del sistema. Los parámetros que definen el “conocimiento” que una red neuronal posee en un momento dado son sus conexiones y el estado de activación de sus unidades de procesamiento [11]. En un sistema conexionista las expresiones lingüísticas o simbólicas no existen como tales. Serían el resultado emergente de la interacción de muchas unidades en un nivel subsimbólico.

Un sistema de procesamiento distribuido en paralelo presenta una serie de ventajas frente a un modelo convencional Von Neumann. Por un lado, se tiene la resistencia al funcionamiento defectuoso de una pequeña parte del sistema. En un modelo conexionista, cada unidad lleva a cabo una computación simple. La fiabilidad de la computación total que el sistema realiza depende de la interacción paralela de un gran número de unidades y, consecuentemente, en la mayoría de casos, el sistema puede continuar su funcionamiento normal, aunque una pequeña parte del mismo haya resultado dañada. En los sistemas convencionales, en cambio, un defecto en un solo paso de una larga cadena de operaciones puede echar a perder la totalidad de la computación. Por otro lado, un modelo conexionista es capaz, en ciertas circunstancias, de reconocer un objeto a pesar de que sólo se le presente como entrada una parte del mismo, o a pesar de que la imagen del objeto esté distorsionada. En cambio, en un sistema convencional el objeto presentado debe corresponderse con una determinada información almacenada en memoria, de lo contrario, no es capaz de reconocer el objeto [12].

Por último, un sistema de RNA no se programa para realizar una determinada tarea a diferencia de una arquitectura Von Neumann, sino que es “entrenado” a tal efecto. Considerar el siguiente ejemplo típico de aprendizaje o formación de conceptos en la estructura de una RNA. Suponga que se presenta a la red dos tipos de objetos, por ejemplo la letra A y la letra E con distintos tamaños y en distintas posiciones. En el aprendizaje de la red neuronal se consigue, tras un número elevado de presentaciones de los diferentes objetos y consiguiente ajuste o modificación de las conexiones del sistema, que la red distinga entre As y Es, sea cual fuere su tamaño y posición en la pantalla. Para ello, se podría entrenar la red neuronal para que proporcionase como salida el valor 1 cada vez que se presente una A y el valor 0 en caso de que se presente una E. El aprendizaje en una RNA es un proceso de ajuste o modificación de los valores o pesos de las conexiones, “hasta que la conducta del sistema acaba por reproducir las propiedades estadísticas de sus entradas” [13]. En el ejemplo, se podría decir que la red ha “aprendido” el concepto de letra A y letra E sin poseer reglas concretas para el reconocimiento de dichas figuras, sin poseer un programa explícito de instrucciones para su reconocimiento [13].

Por tanto, para entrenar a un sistema conexionista en la realización de una determinada clasificación es necesario realizar dos operaciones. Primero, hay que seleccionar una muestra representativa con respecto a dicha clasificación, de pares de entradas y sus correspondientes salidas. Segundo, es necesario un algoritmo o regla para ajustar los valores modificables de las conexiones entre las unidades en un proceso iterativo de presentación de entradas, observación de salidas y modificación de las conexiones.

Las RNA constituyen una línea de investigación en Inteligencia Artificial (IA), la cual tiene como objetivo primario la construcción de máquinas inteligentes [14].

Una característica importante de la cibernética fue la proliferación de distintas perspectivas en torno al problema de las relaciones entre cerebro y máquina. En la segunda mitad de la década de los cincuenta comenzaron a destacar dos de entre estas perspectivas: la IA basada en el procesamiento simbólico, y la investigación en redes neuronales.

La IA simbólica se basó en la expansión del uso de los computadores desde el área de aplicación del cálculo numérico a tareas simbólicas, esto es, al procesamiento de elementos que representan palabras, proposiciones u otras entidades conceptuales. Estos sistemas de IA se basan en las expresiones simbólicas que contienen y en la posibilidad de manipular y transformar dichas expresiones de una manera sensible a la estructura lógico-sintáctica de las mismas. Las estructuras representacionales que contiene un sistema de este tipo son manipuladas y transformadas de acuerdo con ciertas reglas y estrategias (algoritmos y reglas heurísticas), y la expresión resultante es la solución de un determinado problema. En un sistema de este tipo, el procesamiento de la información tiene lugar en el nivel simbólico o representacional y no en el nivel neurobiológico. Los sistemas de IA simbólica simulan procesos mentales y cognitivos humanos por medio de programas ejecutados por un computador del tipo Von Neumann. Entre los investigadores más importantes de esta primera época de investigación en este paradigma se puede destacar a John McCarthy, Allen Newell, Herbert Simon y Marvin Minsky [15].

McCulloch y Pitts (1943) presentaron la estructura y funcionamiento de la unidad elemental de procesamiento de una red conexionista. La neurona de McCulloch-Pitts (ver figura 3), como actualmente se conoce, tiene un funcionamiento muy sencillo: si la suma de entradas excitatorias supera el umbral de activación de la unidad, y además no hay una entrada inhibitoria, la neurona se activa y emite respuesta (representada por el valor 1); en caso contrario, la neurona no se activa (valor 0 que indica la ausencia de respuesta) [16].

Combinando varias neuronas de este tipo con los adecuados umbrales de respuesta, se puede construir una red que compute cualquier función lógica finita.

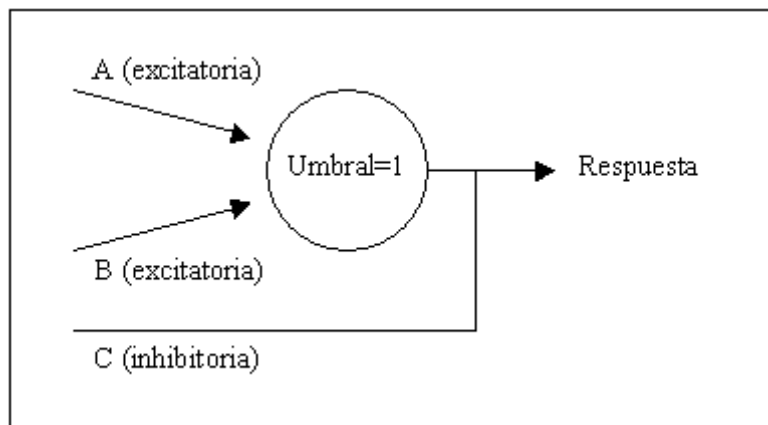


Figura 3. Neurona de McCulloch-Pitts.

Hebb (1949) postuló un sencillo pero potente mecanismo de regulación de las conexiones neuronales, que constituyó la base de las reglas de aprendizaje que más tarde se desarrollarían. La regla de Hebb, en su versión más elemental, se expresa como sigue: “*Cuando un axón de una célula A está bastante cerca para excitar a una célula B y repetida o persistentemente dispara, entonces se produce algún proceso de desarrollo o cambio metabólico de tal forma que la eficiencia del disparo de A hacia B aumenta*” [17]. La propuesta de Hebb es de especial relevancia porque indica que la información necesaria para modificar el valor de una conexión se encuentra localmente disponible a ambos lados de la conexión. En la actualidad existe un gran número de redes neuronales cuyo aprendizaje está basado en la regla de Hebb como las conocidas redes de Hopfield (1982) y algunos modelos de red propuestos por Coñeen (1977).

La evolución de la investigación en redes neuronales desde los años 50 a nuestros días ha estado condicionada por dos grandes acontecimientos: el abandono de esta línea de investigación en la segunda mitad de los 60 debido a las limitaciones observadas en la red perceptrón simple y la emergencia del conexionismo en la segunda mitad de los 80 como paradigma aceptado en IA, gracias, entre otros avances, a la aparición de un algoritmo, denominado *backpropagation error* (propagación del error hacia atrás) o simplemente *backpropagation*, que permite modificar las conexiones de arquitecturas multiestrato.

El teorema de convergencia de la regla de aprendizaje del Perceptrón desarrollado por Rosenblatt (ver figura 4) establecía que, si los parámetros o pesos del sistema eran capaces de realizar una determinada clasificación, el sistema acabaría aprendiéndola en un número finito de pasos, si se modificaban las conexiones de acuerdo con dicha regla de aprendizaje [8]. Más concretamente, la regla de aprendizaje del Perceptrón es un algoritmo de los denominados supervisado por corrección de errores y consiste en ir ajustando de forma iterativa los pesos en proporción a la diferencia existente entre la salida actual de la red y la salida deseada, con el objetivo de minimizar el error actual de la red.

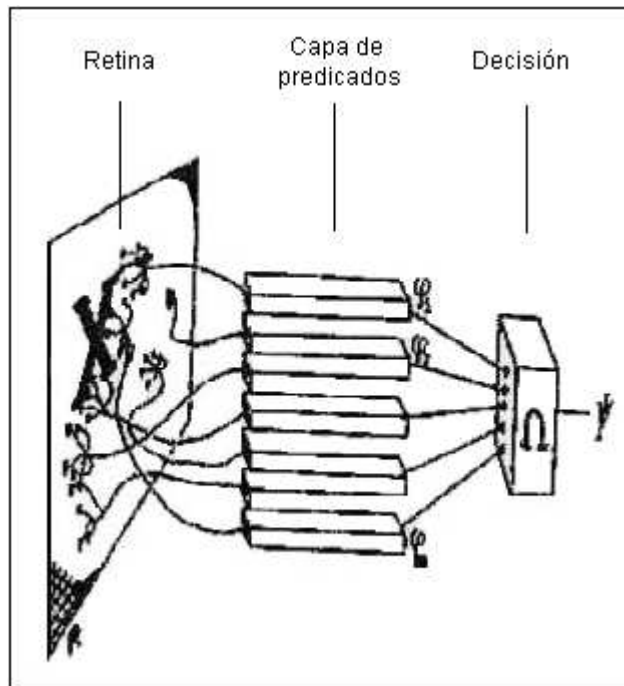


Figura 4. El Perceptrón de Rosenblatt.

La polémica suscitada entre científicos favorables y contrarios al conexionismo fue aumentando en la segunda mitad de los 50 conforme el trabajo de Rosenblatt fue adquiriendo notoriedad. Rosenblatt, un psicólogo de la Universidad de Cornell (Ithaca, Nueva York), fue la figura central del conexionismo de los años cincuenta y sesenta. El Perceptrón, una máquina conexionista diseñada y estudiada teóricamente por Rosenblatt, construida por un grupo de ingenieros del Laboratorio de Aeronáutica de Cornell (CAL, Ithaca, Nueva York) y financiada por la Oficina de Investigación Naval del Ejército de los Estados Unidos (ONR, *Office of Naval Research*), fue una de las contribuciones científicas y tecnológicas más importantes de la primera fase del conexionismo.

Otra importante contribución científica es la aportada por Widrow y Hoff en 1960. Estos autores propusieron un nuevo tipo de unidad de procesamiento, con estructura similar a la del Perceptrón pero con un mecanismo de aprendizaje diferente que permitía también la entrada de información de tipo continuo: la neurona ADALINE (ADAPTative LINEar Elements). La innovación de esta tipología de neurona se halla en su mecanismo de aprendizaje denominado regla delta o regla de Widrow-Hoff, que introduce el concepto de reducción del gradiente del error. La deducción de la regla delta se puede expresar de la siguiente forma: teniendo en cuenta que  $E_p$  (el error que comete la red para un determinado patrón  $p$ ), es función de todos los pesos de la red, el gradiente de  $E_p$  es un vector igual a la derivada parcial de  $E_p$  respecto a cada uno de los pesos. El gradiente toma la dirección del incremento más rápido en  $E_p$ ; la dirección opuesta toma el decremento más rápido en el error. Por

tanto, el error puede reducirse iterativamente ajustando cada peso  $w_i$  en la dirección  $-\frac{\partial E_p}{\partial w_i}$ . Como se verá más adelante, la regla delta basada en la reducción del gradiente del error es la precursora del algoritmo *backpropagation* aplicado a redes de múltiples estratos [18].

Sin embargo, los primeros sistemas conexionistas tenían importantes limitaciones técnicas. Una de las más importantes es que una neurona tipo Perceptrón solamente permite discriminar entre dos clases linealmente separables, es decir, cuyas regiones de decisión pueden ser separadas mediante una única recta o hiperplano (dependiendo del número de entradas). Otra importante limitación era la carencia de técnicas para la modificación de conexiones en sistemas de múltiples estratos. Este problema se puede ilustrar con las conocidas funciones OR y OR-Exclusiva (XOR). En el caso de la función OR, un Perceptrón de una sola capa de conexiones modificables permite solucionar esta función debido a que el problema es linealmente separable (ver figura 5 izquierda). En cambio, en el caso de la función OR-Exclusiva, un Perceptrón de este tipo no permite solucionar esta función debido a que no existe ninguna recta que separe los patrones de una clase de los de la otra. Para ello es necesario que se introduzca una capa intermedia compuesta por dos neuronas que determinen dos rectas en el plano (ver figura 5 derecha).

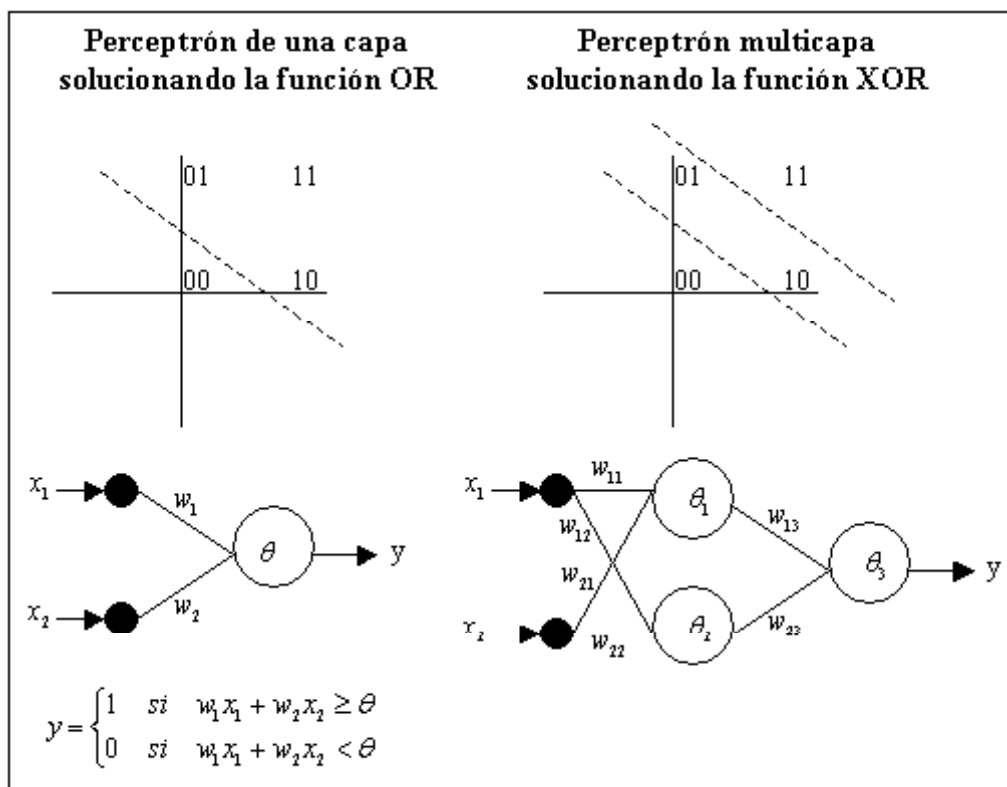


Figura 5. Perceptrones solucionando la función OR y la función XOR.

El declive del primer conexionismo sobrevino cuando Marvin Minsky y Seymour Papert, dos investigadores líderes de la IA simbólica del prestigioso Instituto de Tecnología de Massachusetts



(MIT), publican en 1969 el libro *Perceptrons* (Minsky y Papert, 1969) donde se realizaba una contundente crítica a los modelos de Perceptrón propuestos por Rosenblatt. Las aportaciones principales del estudio de Minsky y Papert pueden agruparse en dos bloques. Por un lado, Minsky y Papert realizaron un estudio, muy elaborado desde un punto de vista matemático, de algunos de los problemas que presentaban las redes de único estrato. En concreto demostraron que el Perceptrón de una capa, actualmente denominado Perceptrón simple, era incapaz de diferenciar entre entradas en distintas partes de la pantalla (triángulo a la derecha, triángulo a la izquierda), ni entre figuras en distintas posiciones de rotación. Tampoco era capaz de computar con efectividad funciones matemáticas como la paridad (dada una cantidad de puntos activos en la retina, reconocer si es un número par o impar), la función topológica de la conectividad (reconocer una figura como una totalidad separada del fondo) y en general funciones no lineales como la mencionada función OR-Exclusiva. Por otro lado, el segundo conjunto de resultados del estudio de Minsky y Papert es el referido a las RNA de múltiples estratos. En este caso dedicaron mucho menos espacio a este problema en su libro, concluyendo que “*el estudio de las versiones de múltiples estratos es estéril*” [19] alegando que sería muy improbable obtener una regla de aprendizaje aplicada a este tipo de arquitecturas.

La polémica de los años setenta entre el simbolismo y el conexionismo terminó con la aceptación por la gran mayoría de los científicos de la IA, del paradigma simbólico como línea de investigación más viable. La credibilidad que la élite de IA simbólica (Herbert Simon, Allen Newell, Marvin Minsky y John McCarthy) consiguió tanto dentro de la comunidad científica (estos investigadores dominaron la disciplina) como fuera de ella (con apoyo económico de DARPA) es un indicativo de la posición favorable en la que estos investigadores quedaron cuando la polémica sobre el Perceptrón se dio por terminada. Ante la situación de crisis, algunos de los principales grupos de RNA abandonaron su investigación. El grupo de Widrow comenzó a aplicar sus técnicas y sistemas de RNA a la ingeniería de las telecomunicaciones, y el grupo de Rosen comenzó un proyecto para la construcción de un robot móvil dentro del paradigma simbólico de IA. Rosenblatt y algunos otros investigadores, en cambio, continuaron con sus investigaciones en RNA. De hecho, la mayoría de los actuales líderes en el campo de las RNA comenzaron a publicar sus trabajos durante la década de los 70. Este es el caso de investigadores como James Anderson, Teuvo Kohonen, Christoph Von Der Malsburg, Kunihiko Fukushima, Stephen Grossberg y Gail Carpenter que se procede a comentar brevemente [20].

Anderson desarrolló un asociador lineal de patrones que posteriormente perfeccionó en el modelo BSB (*Brain-State-in-a-Box*). Simultáneamente, en Finlandia, Kohonen desarrolló un modelo similar al de Anderson; años más tarde, crearía un modelo topográfico con aprendizaje autoorganizado en el que las unidades se distribuyen según el tipo de entrada al que responden. Este modelo topográfico, comúnmente denominado mapa autoorganizado de Kohonen, es una de las redes neuronales más ampliamente utilizadas en la actualidad.

En Alemania, Von Der Malsburg (1973) desarrolló un detallado modelo de la emergencia en la corteza visual primaria de columnas de neuronas que responden a la orientación de los objetos. En

Japón, Fukushima desarrolló el Cognitrón, un modelo de red neuronal autoorganizada para el reconocimiento de patrones visuales. Posteriormente, presentó la red Neocognitrón que permitía superar las limitaciones del primitivo Cognitrón. Por su parte, Grossberg ha sido uno de los autores más prolíficos en este campo.

Klimasauskas (1989) lista 146 publicaciones en las que interviene Grossberg entre 1967 y 1988. Estudió los mecanismos de la percepción y la memoria. Grossberg realizó en 1967 una red, Avalancha, que consistía en elementos discretos con actividad que varía con el tiempo que satisface ecuaciones diferenciales continuas, para resolver actividades tales como reconocimiento continuo del habla y aprendizaje del movimiento de los brazos de un robot (Grossberg, 1982). Sin embargo, la contribución más importante de Grossberg es la Teoría de Resonancia Adaptativa (ART), desarrollada en colaboración con Carpenter (Carpenter y Grossberg, 1985, 1987a, 1987b, 1990). La ART se aplica a modelos con aprendizaje competitivo (denominados ART para la versión no supervisada y ARTMAP para la versión supervisada) en los cuales cuando se presenta cierta información de entrada sólo una de las neuronas de salida de la red se activa alcanzando su valor de respuesta máximo después de competir con las demás neuronas.

En este sentido, uno de los casos más paradigmáticos es el del físico John Hopfield, considerado como uno de los impulsores más importantes del nuevo conexionismo. Hopfield publicó en 1982 un importante artículo en la Academia Nacional de las Ciencias. Este escrito claro y conciso tuvo un importante impacto en el campo por varias razones. En primer lugar, Hopfield era un conocido físico con conexiones institucionales importantes. Su interés y trabajo en redes neuronales legitimó el campo para la comunidad científica. En segundo lugar, impulsó la implementación de los modelos de red mediante dispositivos electrónicos utilizando tecnología VLSI (Muy Alta Escala de Integración). En tercer lugar, Hopfield sugirió una estrecha relación entre los sistemas físicos y las redes neuronales. El concepto clave de las redes propuestas por Hopfield es que considera la fase de ajuste de las conexiones como una búsqueda de valores mínimos en unos paisajes de energía. Según esta idea, cada combinación de pesos de las conexiones de la red tiene asociada una energía, que resulta de evaluar las restricciones determinadas por los datos de entrada y el resultado producido por la red. El intercambio de información entre unidades se mantiene hasta que la entrada y la salida de cada unidad sean iguales, es decir, en términos de Hopfield se ha llegado a un estado de equilibrio energético. A diferencia de las redes Perceptrón y ADALINE [29], las redes utilizadas por Hopfield poseen una arquitectura monocapa cuyas conexiones son modificadas a partir de un algoritmo de aprendizaje basado en la regla de Hebb. Las redes de Hopfield han sido empleadas como memorias autoasociativas, principalmente para el reconocimiento de patrones.

El modelo de Hopfield fue posteriormente desarrollado por Hinton y Sejnowski, dos de los más importantes miembros del grupo de investigación PDP (*Parallel Distributed Processing*) (Universidad de San Diego, California), en su sistema denominado "máquina de Boltzmann" [16]. El algoritmo para la modificación de conexiones del sistema de múltiples estratos de Hinton y Sejnowski fue uno de los

aportes más importantes de la primera fase de la reemergencia del conexionismo de los 80. Era la primera vez que un algoritmo de este tipo encontraba una aceptación considerable en la comunidad científica.

Sin embargo, la contribución más importante en la reemergencia del conexionismo en los años ochenta fue la técnica *backpropagation* desarrollada por Rumelhart, Hinton y Williams, representantes del grupo PDP. Realmente, esta técnica fue desarrollada inicialmente por Paul Werbos (1974) a mediados de los 70, y después independientemente redescubierta por varios grupos de investigadores (Le Cun, 1985; Parker, 1985; Rumelhart, Hinton y Williams, 1986). Es, por tanto, un caso de “descubrimiento múltiple”. Sin embargo, en general se reconoce que fue la versión del grupo PDP la que desató el interés en RNA a mediados de los ochenta y consiguió finalmente forzar la revisión del consenso contrario al conexionismo.

El algoritmo *backpropagation* [21] también recibe el nombre de regla delta generalizada o método de gradiente decreciente, debido a que supone una extensión de la regla propuesta por Widrow y Hoff en 1960 (regla delta) a redes con capas intermedias (ver figura 6). Este tipo de arquitectura recibe el nombre genérico de Perceptrón Multicapa o MLP (*Multilayer perceptrón*). Rosenblatt ya tuvo la idea de utilizar una técnica de este tipo a principios de los sesenta, aunque no pudo desarrollarla de un modo satisfactorio.

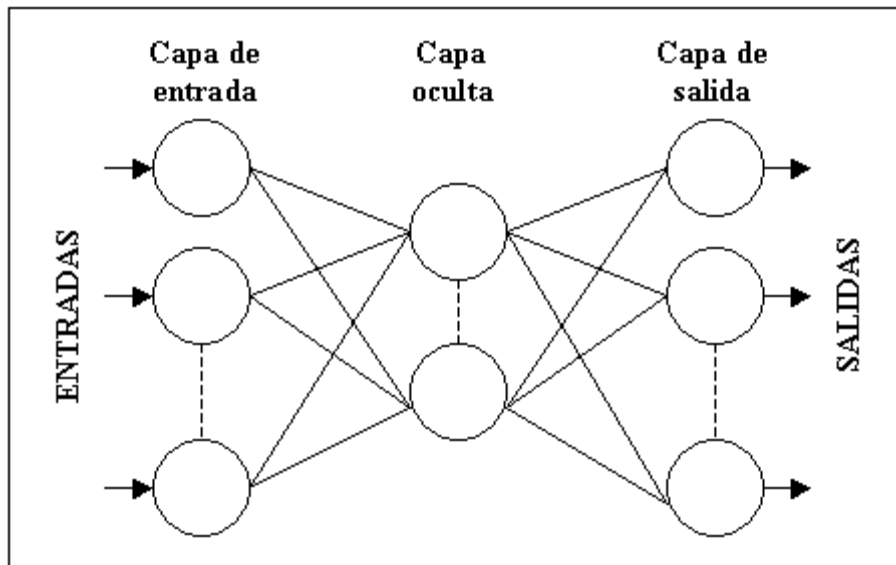


Figura 6. Arquitectura de un Perceptrón multicapa [18].

Como se comentó anteriormente, la falta de un algoritmo para la modificación de conexiones en sistemas de múltiples estratos limitaba considerablemente la capacidad de clasificación de objetos en los sistemas conexionistas de los años 60. En este sentido, el problema principal de la modificación de los valores de las conexiones en una red MLP es hallar el error cometido por las unidades de las capas intermedias. El error cometido por las unidades de salida es inmediatamente visible: es la diferencia

entre la salida producida por dichas unidades y la salida que se desea que produzcan. El objetivo del algoritmo *backpropagation* es propagar los errores cometidos por las unidades de salida hacia atrás, ya que, en un sistema de este tipo, el error cometido por una unidad intermedia depende del error cometido por las unidades de salida a las que dicha unidad intermedia está conectada. Tras conocerse el error cometido por las unidades intermedias, pueden entonces modificarse las conexiones entre unidades de entrada y unidades intermedias. De forma similar a la regla delta, la base matemática del algoritmo *backpropagation* es la técnica de gradiente decreciente, basada en modificar los pesos en la dirección opuesta al gradiente, esto es  $-\frac{\partial E_p}{\partial w_{ij}}$ , en la dirección que determina el decremento más rápido del error.

Una novedad muy importante en el sistema de Rumelhart y sus colegas fue la introducción de funciones de activación continuas en todas las unidades de procesamiento en lugar de la clásica función “escalón” del Perceptrón simple de Rosenblatt. De hecho, el algoritmo *backpropagation* exige la utilización de funciones de activación continuas para poder realizar el cálculo de la derivada parcial del error con respecto a los pesos del modelo.

El proceso de acumulación de resultados e investigaciones y de esfuerzo organizacional por parte del grupo PDP, comenzó a hacer peligrar el consenso anticonexionista con el que terminó la polémica del Perceptrón. Los dos volúmenes PDP, considerados como la “biblia” del conexionismo, son el mayor exponente de este esfuerzo (Rumelhart, McClelland y el grupo de investigación PDP, 1986; McClelland, Rumelhart y el grupo de investigación PDP, 1986). El debate sobre el conexionismo se estaba reabriendo, y esto hizo reaccionar de nuevo a los investigadores críticos con el conexionismo. La reacción fue encabezada, una vez más, por Minsky y Papert que, en el epílogo a la nueva edición de su libro *Perceptrons*, criticaron contundentemente las afirmaciones de Rumelhart y sus colegas acerca de los sistemas de múltiples estratos con el algoritmo *backpropagation*. Minsky y Papert no fueron los únicos en criticar al nuevo conexionismo con vehemencia. Otros científicos líderes en sus áreas de investigación, tales como Poggio (visión), Hillis (computadores paralelos), y Fodor y Pylyshyn (ciencia cognitiva), también realizaron críticas radicales al conexionismo. Sin embargo, esta vez la polémica no acabó con el abandono del conexionismo como ocurriera en la década de los 60.

En el artículo de Horgan (1994) se trata la persona de Marvin Minsky, comentándose algunas de sus opiniones actuales, como, por ejemplo, cómo poco a poco se ha ido apartando de la IA simbólica y su aprobación al actual desarrollo de las RNA.

Gracias al esfuerzo de movilización y acumulación científica y organizacional que el grupo de investigación PDP realizó a lo largo de la década de los ochenta, el conexionismo ha logrado en la actualidad diferenciarse como una especialidad científica aceptada, dentro del marco general de la IA. Este proceso ha culminado con el surgimiento, crecimiento e institucionalización de una comunidad científica diferenciada con su correspondiente sistema de comunicación y control especializado

(publicaciones científicas, congresos, cursos de postgrado, institutos de investigación, programas y becas en las agencias que financian la investigación científica, etc.).

Sin duda, el creciente interés despertado por las RNA, que se manifiesta de forma palpable a principios de los 90, está relacionado con un acontecimiento fundamental en la historia de las RNA comentado al inicio, a saber: la publicación de *Parallel Distributed Processing* (Procesamiento Distribuido en Paralelo o PDP) (Rumelhart, McClelland y el grupo PDP, 1986; McClelland, Rumelhart y el grupo de investigación PDP, 1986), obra que se ha llegado a conocer como la “biblia” del nuevo paradigma conexionista donde se describe, entre otras cosas, el algoritmo *backpropagation* aplicado a redes MLP.

Respecto a la cuestión de cuántos tipos de redes neuronales existen actualmente, se puede decir que se trata de un número inabarcable. Sin embargo, también se puede decir que del total hay aproximadamente 40 modelos que son bien conocidos por la comunidad de investigadores en RNA. La tabla 2 presenta la clasificación de las RNA más conocidas en función del tipo de aprendizaje utilizado: supervisado o no supervisado.

## 1. Supervisado

### 1. Con conexiones feedforward

- Lineales
  - Perceptrón (Rosenblatt, 1958)
  - Adaline (Widrow y Hoff, 1960)
- Perceptrón multicapa (Multilayer perceptrón) (MLP)
  - Backpropagation (Rumelhart, Hinton y Williams, 1986)
  - Correlación en cascada (Cascade correlation) (Fahlman y Lebiere, 1990)
  - Quickpropagation (Quickprop) (Fahlman, 1988)
  - Delta-bar-delta (Jacobs, 1988)
  - Resilient Propagation (RPROP) (Riedmiller y Braun, 1993)
  - Gradiente conjugado (Battiti, 1992)
- Radial Basis Function (RBF) (Broomhead y Lowe, 1988; Moody y Darken, 1989)
  - Orthogonal Least Squares (OLS) (Chen, Cowan y Grant, 1991)
- Cerebellar Articulation Controller (CMAC) (Albus, 1975)
- Sólo clasificación
  - Learning Vector Quantization (LVQ) (Kohonen, 1988)
  - Red Neuronal Probabilística (PNN) (Probabilistic Neural Network) (Specht, 1990)
- Sólo regresión
  - General Regression Neural Network (GRNN) (Specht, 1991)

### 2. Con conexiones feedback

- Bidirectional Associative Memory (BAM) (Kosko, 1992)
- Máquina de Boltzman (Ackley, Hinton y Sejnowski, 1985)
- Series temporales recurrentes
  - Backpropagation through time (Werbos, 1990)

<ul style="list-style-type: none"> <li>- Elman (Elman, 1990)</li> <li>- Finite Impulse Response (FIR) (Wan, 1990)</li> <li>- Jordan (Jordan, 1986)</li> <li>- Real-time recurrent network (Williams y Zipser, 1989)</li> <li>- Recurrent backpropagation (Pineda, 1989)</li> <li>- Time Delay NN (TDNN) (Lang, Waibel y Hinton, 1990)</li> </ul>
<p>3. Competitivo</p> <ul style="list-style-type: none"> <li>- ARTMAP (Carpenter, Grossberg y Reynolds, 1991)</li> <li>- Fuzzy ARTMAP (Carpenter, Grossberg, Markuzon, Reynolds y Rosen, 1992)</li> <li>- Gaussian ARTMAP (Williamson, 1995)</li> <li>- Counterpropagation (Hecht-Nielsen 1987, 1988, 1990)</li> <li>- Neocognitrón (Fukushima, Miyake e Ito, 1983; Fukushima, 1988)</li> </ul>
<p>2. No supervisado</p> <p>1. Competitivo</p> <ul style="list-style-type: none"> <li>- Vector Quantization <ul style="list-style-type: none"> <li>- Grossberg (Grossberg, 1976)</li> <li>- Kohonen (Kohonen, 1984)</li> <li>- Conscience (Desieno, 1988)</li> </ul> </li> <li>- Mapa Auto-Organizado (Self-Organizing Map) (Kohonen, 1982; 1995)</li> <li>- Teoría de la Resonancia Adaptativa (Adaptive Resonance Theory, ART) <ul style="list-style-type: none"> <li>- ART 1 (Carpenter y Grossberg, 1987a)</li> <li>- ART 2 (Carpenter y Grossberg, 1987b)</li> <li>- ART 2-A (Carpenter, Grossberg y Rosen, 1991a)</li> <li>- ART 3 (Carpenter y Grossberg, 1990)</li> <li>- Fuzzy ART (Carpenter, Grossberg y Rosen (1991b)</li> </ul> </li> <li>- Differential Competitive Learning (DCL) (Kosko, 1992)</li> </ul> <p>2. Reducción de dimensionalidad</p> <ul style="list-style-type: none"> <li>- Regla de Oja (Oja, 1989)</li> <li>- Sanger (Sanger, 1989)</li> <li>- Differential hebbian (Kosko, 1992)</li> </ul> <p>3. Autoasociación</p> <ul style="list-style-type: none"> <li>- Autoasociador lineal (Anderson, Silverstein, Ritz y Jones, 1977)</li> <li>- Brain-State-in-a-Box (BSB) (Anderson, Silverstein, Ritz y Jones, 1977)</li> <li>- Red de Hopfield (1982)</li> </ul>

Tabla 2: Clasificación de Redes según Aprendizaje.

### 3.3.1 Ventajas de las Redes Neuronales

Se acaba de ver el desarrollo histórico de los sistemas conexionistas; se ha comprobado que, es una ciencia multidisciplinar donde ingenieros, psicólogos médicos, matemáticos y físicos teóricos han aportado algún elemento a estas teorías, pero, ¿por qué ese interés en estos sistemas? ¿Qué tienen en especial frente a otros que se pudiesen denominar clásicos? en definitiva ¿qué cosas nuevas ofrecen? Se

ha comentado que la potencia computacional de una red neuronal deriva, principalmente, de su estructura de cálculo distribuido paralelo. Esta estructura le permite la resolución de problemas que necesitarían gran cantidad de tiempo en computadores “clásicos”. Pero aparte de este hecho aparecen otras propiedades que las hacen especialmente atractivas para ser usadas en una gran cantidad de problemas prácticos [22]:

a) Son sistemas distribuidos no lineales: Una neurona es un elemento no lineal por lo que una interconexión de ellas (red neuronal) también será un dispositivo no lineal. Esta propiedad permitirá la simulación de sistemas no lineales y caóticos, simulación que, con los sistemas clásicos lineales, no se puede realizar.

b) Son sistemas tolerantes a fallos: Una red neuronal, al ser un sistema distribuido, permite el fallo de algunos elementos individuales (neuronas) sin alterar significativamente la respuesta total del sistema. Este hecho las hace especialmente atractivas frente a los computadores actuales que, por lo general, son sistemas secuenciales de tal forma que un fallo en uno de sus componentes conlleva que el sistema total no funcione.

c) Adaptabilidad: Una red neuronal tiene la capacidad de modificar los parámetros de los que depende su funcionamiento de acuerdo con los cambios que se produzcan en su entorno de trabajo (cambios en las entradas, presencia de ruido, etc.). Con respecto a la capacidad de adaptación hay que tener en cuenta que ésta no puede ser tampoco excesivamente grande ya que conduciría a tener un sistema inestable respondiendo a pequeñas perturbaciones. Este es el problema conocido como el dilema plasticidad-estabilidad.

d) Establecen relaciones no lineales entre datos. Las redes neuronales son capaces de relacionar dos conjuntos de datos mediante relaciones complejas.

e) Posibilidad de implementación en VLSI: Esta posibilidad permite que estos sistemas puedan ser aplicados en sistemas de tiempo real, simulando sistemas biológicos mediante elementos de silicio. Aplicaciones que se pasarán a describir los diferentes modelos que se pueden encontrar.

Todas estas ventajas hacen el uso de las redes neuronales especialmente atractivo en un gran número de aplicaciones.

En el campo de las redes neuronales se conoce con el nombre de arquitectura la forma en la que se unen los diferentes elementos, neuronas, mediante una serie de conexiones, pesos sinápticos. En principio se pueden distinguir tres niveles, en cuanto a arquitectura se refiere, las que se definen como:

- Microestructura: Este nivel hace referencia al elemento más pequeño: la neurona. Este es el nivel más pequeño pero no por ello es el menos importante; aquí se fijan características tan importantes como la función de activación que se explicará a continuación.
- Mesoestructura: Una vez sobrepasado el nivel neuronal se llega a este nivel donde se fija la forma de conexión y la disposición de los elementos explicados anteriormente.
- Macroestructura: Las diferentes redes planteadas en el nivel anterior se pueden combinar entre sí para dar estructuras mayores alcanzándose mejores prestaciones.

### 3.3.2 Modelos Neuronales

En todo modelo artificial de neurona se tienen cuatro elementos básicos [14]:

- Un conjunto de conexiones, pesos o sinapsis que determinan el comportamiento de la neurona. Estas conexiones pueden ser excitadoras (presentan un signo positivo), o inhibitoras (conexiones negativas).
- Un sumador que se encarga de sumar todas las entradas multiplicadas por las respectivas sinapsis.
- Una función de activación no lineal para limitar la amplitud de la salida de la neurona.
- Un umbral exterior que determina el umbral por encima del cual la neurona se activa.

Esquemáticamente, una neurona artificial quedaría representada por el esquema representado en figura 7.

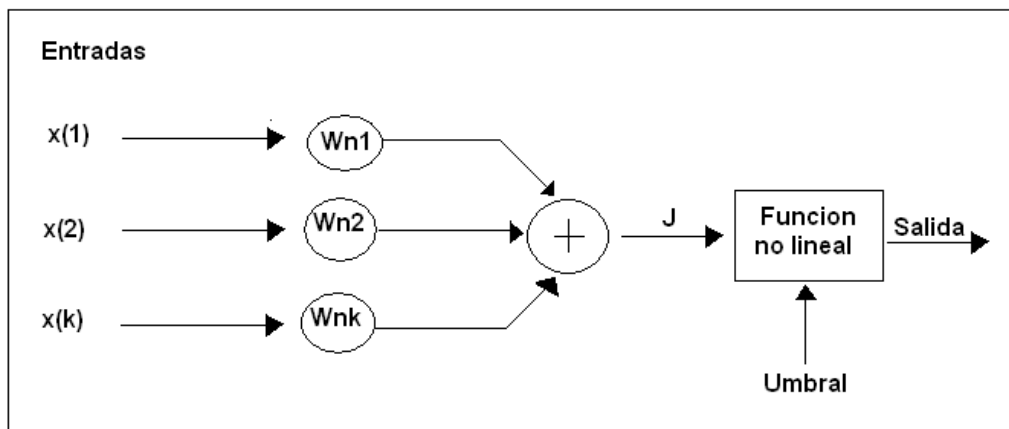


Figura 7: Esquema de un Modelo Neuronal

Matemáticamente las operaciones a realizar serían:

$$U_n = \sum_{j=1}^k W_{nj} x(j)$$

Y



$$\text{Salida} = p(U_n - \text{umbral})$$

Donde  $U_n$  es una función no lineal conocida como función de activación. Las funciones definidas varían entre 0 y 1; se pueden definir a partir de ellas otras funciones que varían entre -1 y 1 simplemente escalando las salidas entre estos límites.

### 3.3.3 Fundamentos biológicos de las Redes Neuronales

El aparato de comunicación neuronal de los animales y del hombre, formado por el sistema nervioso y hormonal, en conexión con los órganos de los sentidos y los órganos efectores (músculos, glándulas), tienen la misión de recoger informaciones, transmitir las y procesarlas; en parte también almacenarlas y enviarlas de nuevo en forma elaborada. El sistema de comunicación neuronal se compone de tres partes:

1. Los receptores, que están en las células sensoriales, recogen las informaciones en forma de estímulos, bien del ambiente, o bien del interior del organismo.
2. El sistema nervioso, que recibe las informaciones, las elabora, en parte las almacena y las envía en forma elaborada a los órganos efectores y a otras zonas del sistema nervioso.
3. Órganos diana o efectores (por ejemplo músculos y glándulas), que reciben la información y la interpretan en forma de acciones motoras, hormonales, etc.

El elemento estructural y funcional más esencial, en el sistema de comunicación neuronal, es la célula nerviosa o neurona. La mayoría de las neuronas utilizan sus productos de secreción como señales químicas (transmisores) para la transmisión de información. Dicha información se envía, entre las distintas neuronas, a través de prolongaciones, formando redes, en las cuales se elabora y almacena información. Además, una parte de las neuronas está en relación con receptores, a través de los cuales llegan comunicaciones procedentes del exterior o el interior del organismo hasta las redes neuronales.

Otra parte conduce la información, elaboradas en forma de órdenes, hacia los efectores. Una de las prolongaciones es la encargada de la conducción de impulsos; se denomina axón.

La misión de las neuronas comprende generalmente cinco funciones parciales:

- Las neuronas recogen la información que llega a ellas en forma de impulsos procedentes de otras neuronas o de receptores.
- La integran en un código de activación propio de la célula.
- La transmiten codificada en forma de frecuencia de impulsos a través de su axón.
- A través de sus ramificaciones el axón efectúa la distribución espacial de los mensajes.
- En sus terminales trasmite los impulsos a las neuronas subsiguientes o a las células efectoras.

### 3.3.4 Modelo Computacional

La gran diferencia entre una máquina confesionista, es decir, una máquina neuronal y los programas de computador convencionales que éstas “elaboran”, en cierta medida, la información de entrada para obtener una salida o respuesta. No se trata de la aplicación ciega y automática de un algoritmo. De hecho, el proceso de elaboración de la información recibida depende de las distintas características, tanto estructurales como funcionales, de la red.

Existen modelos muy diversos de redes de neuronas en los cuales se siguen filosofías de diseño, reglas de aprendizaje y funciones de construcción de las respuestas muy distintas. Una primera clasificación se hace en función del recorrido que sigue la información dentro de la red, y así se distinguen redes alimentadas hacia delante y redes con retro-alimentación.

### 3.3.5 La neurona artificial

La neurona artificial, célula o autómeta, es un elemento que posee un estado interno, llamando nivel de activación, y recibe señales que le permiten, en su caso, cambiar de estado [24].

Si se denomina  $S$  al conjunto de estados posibles de la neurona,  $S$  podrá ser, por ejemplo,  $S = \{0,1\}$ , siendo el 0 el estado inactivo y 1 el activo.  $S$  también podrá tomar un mayor número de valores,  $S = \{0, 1,2,\dots, n\}$  para representar, por ejemplo, una imagen con  $n+1$  niveles de gris, o incluso un intervalo continuo de valores, por ejemplo  $S = [0,1]$ .

Las neuronas poseen una función que les permite cambiar de nivel de activación a partir de las señales que reciben; a dicha función se la denomina función de transición de estado o función de activación. Las señales que recibe cada neurona pueden provenir del exterior o de las neuronas a las cuales está conectada.

El nivel de activación de una célula depende de las entradas recibidas y de los valores sinápticos, pero no de anteriores valores de estados de activación. Para calcular el estado de activación se ha de calcular en primer lugar la entrada total a la célula,  $E_i$ . Este valor se calcula como la suma de todas las entradas ponderadas por ciertos valores. La figura 8 muestra un modelo que representa esta idea.

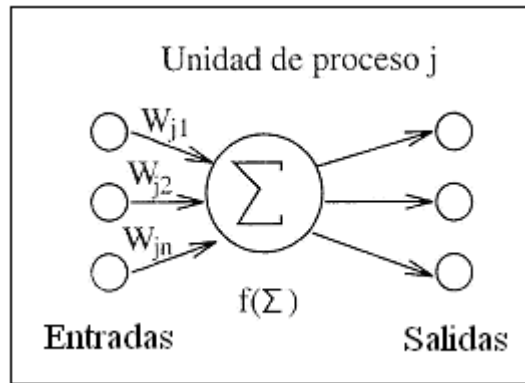


Figura 8: Modelo de Neurona Artificial

Aquí un grupo de entradas  $x_1, x_2, \dots, x_n$  son introducidas en una neurona artificial. Estas entradas, definidas por un vector  $\bar{X}$ , corresponden a las señales de la sinapsis de una neurona biológica. Cada señal multiplica por un peso asociado  $w_1, w_2, \dots, w_n$  antes de ser aplicado al sumatorio etiquetado por  $\Sigma$ . Cada peso corresponde a la fuerza de una conexión sináptica, es decir el nivel de concentración iónica de la sinapsis, y se representa por un vector  $\bar{W}$ .

El sumatorio, que corresponde al cuerpo de la neurona, suma todas las entradas ponderadas algebraicamente, produciendo una salida que se denomina  $E$ , así:

$$E = x_1 w_1 + x_2 w_2 + \dots + x_n w_n$$

Esto puede ser definido en forma vectorial como sigue:

$$E = X^t W$$

Las señales  $E$  son procesadas además por una función llamada función de activación o de salida  $F$ , que produce la señal de salida de la neurona  $S$ . dependiendo de la función  $F$ , habrá distintos modelos de autómatas: por ejemplo:

- Lineal:  $S = KE$  con  $K$  constante.
- Umbral:  $S = 1$  si  $E \geq \Theta$ ,  $S = 0$  si  $E < \Theta$ , siendo  $\Theta$  el umbral constante.
- Cualquier función:  $S = F(E)$ ; siendo  $F$  una función cualquiera de activación de neurona.

Más adelante se hará una descripción más detallada de los tipos de funciones de activación y de su importancia en un modelo de Red Neuronas Artificial.

### 3.3.6 Estructura básica de la red

En la figura 9 se muestra un ejemplo de una unidad típica de proceso de una Red de Neuronas Artificial. Se ve una serie de entradas a la neurona; cada una llega de la salida de otra neurona de la red. Una vez calculada la salida de una neurona, como se ha explicado en el apartado anterior, ésta se propaga, vía conexiones de salida, a las células destino. Todas las conexiones de salida reciben el mismo valor de salida [25].

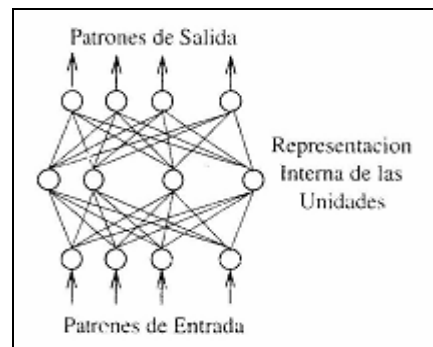


Figura 9: Unidad Típica de proceso de Red Neuronal

A la manera en que las células se conectan entre si se la denomina patrón de conectividad o arquitectura de la red. La estructura de interconexión entre células es lo denominado multicapa. Se trata de una estructura típica de implementación del paradigma conocido como retro-propagación. El primer nivel lo constituyen las células de entrada; estas unidades reciben los valores de unos patrones representados como vectores que sirven de entrada a la red. A continuación hay una serie de capas intermedias, llamadas ocultas, cuyas unidades responden a rasgos particulares que pueden aparecer en los patrones de entrada. Puede haber uno o varios niveles ocultos. El último nivel es el de salida. La salida de estas unidades como salida de toda la red.

Cada interconexión entre unidades de proceso actúa como una ruta de comunicación: a través de estas interconexiones viajan valores numéricos de una célula a otra. Los pesos de las conexiones se ajustan durante la fase de aprendizaje para producir una red de neuronas artificiales final.

Así pues, una red de neuronas artificial podría definirse como un grafo cuyos nodos están constituidos por unidades de proceso idénticas y que propagan información a través de los arco. En este grafo se distinguen tres tipos de nodos: los de entrada, los de salida y los intermedios.

El funcionamiento de la red es simple. Para cada vector de entrada, éste es introducido en la red copiando cada valor de dicho vector en la célula de entrada correspondiente. Cada célula de la red, una vez recibida la totalidad de sus entradas, las procesa y genera una salida que es propagada a través de las conexiones entre células, llegando como entrada a la célula de destino. Una vez que la entrada ha sido completamente propagada por toda la red, se producirá un vector de salida, cuyos componentes son cada uno de los valores de salida de las células de salida.

Así pues, el esquema de funcionamiento de una red de neuronas por capas como la de la figura 9 puede describirse mediante la ecuación:

$$1.1 \vec{S} = k.(K.\vec{X}.W_1).W_2$$

Donde  $W_1$  y  $W_2$  son los pesos de la primera y segunda capa, respectivamente,  $\vec{X}$  es el vector de entrada a la red, y  $S$  es el vector de salida que la red produce.

Obs.:  $W_1$  y  $W_2$  son matrices de conexiones entre las capas de la red (pesos).

Para ilustrar se presenta un ejemplo: sea la red de la figura 9, donde

$$W_1 = \begin{pmatrix} 3 & 2 & 3 \\ 4 & 1 & 2 \\ 5 & 2 & 1 \\ 6 & 1 & 1 \end{pmatrix} \quad W_2 = \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}$$

Si la función de activación es  $F(x) = 2x$  (donde  $F$  es la función de activación idéntica en todas las neuronas) y se introduce el vector  $\vec{x} = \{-3, 1, 4, -2\}$ , según la Ecuación 1.1 la salida será:

$$\vec{A} = \vec{X}\vec{W}_1 = (-3, 1, 4, -2) \cdot \begin{pmatrix} 3 & 2 & 3 \\ 4 & 1 & 2 \\ 5 & 2 & 1 \\ 6 & 1 & 1 \end{pmatrix} = (3, 1, -5)$$

En la figura 10 se muestra un esquema comparativo entre dos redes neuronales, una con sistema multicapas, y el segundo con un esquema monocapa.

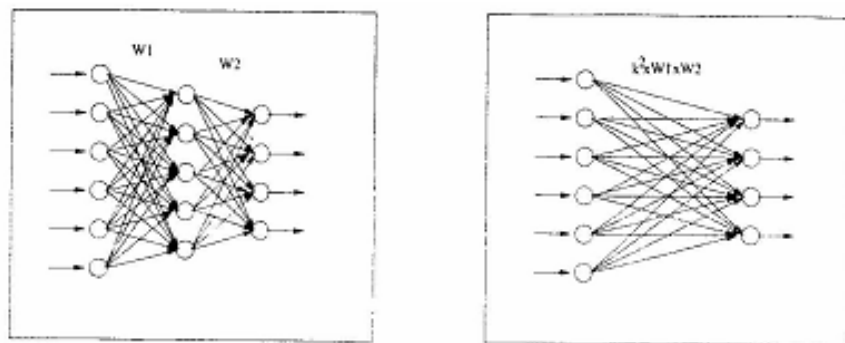


Figura 10: Equivalencia entre redes al utilizar funciones de activación lineal.

$$\vec{S} = F(F(\vec{A}).W_2) = (6,2,-10) \cdot \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 1 & 1 \end{pmatrix} = F((4,12)) = (8,24)$$

La misma salida se obtendría con una sola capa con una matriz de conexiones calculada de la siguiente manera:

$$W_t = k^2 \cdot W_1 \cdot W_2 = 4 \cdot \begin{pmatrix} 3 & 2 & 3 \\ 4 & 1 & 2 \\ 5 & 2 & 1 \\ 6 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 44 & 64 \\ 44 & 64 \\ 52 & 80 \\ 56 & 84 \end{pmatrix}$$

Con una sola salida, para el vector de entrada anterior, se calcularía simplemente como:

$$\vec{S} = \vec{X} \cdot W_t = (-3,1,4,-2) \cdot \begin{pmatrix} 44 & 64 \\ 44 & 64 \\ 52 & 80 \\ 56 & 81 \end{pmatrix} = (8,24)$$

Como se ve, se obtiene la misma salida que anteriormente. Esto quiere decir que si la función de activación es lineal, el introducir más capas en la red es irrelevante; existirá siempre una red con una sola capa equivalente a cualquier otra con un número de capas arbitrario, y su cálculo es trivial.

Todas las redes cuentan con funciones de activación no lineales, que hacen que el potencial de la red para solucionar problemas de forma genérica sea elevado.

A continuación, en la figura 11, se presenta gráficamente el ejemplo presentado anteriormente, para una red con capas ocultas.

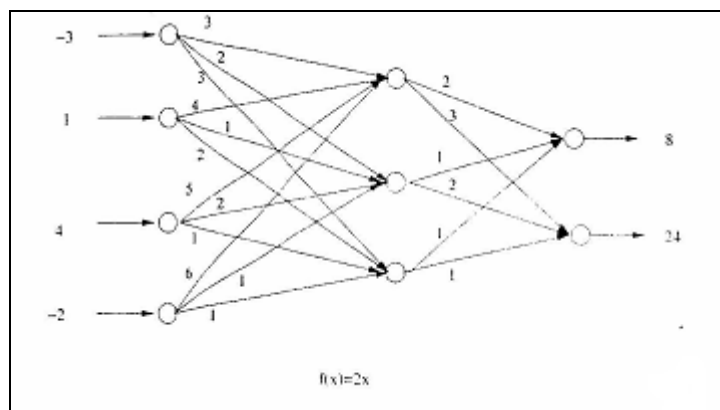


Figura 11: Ejemplo de Red Neuronal con Capas Ocultas

### 3.3.7 Aprendizaje

La parte más importante de una red de neuronas artificial es el aprendizaje. El esquema de aprendizaje de una red es lo que determina el tipo de problemas que será capaz de resolver. Las Redes de Neuronas Artificiales son sistemas de aprendizaje basados en ejemplos. La capacidad de una red para resolver un problema estará ligada de forma fundamental al tipo de ejemplos de que dispone en el proceso de aprendizaje [32]. Desde el punto de vista de los ejemplos, el conjunto de aprendizaje debe poseer las siguientes características:

- Ser significativo. Debe haber un número suficiente de ejemplos. Si el conjunto de aprendizaje es reducido, la red no será capaz de adaptar sus pesos de forma eficaz.
- Ser representativo. Los componentes del conjunto de aprendizaje deberán ser diversos. Si un conjunto de aprendizaje tiene muchos más ejemplos de un tipo que el resto, la red se especializará en dicho subconjunto de datos y no será de aplicación general. Es importante que todas las regiones significativas del espacio de estados estén suficientemente representadas en el conjunto de aprendizaje.

El aprendizaje en una red de neuronas artificial consiste en la determinación de los valores precisos de los pesos para todas sus conexiones, que la capacite para la resolución eficiente de un problema. El proceso general de aprendizaje consiste en ir introduciendo paulatinamente todos los ejemplos del conjunto de aprendizaje, y modificar los pesos de las conexiones siguiendo un determinado esquema de aprendizaje. Una vez introducidos todos los ejemplos se comprueba si se ha cumplido cierto criterio de convergencia; de no ser así se repite el proceso y todos los ejemplos del conjunto vuelven a ser introducidos. La modificación de los pesos puede hacerse después de la introducción de cada ejemplo del conjunto, o una vez introducidos todos ellos.

El criterio de convergencia depende del tipo de red utilizado o del tipo problema a resolver. La finalización del periodo de aprendizaje se puede determinar:

- Mediante un número fijo de ciclos. Se decide a priori cuantas veces será introducido todo el conjunto, y una vez superado dicho número se detiene el proceso y se da por aceptada la red resultante.
- Cuando el error descienda por debajo de una cantidad preestablecida. En este caso habrá que definir en primer lugar una función de error, bien a nivel de patrón individual, bien a nivel de la totalidad del conjunto de entrenamiento. Se decide a priori un valor aceptable para dicho error, y sólo se para el proceso de aprendizaje cuando la red produzca un valor de error por debajo del prefijado. Para este criterio puede suceder que la red jamás consiga bajar por debajo del nivel prefijado, en cuyo caso se debe disponer de un criterio adicional de parada, por ejemplo un número de ciclos, que de

utilizarse por la red significará que ésta no ha convergido. En este caso la red se dice que no ha sido capaz de obtener una solución. será necesario probar cambiando alguno de los parámetros.

- Cuando la modificación de los pesos sea irrelevante. En algunos modelos se define un esquema de aprendizaje que hace que las conexiones vayan modificándose cada vez con menor intensidad. Si el proceso de aprendizaje continuo, llegará un momento en que ya no se producirán variaciones de los valores de los pesos de ninguna conexión, en ese momento se dice que la red ha convergido y se detiene el proceso de aprendizaje.

Dependiendo del esquema de aprendizaje y del problema a resolver, se pueden distinguir tres tipos de esquemas de aprendizaje:

- Aprendizaje supervisado. En este tipo de esquemas, los datos del conjunto de aprendizaje tienen dos tipos de atributos: los datos propiamente dichos y cierta información relativa a la solución del problema. Por ejemplo, si se trata de definir un clasificador para un conjunto de datos, un sistema capaz de distinguir entre caras de diferentes personas, los ejemplos contendrán datos del individuo, una imagen de su cara, e información para modificar las conexiones. La manera más habitual de modificar los valores de los pesos de las conexiones es la representada en la figura 12. Cada vez que un ejemplo es introducido y se procesa para obtener una salida, dicha salida se compara con la salida que debería haber producido, y de la que se dispone al estar incluida dicha información en el conjunto de aprendizaje.

Para este tipo de aprendizaje, se dice que hay un profesor externo encargado de determinar si la red se está comportando de forma adecuada, mediante la comparación entre la salida producida y la esperada, y de actuar en consecuencia modificando apropiadamente los valores de los pesos.

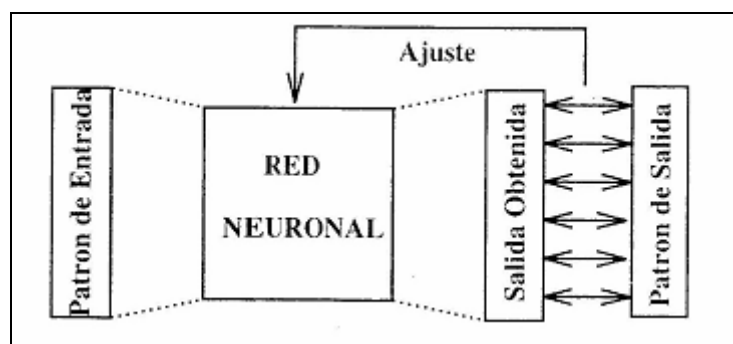


Figura 12: Aprendizaje Supervisado

- Aprendizaje no supervisado. En este aprendizaje los datos del conjunto de aprendizaje sólo tienen información de los ejemplos, y no hay nada que permita guiar en el proceso de aprendizaje. En este caso, no existe profesor externo que determine el aprendizaje. En la figura 13 se representa este



tipo de aprendizaje. La red modificará los valores de los pesos a partir de información interna. Cuando se utiliza aprendizaje no supervisado, la red trata de determinar características de los datos del conjunto de entrenamiento: rasgos significativos, regularidades o redundancias, para que la salida producida sea lo más parecida posible a la deseada. Sin embargo, esto no siempre indica que la red será capaz de solucionar el problema, pues lo importante no es que el sistema dé buenas salidas sobre el conjunto de aprendizaje, que ya son conocidas, sino sobre los datos que puedan presentarse en el futuro, y cuyas salidas se desconocen. A este tipo de modelos se los conoce también como sistemas autoorganizados, debido a que la red se ajusta dependiendo únicamente de los valores recibidos como entrada.

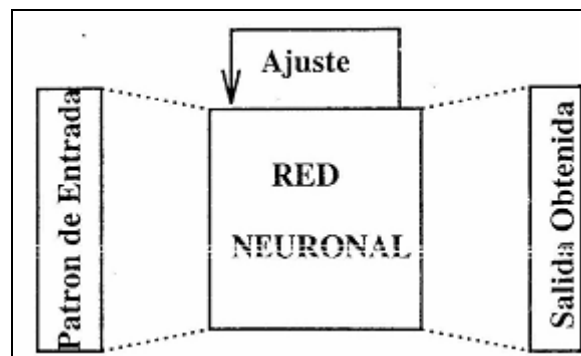


Figura 13: Aprendizaje No Supervisado

- Aprendizaje por refuerzo. Es una variante del aprendizaje supervisado con el que se dispone de información concreta del error cometido por la red, y se determina si la salida producida para dicho patrón es o no adecuada.

Para el caso del aprendizaje por refuerzo, el aprendizaje tiene una serie de características específicas que es importante resaltar. En este caso el conjunto de aprendizaje está compuesto por ejemplos que contienen los datos y sus salidas deseadas. El proceso consiste en modificar los pesos de la red hasta para todos los ejemplos del conjunto de entrenamiento.

Para poder determinar si la red produce salidas adecuadas, se divide el conjunto de entrenamiento en dos conjuntos que se llamarán de entrenamiento y de validación. El conjunto de entrenamiento se utiliza para aprender los valores de los pesos, como se ha descrito anteriormente. La diferencia es que en vez de medirse el error en el conjunto de entrenamiento, se utiliza el de validación. De esta manera para medir la eficacia de la red para resolver el problema, se utilizarán datos que no han sido utilizados para su aprendizaje. Si el error sobre el conjunto de validación es pequeño, entonces quedará garantizada la capacidad de generalización de la red.

Para que este proceso sea eficaz los conjuntos de entrenamiento y validación deben tener las siguientes características:

- El conjunto de validación debe ser independiente del de aprendizaje. No puede haber ningún tipo de sesgo en el proceso de selección de los datos de validación.
- El conjunto de validación debe cumplir las propiedades de un conjunto de entrenamiento, descritas anteriormente.

Además, el conjunto de validación puede utilizarse durante el aprendizaje para guiarlo en conjunto con el de entrenamiento. En este caso el proceso sería el siguiente:

- (1) Asignar a los pesos valores aleatorios.
- (2) Introducir todos los ejemplos del conjunto de entrenamiento, modificando los pesos de acuerdo con el esquema de aprendizaje supervisado elegido.
- (3) Introducir todos los ejemplos del conjunto de validación. Obtener el error producido al predecir dichos ejemplos.
- (4) Si el error calculado en el paso anterior está por encima de cierto valor umbral, ir a (2).
- (5) Acaba el proceso de aprendizaje y dar como salida la red obtenida.

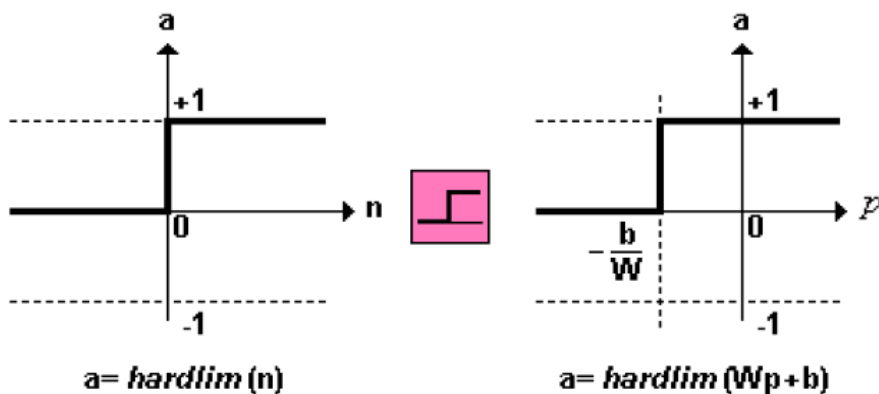
### 3.3.8 Funciones de Transferencia.

1. Función escalón o limitador fuerte (Hardlim)[23] : Para este tipo de función, la transferencia es:

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

Donde n es el campo local inducido de la neurona que es:

$$ak = \sum wk_j n_j + bk$$



Función de Transferencia

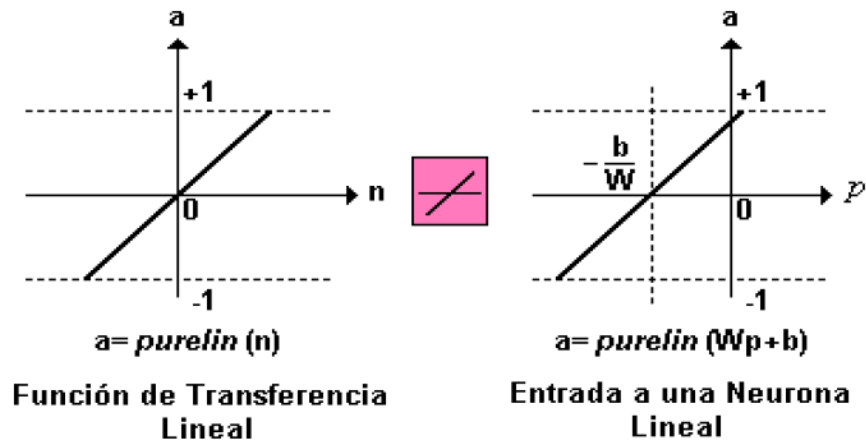
Limitador Fuerte

Entrada a una Neurona

Limitador Fuerte

2. Función Lineal (purelin) [23]. Para esta función se tiene:

$$a = f(n) = K * n$$

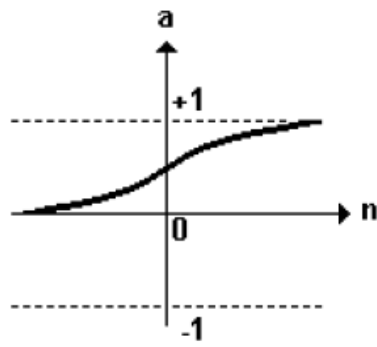


Donde K es el factor de amplificación en la región lineal de operación.

3. Funciones sigmoideas (logsig) [23]: Las funciones de transferencia sigmoideas son las funciones más comúnmente usadas en la construcción de redes neuronales, y están definidas por:

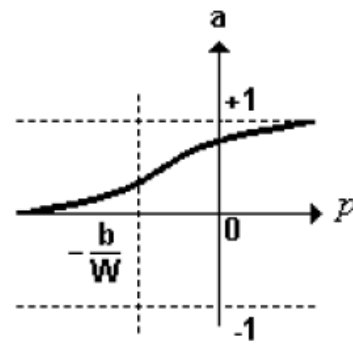
$$a = f(n) = \frac{1}{(1 + e^{-n})}$$

$$a = f(n) = \text{tansig}(n)$$



$$a = \text{logsig}(n)$$

**Función de Transferencia Logarítmica Sigmoidal**



$$a = \text{logsig}(Wp+b)$$

**Entrada a una Neurona Logarítmica Sigmoidal**

La tabla 3 muestra un resumen de las funciones de transferencias, sus relaciones de entradas/salidas y la función de transferencia asociada.

Nombre	Relación Entrada/Salida	Icono de la Función De transferencia	Función de Transferencia
Limitador Fuerte	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$		hardlim
Limitador Fuerte Simétrico	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$		hardlims
Lineal Positiva	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$		poslin
Lineal	$a = n$		purelin
Lineal Saturado	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$		satlin





Lineal Saturado Simétrico	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = +1 \quad n > 1$		satlins
Sigmoidal Logarítmico	$a = \frac{1}{1 + e^{-n}}$		logsig
Tangente Sigmoidal Hiperbólica	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
Competitiva	$a = 1$ <i>Neurona con n max</i> $a = 0$ <i>El resto de neuronas</i>		compet

Tabla 3: Relaciones de Transferencia

### 3.3.9 Modelos de redes neuronales

#### 3.3.9.1 Perceptrón

El perceptrón es una red de alimentación directa, esto es la información fluye desde la capa de entrada hacia la capa de salida. Fue desarrollado por F. Rosenblatt hacia final de la década de los cincuenta basándose en la regla de Aprendizaje de Hebb y de los modelos de neuronas biológicas de McCulloch y Pitts [30].

El perceptrón es un clasificador, asigna a un vector de N valores un valor binario, usando una transformación no lineal. Así cada vector pertenece a una de las particiones que crea el perceptrón.

El perceptrón es una máquina de computación universal y tiene la expresividad equivalente a la lógica binaria ya que se puede crear un perceptrón que tenga el mismo comportamiento que una función booleana NAND y a partir de esta función se puede crear cualquier otra función booleana.

#### 3.3.9.2 Arquitectura del perceptrón

El perceptrón simple es una red que consta de dos capas de neuronas. Esta red admite valores binarios o bipolares como entrada para los sensores y los valores de su salida están en el mismo rango que los de entrada.

La función de la primera capa es hacer de sensor, por ella entran las señales a la red. La segunda capa realiza todo el procesamiento. La manera de interconectar ambas capas es todas con todas, esto es, cada neurona de la primera capa esta unida con todas las de la segunda capa, como se muestra en la figura 14.

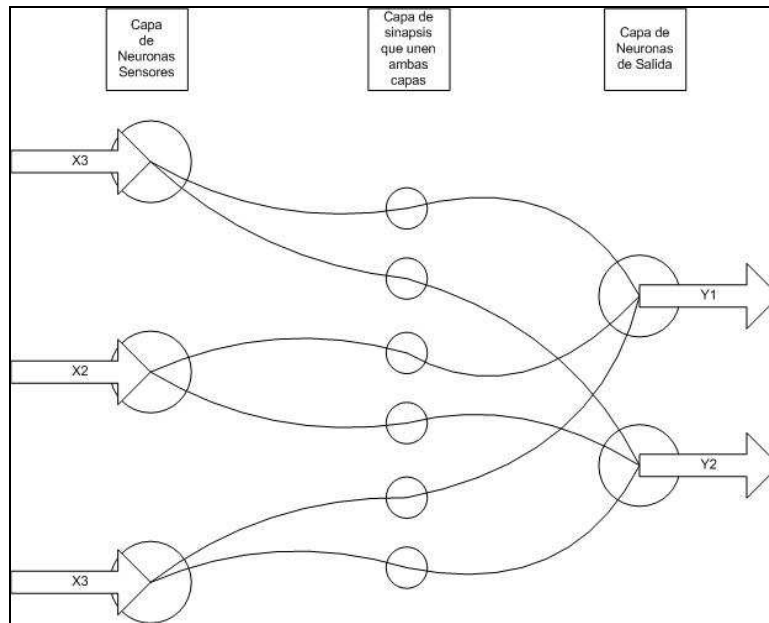


Figura 14: Arquitectura Red Perceptrón

### 3.3.9.3 Aprendizaje del perceptrón

Los pasos para que la red aprenda una lista de patrones son los siguientes:

- 1 Tomar un patrón al azar de la lista.
- 2 Se establece el patrón de entrada en los sensores, la capa de entrada.
- 3 Se establecen los valores deseados en las neuronas de la capa de salida.
- 4 Se actualizan las neuronas de la capa de salida.
- 5 Solicitar que aprendan todas las sinapsis
- 6 Si las sinapsis han cambiado volver al paso 1

Si no han cambiado la red se ha estabilizado y se para el aprendizaje.

Las sinapsis que une las neuronas  $i, j$  aprenderá de la siguiente manera:

Sea

- $E_i$  = el estado de la neurona de la capa de entrada  $i$ ,
- $P_{ij}$  = El peso actual asociado a la sinapsis que une la neurona  $i$  de la capa de entrada y la neurona  $j$  de la capa de salida
- $E_j$  = El estado de la neurona de la capa de salida  $j$
- $S_j$  = El valor deseado para esa neurona

- PijNuevo = El peso Pij actualizado
- TASA = Es una constante entre 0 y 1 que indica cuanto aprende la red

$$P_{ij\text{nuevo}} = P_{ij} + \text{Tasa} * ((E_j - S_j) * E_i)$$

Hay que destacar que el perceptrón aprende sólo cuando se equivoca al clasificar el patrón. Si clasifica correctamente el patrón, entonces, no existe aprendizaje.

### 3.4. Generalización

Una vez finalizada la fase de aprendizaje, la red puede ser utilizada para realizar la tarea para la que fue entrenada. Una de las principales ventajas que posee este modelo es que la red aprende la relación existente entre los datos, adquiriendo la capacidad de generalizar conceptos. De esta manera, una red neuronal puede tratar con información que no le fue presentada durante de la fase de entrenamiento.

Cuando se evalúa una red neuronal no sólo es importante evaluar si la red ha sido capaz de aprender los patrones de entrenamiento. Es imprescindible también evaluar el comportamiento de la red ante patrones nunca antes vistos.

Esta característica de las redes neuronales se la conoce como capacidad de generalización y es adquirida durante la fase de entrenamiento [37].

Es necesario que durante el proceso de aprendizaje la red extraiga las características de las muestras, para poder luego responder correctamente a nuevos patrones.

De lo dicho anteriormente surge la necesidad de evaluar durante la fase de entrenamiento dos tipos de errores. El error de aprendizaje, que indica la calidad de la respuesta de la red a los patrones de entrenamiento, y el error de generalización, que indica la calidad de la respuesta de la red a patrones nunca antes vistos. Para poder obtener una medida de ambos errores es necesario dividir el set de datos disponibles en dos, el set de datos de entrenamiento, y el set de datos de evaluación. El primero se utiliza durante la fase de entrenamiento para que la red pueda extraer las características de los mismos y, mediante el ajuste de sus pesos sinápticos, la red logre una representación interna de la función. El set de evaluación se utiliza para evaluar la capacidad de generalización de la red.

La causa más común de la pérdida de capacidad de generalización es el sobreaprendizaje. Esto sucede cuando la cantidad de ciclos de entrenamientos tiende a ser muy alta. Se observa que la respuesta de la red a los patrones de entrenamiento es muy buena mientras que la respuesta a nuevos patrones tiende a ser muy pobre. Al aumentar el número de ciclos la red tiende a sobreajustar la respuesta a los patrones de entrenamiento, a expensas de una menor capacidad de generalización. La figura 16 muestra una situación idealizada de lo dicho

anteriormente. En la misma se observa que en un determinado punto la red comienza a perder capacidad de generalización como consecuencia del sobreaprendizaje de los patrones de entrenamiento.

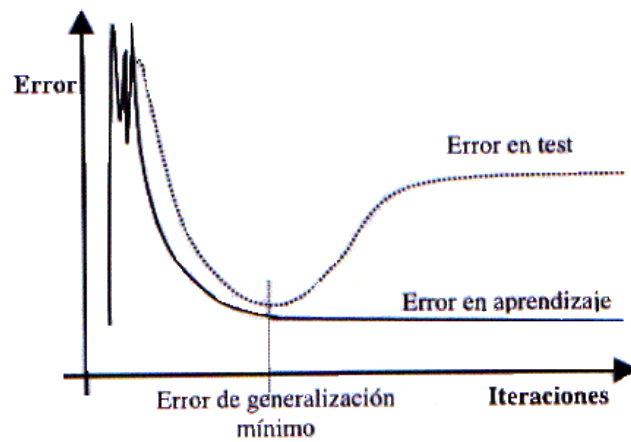


Figura 16: Generalización. Situación ideal

En la figura 17 se muestra una situación más real del mismo caso. A medida que transcurre el proceso de aprendizaje se obtienen varios mínimos sobre el conjunto de evaluación. Existen diversas técnicas de parada temprana (early stopping) aunque en la mayoría de los casos se deja que el proceso de aprendizaje avance hasta alcanzar una cota de error razonable, guardando periódicamente las distintas configuraciones intermedias para luego seleccionar la de menor error de evaluación.

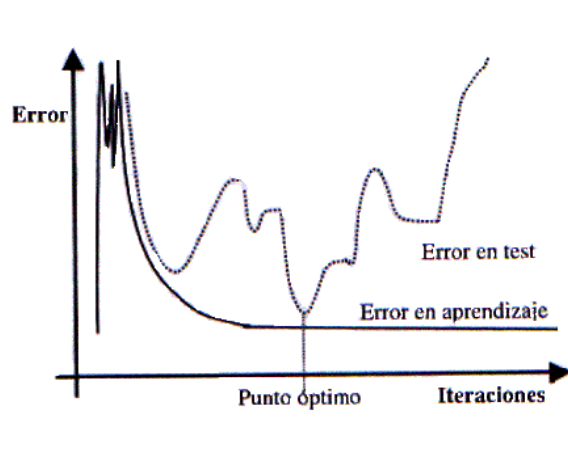


Figura 17. Generalización. Situación real

En ocasiones la pérdida de capacidad de generalización se produce por el uso excesivo de neuronas ocultas en la red neuronal. Esto hace que la red tienda a ajustar con mucha exactitud los patrones de entrenamiento, evitando que la red extraiga las características del conjunto. Este problema se ve agravado cuando los patrones de entrenamiento poseen ruido, ya que la red ajusta también el ruido de los mismos.



## 4. Desarrollo

### 4.1 Codificación de datos

Como se presentó en las etapas anteriores de este de este proyecto, se implementó un modelo de red neuronal que trabaja con datos numéricos, los cuales se entregan como entradas a la red.

En el caso de estudio, se tomó como base los datos entregados por una empresa particular (los cuales sirven como base para la implementación de este proyecto de titulación), en el cual el otorgamiento de créditos depende de las variables sociales presentadas a continuación. Dichos datos poseen variadas codificaciones.

La cantidad total de datos tomados para este proyecto son 500 (ver anexo 1), de los cuales 250 son para entrenamiento, y 250 para validación del entrenamiento.

Para comenzar, se dividieron en dos grandes grupos: Personales, que tiene relación con datos asociados directamente al cliente y su entorno; y el segundo grupo Datos Comerciales, que hacen alusión a datos relacionados con el comportamiento comercial del cliente.

La información originalmente fue digitada en MsExcel pero los subconjuntos necesariamente fueron exportados a dos archivos de texto distintos (\*.txt) para poder utilizarlos.

Los datos elegidos como entradas son las siguientes:

#### I. Datos Personales

- a) Edad: Este dato inicialmente está conformado por el día, mes y año de nacimiento del cliente en cuestión.

Para poder ser trabajado en la red, fueron cambiados a sólo la edad, y en rangos para poder establecer patrones en relación a esta variable. Los rangos utilizados son los siguientes:

- < 20 años = 1
- 21- 30 años = 2
- 31-40 años = 3
- 41-50 años = 4
- 51- 60 años = 5
- > 61 años = 6

- b) Sexo: Esta variable inicialmente es una variable alfanumérica, ya que la base de datos entrega si es hombre o mujer. Para este caso en particular, se codificó de la siguiente manera:

- Hombre = 1
- Mujer = 2

c) Estado Civil: Variable alfanumérica, que detalla el estado civil de la persona. Para este caso se codificó de la siguiente manera:

- Soltero = 1
- Casado = 2
- Separado = 3
- Viudo = 4

d) Profesión u Oficio: Esta variable es muy amplia debido a la gran gama de profesiones existentes. Para facilitar el manejo, la codificación se realizó de la siguiente manera:

- Arquitecto = 1
- Constructor Civil = 2
- Albañil = 3
- Médico = 4
- Sin profesión = 5
- Otro = 6

e) Teléfono: Variable numérica, que fue definida de la siguiente forma:

- Tiene teléfono = 1
- No tiene teléfono = 2

f) Comuna: Comuna de domicilio cliente:

- Quilpué = 1
- Villa Alemana = 2
- Peña Blanca = 3
- Limache = 4
- Viña del Mar = 5
- Valparaíso = 6
- Otras = 7

g) Propietario: Indica si cliente es propietario de algún bien:

- Es Propietario = 1
- No es propietario = 2

## II. Comerciales

h) Rubro: Área a la que se dedica:

- Construcción= 1
- Arquitectura= 2
- Dueña de Casa= 3
- Otros= 4

i) Cliente /Proveedor: Indica si persona es Cliente, Proveedor o ambos:

- Cliente = 1
- Proveedor = 2
- Ambos= 3

j) Primera Compra= Fecha en que se realizó la primera compra:

- Año 2002= 1
- Año 2003 = 2
- Año 2004 = 3
- Año 2005 = 4
- Año 2006 = 5
- Año 2007 = 6
- Año 2008 = 7
- Año 2009 = 8
- Cliente Nuevo = 0

k) Última compra = Fecha en que realizó última compra:

- Año 2002= 1
- Año 2003 = 2
- Año 2004 = 3
- Año 2005 = 4
- Año 2006 = 5
- Año 2007 = 6

- Año 2008 = 7
- Año 2009 = 8
- Cliente Nuevo = 0

l) Ticket promedio. Valor promedio de las compras realizadas:

- Sin compras = 0
- 0- \$ 50.000 = 1
- \$ 50001- \$100.000= 2
- \$ 100001- \$ 200000 = 3
- \$ 200001- \$ 300.000 = 4
- Más de \$ 300.001= 5

m) Productos compra: Indica los productos principalmente asociados a las compras:

- Materiales de Construcción = 1
- Terminaciones = 2
- Cerámicas = 3
- Ferretería = 4

n) Créditos Anteriores: Muestra si anteriormente ha pagado con cheque:

- Si = 1
- No = 2

o) Cheques en Cartera: Indica si posee cheques pendientes de cobro:

- Si = 1
- No = 2

p) Cheques Protestados: Indica si en algún momento se protestó algún cheque:

- Si = 1
- No = 2

q) Informes comerciales: Muestra si la persona posee informes comerciales actuales:

- Si = 1
- No = 2

De esta manera se configuró la planilla de datos para poder determinar las entradas a la red.  
La figura 18 muestra el aspecto de la red.

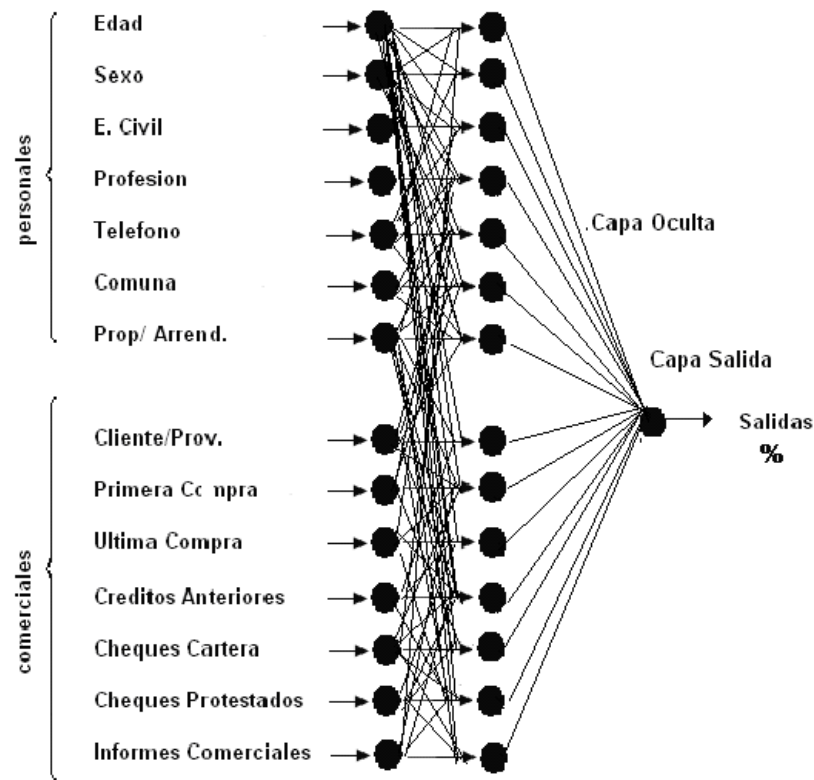


Figura 18: Estructura de la red.

## 4.2 Acerca del Software utilizado

Después de una ardua búsqueda por un software de redes neuronales (ya que casi la totalidad existente requiere de licencia y ésta es muy costosa) se optó por emplear el programa *ThinksPro-Neural Networks for Windows Trial Version 1.05* de *Logical Designs Consulting*.

Este programa de 1995, por antiguo que parezca presta todos los requerimientos necesarios para llevar a cabo el trabajo. Ofrece las facilidades de crear nuevos proyectos, crear redes, editarlas, parametrizarlas, medir el error al iterar, ofrecer y exportar los resultados.

Además se utilizaron alguna de las funcionalidad de MatLAB para pruebas

A continuación se mencionará paso a paso el proceso de creación y entrenamiento de la red.

#### 4.2.1 Proceso de Aplicación del modelo sobre los datos

Los conjuntos de datos tienen por denominación `entrenamiento1.txt` y `entrenamiento2.txt` y se pueden visualizar en los anexos.

#### 4.2.2 Creación de la red

Para crear una red neuronal nueva, en primera instancia se crea un “proyecto nuevo” tal y como se hace con un documento nuevo en un procesador de texto. La primera tarea a cumplir será diseñar y parametrizar una estructura neuro-artificial que concuerde con la lógica del problema en estudio y que utilice la regla de aprendizaje que se ha estudiado.

Como se puede observar en la figura 19, esa es la estructura o “esqueleto” de la red: 3 capas, una de entrada con 15 nodos, una intermedia también con 15 nodos y una de salida con un único nodo. Todos los nodos de la capa anterior se interconectan con todos los nodos de la capa posterior, como es usual en *backpropagation*, la que se explica a continuación.

La simple lógica explica por qué se tienen 15 nodos en la capa de entrada y 1 en la capa de salida: cada nodo de la capa de entrada representará la entrada de cada variable de la hoja de riesgo, esto es  $i_1, i_2, \dots, i_{15}$ ; y el nodo de la capa de salida representa el resultado del análisis crediticio  $dI$ . El número de capas ocultas y nodos en la misma provienen de razones un tanto distintas.

Con base en el número de capas ocultas y en el número de datos entregado como información a la red, se puede decir que en general, si el “tamaño del problema” es pequeño entonces no es necesario sobredimensionar la red con más nodos ni capas ocultas.

En todo caso se considera que el número de capas y nodos en la sección intermedia de la red que se ha utilizado en este problema han sido el máximo posible basados en la experiencia de trabajos de distintos autores.

Después de determinar la estructura se procede a escoger la regla de aprendizaje de la red y sus parámetros. De la hoja de “Network Setup” se elige “*backpropagation*” para la regla de aprendizaje definida anteriormente.

Como la capa de entrada sirve únicamente como “input buffer”, la única parametrización en esa capa es el preprocesamiento de entrada como MAX/MIN, esto sirve para transformar los datos hacia un intervalo decimal fijo considerando los valores máximos y mínimos de cada variable y recodificándolos al hacerlos variar dentro de esos intervalos.

Esto se hace debido a que presta facilidad a los cálculos y se acomoda mejor a la función de transferencia que es umbraloides en la capa oculta.

La figura 19 muestra el diseño de la capa de entrada. En la capa oculta se parametriza conforme lo antes establecido. Se tienen 15 nodos, se utiliza una suma ponderada y una función de transferencia sigmoide (establece una especie de umbral 0-1), para determinar el porcentaje de crédito asignado a cada persona.

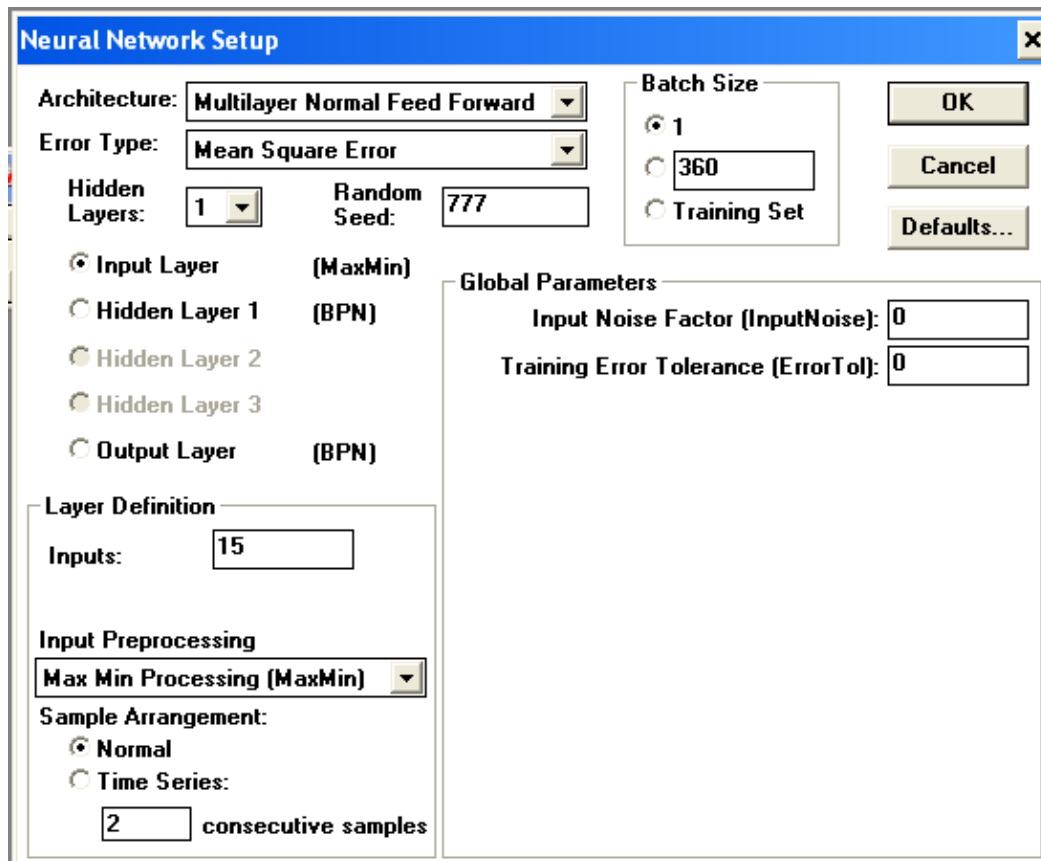


Figura 19: Diseño Capa de entrada

En la figura 20 se muestra además que las tasas de velocidad de aprendizaje y de momento antes discutidas se han elegido con valores de 0.09 y 0.07 respectivamente.

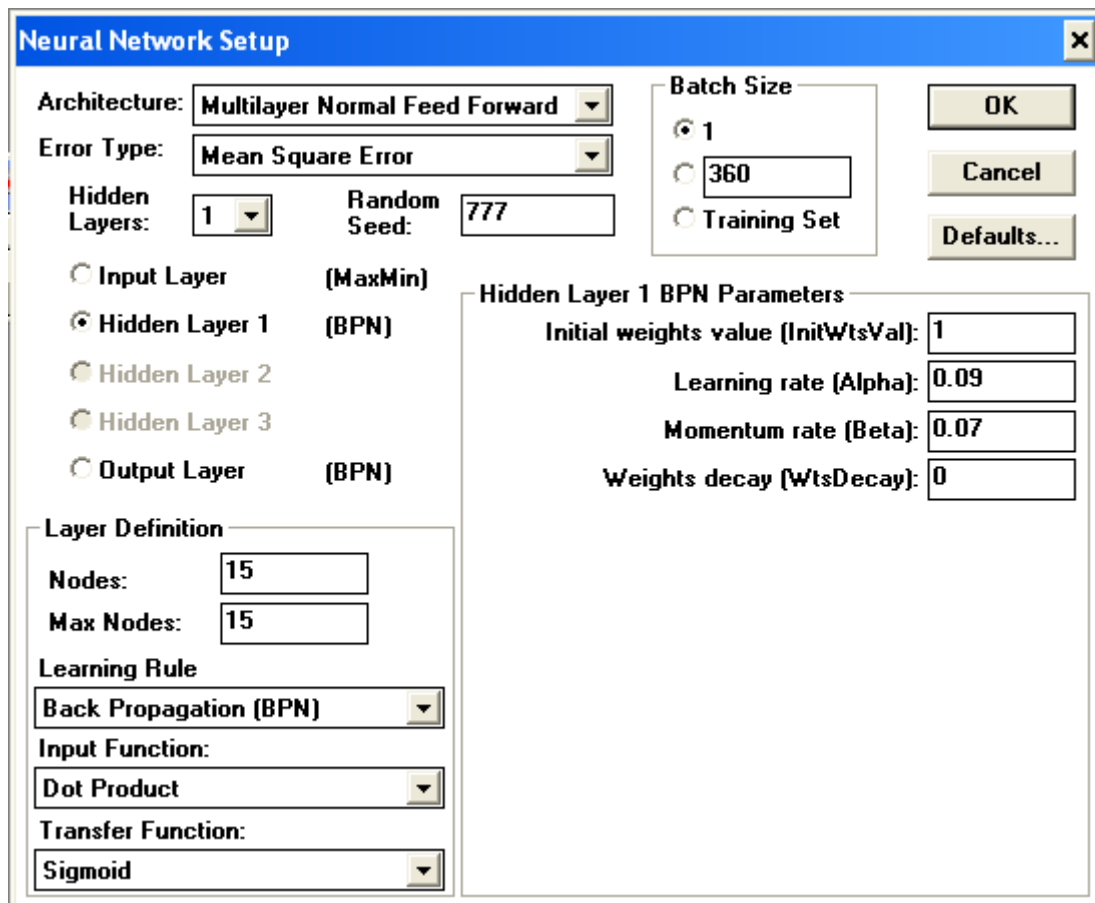


Figura 20: Diseño Capa Intermedia

Estos valores fueron elegidos conforme se hicieron ensayos hasta encontrarlos como los más adecuados al equilibrar la velocidad del aprendizaje y el riesgo de caer en mínimos locales.

En el primer intento por elegir una tasa de aprendizaje se optó por 0.01 y las pruebas indican que era una tasa muy baja ya que el aprendizaje era muy lento.

La mayor parte del aprendizaje la hace la capa oculta en este caso debido al mayor número de nodos que esta tiene con respecto a la capa de salida.

En la figura 21 se observa la parametrización de la capa de salida. Se ve un único nodo cuya función de resultado o transferencia también es una sigmoide. Nótese que en cada capa se puede elegir la regla y parámetros de aprendizaje independientemente de las otras capas, esto es de cierto modo una característica única del programa que se ha utilizado ya que la mayoría no permite tanta interactividad con el diseño de las redes.



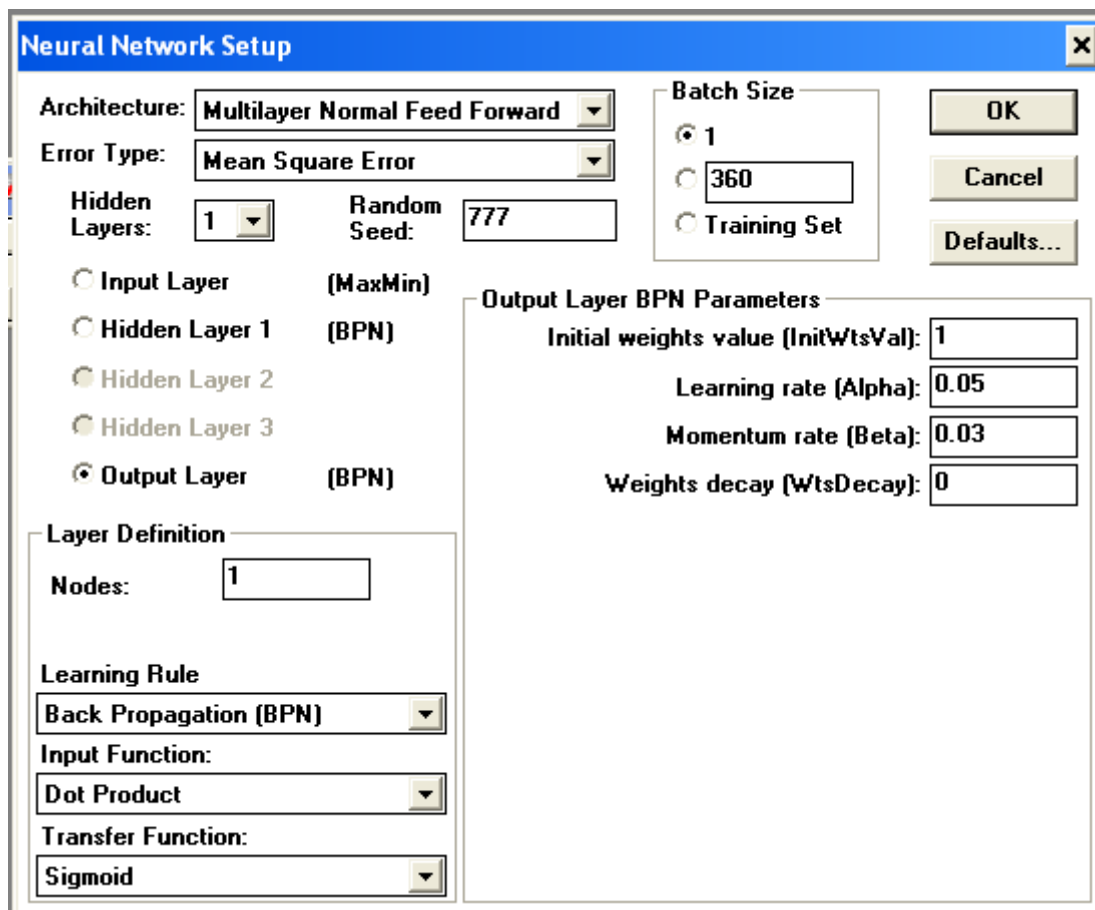


Figura 21: Diseño Capa Salida

Lo anterior se debe a que los creadores de los programas los hacen con el fin de quien tiene muy pocas nociones sobre el diseño de las redes neuronales puedan acceder a esta tecnología por medio de “asistentes virtuales” como los tiene por ejemplo PowerPoint al crear una presentación. Esto obviamente es desventajoso hasta cierto punto para quien desea total interactividad con el diseño debido a su familiaridad con el mismo. La ventaja es el ahorro de tiempo.

De la misma figura 21 nótese que los parámetros de velocidad y momento escogidos fueron 0.05 y 0.03 respectivamente. La razón de ser de estos valores es similar a la expuesta en la capa oculta, con la diferencia que en esta capa la velocidad del aprendizaje no es muy influenciada debido a la existencia de un único nodo.

#### 4.2.3 Importación de los datos

El procedimiento de importación de datos de entrenamiento y prueba es bastante sencillo. Lo que hay que cuidar es digitar la información previamente de tal manera que ThinksPro pueda reconocerla sin ningún problema. Por ejemplo no se debe grabar los datos con el encabezado como nombre de la variable ya que el programa lo hace por omisión.

Simplemente se optó por el comando *File / Import File / Training set* y *File / Import File / Test set* y para importar los archivos entrenamiento1.txt y entrenamiento2.txt que previamente se guardó desde MSEXcel. Después de cada importación la hoja de datos que al principio estuvo vacía se llena automáticamente con los datos importados, también las variables reciben su nombre automáticamente y ThinkPro considera de manera automática que la última variable es la de salida, por eso es necesario que se conserve ese orden al ingresar los datos.

En la figura 22 se observa el caso para los conjuntos de entrenamiento y prueba.

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	D1
1	1	1	1	1	1	3	6	1	2	2	2	1	5	1	2	2
2	1	2	5	1	2	2	2	1	5	1	2	2	1	5	6	2
3	2	2	2	1	5	1	4	2	1	6	7	1	2	2	2	1
4	5	1	1	1	1	5	5	1	2	2	2	1	2	1	5	3
5	1	3	8	1	2	2	2	1	4	1	7	1	1	0	0	2
6	2	2	1	0	5	1	7	2	2	8	8	2	2	2	2	1
7	2	1	7	1	1	6	8	2	1	2	2	1	2	1	2	1
8	1	2	3	2	1	2	2	1	2	1	2	2	1	2	6	2
9	2	2	2	1	2	1	1	2	1	4	6	2	2	2	2	0
10	1	1	1	2	1	5	7	2	2	2	2	1	4	1	7	2
11	1	1	4	2	2	2	2	1	4	1	7	2	2	4	5	1
12	2	2	2	1	1	1	7	2	2	3	4	1	2	1	2	1
13	3	1	5	1	1	1	8	1	1	2	1	1	6	1	5	1
14	1	4	6	1	1	2	1	0	6	2	6	1	1	0	0	2
15	1	2	1	0	6	2	3	2	1	8	8	2	1	2	2	1
16	6	2	2	2	1	3	8	1	1	1	2	1	6	1	2	1
17	2	8	8	2	2	1	2	1	6	2	3	1	2	4	7	2
18	2	1	2	1	2	2	1	2	1	8	8	2	2	2	2	1
19	4	1	1	2	1	8	8	2	2	2	2	1	6	1	7	1
20	1	6	7	2	2	2	2	1	5	1	6	1	1	2	7	2
21	2	2	2	1	6	1	2	1	1	3	6	2	2	2	2	1
22	3	1	5	1	1	4	5	2	2	2	2	1	2	1	6	2
23	1	8	8	1	2	2	2	1	1	1	1	2	1	8	8	1
24	2	2	2	1	5	1	4	2	2	8	8	1	2	2	2	1

Figura 22: Caso de entrenamiento.

#### 4.2.4 Entrenamiento de la red

Inicialmente es posible probar la red o entrenarla (opción abierta que el programa permite) pero suena ilógico probar una red que no ha sido entrenada. Entonces se procede a habilitar la opción de entrenamiento. Para esto se eligió un valor mínimo del error de 0.028 como criterio de convergencia. Se ha definido también que el número de iteraciones sea infinito hasta que se encuentre bajo convergencia.

Se indica *Run* para el conjunto de entrenamiento y se procede a entrenar la red. En la figura 23 se puede ver que el error baja rápidamente hasta que encuentra el criterio de convergencia. La línea inferior representa el error al que se refiere: *ETrn* o “error de entrenamiento”.

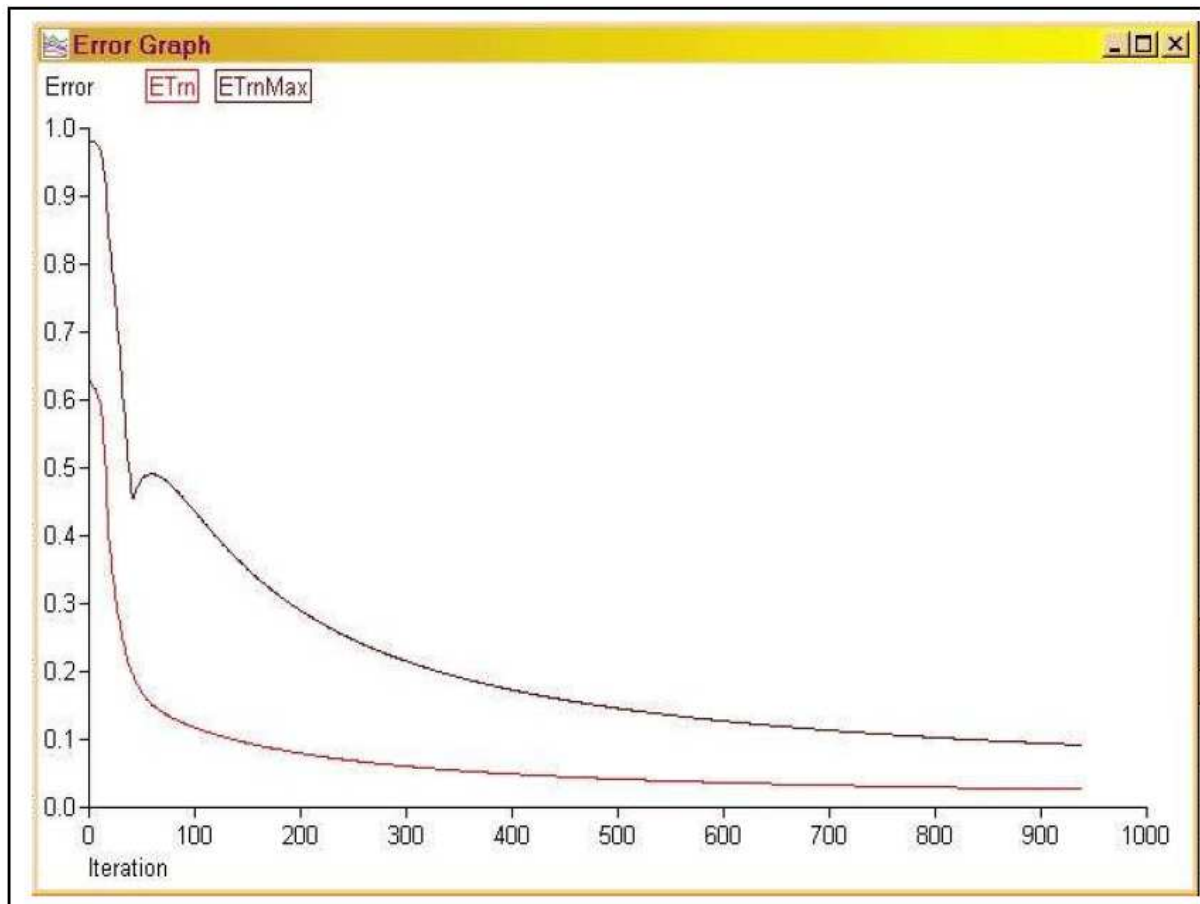


Figura 23: Muestra error criterio de convergencia.

### 4.3 Resultados Obtenidos

Los resultados obtenidos, presentados en anexo B, muestran que luego de la fase de entrenamiento, es posible obtener resultados con los cuales se pueda tomar una buena decisión para crédito, proporcionando una base similar a la entregada por el sistema convencional de análisis crediticio, pero en un tiempo reducido, y con procesos automáticos, que pueden ser controlados por solo una persona, y no por un equipo de trabajo completo, como se necesita para un análisis manual.

El funcionamiento de la red se puede mejorar, si los pesos de ésta se optimizan entrenando la red nuevamente una vez que el número óptimo, o cercanamente óptimo, de nodos de la capa oculta ha sido determinado. El porcentaje de reconocimiento se puede incrementar optimizando los pesos de la red, por medio de un proceso de reentrenamiento de las matrices.

### 4.3.1 Análisis de Sensibilidad

Una de las críticas más importantes que se han lanzado contra el uso de RNA trata sobre lo difícil que es comprender la naturaleza de las representaciones internas generadas por la red para responder ante un problema determinado. A diferencia de los modelos estadísticos clásicos, no es tan evidentes conocer en una red la importancia que tiene cada variable predictora sobre la/s variable/s dependientes/s. Sin embargo, esta percepción acerca de las RNA como una compleja “caja negra”, no es del todo cierta. De hecho, han surgido diferentes intentos por interpretar los pesos o parámetros del modelo, de los que el más ampliamente utilizado es el denominado análisis de sensibilidad.

La sensibilidad es el porcentaje de verdaderos positivos de un instrumento diagnóstico, mientras que el análisis de sensibilidad es el procedimiento para conocer el efecto o influencia de cada variable predictora sobre la/s variable/s dependiente/s.

El método más común para realiza un análisis de sensibilidad, y el utilizado en este proyecto, consiste en fijar el valor de todas las variables de entrada a su valor medio e ir variando el valor de una de ellas a lo largo de todo su rango, con el objeto de observar el efecto que tiene sobre la salida de la red. Siguiendo este método, se fue registrando los cambios que se producían en la salida de la red, cada vez que se aplicaba un pequeño incremento n-incrementos de un 2%, en una variable de entrada. Se propuso como objetivo cuantificar la influencia que tiene cada variable de entrada. La simple suma de los cambios producidos proporcionara una medida intuitiva de sensibilidad. Esta medida representa el efecto relativo que tiene una variable de entrada sobre la salida de la red. Así, un valor cercano a 0 indicaría poco efecto o sensibilidad; a medida que se fuese alejando de 0, indicaría que el efecto va aumentando. Esta medida de sensibilidad se obtuvo mediante la siguiente expresión:

$$S_{ik} = \sum_n^N |X_{kn} - X_{k \min}|$$

Donde

$S_{ik}$ = medida de sensibilidad de la variable de entrada i sobre la salida k.

$X_{kn}$ = valor de la salida k obtenido con el incremento n en la variable de entrada i.

$X_{kmin}$ = valor de la salida k obtenido con el valor mínimo posible de la variable de entrada i.

A partir del modelo general inicialmente entrenado, se obtuvo el valor de la medida de sensibilidad para cada variable predictora de riesgo crediticio. En la tabla se presentan los valores de mayor a menor. Así, los primeros valores de la tabla corresponden a las variables de entrada con más influencia o relación con la salida de la red.

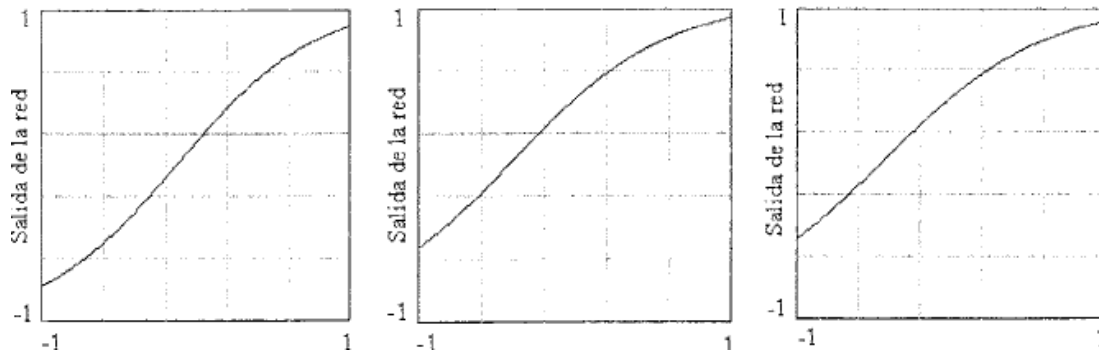
En la tabla 4 se pueden observar todas las variables asociadas a la sensibilidad de riesgo crediticio, ordenadas en relación de influencia mayor a menor.

<b>Medida de sensibilidad de las variables asociadas a riesgo crediticio</b>	
<b><i>Variable</i></b>	<b><i>Sensibilidad</i></b>
Informes Comerciales	58.93
Créditos Anteriores	43.18
Cheques Protestados	26.50
Cheques en Cartera	22.47
Profesión	16.94
Edad	12.31
Propietario /Arrendatario	11.03
Teléfono	9.42
E. Civil	6.84
Ultima Compra	4.93
Primera Compra	3.89
Cliente/Proveedor	2.80
Sexo	2.77
Comuna	1.08

Tabla 4: Medida de Sensibilidad de las variables

Cabe destacar, que según la información contenida en la tabla 4, las variables que tienen mayor influencia en el otorgamiento o no del crédito son aquellas variables indicativas de antecedentes comerciales, y no los personales propiamente tal, como era de esperarse. Es por esta razón, que se debe poner especial interés en estos aspectos, y no descuidarlos al momento del análisis.

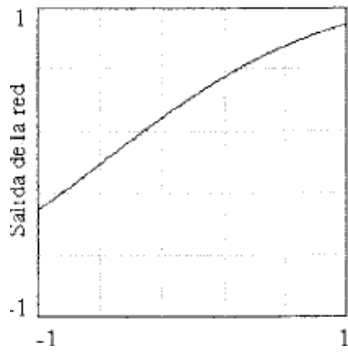
Por último, para obtener una información más completa, la medida de sensibilidad debe ir acompañada de la correspondiente representación gráfica. A continuación se muestra la representación grafica del análisis de sensibilidad de las 14 variables predictoras de la muestra.



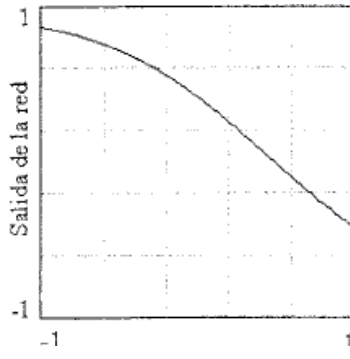
Edad

Sexo

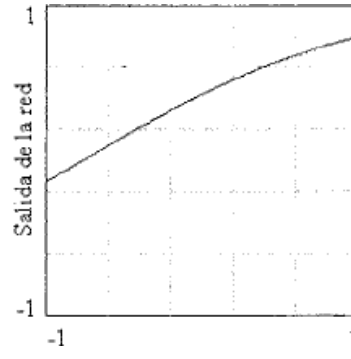
E. civil



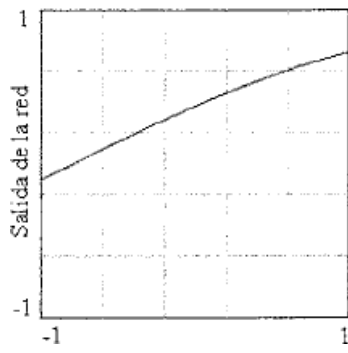
Profesión



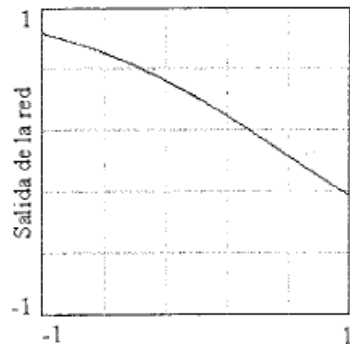
Teléfono



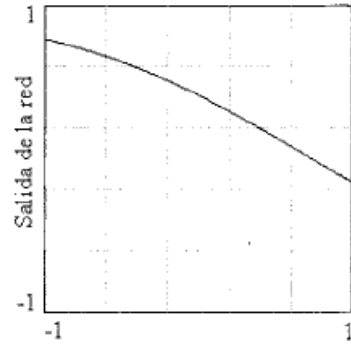
Comuna



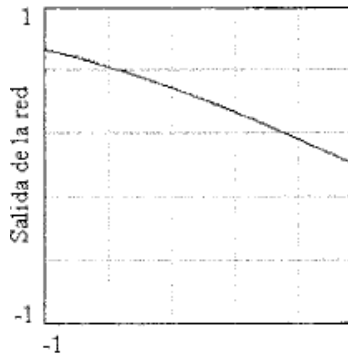
Prop./ Arrend.



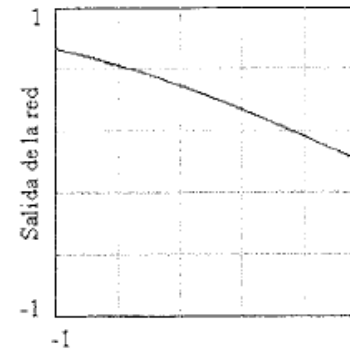
Cliente/proveed.



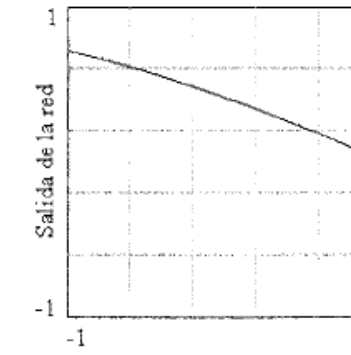
Primera Compra



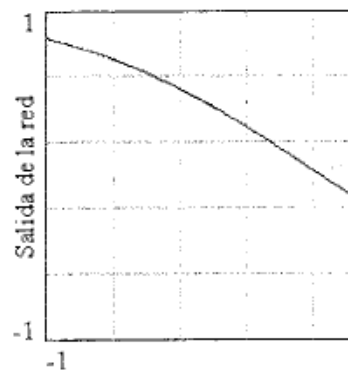
Ultima Compra



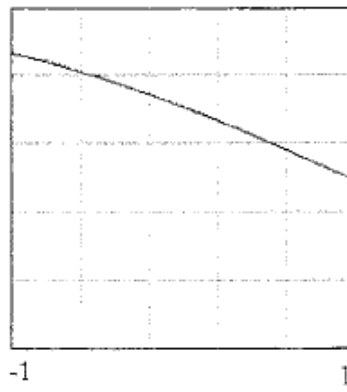
Cheq. Anteriores



Cheq. Cartera



Cheq. Protestados



Inform. Comerciales

En base a la información presentada anteriormente, podemos definir un perfil, para el cual la red asigna crédito sin problemas, y un perfil para el cual no se debiese asignar créditos.

### **Perfil Apto Crédito.**

Primeramente, se debe considerar que el comportamiento anterior de la persona que requiere crédito es de vital relevancia, ya que si la persona posee algún tipo de informes comerciales negativos, cheques protestados, o cheques pendientes de pago, de inmediato las opciones de asignación de crédito disminuyen.

Bajo esta premisa, podemos definir algunas condiciones que pueden servir de patrón al momento de identificar, como primera impronta, si la persona es apta para ser analizada como posible receptor de crédito:

- Persona responsable en temas comerciales: sin antecedentes comerciales ni cheques protestados.
- Profesional: una persona con un título profesional presenta una base más sólida que permite asumir una cierta solvencia económica que permitiría el pago del crédito.
- Propietario de algún bien: este hecho también puede ser considerado como una base de solvencia para futuras respuestas a créditos. Además representa un bien sobre el cual se puede solicitar prenda para pago.

### **Perfil No Apto Crédito**

Se debe considerar que personas con un mal comportamiento crediticio anterior, automáticamente restan muchas posibilidades de asignación de crédito.

Condiciones que calificarían a persona como no apto para crédito.

- Persona que posea antecedentes comerciales negativos en cuanto a comportamiento crediticio anterior, ya sea con la empresa a la que se solicita el crédito actual, o en cualquier otra.
- Persona que no tenga un trabajo estable o profesión en ejercicio: ya que se puede considerar que no poseerá la solvencia necesaria para cumplir con los compromisos asumidos.
- Persona que no tenga antecedentes actualizados: si no se tienen antecedentes actuales, debido a que hace mucho tiempo que no tiene movimiento la cuenta, también es un indicio que se debiese solicitar antecedentes actuales para poder analizar.

## 5. Conclusiones.

A lo largo del desarrollo de este proyecto queda claro que el proceso de riesgo crediticio no es un proceso sencillo, se deben considerar diferentes factores, que sobre todo en estos tiempos son de gran importancia al momento de la toma de decisión de otorgar o no el crédito.

En cuanto a las redes neuronales, se logró comprender que poseen un grado de dificultad que de manera inicial no es sopesado; pero en conjunto con la dificultad presentan un gran potencial para la solución de un sin número de problemáticas en diferentes áreas de investigación.

En términos técnicos, se pudo probar que la modificación de los pesos entregan un acercamiento acertado de la influencia de cada variable en el modelo. En este caso en particular que se analizó, se pueden generar patrones de comportamiento adecuados para predecir el comportamiento de los sujetos con características acordes a las entregadas a la red.

Las redes neuronales poseen aristas que eran desconocidas de manera inicial, pero con el estudio realizado, queda claro que era absolutamente necesario este proceso.

Estas técnicas vienen a apoyar y a automatizar todo el proceso realizado antes manualmente en base a métodos estadísticos e inferencia de datos.

En relación al análisis del proyecto como tal, cabe mencionar como tema personal que fue un proceso largo, y ratos un tanto desmotivado al no lograr comprender cabalmente el tema de la inmensa capacidad de las redes neuronales.

En el análisis final del proyecto, aparecieron sorpresas tales como la sensibilidad de las variables, que en primera instancia, y sólo por un tema de percepción, se creyó que serían diferentes, pero las herramientas utilizadas lograron ser un gran aporte al tema, ya que para personas que están recién comenzando en estas áreas entregan ayuda clara y concisa de los pasos a seguir para su utilización.

Como conclusión general se puede mencionar que las redes neuronales, sin duda representan una gran oportunidad para las empresas de hoy en día, que pueden tomar decisiones sobre datos más certeros y concretos, y que facilitan de gran manera el trabajo de la alta gerencia de las empresas.



## 6. Referencias

- [1] Libro: "Introducción al Análisis de Decisiones Financieras"  
Autor: Ricardo Pascale  
Editorial: Ed. Contabilidad Moderna, 1985.
  
- [2] Libro: "Análisis de Riesgo Crediticio"  
Autor: José Miguel Bulnes.  
Editorial: McGraw-Hill, 1995.
  
- [3] Paper: "Riesgo Crediticio"  
Autor: Manfredo Añez  
Santa Cruz- Bolivia, 2001.
  
- [4] Paper: "Effects of economic interactions in credit risk"  
Autor: J P L Hatchett and R Kühn  
Londres- Inglaterra, 2006.
  
- [5] Paper: "Métodos de la Ingeniería Informática Avanzada".  
Autor: R. García Martínez.  
Instituto de Tecnología, Buenos Aires- Argentina, 2003.
  
- [6] Libro: "Introduction to Neural Networks.  
Autor: Kevin Gurney.  
Editorial: Routledge, 1997.
  
- [7] Paper: "Entrenamiento de Redes Neuronales basadas en algoritmos evolutivos".  
Autor: Luis Federico Bertona.  
Universidad de Buenos Aires- Argentina, 2005.
  
- [8] Tesis: "Las Redes Neuronales artificiales como una alternativa o complementaria a los métodos de clasificación tradicionales".  
Autor: Alma Aguilar Sánchez.  
Universidad: Universidad Nacional Autónoma de México.- Agosto 2008.
  
- [9] Libro: "Redes Neuronales y Sistemas Borrosos".  
Autor: Martín del Brio.  
Editorial: Ra-Ma, 2006.

- [10] Libro: "Fundamentals of neural network".  
Autor: L. Fausset.  
Editorial: Prentice-Hall, 1994.
- [11] Paper: "Diseño de redes neuronales artificiales mediante computación evolutiva".  
Autor: P. Gutiérrez..  
Universidad de Granada- España, 2007.
- [12] Paper: "Transactions of neural network"  
Autor: Derong Liu.  
Institute of Automatin Chinese Academy of Sciences- Beijing, 2008.
- [13] Paper: "Configuration of Continuous Piecewise-Linear Neural Networks".  
Autor: Shuning Wang, Xiaolin Huang, and Khan M. Junaid.  
Tsinghua Univ., Beijing , 2008.
- [14] Paper: "Connectionism and Cognitive Architecture: A Critical Analysis".  
Autor: Jerry A. Fodor and Zenon W. Pylyshyn  
Rutgers University, NewYork- EEUU, 2003.
- [15] Paper: "Artificial Neural Networks: A Tutorial".  
Autor: Anil K Jain, Jianchang Mao  
Michigan State University, 1996
- [16] Paper: "An optimization methodology for neural network weights and architectures".  
Autor: T. B. Ludermir, A. Yamazaki, and C. Zanchettin  
Univ. Fed. de Pernambuco, 2006.
- [17] Libro: "Redes neuronales artificiales: fundamentos, modelos y aplicaciones".  
Autor: J. R. Hilera, V.J. Martínez.  
Editorial: Ra-Ma, 1995.
- [18] Libro: "Introduction to artificial intelligence and expert systems"  
Autor: D. Patterson.  
Editorial: Prentice-Hall, 1990.
- [19] Paper: "Artificial neural networks: past, present and future: an overview of the structure and training of artificial learning systems",  
Autor: Johnson, J.D. y Whinston, A.B.  
Johnson, J.D. y Whinston, A.B, 1994.

- [20] Paper: "A learning algorithm for Boltzmann machines. *Cognitive Science*"  
Autor: Ackley, D.H., Hinton, G.E. y Sejnowski, T.J.  
The Johns Hopkins University, 1985.
- [21] Paper: "Optimal convergence of on-line backpropagation"  
Autor: M. Gori and M. Maggini,  
*IEEE Trans. Neural Netw.*, vol. 7, no. 1, pp. 251–254, Jan. 1996.
- [22] Paper: "Global Convergence and Limit Cycle Behavior of Weights of Perceptron"  
Autor: Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Hak-Keung Lam, and Muhammad H. U. Nasir  
Faculty of Technology ePrints Service, 2009.
- [23] Paper: "Neural networks and related methods for classification".  
Autor: Ripley, B.D  
*Journal of the Royal Statistical Society*, 1994.
- [24] Libro: "Redes Neuronales: Algoritmos, Aplicaciones y Técnicas de Propagación."  
Autor: Freeman J, and Skapura D.  
Editorial: Díaz de Santos, 1993.
- [25] Libro: "Artificial Neural Systems, Foundations, Paradigms, Applications and Implementations".  
Autor: Patrick K.Simpson,.  
Editorial: Mc Graw-Hill, 1990.
- [26] Libro: "*Introduction to the theory of neural computation.Neurocomputing*".  
Autor: Addison-Wesley, Hertz, J., Krogh, A. y Palmer, R.  
Editorial: Westview Press, 1991.
- [27] Paper: "An adaptative "Adaline" neural".  
Autor: Widrow,  
Dept. of Math. & Comput. Sci., R. Mil. Coll. of Canada, Kingston, Ont., Canada, 1990
- [28] Paper: "Multilayer Perceptrons: Approximation Order and Necessary Number of Hidden Units"  
Autor: Stephan Trenn,  
2008
- [29] Paper: "Dynamics of Learning in Multilayer Perceptrons Near Singularities"  
Autor: Florent Cousseau, Tomoko Ozeki, and Shun-ichi Amari.  
*Life Fellow, IEEE*, 2008

[30] Paper: “Beyond Feedforward Models Trained by Backpropagation: A Practical Training Tool for a More Efficient Universal Approximator”

Autor: Roman Ilin, Robert Kozma.

Dept. of Comput. Sci., Univ. of Memphis, Memphis, TN

[31] Tesis: “Redes Neuronales Artificiales aplicadas al Análisis de Datos”.

Autor: Juan José Montaña Moreno.

Universitat de les Illes Balears, 2002.

## Anexo A: Datos

profesión	oficio	teléfono	comuna	prop.	cliente/ proveedor	primera compra	última compra	cred. anter.	cheq. cartera	cheq. protest.	inform. comerc.	crédito
	1	1	1	1	1	3	6	1	2	2	2	1
	5	1	2	2	1	2	5	1	2	2	2	1
	5	1	2	2	1	5	6	2	2	2	2	1
	5	1	4	2	1	6	7	1	2	2	2	1
	5	1	1	1	1	5	5	1	2	2	2	1
	2	1	5	3	1	3	8	1	2	2	2	1
	4	1	7	1	1	0	0	2	2	2	1	0
	5	1	7	2	2	8	8	2	2	2	2	1
	2	1	7	1	1	6	8	2	1	2	2	1
	2	1	2	1	1	2	3	2	1	2	2	1
	2	1	2	2	1	2	6	2	2	2	2	1
	2	1	1	2	1	4	6	2	2	2	2	0
	1	1	1	2	1	5	7	2	2	2	2	1
	4	1	7	2	1	1	4	2	2	2	2	1
	4	1	7	2	2	4	5	1	2	2	2	1
	1	1	7	2	2	3	4	1	2	1	2	1
	3	1	5	1	1	1	8	1	1	2	1	1
	6	1	5	1	1	4	6	1	1	2	1	0
	6	2	6	1	1	0	0	2	1	2	1	0
	6	2	3	2	1	8	8	2	1	2	2	1
	6	2	2	2	1	3	8	1	1	1	2	1
	6	1	2	1	2	8	8	2	2	1	2	1
	6	2	3	1	2	4	7	2	2	1	2	1
	2	2	1	2	1	8	8	2	2	2	2	1
	4	1	1	2	1	8	8	2	2	2	2	1
	6	1	7	1	1	6	7	2	2	2	2	1
	2	2	6	1	1	2	7	2	2	2	2	1
	6	1	2	1	1	3	6	2	2	2	2	1
	3	1	5	1	1	4	5	2	2	2	2	1
	2	1	6	2	1	8	8	1	2	2	2	1
	1	1	1	2	1	8	8	1	2	2	2	1
	5	1	4	2	2	8	8	1	2	2	2	1
	5	2	7	2	2	4	6	1	2	2	2	1
	5	1	4	2	1	3	8	1	2	2	1	1
	5	2	7	2	1	2	8	1	2	2	2	1
	2	2	1	1	1	3	4	1	2	2	2	0
	4	1	1	2	1	7	7	2	2	2	2	1
	5	1	2	2	1	0	0	2	2	2	2	1
	2	1	2	2	1	4	8	2	2	2	2	1
	2	1	5	1	1	3	4	2	2	1	2	1
	2	2	4	1	1	8	8	2	1	1	1	1
	2	1	2	1	1	6	8	1	1	1	1	1

1	1	1	1	1	5	6	1	1	1	2	1
4	2	1	2	1	4	7	1	2	1	2	1
4	2	1	2	1	6	8	1	1	1	2	0
1	1	1	2	1	7	8	1	1	2	2	1
3	2	2	1	1	8	8	1	2	2	2	1
6	1	2	2	1	3	7	2	1	2	2	1
6	1	6	2	1	8	8	2	1	2	2	1
6	1	4	1	3	4	7	1	1	2	2	1
6	1	3	1	1	8	8	1	1	2	2	1
6	1	3	1	1	3	8	2	2	2	2	1
6	1	2	2	1	7	7	2	2	1	2	0
2	1	2	2	1	6	7	1	2	2	2	0
4	1	2	2	1	5	6	2	2	2	2	0
6	1	2	1	1	3	8	1	2	2	2	1
5	2	2	2	1	4	5	2	2	1	1	0
6	1	2	1	1	4	8	1	2	2	1	0
3	2	7	1	1	0	0	2	2	2	2	0
2	2	7	1	1	4	5	1	2	2	2	1
1	1	7	1	1	7	7	1	2	2	2	1
5	1	7	2	3	3	5	1	2	2	2	1
5	1	4	2	1	7	8	1	2	2	2	0
5	1	1	2	1	8	8	1	2	2	2	1
5	1	4	2	1	3	5	1	2	2	2	1
2	1	6	2	2	5	6	1	2	2	2	1
4	2	5	2	1	6	8	2	2	2	2	1
5	1	1	2	1	5	8	2	2	2	2	1
2	1	1	2	1	3	4	2	2	2	1	1
2	2	2	2	1	7	7	1	2	2	1	0
2	1	2	2	1	8	8	1	2	2	1	1
2	1	4	2	1	1	5	1	2	2	2	1
1	1	1	2	2	2	3	1	2	1	2	1
4	1	5	1	2	8	8	1	2	2	2	1
4	1	7	1	1	1	8	2	2	2	2	0
1	1	7	1	1	2	2	2	2	2	2	1
3	1	7	1	1	5	6	2	2	1	2	1
6	1	2	2	1	7	8	2	2	1	2	1
6	1	2	2	2	7	8	2	1	2	2	1
6	1	1	2	1	7	7	2	2	2	2	1
6	2	1	2	1	7	7	2	2	2	2	0
6	1	7	2	1	1	4	2	2	2	2	0
6	1	7	2	1	2	6	2	2	2	2	1
2	1	7	2	1	8	8	2	2	2	2	1
4	1	5	1	1	4	6	2	2	1	2	1
6	1	5	1	1	7	8	2	2	2	1	1
5	1	6	1	1	0	0	2	2	1	2	1
6	2	3	2	1	1	7	2	2	2	2	1
3	2	2	2	2	7	8	2	1	2	2	1

2	2	2	2	2	2	2	1	1	2	2	1
1	1	3	2	2	3	6	1	1	2	2	1
5	1	1	3	2	2	3	1	2	2	2	1
5	1	1	3	1	8	8	1	2	2	2	1
5	1	7	2	1	6	8	2	2	2	2	1
5	1	6	1	1	7	8	1	2	2	2	1
2	1	2	1	1	4	4	2	2	2	1	1
4	1	5	2	1	6	7	1	2	2	1	1
5	1	6	2	1	1	5	2	2	2	1	1
2	1	1	2	1	1	3	2	2	2	2	0
2	1	4	1	1	6	8	2	2	2	2	1
2	1	7	3	1	1	1	2	2	2	2	1
2	2	4	1	1	2	2	1	2	2	2	1
1	2	7	2	1	7	7	1	2	2	2	1
4	1	1	1	1	3	4	1	2	2	2	1
4	1	1	1	1	4	8	1	1	2	2	1
1	1	2	2	1	0	0	1	2	2	2	1
3	2	2	2	3	5	7	2	1	2	2	1
6	1	5	2	1	2	4	2	2	1	2	0
6	1	4	2	1	8	8	2	2	2	2	1
6	2	2	2	1	3	3	2	2	2	2	1
6	1	1	2	3	1	6	1	2	2	2	1
6	1	1	1	1	2	7	1	2	2	2	1
6	1	1	1	1	2	4	1	2	1	1	1
2	1	1	1	1	4	8	1	2	1	2	1
4	1	2	2	1	3	4	1	2	1	2	1
6	2	2	2	1	2	3	2	2	2	2	0
5	1	6	1	1	4	7	2	2	2	2	0
6	1	4	1	1	2	2	2	2	2	2	0
3	1	3	2	1	7	8	2	2	2	2	1
2	1	3	2	1	6	6	2	2	2	1	0
1	1	2	1	1	8	8	2	2	2	1	0
4	1	2	1	1	2	8	2	2	2	2	0
4	1	2	1	1	0	0	2	2	2	2	1
1	1	2	1	1	0	0	2	2	2	2	1
3	1	2	2	1	1	3	2	1	2	2	1
6	2	2	2	1	2	8	1	1	2	2	0
6	1	7	2	1	7	8	1	1	2	2	1
6	2	7	2	1	7	7	1	2	2	2	1
6	1	7	2	1	6	7	1	1	2	2	1
6	1	7	2	1	4	4	1	1	2	2	1
6	1	4	1	2	1	1	1	1	2	2	1
2	2	1	2	1	1	1	1	2	1	2	1
4	1	4	2	1	2	5	1	2	1	2	0
6	1	6	2	2	8	8	2	2	1	2	1
5	1	5	1	2	7	8	2	2	1	2	1
6	1	1	1	1	5	7	2	2	1	1	1

3	2	1	1	1	4	7	2	1	1	1	1
2	1	2	1	1	6	7	2	2	2	2	0
1	1	2	2	1	4	8	2	2	2	2	1
5	1	4	2	2	8	8	1	2	2	2	1
5	1	1	2	1	6	7	1	2	2	2	1
5	1	5	1	1	6	6	1	1	2	2	1
5	1	7	2	1	6	7	1	1	2	2	1
2	1	7	2	1	2	4	1	2	2	2	0
4	1	7	1	1	0	0	1	2	1	2	0
5	1	2	1	1	1	2	1	2	2	2	1
2	1	2	1	1	8	8	2	2	2	2	1
2	1	1	2	1	2	2	2	2	2	1	1
2	1	1	2	1	6	6	2	2	1	1	1
2	1	7	2	2	4	8	2	2	2	1	1
1	1	7	1	2	7	8	2	2	2	2	1
4	1	7	2	2	4	4	1	2	2	2	1
4	1	5	1	2	8	8	1	2	2	2	1
1	1	5	1	1	4	7	1	2	2	2	1
3	1	6	1	1	3	4	1	2	2	2	1
6	1	3	1	3	7	7	1	2	2	2	1
6	1	2	2	1	8	8	1	2	2	2	1
6	2	2	2	1	0	0	1	2	2	2	1
6	1	3	2	1	4	5	2	2	2	2	1
6	2	1	2	1	0	0	2	2	2	2	1
6	1	1	2	3	5	8	2	2	2	2	1
2	2	7	2	1	5	6	2	2	2	2	0
4	2	6	2	1	4	8	2	1	2	2	1
6	1	2	2	1	6	6	1	1	2	2	1
5	1	5	2	1	8	8	1	1	1	1	1
6	1	6	2	3	5	5	2	2	2	2	1
3	2	1	2	1	4	5	2	2	2	2	1
2	1	4	2	1	2	6	1	2	2	2	1
1	1	7	1	1	8	8	1	2	1	2	1
4	1	4	1	3	5	6	1	2	1	2	1
4	1	7	1	1	3	3	2	2	2	2	0
1	1	1	1	1	7	8	2	2	2	2	1
3	1	1	2	1	5	5	2	2	2	2	1
6	2	2	2	1	0	0	2	2	2	2	1
6	1	2	2	1	4	8	1	2	2	1	1
6	1	5	2	1	1	1	1	2	2	1	1
6	1	4	2	3	0	0	1	2	1	1	1
6	2	2	2	3	7	7	1	1	2	2	1
6	1	1	2	3	5	5	1	1	1	2	0
2	1	1	1	1	1	5	1	2	2	2	0
4	1	1	1	1	3	3	1	2	2	2	0
6	1	1	1	1	7	8	1	2	2	2	1
5	1	2	2	1	2	4	1	2	2	2	0



6	1	2	2	1	7	7	2	2	2	2	0
3	2	6	2	1	7	8	2	2	2	2	0
2	2	4	2	1	7	7	2	1	2	2	1
1	1	3	1	1	7	7	2	1	2	2	1
5	2	3	1	1	0	0	1	1	2	2	1
5	2	2	2	1	8	8	2	1	2	2	0
5	1	2	1	1	2	8	1	1	2	2	1
5	2	2	1	1	1	2	1	2	2	2	1
2	2	2	2	2	7	8	1	2	2	1	1
4	1	2	2	1	4	8	2	2	2	2	1
5	1	2	2	1	7	8	1	2	2	2	1
2	2	7	1	2	5	5	1	2	2	2	1
2	1	7	3	2	3	4	1	2	2	2	0
2	2	7	1	1	6	7	1	2	2	2	1
2	1	7	2	1	6	7	2	2	2	2	1
1	2	4	1	1	6	8	2	2	2	1	1
4	1	1	1	1	7	8	2	2	1	1	1
4	1	4	2	2	3	5	2	2	2	2	0
1	1	6	2	1	7	7	2	2	2	2	1
3	1	5	2	1	6	6	2	2	2	2	1
6	2	1	2	1	0	0	2	2	2	2	1
6	1	1	2	1	4	6	1	2	1	2	1
6	1	2	2	1	6	8	1	2	1	2	1
6	1	2	1	1	6	8	1	2	1	2	0
6	1	4	1	1	2	4	1	2	2	2	0
6	1	1	1	1	2	5	2	2	2	2	1
2	1	5	2	1	5	7	2	1	2	2	1
4	2	7	2	2	3	5	1	1	2	2	1
6	1	7	1	2	7	8	2	1	2	2	1
5	2	7	1	2	5	7	2	2	2	2	1
6	1	2	2	2	2	8	2	1	2	2	1
3	2	2	2	1	1	2	2	1	2	1	1
2	1	1	1	1	7	7	2	2	2	1	1
6	1	1	1	1	2	8	2	1	2	2	1
6	1	7	1	1	0	0	2	1	2	2	1
6	2	7	1	1	3	3	2	1	2	2	1
6	1	7	2	1	4	6	1	1	2	2	1
6	1	5	2	1	6	8	1	2	2	2	1
2	2	5	2	1	8	8	1	2	2	2	1
4	1	6	2	1	2	2	1	2	2	2	1
6	2	3	2	1	8	8	1	2	1	2	1
5	1	2	2	3	2	3	1	2	1	2	0
6	1	2	1	1	5	8	1	2	1	2	1
3	2	3	2	1	6	7	2	2	1	1	1
2	1	1	2	1	8	8	2	2	1	1	1
1	1	1	2	1	1	8	2	2	1	1	1
4	2	7	1	1	4	4	2	2	2	2	1

4	1	6	1	1	6	8	2	2	2	2	1
1	1	2	1	1	8	8	1	2	2	2	1
3	1	5	1	3	7	8	1	2	2	2	1
6	1	6	2	1	0	0	1	2	2	2	0
6	1	1	2	1	7	8	1	2	2	2	1
6	1	4	2	1	1	2	1	2	2	2	1
6	2	7	1	3	1	7	1	2	1	2	1
6	1	4	2	1	7	8	2	2	2	2	1
6	1	7	2	1	2	3	2	2	2	2	1
2	2	1	1	1	4	5	1	2	2	2	1
4	1	1	1	1	4	4	1	2	1	2	1
6	2	2	1	1	2	7	2	2	2	2	0
5	1	2	2	1	3	5	2	2	2	2	0
6	1	5	2	3	2	7	1	2	2	1	0
3	1	4	2	3	6	7	2	2	2	2	1
2	1	2	1	1	8	8	1	2	2	2	0
1	1	1	2	1	0	0	2	2	2	2	0
5	1	1	1	1	2	6	1	1	2	2	0
5	2	1	1	1	5	8	2	2	2	2	1
5	1	1	1	1	1	2	1	2	2	2	1
5	1	2	1	1	2	4	1	2	2	2	1
2	2	2	2	1	8	8	1	2	2	2	0
4	1	6	2	2	3	4	1	2	2	2	1
5	1	4	2	1	5	5	1	2	2	1	1
2	1	3	2	1	6	6	1	2	2	1	1
2	1	3	2	2	1	1	1	2	2	1	1
2	1	2	2	2	3	7	2	2	1	2	1
2	1	2	2	1	7	7	2	1	2	2	1
1	1	2	2	1	7	7	2	1	2	2	0
4	2	2	2	1	5	5	1	1	2	2	1
4	1	2	2	1	4	8	1	2	1	2	1
1	1	2	2	2	6	7	1	2	1	2	1
3	2	7	2	1	8	8	1	2	2	2	1
6	1	7	1	1	8	8	1	2	2	2	0
6	1	7	1	3	4	4	2	2	2	2	1
6	1	7	1	1	8	8	2	2	2	2	1
6	1	4	1	1	2	4	2	2	2	2	1
6	2	1	2	1	5	8	2	2	2	2	1
6	1	4	2	1	5	6	2	2	1	2	1
2	1	6	2	1	6	8	2	2	2	2	0
4	1	5	2	1	6	6	2	2	1	1	0
6	1	1	2	2	7	7	2	2	2	2	1
5	1	1	2	2	7	7	2	2	2	2	1
6	1	2	2	2	2	6	2	1	2	2	1
3	1	2	1	3	4	4	2	2	2	2	1
2	1	4	1	1	4	4	2	1	2	2	1
6	2	1	1	1	4	4	2	2	2	2	1

6	2	5	2	1	3	5	2	2	2	1	1
6	1	7	2	1	8	8	2	2	2	1	1
6	1	7	2	1	0	0	1	2	2	2	1
6	2	7	2	1	1	4	1	2	2	2	1
2	1	2	3	1	8	8	1	2	2	2	1
4	1	2	3	3	1	1	1	2	2	2	1
6	1	1	2	1	5	6	2	2	2	2	1
5	1	1	1	1	2	5	1	2	2	2	1
6	1	7	1	1	5	5	2	2	2	2	1
3	2	7	2	1	6	7	1	2	2	2	1
2	2	7	2	1	6	7	2	2	2	2	0
1	1	5	2	3	5	7	2	2	2	2	1
4	1	5	1	1	3	7	2	2	2	2	1
4	1	6	3	1	3	3	2	2	2	2	1
1	2	3	1	1	5	7	1	2	1	2	1
3	2	2	2	1	4	6	1	2	2	2	1
6	1	2	1	3	2	4	1	1	2	1	1
6	1	3	1	1	2	8	1	1	2	1	1
6	2	1	2	1	2	8	1	1	2	2	1
6	1	1	2	1	8	8	2	2	1	2	0
6	1	7	2	3	7	8	2	1	1	2	1
6	1	6	2	1	2	3	2	1	1	2	1
2	1	2	2	1	1	4	2	1	2	2	1
4	1	5	2	1	6	8	1	2	2	2	1
6	1	6	1	1	7	8	1	2	2	2	1
5	2	1	1	1	8	8	1	2	2	2	1
6	1	4	1	1	6	6	1	2	2	2	1
3	1	7	2	1	2	2	1	2	2	2	0
2	1	4	2	1	3	6	2	1	2	1	0
1	1	7	1	1	3	6	2	2	2	1	0
5	1	1	1	1	8	8	2	2	2	1	1
5	1	1	2	3	2	7	2	2	2	2	0
6	1	2	2	1	8	8	2	2	2	2	0
6	2	2	1	1	4	5	2	1	2	2	0
2	2	5	1	1	4	7	2	1	2	2	1
4	2	4	1	1	4	8	2	2	2	2	1
5	1	2	1	2	6	8	2	2	2	2	1
2	2	1	2	1	1	4	2	2	2	2	0
6	1	1	2	1	2	4	1	2	1	2	1
2	2	1	2	2	6	6	1	2	1	2	1
2	1	1	2	2	7	7	1	2	1	2	1
1	1	2	2	1	1	8	1	2	1	2	1
4	1	2	2	1	2	3	1	2	1	2	1
4	1	6	1	1	1	5	1	2	1	2	1
1	1	4	2	1	0	0	1	2	2	2	0
3	1	3	2	2	4	4	1	2	2	1	1
6	2	3	2	1	1	8	2	2	2	2	1

6	1	2	1	1	5	7	2	2	2	2	1
6	1	2	1	3	6	7	2	2	2	2	1
6	2	2	1	1	4	6	2	2	2	2	0
6	1	2	1	1	1	4	2	2	2	2	1
6	1	2	2	1	8	8	2	2	1	2	1
2	1	2	2	1	8	8	1	2	2	2	1
4	1	7	2	1	4	7	1	2	2	2	1
6	2	7	1	1	4	5	1	1	2	2	1
5	2	7	2	2	4	4	1	1	1	1	0
6	1	7	2	2	6	6	1	1	2	1	0
3	2	4	1	2	4	7	1	2	2	1	1
2	1	1	1	2	8	8	1	2	2	2	1
6	1	4	1	1	6	8	2	2	2	2	1
6	1	6	2	1	1	2	2	2	2	2	1
6	1	5	2	1	1	4	2	2	2	2	1
6	1	1	2	1	0	0	2	2	2	2	1
6	1	1	1	1	5	6	2	2	2	2	1
2	2	2	2	1	8	8	1	2	2	2	1
4	1	2	1	1	3	3	1	2	2	2	1
6	1	4	1	1	7	8	1	2	2	2	1
5	1	1	1	1	2	2	1	2	2	2	1
6	2	5	1	1	1	6	1	2	2	2	1
3	2	7	2	1	3	5	1	1	2	2	1
2	1	7	2	1	4	5	1	1	2	2	1
1	1	7	2	1	7	8	2	2	1	2	1
4	1	2	2	3	7	7	2	2	2	1	1
4	1	2	2	1	8	8	2	2	2	2	0
1	1	1	2	1	8	8	2	2	2	2	1
3	1	1	2	1	3	4	2	2	1	2	1
6	1	7	2	1	7	8	1	2	1	2	1
6	1	7	2	1	6	8	1	1	2	2	1
6	1	7	2	1	6	7	2	1	2	2	1
6	1	5	2	1	8	8	2	1	2	1	1
6	1	5	2	1	8	8	1	1	2	1	1
6	1	6	1	1	5	6	1	1	2	2	1
2	2	3	1	1	1	6	1	2	2	2	0
4	2	2	1	1	4	5	2	2	1	2	1
6	2	2	1	1	5	8	2	2	2	2	1
5	1	3	2	1	7	8	2	2	1	2	1
6	1	1	2	1	2	7	2	2	2	2	1
3	1	1	2	1	2	2	1	2	2	2	1
2	1	7	2	1	1	5	1	2	2	2	1
1	1	6	2	1	6	8	1	2	2	2	1
5	1	2	2	1	6	6	1	2	2	2	0
5	1	5	2	1	6	6	1	2	2	2	0
5	2	6	1	1	1	7	1	2	2	2	0
6	1	1	1	1	3	5	1	2	2	2	1

6	1	4	1	1	8	8	1	2	2	2	0
4	2	7	2	1	0	0	1	2	2	1	0
5	2	4	2	1	1	4	2	2	2	1	0
2	1	7	2	2	4	5	2	2	2	2	1
2	2	1	2	1	2	8	2	2	2	2	1
2	1	1	1	1	6	8	2	2	2	2	1
1	2	2	1	2	8	8	1	2	2	2	0
6	1	2	2	2	3	5	2	1	2	2	1
4	1	5	1	1	1	2	1	1	2	2	1
4	1	4	1	1	6	7	1	1	2	2	1
1	1	2	2	1	8	8	1	2	2	2	1
3	1	1	2	1	2	5	2	1	2	2	1
6	1	1	2	2	6	8	1	1	1	2	1
6	1	1	1	1	1	4	1	2	2	1	0
6	1	1	3	1	7	7	1	1	2	1	1
6	1	2	1	1	1	7	1	1	2	1	1
6	1	2	2	1	3	5	2	1	2	2	1
6	2	6	1	1	3	4	2	1	1	2	1
2	1	4	1	1	5	7	2	2	1	2	0
4	1	3	2	1	3	4	2	2	1	2	1
6	1	3	2	1	7	7	2	2	2	2	1
5	2	2	2	1	4	4	2	2	2	2	1
6	1	2	2	2	2	6	2	2	2	2	1
3	1	2	2	3	7	7	1	2	2	2	1
2	1	2	2	2	3	3	1	2	2	2	0
6	2	2	1	3	1	1	1	2	2	2	0
6	2	2	1	1	8	8	1	2	2	2	1
6	1	7	1	1	6	8	2	2	2	2	1
6	2	7	2	1	7	7	2	2	2	2	1
6	1	7	2	1	6	7	1	2	2	2	1
2	1	7	1	3	1	5	2	2	2	1	1
4	1	4	1	1	7	7	2	2	2	2	1
6	2	1	2	1	3	7	2	2	2	2	1
5	1	4	2	1	6	8	2	2	2	2	1
6	2	6	1	1	6	6	2	2	2	2	1
3	2	5	1	1	4	7	2	2	2	2	1
2	2	1	1	1	3	7	2	2	1	2	1
1	1	1	1	1	6	7	2	2	1	2	1
4	1	2	2	1	8	8	1	2	1	2	1
4	1	2	2	1	7	7	1	2	1	2	1
1	1	4	2	1	0	0	1	2	1	1	1
3	1	1	2	1	6	8	1	2	1	1	1
6	1	5	2	1	7	7	1	2	2	1	0
6	1	7	2	1	5	5	1	2	2	2	1
6	1	7	1	3	7	7	1	2	2	2	1
6	1	7	2	1	3	5	2	1	2	2	1
6	1	2	2	1	1	8	2	2	2	2	1

6	1	2	2	1	8	8	2	2	2	2	1
2	1	1	1	1	7	8	2	2	2	2	1
4	1	1	1	1	7	7	2	2	1	2	1
6	1	7	1	3	4	8	1	2	2	2	1
5	1	7	1	1	6	8	1	2	2	2	0
6	1	7	2	1	3	6	1	2	2	2	1
3	1	5	2	1	6	6	1	2	1	2	1
2	1	5	2	3	8	8	1	2	2	2	1
1	1	6	1	1	8	8	1	1	2	2	1
5	1	3	2	1	3	4	2	1	2	2	1
5	1	2	2	1	3	6	2	1	2	1	1
5	1	2	1	1	1	1	1	2	2	2	1
5	1	3	1	1	6	8	1	2	2	2	0
2	1	1	1	1	4	8	2	2	2	2	0
4	1	1	2	1	3	8	2	2	2	2	0
5	1	7	2	1	1	3	1	2	2	2	1
2	1	6	2	1	2	3	2	2	2	2	0
6	1	2	1	2	2	5	1	2	2	1	0
6	1	5	2	1	2	8	2	2	2	1	0
2	1	6	1	1	6	7	1	2	2	2	1
1	2	1	1	2	3	8	2	2	2	2	1
4	1	4	1	2	6	8	1	2	2	2	1
4	1	7	1	1	5	8	1	2	1	2	0
1	1	4	2	1	2	5	1	2	2	2	1
3	2	7	2	1	8	8	1	1	2	2	1
6	1	1	2	1	4	7	1	2	2	2	1
6	2	1	2	2	2	2	1	1	1	2	1
6	1	2	2	1	7	7	1	2	1	2	1
6	1	2	2	1	0	0	2	2	2	2	1
6	1	5	2	1	1	7	2	2	2	2	0
6	1	4	2	1	4	8	2	2	2	2	1
2	1	2	2	1	6	6	1	2	2	2	1
4	1	1	2	1	1	8	1	2	2	2	1
6	1	1	2	1	4	5	1	2	2	1	1
5	1	1	2	3	8	8	1	2	1	1	0
6	1	1	1	1	8	8	1	2	2	2	1
3	2	2	1	2	6	7	2	2	1	2	1
2	1	2	1	2	6	6	2	2	2	2	1
6	2	6	1	3	7	8	2	2	2	2	1
6	1	4	2	2	4	7	2	2	2	2	1
6	1	3	2	1	6	8	2	2	2	2	0
6	1	3	2	1	6	7	2	2	2	2	0
6	1	2	2	1	3	4	2	2	2	2	1
2	1	2	2	1	3	4	2	2	2	2	1
4	1	2	2	1	0	0	2	1	2	2	1
6	1	2	2	1	4	8	2	1	2	1	1
5	1	2	1	1	3	3	2	1	2	1	1

6	1	2	1	1	6	8	2	2	2	1	1
3	1	7	1	1	4	7	2	1	2	2	1
2	2	7	2	1	8	8	2	1	2	2	1
1	1	7	2	3	4	8	2	1	2	2	1
4	2	7	2	1	3	4	1	2	2	2	1
4	1	4	2	1	8	8	1	2	2	2	1
1	1	1	3	1	1	5	1	2	2	2	1
3	1	4	3	1	5	8	1	2	2	2	1
6	1	6	2	1	5	8	2	2	2	2	1
6	1	5	1	1	6	8	1	1	2	2	1
6	2	1	1	1	8	8	2	2	1	2	1
6	1	1	2	3	3	8	1	2	2	2	0
6	1	2	2	1	3	3	2	2	2	2	1
6	1	2	2	3	1	7	2	2	2	2	1
2	1	4	1	1	7	8	2	1	2	2	1
4	1	1	3	1	7	7	2	1	1	1	1
6	1	5	1	1	8	8	1	2	1	2	1
5	1	7	2	3	4	5	1	2	1	2	1
6	1	7	1	1	7	7	1	2	2	2	1
3	1	7	1	1	0	0	1	2	2	2	1
2	1	2	2	1	6	8	1	2	2	2	0
1	1	2	2	1	7	7	2	2	2	2	1
5	1	1	2	1	7	8	2	2	2	2	1
5	1	1	2	1	0	0	2	2	2	2	1
6	2	7	2	1	6	8	2	2	2	2	1
5	2	7	2	1	3	7	1	2	2	1	1
2	1	7	1	3	7	8	1	2	2	1	1
4	1	5	1	1	5	8	1	2	2	1	1
5	1	5	1	1	5	6	1	2	2	2	0
2	1	6	2	1	3	3	1	2	2	2	0
6	1	3	2	1	5	8	2	2	2	2	0
2	1	2	1	2	7	7	2	2	2	2	1
2	1	2	1	2	7	8	2	2	2	2	0
1	1	3	2	2	7	7	2	2	2	2	0
4	1	1	2	1	8	8	2	2	1	2	0

## Anexo B: Resultados Obtenidos

1	1	1	1	1	3	6	1	2	2	2	1	5	1	2	2	1
1	2	5	1	2	2	2	1	5	1	2	2	1	5	6	2	1
2	2	2	1	5	1	4	2	1	6	7	1	2	2	2	1	0.987155
5	1	1	1	1	5	5	1	2	2	2	1	2	1	5	3	1
1	3	8	1	2	2	2	1	4	1	7	1	1	0	0	2	0.799996
2	2	1	0	5	1	7	2	2	8	8	2	2	2	2	1	0.85914
2	1	7	1	1	6	8	2	1	2	2	1	2	1	2	1	0.62999
1	2	3	2	1	2	2	1	2	1	2	2	1	2	6	2	1
2	2	2	1	2	1	1	2	1	4	6	2	2	2	2	0	0.00979823
1	1	1	2	1	5	7	2	2	2	2	1	4	1	7	2	1
1	1	4	2	2	2	2	1	4	1	7	2	2	4	5	1	1
2	2	2	1	1	1	7	2	2	3	4	1	2	1	2	1	0.995466
3	1	5	1	1	1	8	1	1	2	1	1	6	1	5	1	1
1	4	6	1	1	2	1	0	6	2	6	1	1	0	0	2	1
1	2	1	0	6	2	3	2	1	8	8	2	1	2	2	1	0.0599851
6	2	2	2	1	3	8	1	1	1	2	1	6	1	2	1	0.999999
2	8	8	2	2	1	2	1	6	2	3	1	2	4	7	2	1
2	1	2	1	2	2	1	2	1	8	8	2	2	2	2	1	0.997879
4	1	1	2	1	8	8	2	2	2	2	1	6	1	7	1	1
1	6	7	2	2	2	2	1	5	1	6	1	1	2	7	2	1
2	2	2	1	6	1	2	1	1	3	6	2	2	2	2	1	0.0699924
3	1	5	1	1	4	5	2	2	2	2	1	2	1	6	2	1
1	8	8	1	2	2	2	1	1	1	1	2	1	8	8	1	0.709994
2	2	2	1	5	1	4	2	2	8	8	1	2	2	2	1	0.991711
5	2	7	2	2	4	6	1	2	2	2	1	5	1	4	2	1
1	3	8	1	2	2	1	1	5	2	7	2	1	2	8	1	1
2	2	2	1	2	2	1	1	1	3	4	1	2	2	2	0	0.0145308
4	1	1	2	1	7	7	2	2	2	2	1	5	1	2	2	1
1	0	0	2	2	2	2	1	2	1	2	2	1	4	8	2	1
2	2	2	1	2	1	5	1	1	3	4	2	2	1	2	1	1
2	2	4	1	1	8	8	2	1	1	1	1	2	1	2	1	1
1	6	8	1	1	1	1	1	1	1	1	1	1	5	6	1	0.399784
1	1	2	1	4	2	1	2	1	4	7	1	2	1	2	1	1
4	2	1	2	1	6	8	1	1	1	2	0	1	1	1	2	1
1	7	8	1	1	2	2	1	3	2	2	1	1	8	8	1	1
2	2	2	1	6	1	2	2	1	3	7	2	1	2	2	1	0.269999
6	1	6	2	1	8	8	2	1	2	2	1	6	1	4	1	1
3	4	7	1	1	2	2	1	6	1	3	1	1	8	8	1	1
1	2	2	1	6	1	3	1	1	3	8	2	2	2	2	1	1
6	1	2	2	1	7	7	2	2	1	2	0	2	1	2	2	0.448712
1	6	7	1	2	2	2	0	4	1	2	2	1	5	6	2	1
2	2	2	0	6	1	2	1	1	3	8	1	2	2	2	1	0.993368
5	2	2	2	1	4	5	2	2	1	1	0	6	1	2	1	1
1	4	8	1	2	2	1	0	3	2	7	1	1	0	0	2	0.969998



2	2	2	0	2	2	7	1	1	4	5	1	2	2	2	1	0.795308
1	1	7	1	1	7	7	1	2	2	2	1	5	1	7	2	1
3	3	5	1	2	2	2	1	5	1	4	2	1	7	8	1	1
2	2	2	0	5	1	1	2	1	8	8	1	2	2	2	1	0.990674
5	1	4	2	1	3	5	1	2	2	2	1	2	1	6	2	1
2	5	6	1	2	2	2	1	4	2	5	2	1	6	8	2	1
2	2	2	1	5	1	1	2	1	5	8	2	2	2	2	1	0.77965
2	1	1	2	1	3	4	2	2	2	1	1	2	2	2	2	0.569978
1	7	7	1	2	2	1	0	2	1	2	2	1	8	8	1	1
2	2	1	1	2	1	4	2	1	1	5	1	2	2	2	1	0.999714
1	1	1	2	2	2	3	1	2	1	2	1	4	1	5	1	1
2	8	8	1	2	2	2	1	4	1	7	1	1	1	8	2	0.259863
2	2	2	0	1	1	7	1	1	2	2	2	2	2	2	1	1
3	1	7	1	1	5	6	2	2	1	2	1	6	1	2	2	1
1	7	8	2	2	1	2	1	6	1	2	2	2	7	8	2	1
1	2	2	1	6	1	1	2	1	7	7	2	2	2	2	1	0.9989
6	2	1	2	1	7	7	2	2	2	2	0	6	1	7	2	1
1	1	4	2	2	2	2	0	6	1	7	2	1	2	6	2	1
2	2	2	1	2	1	7	2	1	8	8	2	2	2	2	1	0.993109
4	1	5	1	1	4	6	2	2	1	2	1	6	1	5	1	1
1	7	8	2	2	2	1	1	5	1	6	1	1	0	0	2	1
2	1	2	1	6	2	3	2	1	1	7	2	2	2	2	1	0.990491
3	2	2	2	2	7	8	2	1	2	2	1	2	2	2	2	1
2	2	2	1	1	2	2	1	1	1	3	2	2	3	6	1	1
1	2	2	1	5	1	1	3	2	2	3	1	2	2	2	1	0.999557
5	1	1	3	1	8	8	1	2	2	2	1	5	1	7	2	1
1	6	8	2	2	2	2	1	5	1	6	1	1	7	8	1	1
2	2	2	1	2	1	2	1	1	4	4	2	2	2	1	1	0.991772
4	1	5	2	1	6	7	1	2	2	1	1	5	1	6	2	1
1	1	5	2	2	2	1	1	2	1	1	2	1	1	3	2	1
2	2	2	0	2	1	4	1	1	6	8	2	2	2	2	1	0.997366
2	1	7	3	1	1	1	2	2	2	2	1	2	2	4	1	1
1	2	2	1	2	2	2	1	1	2	7	2	1	7	7	1	0.999951
2	2	2	1	4	1	1	1	1	3	4	1	2	2	2	1	0.989143
4	1	1	1	1	4	8	1	1	2	2	1	1	1	2	2	1
1	0	0	1	2	2	2	1	3	2	2	2	3	5	7	2	1
1	2	2	1	6	1	5	2	1	2	4	2	2	1	2	0	0.0023989
6	1	4	2	1	8	8	2	2	2	2	1	6	2	2	2	1
1	3	3	2	2	2	2	1	6	1	1	2	3	1	6	1	1
2	2	2	1	6	1	1	1	1	2	7	1	2	2	2	1	0.999985
6	1	1	1	1	2	4	1	2	1	1	1	2	1	1	1	1
1	4	8	1	2	1	2	1	4	1	2	2	1	3	4	1	1
2	1	2	1	6	2	2	2	1	2	3	2	2	2	2	0	0.00900701
5	1	6	1	1	4	7	2	2	2	2	0	6	1	4	1	1
1	2	2	2	2	2	2	0	3	1	3	2	1	7	8	2	1
2	2	2	1	2	1	3	2	1	6	6	2	2	2	1	0	0.00737278
1	1	2	1	1	8	8	2	2	2	1	0	4	1	2	1	1

1	2	8	2	2	2	2	0	4	1	2	1	1	0	0	2	1
2	2	2	1	1	1	2	1	1	0	0	2	2	2	2	1	1
3	1	2	2	1	1	3	2	1	2	2	1	6	2	2	2	0.999981
1	2	8	1	1	2	2	0	6	1	7	2	1	7	8	1	1
1	2	2	1	6	2	7	2	1	7	7	1	2	2	2	1	0.999789
6	1	7	2	1	6	7	1	1	2	2	1	6	1	7	2	1
1	4	4	1	1	2	2	1	6	1	4	1	2	1	1	1	0.999998
1	2	2	1	2	2	1	2	1	1	1	1	2	1	2	1	0.999202
4	1	4	2	1	2	5	1	2	1	2	0	6	1	6	2	1
2	8	8	2	2	1	2	1	5	1	5	1	2	7	8	2	1
2	1	2	1	6	1	1	1	1	5	7	2	2	1	1	1	1
3	2	1	1	1	4	7	2	1	1	1	1	2	1	2	1	0.999993
1	6	7	2	2	2	2	0	1	1	2	2	1	4	8	2	1
2	2	2	1	5	1	4	2	2	8	8	1	2	2	2	1	0.991713
5	1	1	2	1	6	7	1	2	2	2	1	5	1	5	1	1
1	6	6	1	1	2	2	1	5	1	7	2	1	6	7	1	0.999876
1	2	2	1	2	1	7	2	1	2	4	1	2	2	2	0	0.00683393
4	1	7	1	1	0	0	1	2	1	2	0	5	1	2	1	1
1	1	2	1	2	2	2	1	2	1	2	1	1	8	8	2	1
2	2	2	1	2	1	1	2	1	2	2	2	2	2	1	1	0.993917
2	1	1	2	1	6	6	2	2	1	1	1	2	1	7	2	1
2	4	8	2	2	2	1	1	1	1	7	1	2	7	8	2	0.999999
3	1	7	1	1	5	6	2	2	1	2	1	6	1	2	2	1
1	7	8	2	2	1	2	1	6	1	2	2	2	7	8	2	1
1	2	2	1	6	1	1	2	1	7	7	2	2	2	2	1	0.9989
6	2	1	2	1	7	7	2	2	2	2	0	6	1	7	2	1
1	1	4	2	2	2	2	0	6	1	7	2	1	2	6	2	1
2	2	2	1	2	1	7	2	1	8	8	2	2	2	2	1	0.993109
1	1	2	1	2	2	2	1	2	1	2	1	1	8	8	2	1
2	2	2	1	2	1	1	2	1	2	2	2	2	2	1	1	0.993917
2	1	1	2	1	6	6	2	2	1	1	1	2	1	7	2	1
2	4	8	2	2	2	1	1	1	1	7	1	2	7	8	2	0.999999
3	1	7	1	1	5	6	2	2	1	2	1	6	1	2	2	1
5	1	6	1	1	4	7	2	2	2	2	0	6	1	4	1	1
1	2	2	2	2	2	2	0	3	1	3	2	1	7	8	2	1
2	2	2	1	2	1	3	2	1	6	6	2	2	2	1	0	0.00737278
1	1	2	1	1	8	8	2	2	2	1	0	4	1	2	1	1
1	2	8	2	2	2	2	0	4	1	2	1	1	0	0	2	1
2	2	2	1	1	1	2	1	1	0	0	2	2	2	2	1	1
3	1	2	2	1	1	3	2	1	2	2	1	6	2	2	2	0.999981
1	2	8	1	1	2	2	0	6	1	7	2	1	7	8	1	1
1	2	2	1	6	2	7	2	1	7	7	1	2	2	2	1	0.999789
6	1	7	2	1	6	7	1	1	2	2	1	6	1	7	2	1
1	4	4	1	1	2	2	1	6	1	4	1	2	1	1	1	0.999998

## Anexo C. Aplicación Ejemplo

A modo de aplicación de red neuronal, se tomó un caso de consumidores de drogas [17].

La muestra estaba formada por dos grupos de sujetos, 148 consumidores de éxtasis, y 148 no consumidores. El muestreo fue intencional, encuestándose a los jóvenes en los lugares recreativos donde acudían, y se realizó en cinco países de la comunidad Europea: España, Francia, Holanda, Italia y Portugal [29]. A su vez la muestra se podía dividir en función del lugar donde se había realizado el cuestionario: un grupo de asistentes a discotecas y otro grupo de estudiantes universitarios. En la tabla C1 se presentan las características demográficas de los sujetos consumidores y no consumidores.

Características demográficas de los sujetos		
	Consumidores= 148	No Consumidores=148
Sexo		
Hombre	58	59
Mujer	90	89
Edad	22.38 (media)	22.82 (media)
País		
España	48	34
Francia	21	29
Holanda	35	24
Italia	8	18
Portugal	36	43
Lugar		
Discoteca	108	69
Universidad	40	79

Tabla C1. características demográficas de los sujetos consumidores y no consumidores de éxtasis

El grupo de consumidores se caracteriza por ser consumidores habituales- consumían éxtasis mas de una vez al mes. En general, los sujetos que forman esta categoría eran además consumidores de otras sustancias como: marihuana (118), cocaína (70), anfetaminas (51), LSD (44) y Heroína (7). Por su parte, el grupo de no consumidores que sirvió como grupo de control, se caracteriza por no haber consumido nunca éxtasis ni ninguna otra sustancia ilegal.

### C.1 Datos

Para determinar las características predoctoras del consumo de éxtasis, se realizó una encuesta basada en 25 ítems, dividida en cinco categorías:

- a) Demografía, relación con los padres y creencias religiosas.
- b) Ocio
- c) Consumo
- d) Opinión sobre el éxtasis
- e) Personalidad.

## C.2 Selección de variables relevantes y preprocesamiento de datos

Se seleccionaron 25 variables de entrada las que permiten evaluar diferentes aspectos del sujeto, susceptibles de poder predecir el consumo de éxtasis. En la tabla C2 se proporciona una descripción de las variables predoctoras utilizadas y la variable dependiente.

Variable	Alternativas de respuestas
<b>VARIABLES PREDOCTORAS</b>	
Demografía, padres y religión Estado civil Nivel de Estudios Ocupación  Estatus económico ¿con quién vives?  Relaciones con los padres  ¿Eres creyente?	1:soltero/a 2:casado/a 3:vivo en pareja 4:otros 1:primarios 2:bachiller 3:superiores 1:estudio 2:estudio y trabajo 3:trabajo eventual 4: trabajo fijo 5: servicio militar 6:parado 7:otros 1:bajo 2:medio/bajo 3:medio 4:medio/alto 5:alto 1:padres/familia 2:cónyuge/pareja 3:amigos 4:colegio/residencia 5:solo 6:otros 1:muy malas 2:bastantes malas 3:regulares 4:bastante buenas 5:muy buenas 1:si 2:no
Ocio ¿Vas a Bares?	1: nunca 2: a veces 3: a menudo 4: casi siempre 1: nunca 2: a veces 3: a menudo 4: casi siempre 1: nunca 2: a veces 3: a menudo 4: casi siempre 1: nunca 2: a veces 3: a menudo 4: casi siempre 1: nunca 2: a veces 3: a menudo 4: casi siempre 1: nunca 2: a veces 3: a menudo 4: casi siempre 1: house-backload 2:hardcore 3:hardcore-house 4:mellow-house 5:rock 6:pop 7:otros
Consumo ¿Cuántos amigos toman éxtasis? ¿Has consumido alcohol este último mes? ¿Has consumido tabaco este último mes? ¿Te has emborrachado este último mes?	1: ninguno 2: pocos 3: la mitad 4: casi todos 5: todos 1:si 2:no 1:si 2:no 1:no 2:una vez al mes 3:varias veces al mes 4: alguna vez por semana 5: una vez por semana 6:cada día
Opinión sobre el éxtasis ¿Crees que el éxtasis puede crear problemas?	1:no 2: si, es ilegal 3:si, después mal 4: si, crea adicción 5: si, mis amigos no toman 6: si, efectos imprevisibles

¿Cuál crees que es la razón para consumir éxtasis?	7:si, adulteración 8:si, problemas con familia 9:otros 1:relajarse 2: disfrutar de bailar 3: bailar mas tiempo 4:estar mejor con otros 5: olvidar los problemas 6: sentirse bien 7: mejor sexo 8: estimular los sentidos
Variable dependiente	1: consumidor (mas de una vez al mes) 2: no consumidor
Estatus de consumo de éxtasis	

Tabla C2: Variables Dependiente e Independientes

Hasta acá, los datos fueron tomados de un paper donde describía el contexto en que se realizo el trabajo con redes neuronales. En este paper también se comentaban los resultados obtenidos, los cuales serán comparados con los nuestros en las siguientes etapas.

### C.3 Preprocesamiento de datos

Una vez que se seleccionan las variables que iban a ser parte del modelo, se realiza el preprocesamiento de datos. En este caso, se trabajo con el modelo de red neuronal conocido como backpropagation y del cual se explica su funcionamiento en secciones anteriores.

Las variables deberían seguir una distribución normal o uniforme en tanto que el rango de posibles valores debería ser aproximadamente el mismo y acotado dentro del intervalo de trabajo de la función de activación empleada en las capas ocultas y de salida de la red neuronal.

Para adaptar los datos obtenidos de la muestra a las condiciones descritas anteriormente se aplico una transformación logarítmica en las variables continuas que no seguían una distribución normal. A continuación se acoto los valores de todas las variables previctorias al rango [-1, 1], limites de la función de activación que es utilizada por las neuronas de las capas ocultas. Por otro lado, la variable dependiente, fue codificada como: -1 = no consumidor, 1= consumidor.

Durante el aprendizaje la red va modificando los pesos en función de los datos de entrenamiento y de forma alternada se va obteniendo el error que comete la red ante los datos de validación.

Para este estudio, se agruparon los datos mediante una asignación aleatoria de los 296 sujetos que formaban la muestra. Así se contó con 176 sujetos de entrenamiento, de los cuales 88 eran no consumidores de éxtasis y 88 eran consumidores. Además se consideraron 60 sujetos de validación, de los cuales 30 eran consumidores y 30 no consumidores- y 60 sujetos de test, de los cuales 30 eran consumidores y 30 no consumidores.

La red conseguida tiene el aspecto mostrado en la figura C1:

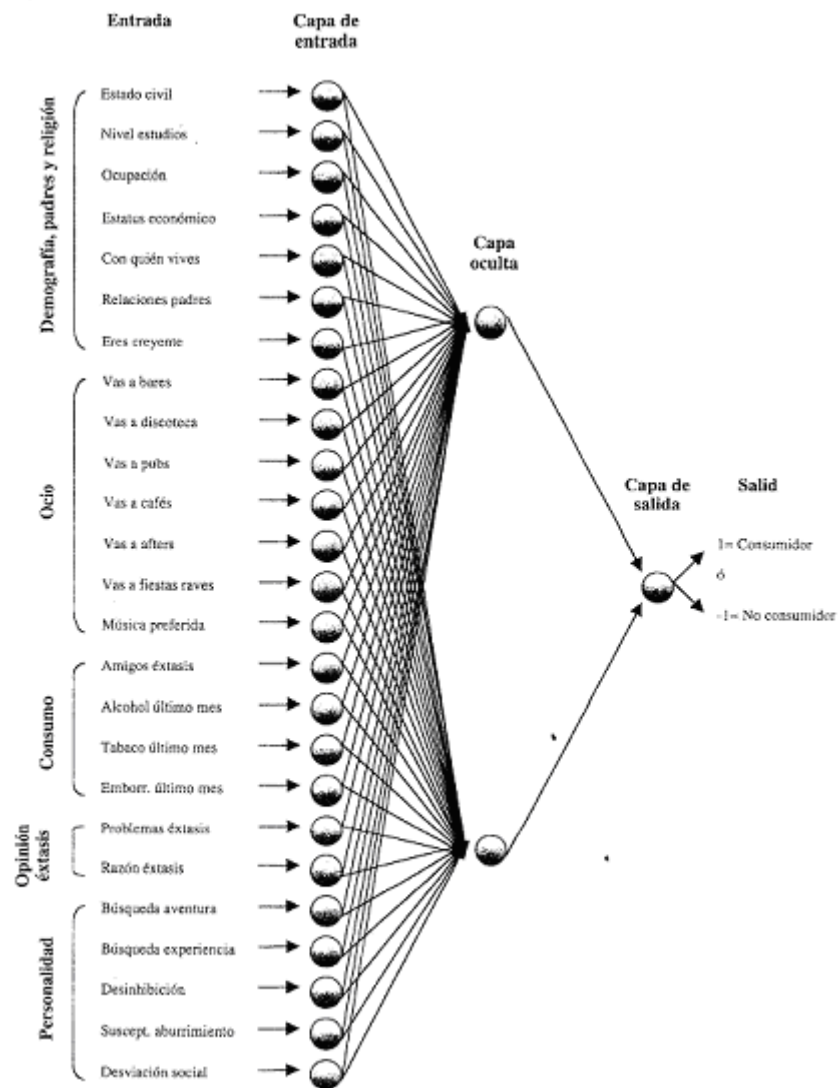


Figura C1: Modelo Red Neuronal Utilizado en el estudio

## C.4 Entrenamiento de la red neuronal

Como se menciona anteriormente, como modelo de red neuronal se ha empleado una arquitectura de red perceptrón multicapa entrenada mediante la regla de aprendizaje backpropagation.

El funcionamiento de una red de este tipo consiste en el aprendizaje de un conjunto de pares de entradas y salidas de información dados como ejemplos, empleando un ciclo de propagación-adaptación compuesto por dos fases. En este caso particular, la red debe aprender a relacionar los valores de las variables predictoras con el correspondiente estatus de consumo del sujeto. En la fase de propagación, se presenta a la capa de entrada de la red los valores de las 25 variables predictoras correspondientes a un

sujetos de entrenamiento, esta información se va propagando a traves de todas las capas superiores hasta generar una salida, se compara el resultado obtenido con la salida que se desea obtener, -1 si el sujeto no es consumidor y 1 si el sujeto es consumidor-, y se calcula el error que comete la neurona de la capa de salida. En la fase de adaptación, este error se propaga hacia atrás, capa por capa, recibiendo cada neurona un error que describe su aportación relativa al error global que comete la red. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presenten los valores del mismo sujeto, la salida esté mas cerca de la deseada, es decir, el error disminuya.

Los valores de la tasa de aprendizaje ( $\epsilon$ ) y el momento ( $\eta$ ) tienen un papel crucial en el proceso de entrenamiento de una red neuronal, ya que controlan el tamaño del cambio de los pesos en cada iteración. Se deben evitar dos extremos: un ritmo de aprendizaje demasiado pequeño puede ocasionar una disminución importante en la velocidad de convergencia y la posibilidad de acabar con una configuración de pesos poco eficientes; en cambio, un ritmo de aprendizaje demasiado grande puede conducir a inestabilidades en la función de error o a saturar las neuronas de la red. Por lo tanto, se debe elegir un ritmo de aprendizaje lo mas grande posible sin que provoque grandes oscilaciones.

En general, el valor de la tasa de aprendizaje suele estar comprendida entre 0.05 y 0.5, mientras que el valor del momento suele ser aproximadamente igual a 0.9. En el caso de nuestro estudio, los mejores resultados se obtuvieron con los valores de  $\epsilon=0.3$  y  $\eta= 0.8$ ; esta configuración de valores permitió alcanzar la convergencia, es decir, hasta que el valor de los pesos permanece estable, en 1200 ciclos de iteraciones o ciclos de aprendizaje, momento en que se decidió parar el entrenamiento.

Finalmente, la función de activación exigida por backpropagation debe ser continua, y por tanto, derivable para poder obtener el error o valor delta de las neuronas ocultas y de salida. En este caso en particular los mejores resultados se obtuvieron utilizando la función sigmoideal tangente hiperbólica (con límites entre -1 y 1) para las neuronas de la capa oculta y la función lineal para la neurona de la capa de salida.

## **C.5 Evaluación del rendimiento de la red.**

En este sentido, existe un amplio abanico de medidas de rendimiento: media cuadrática del error, funciones de coste, matrices de confusión, índices de sensibilidad y especificidad, etc. Para nuestro estudio, la evaluación de rendimiento se realizo a partir de los índices de sensibilidad, especificidad y eficacia.

La sensibilidad es en nuestro caso, el porcentaje de consumidores que son clasificados correctamente, verdaderos positivos. Por otra parte, la especificidad es el porcentaje de no consumidores que con clasificado correctamente, verdaderos negativos. Por ultimo a raíz de los dos índices anteriores, la eficacia es el porcentaje de sujetos (consumidores y no consumidores) correctamente clasificados.

## C.6 Análisis de Sensibilidad

El método mas común para realizar un análisis de sensibilidad consiste en fijar el valor de todas las variables de entrada a su valor medio e ir variando el valor de una de ellas a lo largo de todo su rango, con el objeto de observar el efecto que tiene sobre la salida de la red. Siguiendo este método, se fue registrando los cambios que se producían en la salida de la red cada vez que se aplicaba un pequeño incremento, en este caso de un 2%, en una variable de entrada.

## C.7 Resultados

A continuación se presenta una comparación entre los resultados obtenidos entre nuestro desarrollo de la red neuronal y la entregada en el paper antes mencionado.

Se obtuvieron resultados bastante satisfactorios a partir del grupo de test. Así, estableciendo un punto de corte igual a cero en la salida de la red – las salidas negativas eran consideradas como “no consumidores” y las positivas como “consumidores”. De los resultados obtenidos, solo dos sujetos fueron incorrectamente clasificados, en comparación con solo un valor mal clasificado del modelo original. Esto puede deberse a la cantidad de entrenamiento de la red, puede ser que la cantidad de iteraciones realizadas en el modelo original, sean mayores; pero esta información no era contemplada en el paper.

Ahora mas en detalle, se midió la sensibilidad, especificidad y eficacia de la red, los que originalmente eran de 97,66%, en nuestro caso entrego valores de 93,74% para todos los valores.

A continuación se presenta una tabla con los índices de rendimiento esperados y los obtenidos

La Tabla C3 muestra el índice de rendimiento de los cinco submodelos de red a partir del grupo test.

Categoría	Sensibilidad	Especificidad	Eficacia
Demografía, padres y religión	80.00	66.66	73.33
Ocio	90.00	93.33	91.66
Consumo	88.66	80.00	85.00
Opinión sobre el éxtasis	46.66	96.66	70.00
Personalidad	90.00	70.00	80.00

Tabla C3: Índice de rendimiento de los cinco submodelos (paper)



La Tabla C4 muestra los índices de rendimientos obtenidos de los cinco submodelos de red a partir de grupo de test.

Categoría	Sensibilidad	Especificidad	Eficacia
Demografía, padres y religión	81.33	66.66	73.33
Ocio	89.00	93.33	90.00
Consumo	80.00	81.33	85.00
Opinión sobre el éxtasis	45.00	93.33	70.00
Personalidad	90.00	70.00	80.00

Tabla C4: Índice de rendimientos obtenidos