

12  
Lej



# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Escuela Nacional de Estudios  
Profesionales "Acatlán"

MODELO PREDICTIVO DE CONCENTRA-  
CION DE CONTAMINANTES EN LA  
ATMOSFERA BASADO EN REDES  
NEURONALES

T E S I S

Que para obtener el título de:

LICENCIADO EN MATEMATICAS APLICADAS  
Y COMPUTACION

p r e s e n t a n

ALEJANDRO MONDRAGON HERNANDEZ

OSCAR PASCUAL CUEVAS TORRES



Director de Tesis:

Ing. Eduardo Vázquez Zamora

Naucalpan de Juárez, Edo. de Méx. Enero 1999

TESIS CON  
FALLA DE ORIGEN

270591



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

---

# **Modelo predictivo de concentración de contaminantes en la atmósfera basado en redes neuronales**



p

/

D

## **AGRADECIMIENTOS**

**A mis Padres... por su apoyo incondicional,  
comprensión y confianza.**

**Al Ing. Eduardo Vázquez... por su paciencia  
y sus sabios consejos.**

**OSCAR Y ALEJANDRO**

# Contenido

<b>I. Introducción</b>	<b>3</b>
1.1 Problemática	3
<b>II. Dinámica ambiental</b>	<b>8</b>
2.1. Contaminación del aire	8
2.2. Interacción entre contaminantes y factores meteorológicos	9
2.3. Condiciones ambientales que influyen en la disminución de la concentración de contaminantes atmosféricos	11
2.3.1. Temperatura	11
2.3.2. Viento	12
2.3.3. Humedad	14
2.4. Estabilidad Atmosférica	14
<b>III. Antecedentes de la contaminación en la Ciudad de México</b>	<b>15</b>
3.1. Antecedentes	15
3.2. Principales Contaminantes Atmosféricos	17
3.3. Indicadores Ambientales	18
3.4. Índices Ambientales	18
<b>IV. Generalidades de las Redes Neuronales</b>	<b>20</b>
4.1. Neuronas biológicas	20
Fisiología de una neurona individual	21
4.2. Sistemas neuronales artificiales	27
4.2.1. Utilidad de las redes neuronales	28
4.2.2. Áreas de actividad de los sistemas neuronales artificiales	28
4.2.3. Relación entre los sistemas neuronales artificiales y las neurociencias	30
4.3. Tecnología de redes neuronales	33
4.3.1. Propiedades de las Redes Neuronales	35
4.3.2. La unidad de cálculo en una red neuronal	36
<b>V. El modelo de Propagación</b>	<b>40</b>
5.1. Fundamentos	40
5.2. Antecedentes del modelo de propagación hacia atrás	41
5.2.1. Gradiente descendente	41
5.2.2. Cálculo del gradiente	45
5.3. Propagación hacia atrás	50
5.3.1. Descripción	50
5.3.2. Generalización de la regla delta	52
5.4. Algoritmo completo de propagación hacia atrás	57
5.5. Consideraciones	58
5.5.1. Datos de entrenamiento	58
5.5.2. Dimensión de la red	59
5.5.3. Funciones de salida	59
5.5.4. Pesos	59

5.5.5. Parámetros de aprendizaje.....	60
5.6. Red neuronal (Programa).....	62
5.6.1. Interfaz gráfica de usuario.....	62
5.6.2. Base de datos.....	62
<b>VI. Alimentación de la red y generación de resultados</b>	<b>63</b>
6.1. Alimentación de la red.....	63
6.1.1. Fuente de los datos.....	63
6.1.2. Datos que se consideraron como relevantes en la concentración de contaminantes.....	64
6.1.3. Rangos de medición.....	64
6.1.4. Explicación de la elección de variables.....	65
6.2. Resultados (Corrida del programa).....	67
6.2.1. Carga de datos.....	67
6.2.2. Parámetros de aprendizaje.....	67
6.2.3. Resultados.....	67
<b>VII. Conclusiones</b>	<b>70</b>
<b>Apéndice A</b>	<b>74</b>
Indice Metropolitano de Calidad del Aire (IMECA).....	74
<b>Apéndice B</b>	<b>79</b>
Código fuente de las funciones de propagación hacia atrás.....	79
f_propagacion_hacia_atrás.....	79
f_calcula_nodos_ocultos.....	82
f_calcula_sigmoide.....	82
f_derivada_sigmoide.....	82
f_inicializa_pesos_y_deltas.....	82
<b>Apéndice C</b>	<b>83</b>
Manual de operación.....	83
¿Qué es el propagador neuronal?.....	83
¿Cómo funciona?.....	83
Menú Principal.....	84
Menú Aprendizaje.....	86
<b>Apéndice D</b>	<b>88</b>
Pronósticos calculados por el programa computarizado.....	88
<b>Bibliografía</b>	<b>95</b>
<b>Glosario de Términos</b>	<b>98</b>

# I. Introducción

---

## 1.1 Problemática

El crecimiento y el desarrollo industrial de la Ciudad de México, han provocado una enorme variedad de problemas, de entre los cuales sobresalen por su complejidad y magnitud: la sobrepoblación, el consumo irracional de los recursos naturales y la contaminación. El crecimiento desmedido en el número de habitantes del Valle de México, trae consigo un aumento en la demanda de bienes y servicios que satisfagan sus necesidades. Esto se refleja en el aumento de vehículos de transporte (tanto de pasajeros como de carga) y en la creación de nuevas industrias a las cuales hay que dotar de un mayor volumen de combustibles e insumos para poder llevar a cabo sus actividades cotidianas; trayendo consigo desechos sólidos, líquidos, gaseosos y partículas que son liberados al medio ambiente, formando lo que conocemos como contaminación.

La contaminación no sólo considera desechos sólidos o líquidos. El aire también es afectado por una gran diversidad de agentes contaminantes, producto de procesos químicos y de combustión. Dichos agentes en

concentraciones elevadas producen diversas reacciones en plantas, animales y personas, las cuales pueden ir de leves a graves e incluso bajo ciertas condiciones ambientales, causar la muerte. *De ahí la importancia de conocer con anticipación, a través del uso de modelos matemáticos, su comportamiento durante una determinada estación del año (representada ésta por sus condiciones climáticas) en el Valle de México, para adoptar medidas preventivas o correctivas que eviten el deterioro del medio ambiente y de nuestra salud.*

En la actualidad, el afán de simular los procesos de razonamiento humano, ha dado como uno de sus frutos lo que se denomina Neurocomputación y sobre todo, los modelos denominados “*redes neuronales*”, cuya aplicación ha crecido en los últimos años. Los éxitos logrados hasta ahora se han dado fundamentalmente en los campos de ingeniería química y biología, pero su enorme flexibilidad los hace fáciles de aplicar a diversas áreas del conocimiento humano, convirtiéndose de esta manera, en una herramienta útil, pues proporcionan un amplio rango de aplicaciones y manejan en forma sencilla problemas complejos y no lineales. Las redes neuronales pueden identificar relaciones entre datos de entrada y sus respectivos valores de salida, de tal forma que pueden ser usadas para pronosticar o predecir datos futuros a partir de nuevos datos de entrada.

El presente estudio delimita el campo de trabajo sólo a los niveles de contaminación registrados en el Valle de México y a dos agentes contaminantes (el monóxido de carbono, por ser el que se produce en los vehículos

automotores y tener mayor toxicidad que el ozono, y el ozono por ser el que registra una mayor concentración). Los objetivos de la tesis son:

1. Aplicar la teoría de redes neuronales a problemas de tipo ambiental, particularmente, a la contaminación del aire en el Valle de México, *mostrando que el algoritmo de propagación hacia atrás puede presentar una nueva opción en la predicción de concentración de contaminantes. Esto, debido a que puede efectuar reconocimiento de tramas complejas y establecer funciones de correspondencia para nuevos casos de entrada.*
2. Generar un modelo capaz de pronosticar la concentración de alguno de los principales contaminantes atmosféricos, tomando como base:
  - Aquellos agentes ambientales que influyen directamente sobre los niveles de concentración de los contaminantes.
  - Los índices de contaminación registrados durante los doce meses de 1997 y el primer trimestre de 1998 para reflejar la actualidad del problema.

El texto se divide en seis capítulos que se describen brevemente a continuación:

***Dinámica ambiental.*** Presenta un análisis de las condiciones ambientales que favorecen la dispersión de contaminantes.

***Antecedentes de la contaminación en la ciudad de México.*** Se presenta una descripción del problema de la contaminación atmosférica en la ciudad de México, destacando los principales contaminantes e indicadores (índices ambientales).

***Generalidades de las redes neuronales.*** Breve exposición sobre los conceptos fundamentales y aplicaciones de redes neuronales en diversos campos de la Ciencia; así como la relación existente entre las neurociencias y los sistemas neuronales artificiales (neurocomputación).

***El modelo de propagación.*** Descripción del modelo de red neuronal que da solución al problema planteado. Elaboración de un programa computarizado utilizando el algoritmo de propagación hacia atrás (back propagation).

***Alimentación de la red y generación de resultados.*** Ejecución del programa computarizado utilizando los datos proporcionados por la Red Automática de Monitoreo Ambiental (RAMA) y Servicio Meteorológico Nacional; así como una breve descripción de las variables que intervienen.

Generación de los resultados del programa: arquitectura de red, pesos de conexión (capa de entrada y oculta) y resultados obtenidos.

**Conclusiones.** Evaluación del alcance de los objetivos en base a los resultados obtenidos.

## II. Dinámica ambiental

---

### 2.1. Contaminación del aire

Normalmente la atmósfera está constituida por gases y aerosoles. Estas sustancias vienen a ser contaminantes cuando su concentración se incrementa a niveles que amenazan la armonía de los seres vivientes, rompiendo con los procesos físicos y biológicos.

La contaminación atmosférica se ha convertido en una creciente amenaza para la salud debido al aumento sostenido de la emisión de contaminantes atmosféricos. Un adulto diariamente inhala 14 kg. de aire aproximadamente; en toda una vida esta cantidad podría alcanzar un volumen de aire suficiente para llenar un estadio de fútbol cubierto. Si uno pudiera imaginarse un edificio lleno de aire contaminado cuya totalidad hubiera circulado por los pulmones de una persona sería fácil comprender la necesidad de disponer de aire limpio.

Desde que comenzó la época industrial, la humanidad ha retado a la naturaleza vertiendo una gran cantidad de residuos tóxicos y partículas a la atmósfera.

La contaminación del aire es especialmente peligrosa cuando las condiciones atmosféricas favorecen su concentración. Cuando los contaminantes son emitidos en la atmósfera, su concentración usualmente comienza a declinar. El rango de concentración disminuye, o se diluye, en parte por la dispersión o mezcla de contaminantes con el aire limpio. Cuando las condiciones en la atmósfera favorecen una rápida dilución, el impacto de la polución del aire es menor. En otras ocasiones, llamados "episodios de la contaminación del aire", las condiciones atmosféricas minimizan la dilución, y el impacto de los contaminantes del aire puede ser severo, particularmente en la salud humana.

---

## **2.2. Interacción entre contaminantes y factores meteorológicos**

La mayoría de los episodios con mayor contaminación han sido asociados con extremos meteorológicos, los cuales han confundido la aseveración del "papel casual" en la generación de mortalidad y morbilidad.

El efecto independiente de temperaturas bajas sobre la mortalidad ha sido confirmada muchas veces. Estudios recientes en el Reino Unido han indicado que la mortalidad está más relacionada al frío y humedad que a la frecuencia de bruma. Evidencia epidemiológica indica que los efectos de los contaminantes del aire en la salud son exagerados en climas húmedos y fríos.

Las variables meteorológicas pueden contribuir a que los contaminantes afecten la salud básicamente en dos formas. Específicamente, debilitan los mecanismos de defensa naturales llevando toda clase de contaminantes contenidos en el aire. Los días fríos y húmedos son conocidos por decrementar la resistencia pulmonar, más aún, las infecciones respiratorias son también asociadas con el clima frío. Las variables meteorológicas tales como temperatura y humedad relativa interactúan con los contaminantes aumentando su efecto tóxico. Como un ejemplo, estudios recientes han mostrado que la alta humedad relativa promueve la formación de aerosoles ácidos.

Aunque la mayoría de las contingencias ambientales han sido asociadas con el clima frío, en la ciudad de Los Angeles California las alertas por efectos de los oxidantes fotoquímicos ocurren durante las épocas de calor. Estudios en ésta ciudad han notado un marcado incremento en la mortalidad durante períodos con elevados niveles de oxidantes fotoquímicos, sin embargo, la mayoría de los investigadores han concluido que la excesiva mortalidad observada fue producto del calor más que ocasionada por los contaminantes. En general, las ondas de calor son una de las más grandes preocupaciones en la gente mayor y enfermiza. No obstante, hay una posibilidad de que exista una relación entre el calor y los altos niveles de oxidantes. Estudios de laboratorio han mostrado que las temperaturas elevadas aumentan la toxicidad de algunos gases. Por ejemplo, animales expuestos a un incremento de

temperatura y al ozono tienden a aumentar su índice de mortalidad [Sloane, 1991].

---

## **2.3. Condiciones ambientales que influyen en la disminución de la concentración de contaminantes atmosféricos**

### **2.3.1. Temperatura**

La mayoría de los contaminantes del aire están contenidos en gases que, al producirse por medio de la combustión, son mucho más calientes que el aire circundante; como consecuencia, son ligeros y tienden a elevarse, los gases calientes se hacen menos densos y por consiguiente son más ligeros que los fríos. Conforme los gases suben, se mezclan con el aire que los rodea, se vuelven progresivamente más fríos y por lo tanto se elevan más despacio. El ascenso de la mezcla del aire y los gases de desecho que contienen los contaminantes depende, entonces, de los cambios en la temperatura del aire con el incremento de la altura.

Si una porción de aire pasa de un nivel bajo a una altitud mucho más elevada y no intercambia calor con el aire de su contorno, se expande conforme la presión decrece y se enfría. Esta disminución de la temperatura se conoce como tasa de declinación adiabática, la cual es de 1°C por 100 m de elevación para aire seco. A la inversa, si la proporción de aire se comprime conforme cae, la temperatura se incrementará a una tasa similar.

Mientras que la disminución teórica de la temperatura con la altura es la tasa de declinación adiabática, los cambios efectivos pueden ser muy diferentes debido al efecto de los vientos. Si la tasa de disminución de la temperatura es más grande que la tasa de declinación adiabática, un cuerpo de aire caliente se elevará rápidamente como si fuera una pluma de humo; en éste caso la situación se considera inestable. Bajo estas condiciones, los contaminantes del aire se dispersan rápidamente. Sin embargo, si la disminución de temperatura es menor que la tasa adiabática un cuerpo de aire caliente ascenderá mucho más despacio, y después de alguna dilución y enfriamiento por expansión se estabilizará. Tales condiciones se consideran estables y los contaminantes del aire no se dispersan.

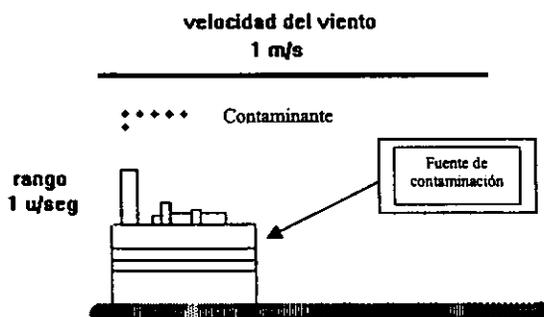
### **2.3.2. Viento**

Intuitivamente, sabemos que el aire es una mezcla más densa en un día calmado que un día con viento. Cuando sopla el viento fuertemente, las corrientes arremolinadas se encargan de mezclar los contaminantes limpiando el aire y provocando una acelerada dilución, pero cuando el viento está calmado la dilución está determinada por un lento proceso de difusión molecular. Como regla general, cuando la velocidad del viento se duplica, la concentración de contaminantes en el aire baja en un cincuenta por ciento.

Los días de contaminación alta no quieren decir necesariamente que haya habido un aumento de las emisiones de contaminantes, sino que el aire contaminado

no está siendo renovado por el viento. Esto se da cuando sucede una inversión térmica. El aire caliente se extiende sobre el aire más frío como una tapadera que impide el movimiento ascendente de las corrientes del aire y retiene los contaminantes contra el superficie terrestre.

La velocidad del viento es influenciada no sólo por la presión horizontal del aire, sino también por la fricción contra la superficie. En una ciudad, los vientos son bajos por la superficie áspera creada por la acanalada topografía de los altos edificios y estrechas calles. En promedio la velocidad del viento puede ser 25 por ciento menor dentro de una ciudad que en sus inmediaciones. Cuando los vientos regionales son ligeros (menores a 15 km por hora), en contraste entre la ciudad y los pueblos, la diferencia se acentúa reduciéndose en un 30%. La dilución de contaminantes en el aire debido al viento, se ve impedida en las zonas urbanas y es ahí donde la mayoría de los agentes contaminantes se generan.



*Figura 1 Como regla general, cuando la velocidad del viento se duplica, la concentración de contaminantes en el aire se baja en un cincuenta por ciento.*

### **2.3.3. Humedad**

La humedad atrapa a los contaminantes y los mantiene a nivel del suelo. Sin embargo, ciertas corrientes ambientales producen ligeros vientos que inhiben la acción dispersante. En cambio, la lluvia y la nieve asociada con ciclones limpian el aire de contaminantes sobre la tierra.

---

## **2.4. Estabilidad Atmosférica**

La convección y turbulencia son incrementadas cuando la temperatura del aire está inestable y desaparecen cuando el aire está estable. La estabilidad del aire influye en el rango de combinación de los contaminantes con el aire limpio. Los contaminantes emitidos dentro de aire estable experimentan mayor combinación, mientras que los producidos en temperaturas inestables impiden el flujo de mezcla, y algunas capas de aire actúan como capa que atrapa los contaminantes. La continua emisión de contaminantes dentro de temperaturas estables da como resultado la acumulación y concentración de contaminantes.

La convección tiende a aumentar en las tardes más que en las mañanas, durante el día que en la noche, y en verano que en invierno [Moran, 1991].

# III. Antecedentes de la contaminación en la Ciudad de México

---

## 3.1. Antecedentes

La década de los años cuarenta marca el inicio del desarrollo acelerado de la industria en México. Los efectos que provocó tal modelo de desarrollo son:

- > Las actividades económicas del país se concentraron en los centros de población o ciudades.
  
- > Intensificación de las migraciones del campo a la ciudad.

Estos fenómenos influyeron en la creación y aumento en la demanda de: sistemas de transporte, industrias de elaboración de bienes de consumo, nuevas redes de distribución de combustibles e industrias de servicios; los cuales, en su forma actual, propician la degradación del medio ambiente. Podemos decir que un aumento de población origina un mayor consumo de recursos naturales

y energéticos e incrementa de esta manera las descargas de desechos tóxicos tanto en el agua, suelo y aire.

Además, por mucho tiempo se consideró en México a *"la contaminación del medio ambiente como un costo necesario del proceso de industrialización"*. Lo cual se refleja en la falta de conciencia de los industriales en materia de conservación y preservación del medio ambiente. Es necesario remarcar que esta actitud se vio reforzada por el hecho de que muchas empresas paraestatales de transformación (PEMEX, CFE, FERTIMEX por citar algunas) eran las más contaminantes del país [SEDUE, 1988].

La zona metropolitana de la Ciudad de México es considerada como una de las más pobladas en el mundo. En ella se han asentado aproximadamente 30 mil industrias, de las cuales el 5% son consideradas de gran magnitud, el 10% mediana industria y el 85% pequeña industria, aunado a esto, la ciudad cuenta con un parque vehicular que se estima en 3.1 millones de unidades. Además, diferentes factores físicos (topografía, altitud, vientos, temperatura, entre otros), contribuyen al deterioro de la calidad del aire<sup>1</sup>.

La contaminación del aire es un problema que preocupa a las autoridades y a la sociedad, por lo tanto, se ha requerido de una atención especial para buscar alternativas que permitan mejorar la calidad del aire en el Valle de México.

---

<sup>1</sup> Fuente: Secretaría de Desarrollo Social (SEDESOL).

---

## 3.2. Principales Contaminantes Atmosféricos

En la actualidad en el Valle de México se produce una gran cantidad de contaminantes atmosféricos de los cuales destacan por su producción, concentración o toxicidad los siguientes:

**OZONO(O<sub>3</sub>)**

**MONÓXIDO DE CARBONO (CO)**

**ÓXIDOS DE NITRÓGENO (NO<sub>x</sub>)**

**BIÓXIDO DE AZUFRE (SO<sub>2</sub>)**

**PARTICULAS SUSPENDIDAS TOTALES (PST)**

Los efectos causados por cada uno de los contaminantes se detallan en la siguiente tabla:

<b>Contaminante</b>	<b>Efectos sobre el organismo</b>
Ozono	Irrita los ojos, ataca la nariz y garganta, y disminuye la capacidad para realizar ejercicio físico; también ocasiona dolores de cabeza
Monóxido de Carbono	Debilita las contracciones del corazón y altera el funcionamiento del sistema nervioso; se disuelve más rápidamente que el oxígeno resultando en la formación de carboxihemoglobina, y provocando la desoxigenación de la sangre.
Oxidos de Nitrógeno	Irritan los pulmones y bajan la resistencia a infecciones. Pueden desencadenar pulmonía o bronquitis

Contaminante	Efectos sobre el organismo
Dióxido de Azufre	Forma el ácido sulfúrico en combinación con el agua, causando bronquitis crónica, daños cerebrales o problemas pulmonares.
Partículas Suspendidas Totales	Producen diversos trastornos del aparato respiratorio

[Stoker, 1990]

---

### 3.3. Indicadores Ambientales

Los indicadores ambientales representan una señal, síntoma o manifestación de una variable ambiental que permite describirla o caracterizarla. Los indicadores se clasifican en dos tipos:

**CUALITATIVOS.** Ejemplo, Clasificar el aire como: limpio o muy limpio, sucio o muy sucio.

**CUANTITATIVOS.** Ejemplo, Clasificar la calidad del aire en función de una concentración de los contaminantes presentes en la atmósfera.

---

### 3.4. Índices Ambientales

Los índices agrupan a uno o más indicadores ambientales, y por lo tanto, caracterizan de una manera más completa a la variable ambiental que nos interesa. La mayoría de los índices contienen una diversidad de subíndices que no necesariamente se expresan en las mismas unidades, por lo que se registran sin unidades, por ejemplo 50, 100, 400 o más [SEDUE, 1988].

La Red Automática de Monitoreo Atmosférico (RAMA) registra la concentración de contaminantes atmosféricos en partes por millón, unidad que se utiliza técnicamente para reportar la concentración de contaminantes gaseosos como el ozono, monóxido de carbono, bióxido de nitrógeno y bióxido de azufre. Las partículas suspendidas totales y su fracción respirable se reportan en microgramos por metro cúbico ( $\mu\text{g} / \text{m}^3$ ).

Con el fin de facilitar la comunicación de los niveles de contaminación hacia el público general y evitar el manejo de diferentes unidades de concentración, se emplea en la Ciudad de México el IMECA (apéndice A). Esta unidad permite comparar las magnitudes de los diversos contaminantes en una escala homogénea.

# IV. Generalidades de las Redes Neuronales

---

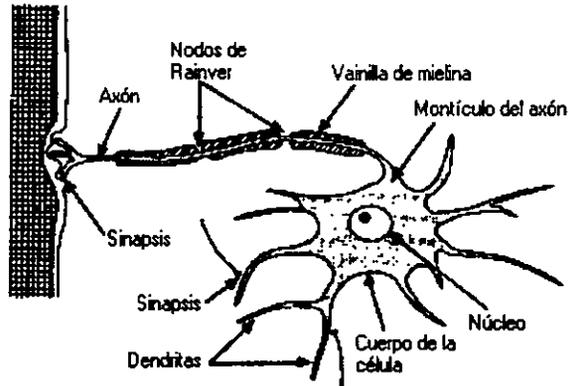
## 4.1. Neuronas biológicas

El cerebro está compuesto de cien diferentes clases de células cerebrales. El número de neuronas en el cerebro está estimado en un rango de 50 billones a 100 billones. Las neuronas están interconectadas en grupos llamados redes que proveen funciones especiales y cada grupo contiene algunos miles de neuronas altamente interconectadas unas con otras. Así, el cerebro puede ser visto como una conexión de redes neurales.

El pensamiento y la inteligencia son controlados por el cerebro y el resto del sistema nervioso central. La habilidad para reaccionar a los cambios del medio ambiente y aprender de ellos requiere inteligencia. Un ejemplo es el patrón óptico de sistemas visuales. Los estímulos externos son conectados vía células dentro de signos que conectan características de la imagen visual dentro de la memoria interna. La inteligencia humana es entonces utilizada para

entender las varias características visuales que son extraídas y almacenadas en la memoria

### Fisiología de una neurona individual



*Figura 2. Las neuronas son células altamente especializadas que generalmente tienen una forma estrellada, con un gran número de ramificaciones cortas (dendritas) y una sola más larga llamada axón.*

La figura 2 representa los componentes principales de una célula nerviosa típica perteneciente al sistema nervioso central. La membrana de la neurona separa el plasma intracelular del fluido intersticial que se encuentra fuera de la célula. La membrana es permeable para ciertas especies iónicas, y actúa de tal forma que se mantenga una diferencia de potencial entre el fluido intracelular y el fluido extracelular. Este efecto se consigue primordialmente por la acción de una bomba de sodio-potasio. También están presentes otras especies iónicas, como son los iones cloruro e iones orgánicos negativos.

Todas las especies iónicas se pueden difundir a través de la membrana, con la excepción de los iones orgánicos que son demasiado grandes. Dado que los iones orgánicos no pueden salir de la célula por difusión, su carga negativa neta dificulta la entrada en la célula de iones cloro por difusión; por tanto, habrá una concentración más alta de iones cloro fuera de la célula. La bomba de sodio-potasio determina una concentración más alta de potasio dentro de la célula y una concentración más alta de sodio fuera de ella.

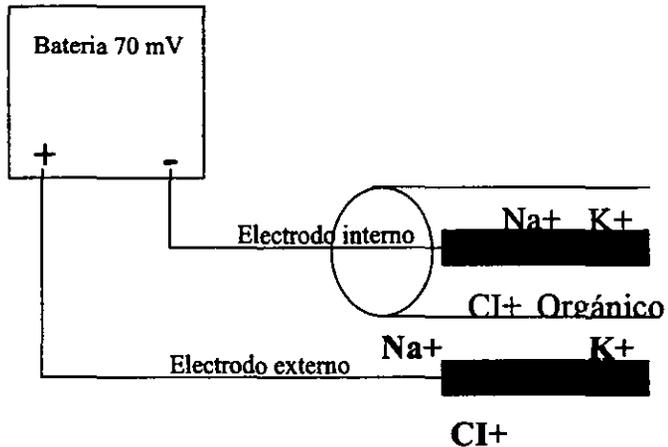
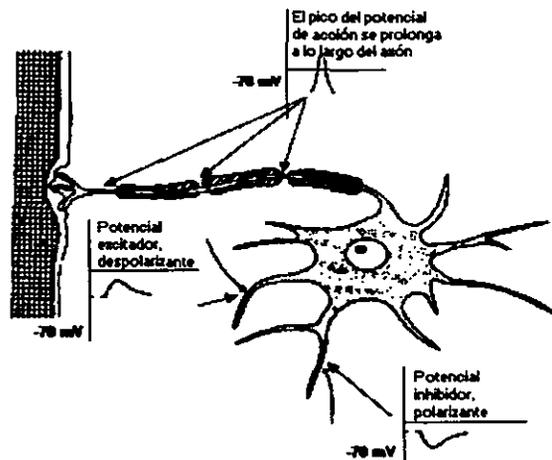


Figura 3. Esta figura ilustra el potencial de reposo que se establece a ambos lados de la membrana de una neurona. Los tamaños relativos de los rútilos de las especies iónicas denotan aproximadamente la concentración relativa de cada especie en las regiones internas y externas de la célula.

La membrana celular es selectivamente más permeable para los iones de potasio que para los iones de sodio. El gradiente químico del potasio tiende a hacer que los iones de potasio salgan de la célula por difusión, pero la fuerte atracción de los iones orgánicos negativos tiende a

mantener dentro el potasio. El resultado de estas fuerzas opuestas es que se alcanza un equilibrio en el cual hay más iones de sodio y cloro fuera de la célula, y más iones orgánicos y de potasio dentro de ella. Además, el equilibrio resultante produce una diferencia de potencial a través de la membrana de la célula de unos 70 a 100 milivoltios (mV), siendo el más negativo el fluido intracelular. Este potencial, que se denomina **potencial de reposo** de la célula, se ha representado esquemáticamente en la Figura 3.



*Figura 4. Las conexiones que llegan a una neurona procedentes de otras producen en distintos lugares de la célula, conocidos con el nombre de sinapsis. Los impulsos nerviosos que pasan a través de estas neuronas que están conectadas pueden dar lugar a cambios locales en el potencial del cuerpo celular de la neurona receptora. Estos potenciales, denominados potenciales graduados o potenciales de entrada, pueden propagarse por el cuerpo principal de la célula. Pueden ser o bien excitatorios (que hacen disminuir la polarización de la célula) o bien inhibitorios (que incrementan la polarización de la célula). Los potenciales de entrada se suman en el montículo del axón. Si la cantidad de despolarización del montículo del axón es suficiente, se genera un potencial de acción, este potencial viaja a lo largo del axón, alejándose del cuerpo principal de la célula.*

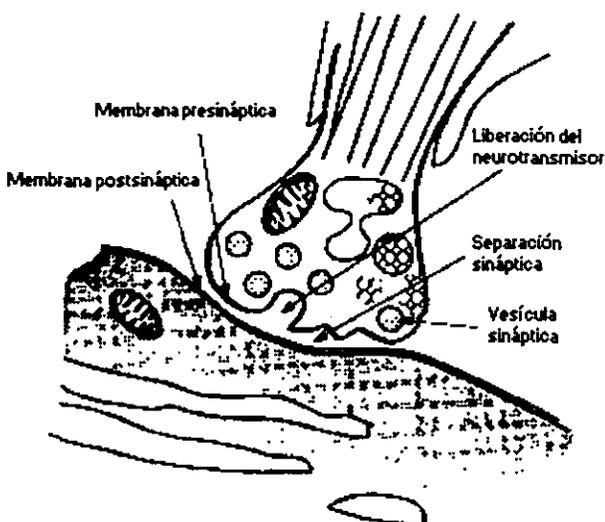
La Figura 4 ilustra una neurona con varias conexiones de entrada y los potenciales que se tienen en distintas

posiciones. La figura muestra el axón con una cubierta que se denomina **vaina de mielina**. Esta capa aislante es interrumpida en varios puntos por los **nodos de Ranvier**.

Las entradas excitatorias que llegan a la célula reducen la diferencia de potencial que existe entre los dos lados de la membrana celular. La despolarización resultante en el montículo del axón altera la permeabilidad de la membrana celular a efectos de los iones de sodio. Como resultado hay un fuerte flujo entrante de iones de sodio positivos, que penetran en la célula contribuyendo aún más a la despolarización. Este efecto autogenerado da lugar al potencial de acción.

Las fibras nerviosas no son buenos conductores de electricidad. La transmisión del potencial de acción a lo largo del axón es el resultado de una serie de despolarizaciones que tienen lugar en los nodos de Ranvier. Cuando uno de los nodos se despolariza, se desencadena la despolarización del siguiente nodo. El potencial de acción viaja a lo largo de la fibra en forma discontinua, de un nodo a otro. Una vez que un potencial de acción ha pasado por un cierto punto, ese punto no puede volver a ser excitado aproximadamente durante un milisegundo, que es el tiempo que tarda en volver a su potencial de reposo. Este período refractario limita la frecuencia de transmisión de los impulsos nerviosos a unos 1,000 por segundo.

## La unión sináptica



*Figura 5. Los neurotransmisores están almacenados en vesículas próximas a la membrana presináptica. Estas sustancias se liberan en la separación sináptica y se difunden hacia la membrana postsináptica, en la cual son absorbidas posteriormente.*

Examinemos brevemente la actividad que se desarrolla en unión existente entre dos neuronas, que se denomina unión sináptica o **sinapsis**. La comunicación entre las neuronas tiene lugar como resultado de la liberación de unas sustancias por parte de la célula presináptica, llamadas **neurotransmisores**, que son absorbidas por la célula postsináptica. La Figura 5 muestra esta actividad. Cuando el potencial de acción llega a la membrana presináptica, los cambios de permeabilidad de la membrana dan lugar a un flujo entrante de iones de calcio. Estos iones dan lugar a que las vesículas que contienen los neurotransmisores se fundan con la membrana sináptica,

liberando así sus neurotransmisores en la separación sináptica.

Los neurotransmisores se difunden a través de la unión y se unen a la membrana postsináptica en ciertos lugares llamados receptores. La acción química que se produce en los receptores da lugar a cambios de permeabilidad de la membrana postsináptica para ciertas especies iónicas. Un flujo entrante de especies positivas hacia la célula tenderá a despolarizar el potencial de reposo; este efecto es excitatorio. Si entran iones negativos, se producirá un efecto hiperpolarizante; este efecto es inhibitorio. Estos dos efectos son locales y actúan tan sólo a lo largo de una pequeña distancia hacia el interior de la célula; se suman en el montículo del axón. Si la suma es mayor que un cierto valor umbral se genera un potencial de acción.

Una red neuronal artificial (RNA) es un modelo que emula una red neural biológica. Los nodos en una RNA están basados en principios matemáticos de la forma que se piensa que ellas funcionan. Actualmente se ocupa un conjunto limitado de conceptos de los sistemas biológicos neurales para implementar programas simuladores de procesos paralelos que involucran elementos de procesamiento llamados neuronas o neurodos, interconectados en una arquitectura de red. El neurodo es análogo a la neurona biológica, recibe entradas que representan impulsos eléctricos en comparación con las neuronas. La salida de un neurodo corresponde a una señal de salida enviada desde una neurona biológica. La salida de

una neurona biológica afecta a las entradas de otras, y los impulsos son transmitidos sobre sinapsis. Una sinapsis es capaz de incrementar o decrementar su energía, afectando así el nivel de propagación de señales, con lo que puede causar excitación o inhibición de una neurona subsecuente [Freeman/Skapura, 1994].

---

## 4.2. Sistemas neuronales artificiales

El interés en los procesos cognoscitivos y la inteligencia, tanto artificiales como naturales, es casi universal. Sólo hizo falta que se dieran algunos avances en ciertas disciplinas para convertir el interés en investigación. Varios grupos de investigadores han intentado crear un modelo computarizado que reproduzca la funcionalidad cerebral, comenzando con formas muy simplificadas; a esto se le puede llamar *sistemas neuronales artificiales*.

Muchos factores han contribuido al rápido desarrollo de esta disciplina, pero tres de ellos son particularmente importantes:

- ① Primero, aplicaciones industriales en áreas como el procesamiento del conocimiento, control en robótica, reconocimiento y clasificación de patrones o tramas, reconocimiento y síntesis de voz y visión computarizada han encontrado problemas que sobrepasan la capacidad de procesamiento simbólico, característico del enfoque tradicional en inteligencia artificial.

- ② Segundo, han habido avances significativos en los métodos experimentales de las neurociencias, así como en la aplicación de herramientas computacionales para el análisis, simulación de datos y procesos biológicos.
- ③ Tercero, publicaciones con amplia difusión han logrado motivar a investigadores de diversas disciplinas para llevar a cabo proyectos relacionados con el área de redes neuronales.

#### **4.2.1. Utilidad de las redes neuronales**

Las redes neuronales son muy útiles en el reconocimiento de patrones o modelos, tales como procesos de reconocimiento de imágenes, robótica, reconocimiento de caracteres y otros tipos de rastreo o detección.

Se han diseñado redes neuronales que simulan el pensamiento de expertos humanos en *problemas de clasificación y diagnóstico, análisis de crédito, pronósticos, otros servicios de asesoría y en general problemas en los cuales los sistemas expertos han sido utilizados*. También son muy utilizadas en aquellos problemas donde el *análisis de regresión* ha sido aplicado.

#### **4.2.2. Areas de actividad de los sistemas neuronales artificiales**

Los sistemas neuronales artificiales comprenden tres áreas principales de actividad:

- > Arquitecturas de redes neuronales
- > Desarrollo de software neuronal (neurosoftware)
- > Aplicaciones de los sistemas neuronales artificiales (implantación)

**Arquitecturas de redes neuronales.** En los sistemas neuronales artificiales, la palabra arquitectura es usada para indicar la definición matemática formal "independiente de la máquina" de una estructura neuronal artificial. El área de las arquitecturas de los sistemas neuronales artificiales se interesa en el estudio y desarrollo de estructuras de redes neuronales y de la operación de las mismas.

**Desarrollo de software neuronal (neurosoftware).** Implica la investigación y el desarrollo de los medios para implantar arquitecturas neuronales artificiales en los equipos de computo existentes. Esto incluye el diseño y desarrollo de software (lenguajes para describir arquitecturas de redes neuronales de manera independiente a la máquina a utilizarse).

**Aplicaciones de los sistemas neuronales artificiales (implantación).** El área con mayor actividad en los sistemas neuronales artificiales es la búsqueda de problemas que puedan ser resueltos exitosamente a través de esta tecnología. Mucha gente con antecedentes académicos diversos está aplicando los sistemas neuronales artificiales a una amplia variedad de problemas. La más exitosa metodología parece ser el enfoque de la "solución en busca

de problemas", en el que los expertos en algún dominio de aplicación aprenden sobre la capacidad y limitaciones de múltiples arquitecturas de redes neuronales, para después identificar en su área de conocimiento algún problema importante que pueda ser resuelto con ayuda de los sistemas neuronales artificiales.

### **4.2.3. Relación entre los sistemas neuronales artificiales y las neurociencias**

Es claro que existe una relación entre los sistemas neuronales artificiales y las neurociencias. Estas últimas, de manera general, pueden ser definidas como las disciplinas involucradas con el entendimiento del cerebro. Se considera que el cerebro está compuesto por complejas redes de neuronas, las cuales son mucho más complejas que los elementos utilizados en los sistemas neuronales artificiales y sus funciones son no comprendidas en su totalidad todavía.

Como en cualquier ciencia, el proceso en las neurociencias se realiza al crear conceptos funcionales y modelos basados en resultados experimentales, para después refinar o refutar tales conceptos y modelos al llevar a cabo más experimentos. Hasta ahora, esencialmente todos los modelos que se han presentado están excesivamente simplificados. Pero la producción de conceptos funcionales y modelos en las neurociencias es continua y prodigiosa. Eso provee una fuente de nuevos conceptos y principios para ser utilizados en los sistemas neuronales artificiales, dada la relación que existe entre ambas.

Naturalmente, dado que estos conceptos y modelos de las neurociencias no son representaciones precisas del funcionamiento del cerebro, como los sistemas neuronales artificiales basados en estas ideas, no pueden ser descritos como "basados en la operación del cerebro humano". Tal vez sea más conveniente decir que operan "al estilo del cerebro". Sin embargo, el beneficio que obtienen los sistemas neuronales artificiales a través de los conceptos y modelos desarrollados por las neurociencias es invaluable, no obstante las limitaciones que estas puedan tener.

El fundamento de los sistemas neuronales artificiales es una arquitectura inspirada en las redes de neuronas biológicas. Generalmente, una red neuronal artificial se basa en un modelo matemático que consiste de un cierto número de unidades simples tipo neurona, llamados elementos de procesamiento, interconectados de alguna manera. Dado que los sistemas neuronales artificiales están abstractamente basados en lo que se conoce del funcionamiento del cerebro, existe una relajada correspondencia entre los componentes de un sistema neurofisiológico y los sistemas neuronales artificiales. La estructura y principios fundamentales de los sistemas neuronales artificiales tiene su fundamento en el funcionamiento, a un *nivel muy simplificado*, de los componentes del proceso fisiológico del *disparo de las neuronas*.

De manera muy superficial, sumamente simplificada, tal proceso puede ser descrito como sigue. Primero, las señales de entrada inciden en la neurona, provenientes de ella misma o de otras neuronas, a través de las sinapsis, que

regulan la cantidad de información que pasa desde las terminales nerviosas incidentes hacia la neurona, pudiendo ser excitadoras o inhibidoras. Las señales son pasadas entonces al área central de activación de la neurona, donde se suman. Si este nivel total de energía de entrada supera un cierto umbral, que determine el mínimo nivel que se debe alcanzar para que la neurona dispare, entonces la neurona emite energía a través del axón por donde es transmitida a otra sinapsis y puede ser realimentada a la sinapsis original; pero si la suma es menor al umbral, nada relevante puede suceder.

*Un proceso similar ocurre dentro de cada elemento de procesamiento de una red neuronal típica. Las señales llegan al elemento procesador a través de conexiones ponderadas, pudiendo ser positivas o negativas. Estas señales son sumadas a través de una función de adición. El resultado de esta suma es suministrado como entrada a una función de transferencia, la cual determina la salida total como una función de la suma de las señales ponderadas. Esta salida es generada dependiendo de las condiciones de la función de transferencia, que puede ser lineal, no lineal o lineal a trazos.*

Por otro lado es importante mencionar que comienza a haber un flujo de ideas de los sistemas neuronales artificiales a las neurociencias. Miembros de la comunidad de los sistemas neuronales artificiales están constantemente desarrollando nuevas arquitecturas de redes neuronales, así como nuevos conceptos y teorías para explicar la operación de estas arquitecturas. Muchos de estos desarrollos pueden

ser usados por neurocientíficos como nuevos paradigmas para la construcción de conceptos funcionales y modelos a las neurociencias.

---

### 4.3. Tecnología de redes neuronales

La posibilidad de que una computadora sea capaz de aprender a partir de la experiencia, en vez de interpretar de manera indefinida un conjunto de instrucciones generadas por un programador ha dado lugar a la creación de nuevas tecnologías conocidas bajo el nombre de *procesamiento paralelo*.

Las computadoras actuales funcionan de manera *secuencial* y requieren de que un experto humano programe las reglas mediante las cuales debe resolverse un problema. Esto en ocasiones no siempre es posible, ya que existen obstáculos que no permiten alcanzar la solución a través de sistemas secuenciales debido a la complejidad de su naturaleza.

Para poder resolver problemas no secuenciales, se han desarrollado una serie de técnicas que toman prestadas algunas características fisiológicas del cerebro y que se conocen como *redes neuronales*.

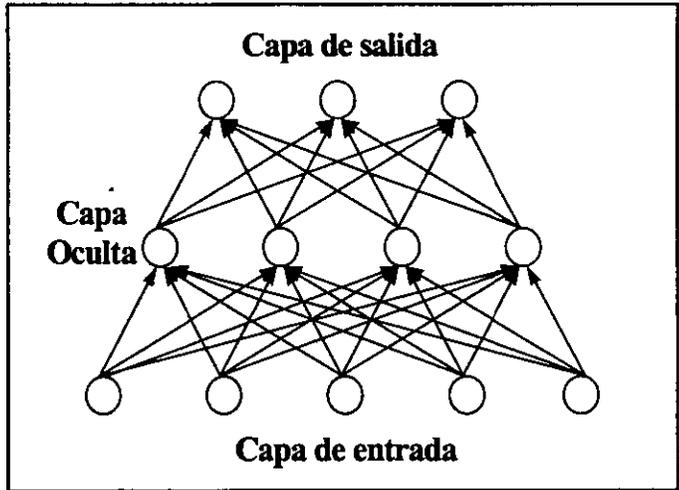


Figura 6. Diagrama de una red neuronal con tres capas.

Estructuralmente las redes neuronales son colecciones de **procesadores paralelos** conectados unos a otros formando un grafo dirigido, organizado de tal manera que la red resultante sea adecuada para resolver el tipo de problema que sea planteado de antemano. La Figura 6 nos ilustra una red en la que cada **elemento de procesamiento** esta representado por un **nodo** y las conexiones en los nodos en forma de **arcos**. La dirección del flujo dentro de la red se indica mediante flechas en los extremos de los arcos. Los nodos de entradas y salidas se agrupan formando capas las cuales están conectadas mediante el empleo de conexiones ponderadas.

Una ventaja del empleo de las redes neuronales consiste en que no necesariamente debe tener un proceso bien definido para convertir algorítmicamente una entrada

en una salida. En cambio, lo que necesita la mayoría de las redes neuronales es un conjunto de datos de ejemplos representativos que nos lleven a la solución deseada. De esta manera la red aprende a proporcionar una salida deseada cuando se le presenta una entrada dada como ejemplo. También responderá cuando se le presente una entrada que no había sido suministrada anteriormente, tales como tramas que contengan ruido. Esta capacidad de reconocer tramas ruidosas o distorsionadas hace a las redes neuronales más útiles que los métodos algorítmicos tradicionales<sup>2</sup>. El proceso de entrenar la red es solamente cuestión de modificar los pesos de conexión sistemáticamente, para codificar las relaciones de entrada-salida deseadas.

#### **4.3.1. Propiedades de las Redes Neuronales**

A pesar de que es posible demostrar teoremas que garantizan ciertas capacidades y propiedades de las redes neuronales (tales como tolerancia a fallas, robustez y la habilidad de implantar una clase particular de transformaciones), estos resultados sólo aseguran que las redes neuronales pueden ser empleadas útilmente y con seguridad.

Existe una serie de características, comúnmente compartidas por las redes neuronales, que las hacen únicas como sistemas procesadores de información; en general tienen la capacidad de adaptarse, de aprender, de auto-

---

<sup>2</sup> La potencia de la aproximación de las redes neuronales consiste en la generalidad de la red que es capaz de hallar sus propias soluciones para problemas concretos, dándose únicamente ejemplos del comportamiento deseado.

organizarse, de generalizar, de ser entrenadas, de proveer de memoria asociativa y distribuida y suelen ser resistentes a fallas parciales. Es por estas características que las redes neuronales son muy adecuadas para resolver problemas específicos, con propiedades muy particulares, especialmente aquellos que resultan difíciles de resolver utilizando modelos computacionales tradicionales.

### 4.3.2. La unidad de cálculo en una red neuronal

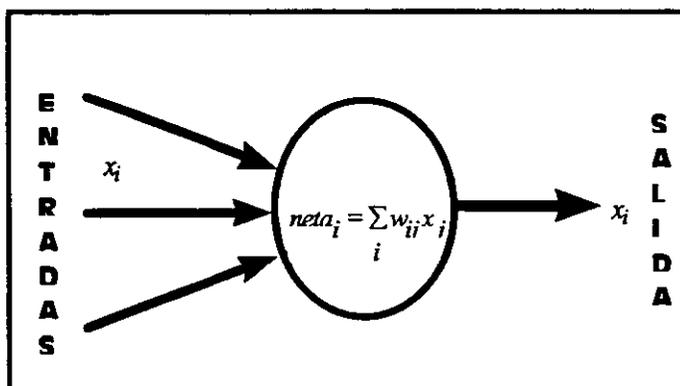


Figura 7. Representación de un elemento de procesamiento.

La mayoría de los modelos de redes neuronales tienen como unidad de cálculo a los nodos también llamados unidades o elementos de procesamiento. La Figura 7 nos muestra gráficamente el modelo general<sup>3</sup> de un elemento de procesamiento.

El elemento de procesamiento tiene muchas entradas pero una sola salida, que se aplica a otros elementos de

<sup>3</sup> Conocido como ADALINE el cual fue desarrollado por los neurólogos Widrow y Hoff.

procesamiento dentro de la red. La entrada que recibe el  $i$ -ésimo nodo procedente del  $j$ -ésimo nodo se indica de la forma  $x_{ij}$ . Cada conexión con el  $i$ -ésimo nodo tiene asociada una cantidad llamada **peso de conexión**. El peso de la conexión proveniente del  $j$ -ésimo nodo y que llega al  $i$ -ésimo nodo se denota por  $w_{ij}$ . Todas las cantidades tienen sus analogías con las neuronas: la salida del nodo se corresponde con la frecuencia de disparo de la neurona, y los pesos se corresponden a la intensidad de las conexiones sinápticas entre neuronas.

La conexión de entrada puede ser excitatoria o inhibitoria. Las conexiones excitatorias tienen pesos positivos, y las conexiones inhibitorias tienen pesos negativos. Cada nodo calcula un valor de entrada neto mediante todas las conexiones de entrada; por lo general el valor de entrada neto se calcula sumando los valores de entrada, multiplicados por sus pesos correspondientes. La entrada para el  $i$ -ésimo nodo será

$$neta_i = \sum_j x_j w_{ij}$$

en donde el índice  $j$  recorre todas las conexiones que tenga el nodo. Hay que considerar que la excitación y la inhibición se obtienen automáticamente mediante el signo de sus pesos. Dado que el número de interconexiones en la red es muy elevado, la velocidad con la que se puede llevar a cabo este cálculo suele ser determinante para el rendimiento de la simulación de cualquier red.

Una vez que el valor neto de la entrada es calculado, se transforma el **valor de activación**, o **activación** simplemente, para ese nodo. Se puede escribir ese valor de activación en la forma

$$a_i = F_i(a_i(t-1), \text{net}_i(t))$$

que denota que la activación es una función explícita del valor neto de entrada. La activación actual puede depender del valor anterior de la activación  $a(t-1)$ . Como el valor de activación y el valor neto de entrada en la mayoría de los casos suelen ser idénticos se emplean de forma indistinta.

Para determinar el valor de salida una vez que se cuenta con el valor de activación, se aplica la **función de salida**:

$$x_i = f_i(a_i)$$

Como  $a_i = \text{net}_i$ , la función de salida puede ser escrita como:

$$x_i = f_i(\text{net}_i)$$

Las redes neuronales suelen ser vistas como sistemas dinámicos, es decir, como sistemas que evolucionan a través del tiempo. Haciendo uso de ecuaciones diferenciales para describir el ritmo de cambio de las salidas de los distintos nodos, se puede escribir un sistema de ecuaciones diferenciales para los pesos,  $w'_{ij} = G_j(w_{ij}, x_i, x_j, \dots)$  en donde  $G$  denota la **ley de aprendizaje**. El proceso de aprendizaje

consiste en hallar los pesos que describen ese conocimiento que deseamos que aprenda el sistema [Freeman/Skaupra, 1994].

# V. El modelo de Propagación

---

## 5.1. Fundamentos

Una red neuronal está constituida de elementos de procesamiento que están conectados entre sí. Cada elemento de procesamiento recibe un número de entradas,  $x$ , los cuales son evaluados con pesos,  $w$ . Mediante cálculo del peso total de la entrada, los elementos de procesamiento calculan una salida simple,  $y$ .

El aprender implica que el elemento de procesamiento cambia su conducta en respuesta al medio. Por ejemplo, si el procesamiento original da una salida de +1 en respuesta a una señal original de entrada, después de que el aprendizaje tiene su lugar, el elemento de procesamiento podría tener una salida de -1 para el mismo patrón de entrada. Entonces se diría que el elemento de procesamiento ha cambiado su pensamiento acerca de la respuesta correcta.

La salida de las neuronas es calculada como el resultado de una función de transferencia de la entrada evaluada. El aprendizaje en una red neuronal está acompañado por cambios de pesos en las entradas. La regla

de entrenamiento para la red debe definir precisamente como cambiar los pesos en respuesta a algún estímulo, para dar un par de entrada y salida<sup>4</sup>.

---

## 5.2. Antecedentes del modelo de propagación hacia atrás

### 5.2.1. Gradiente descendente

Widrow and Hoff fueron originalmente quienes propusieron esta técnica en el contexto de proporcionar un dispositivo de conducta llamado ADALINE<sup>5</sup>.

La dirección del descenso acelerado es sencilla de calcular, y es llamado *gradiente*.

La derivada de una función de una variable,  $f(x)$ , envuelve la evaluación de la función en  $f(x)$  y  $f(x + \Delta x)$ , substrayendo los dos valores, formando el cociente, y tomando el límite

$$\limite_{\Delta x \rightarrow 0} \left( \frac{f(x + \Delta x) - f(x)}{\Delta x} \right) = \frac{dy}{dx}$$

La derivada es una medida de que tan rápido la función cambia. Cuando la función tiene más de una dimensión, la derivada no es tan fácil de calcular. Por ejemplo la función de dos dimensiones,  $f(x, y)$ , da diferentes valores para  $\Delta f(x, y)$  dependiendo de la dirección en que se mueva. Puede definirse una derivada

---

<sup>4</sup> Entrenamiento es un proceso externo aplicado a la red que provoca que el aprendizaje ocurra.

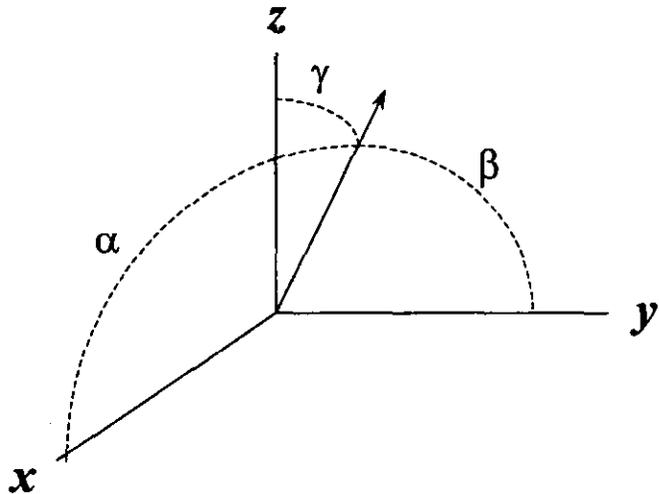
<sup>5</sup> Modelo de red neuronal consistente de un único elemento de procesamiento.

direccional,  $dw/ds$ , por analogía con la derivada de una sola dimensión. Sin embargo, la función es más compleja  $\Delta w$  es el cambio de la función  $f(x, y)$ , y  $\Delta s$  es la cantidad movida en la dirección dada por el ángulo  $\phi$ .

$$\lim_{\Delta s \rightarrow 0} \left( \frac{f(w_1) - f(w_0)}{(\Delta x^2 + \Delta y^2)^{1/2}} \right) = \frac{dw}{ds}$$

En este caso puede representarse  $dw/ds$  sin el uso la notación de límites de la siguiente manera:

$$\frac{dw}{ds} = \frac{\delta f}{\delta x} \cos \phi + \frac{\delta f}{\delta y} \sin \phi$$



Una similar expresión puede ser obtenida para la derivada direccional en tres dimensiones, usando el coseno, y por extensión para dimensiones mayores. En tres dimensiones,

supóngase que los ángulos para los ejes  $x$ ,  $y$  y  $z$  son  $\alpha$ ,  $\beta$  y  $\gamma$ . La ecuación anterior puede representarse como

$$\frac{dw}{ds} = \frac{\partial f}{\partial x} \cos \alpha + \frac{\partial f}{\partial y} \cos \beta + \frac{\partial f}{\partial z} \cos \gamma$$

La derivada direccional no es vector cuantitativo, pero sus cálculos contienen componentes que son vectores cuantitativos. Una forma más sencilla de calcular la derivada direccional podría ser romper el cálculo dentro de su parte de vector natural. La primera parte es la dirección de la derivada direccional. Una notación estándar en ingeniería y física es definir vectores unitarios  $i$ ,  $j$  y  $k$  en las direcciones de los ejes  $x$ ,  $y$  y  $z$ . Entonces nosotros podemos definir un vector unitario<sup>6</sup>  $u$ , en la dirección de la derivada, como

$$u = i \cos \alpha + j \cos \beta + k \cos \gamma + \dots$$

También puede definirse un vector de derivadas parciales

$$v = i \frac{\partial f}{\partial x} + j \frac{\partial f}{\partial y} + k \frac{\partial f}{\partial z}$$

La razón de desarrollar estos dos vectores es porque necesitamos calcular la derivada direccional como el producto interno entre dos vectores. Considerando el producto interno  $[u, v]$  se obtiene

---

<sup>6</sup> Un vector unitario es aquel vector cuya longitud es igual a 1.

$$[u, v] = \frac{\partial f}{\partial x} \cos \alpha + \frac{\partial f}{\partial y} \cos \beta + \frac{\partial f}{\partial z} \cos \gamma \dots = \frac{dw}{ds}$$

El vector  $v$  tiene algunos nombres:

$v$  = gradiente de  $f$  en el punto  $p_0$

= grad  $f$

=  $\nabla f$  = "del"  $f$

Regresando al problema original: cómo encontrar la dirección en la cual moverse para llegar al punto mínimo basado solamente en la información local. Esto significa que deseamos encontrar el valor más grande de  $dw/ds$ , o de  $-dw/ds$ , dependiendo si deseamos llegar al punto máximo o descender al punto mínimo del valle. Esto es fácil, en términos del producto interno. El punto  $p_0$ , puede escribirse como el producto punto interno

$$[u, v] = \frac{df}{ds} = |\nabla f| |u| \cos \theta,$$

donde  $\theta$  es el ángulo entre la dirección y el gradiente. Esta expresión tiene un máximo cuando el coseno de  $\theta$  es de +1 y un mínimo cuando el coseno de  $\theta$  es de -1. Así la función se incrementa más rápidamente cuando la dirección del movimiento está en la misma dirección del gradiente. La función decrementa más rápidamente cuando la dirección del movimiento está en la dirección del gradiente negativo.

El gradiente indica la dirección del movimiento para obtener el punto mínimo de manera más rápida. Esto es, ajustando los pesos para que cambien el sentido del movimiento hacia la dirección del gradiente negativo, ocasionando que el sistema converja a un mínimo local.

### 5.2.2. Cálculo del gradiente

El objetivo del aprendizaje es minimizar el error total de un conjunto de tramas de entrenamiento. Sin embargo, en la práctica el aprendizaje se produce por la corrección del error en una sola trama, entonces corregir el error en otra puede resultar difícil. El corregir el error en una trama puede ocasionar que el error de otra se incremente, pero en promedio la red deberá seguir la dirección del gradiente.

Considerando un vector de pesos  $(w_1, \dots, w_n)$ , y un peso de tendencia  $w_0$ , siempre conectado al valor +1. Asumiendo que  $f$  representa el aprendizaje de una trama de entrada. Suponiendo que el valor escalar  $t$  es valor real de salida<sup>7</sup> (solamente estamos analizando una unidad de procesamiento). Asumiendo que solamente existen dos pesos  $w_1$  y  $w_2$ , y un peso de tendencia  $w_0$ . El error  $e$  puede representarse de la siguiente manera

$$e = t - w_0 - w_1 f_1 - w_2 f_2$$

Minimizando el cuadrado del error se obtiene

---

<sup>7</sup> Este valor también es conocido como supervisor.

$$e^2 = t^2 + w_0^2 + (w_1 f_1)^2 + (w_2 f_2)^2 - 2tw_0 - 2w_1 f_1 - 2w_2 f_2 + 2f_1 w_0 w_1 + 2f_2 w_0 w_2 + 2f_1 f_2 w_1 w_2$$

Ahora para calcular el gradiente, es necesario calcular las derivadas parciales de cuadrado del error, esto es

$$\frac{\delta(e^2)}{\delta w_0}, \frac{\delta(e^2)}{\delta w_1} \text{ y } \frac{\delta(e^2)}{\delta w_2}.$$

Calculando solamente una de esta derivadas, se tiene

$$\frac{\delta(e^2)}{\delta w_0} = f_1(-2t + 2w_0 + 2w_1 f_1 + 2w_2 f_2).$$

Lo que esta dentro del paréntesis es el termino  $e$  multiplicado por -2.

En un sistema de aprendizaje supervisado, se mide el término de error calculando la diferencia entre el valor real y el valor estimado de salida. De esta forma pueden calcularse todas las derivadas parciales del término del error  $e$ , por

$$\frac{\delta(e^2)}{\delta w_j} = -2f_j e.$$

Para minimizar el error debe usarse el gradiente negativo, esto es, cambiar el peso mediante una cantidad proporcional al error del peso de entrada, es decir

$$\Delta w = \eta f_j e$$

donde  $\eta$  es una constante de aprendizaje. La expresión anterior forma la regla de aprendizaje Hebbian, con la salida real  $t$  remplazada por el error de salida. Debido a que se está aprendiendo de la diferencia ( $\Delta$ ) entre el valor real y estimado de las salidas, esta técnica es conocida como “*regla delta*” [Anderson, 1995].

Mostrando geoméricamente que el agregado de los errores promedios elevados al cuadrados es una función cuadrática del vector de peso, al dibujar el error promedio elevado al cuadrado contra todos los posibles vectores de peso obtendríamos una parábola (un paraboloides, si el problema existe en  $n$  dimensiones en el espacio). Una característica de las parábolas es que tienen un punto mínimo, en este caso representa el mínimo error promedio elevado al cuadrado. El vector de peso que corresponde a este mínimo error es el vector de peso ideal, es decir, el mejor vector de peso que posiblemente el sistema puede usar para este patrón de entrada. La regla Delta mueve el vector de peso desde el fondo hasta el punto máximo de esta parábola o tazón, completando esta tarea para mover sólo el gradiente negativo de la parábola, la cual es la ruta más directa al fondo del tazón.

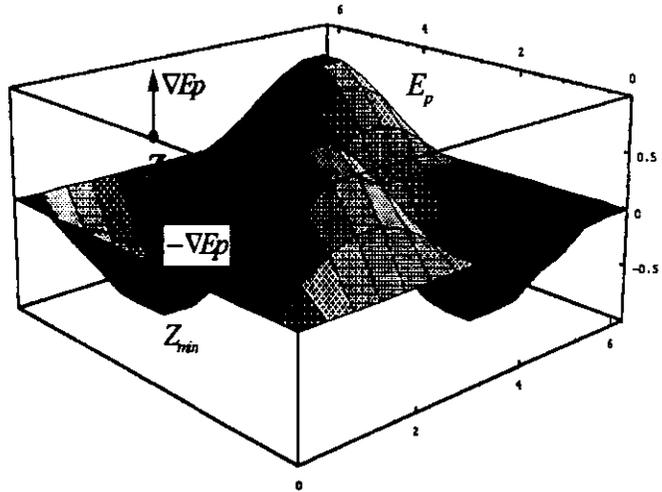


Figura 8. Si dibujáramos el error medio cuadrado contra todos los posibles vectores de peso, obtendríamos una parábola (un paraboloido, si el problema existe en n-dimensiones en el espacio). Una característica de las parábolas es que tienen un punto mínimo. Este punto representa el mínimo error cuadrado-medio.

El gradiente de una curva siempre apunta a la dirección de la cima de una curva, y el gradiente negativo ( $-\nabla E_p$ ) es la curva descendente a algún punto dado, lo cual significa que este algoritmo de aprendizaje siempre tome la ruta más eficiente desde la actual posición del vector de peso a la posición ideal, basada sobre el actual patrón de entrada. Así, la regla delta no solo minimiza el error promedio elevado al cuadrado, sino obtiene el valor más eficiente.

Ahora, el valor de  $\eta$  ciertamente debe ser positivo, si fuera negativo entonces la dirección del vector delta hacia el vector ideal sería equivocada, no la mejor. También, matemáticamente  $\eta$  debe ser menor que 2, o la red no puede

ser estable. Pero se recomienda que se guarde siempre un valor menor o igual a uno para evitar que los resultados se disparen.

El parámetro de aprendizaje  $\eta$  se interpreta como una medida de la velocidad de convergencia del vector de peso.

Una derivación de la regla delta es conocida como propagación hacia atrás (back propagation). Esta usa la regla de la cadena para derivadas parciales. Su razonamiento es

$$\frac{\delta(e^2)}{\delta w_j} = -2ef_j$$

y

$$e = (t - \sum_j w_j f_j)$$

---

## 5.3. Propagación hacia atrás

### 5.3.1. Descripción

La red aprende un conjunto predefinido de pares de entradas y salidas dados. Una vez que se ha aplicado una trama de entrada como estímulo para la primera capa de unidades de la red, ésta se va propagando a través de todas las capas superiores hasta generar una salida. La señal de salida se compara entonces con la salida deseada, y se calcula una señal de error para cada unidad de salida.

Las señales de error se transmiten entonces hacia atrás, partiendo de la capa de salida, hacia todos los nodos de la capa intermedia que contribuyan directamente a la salida. Sin embargo, las unidades de la capa intermedia sólo reciben una fracción de la señal total de error, basándose en la contribución relativa aproximada que haya aportado la unidad a la salida original. Este proceso se repite capa por capa, hasta que todos los nodos de la red hayan recibido una señal de error que describa su contribución relativa al error total. Basándose en la señal de error percibida, se actualizan los pesos de conexión de cada unidad para hacer que la red converja hacia un estado que permita codificar todas las tramas de entrenamiento.

La importancia de este proceso consiste en que, a medida que se entrena la red, los nodos de las capas intermedias se organizan por sí mismos de tal modo que los distintos nodos aprenden a reconocer distintas características del espacio total de entradas. Después del

entrenamiento, cuando se les presente una trama arbitraria de entrada que contenga ruido o que esté incompleta, las unidades en las capas ocultas de la red responderán con una salida activa si la nueva contiene una trama que se asemeje a aquella característica que las unidades individuales hayan aprendido a reconocer durante su entrenamiento. Y a la inversa, las unidades de las capas ocultas tienden a inhibir sus salidas si la trama de entrada no contiene las características que han aprendido a reconocer.

A medida que las señales se propagan a través de las diferentes capas de la red, la trama de actividad que está presente en todas las capas superiores se puede ver como una trama con características que son reconocidas por las unidades de la capa subsiguiente. La trama de salida que se genera, se puede ver como un mapa de características que ofrece una indicación de la presencia o ausencia de muchas combinaciones distintas de características dentro de la entrada. El resultado global de este comportamiento es que la red constituye un medio eficiente para permitir a un sistema de computadoras examinar tramas de datos que pueden ser ruidosas, o estar incompletas, y reconocer tramas sutiles a partir de tramas parciales.

Las redes de propagación hacia atrás (RPHA) tienden a desarrollar relaciones internas entre nodos con el fin de organizar datos de entrenamiento en clases de tramas. Esta tendencia se puede extrapolar, para llegar a la hipótesis de que todas las unidades de capas ocultas de una RPHA están asociadas de alguna manera a características específicas de la trama de la entrada como consecuencia del

entrenamiento. La red encuentra una representación interna que le permite generar las salidas deseadas cuando se le dan las entradas de entrenamiento. Esta representación interna se puede aplicar a entradas que no fueron utilizadas durante el entrenamiento. La RPHA clasificará estas entradas que no había visto hasta el momento según las características que compartan con los ejemplos de entrenamiento.

### 5.3.2. Generalización de la regla delta.

La propagación hacia atrás, no es otra cosa más que la generalización de la regla delta, que es el algoritmo de aprendizaje de la red.

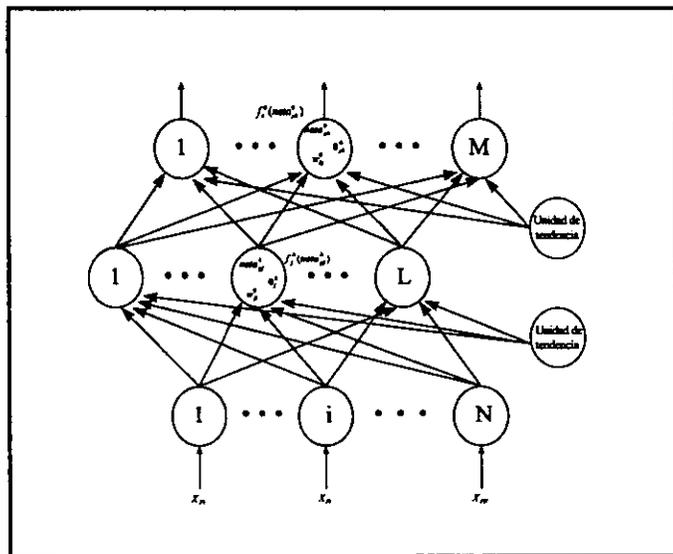


Figura 9. Representación gráfica de la red de propagación hacia atrás.

La RPHA es una red formada por tres capas, con propagación hacia adelante, completamente conectadas entre sí. Por tanto, no hay conexiones de realimentación ni conexiones que salten una capa para ir directamente hacia una capa superior. Aunque en nuestro tratamiento solo se utilizan tres capas, se puede admitir más de una capa oculta (Figura 9).

Los pesos de tendencia proporcionan un valor de entrada ficticio igual a 1 en una conexión de peso de tendencia. Se puede tratar entonces el peso del elemento de procesamiento, o peso simplemente, como cualquier otro peso contribuye al valor de entrada de la red y participa en el proceso de aprendizaje.

Se dice que una red neuronal es una *red de correspondencia* si se demuestra capaz de calcular alguna relación funcional entre su entrada y salida. Por ejemplo, si la entrada de la red es el valor de un ángulo  $\theta \rightarrow \cos(\theta)$ .

Supongamos que se tiene un conjunto de  $p$  pares de vectores  $(x_1, y_1), (x_2, y_2), \dots, (x_p, y_p)$ , que son ejemplos de una correspondencia funcional  $y = \phi(x) \mid x \in R^N, y \in R^M$ . Deseamos entrenar a la red para que aprenda que  $O = y' = \phi(x)$ . Vamos a derivar un método para hacer este entrenamiento que suele funcionar, suponiendo que los pares de vectores de entrenamiento se hayan seleccionado adecuadamente y que haya un número suficiente de ellos. Hay que recordar que el aprendizaje, en una red neuronal, significa hallar un conjunto adecuado de pesos. La técnica

de aprendizaje que aquí se describe, se asemeja al problema de hallar la ecuación de una línea que sea la que mejor se ajuste a un cierto número de pesos conocidos; además es una generalización de la regla Delta. Para un problema de ajuste de líneas es probable que se utilice una aproximación de mínimos cuadrados. Dado que la relación que pretendemos estudiar probablemente será no lineal, además de multidimensional, empleamos una versión iterativa del sencillo método de mínimos cuadrados denominada técnica del *descenso más pronunciado o gradiente descendente*.

Para empezar, revisemos las ecuaciones para el procesamiento de información que hay en la red de tres capas de la figura 9. Se aplica un vector de entrada,  $x_p = (x_{p1}, x_{p2}, \dots, x_{pn})$  en la capa de entrada de la red. Las unidades de entrada distribuyen los valores a las unidades de la capa oculta. La entrada neta de la  $j$ -ésima unidad oculta es

$$neta_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h$$

en donde  $w_{ji}^h$  es el peso de la conexión procedente de la  $i$ -ésima unidad de entrada, y  $\theta_j^h$  es el término de tendencia. El índice  $h$  se refiere a magnitudes de la capa oculta. Se supone que la activación de este nodo es igual a la entrada neta; entonces, la salida de este nodo es

$$i_{pj} = f_j^h(neta_{pj}^h)$$

Las ecuaciones para los nodos de salida son las que siguen:

$$neta_{pj}^o = \sum_{ij=1}^L w_{kj}^o i_{pj} + \theta_k^o$$

$$O_{pk} = f_k^o(neta_{pk}^o)$$

en donde el índice  $p$  se refiere a magnitudes de la capa de salida.

El conjunto inicial de valores de pesos representa una primera aproximación de los pesos correctos para el problema. A diferencia de otros métodos, la técnica que empleamos aquí no depende de hacer una buena aproximación. Existen, sin embargo, líneas generales para seleccionar los pesos iniciales.

El procedimiento básico para entrenar la red es el siguiente:

1. Se aplica el vector de entrada a la red y se calculan los correspondientes valores de salida.
2. Se comparan las salidas obtenidas con las correctas y se determina una medida de error.
3. Se determina la dirección en que debe cambiar cada peso con objeto de reducir el error.

4. Se determina la cantidad en que es preciso cambiar cada peso.
5. Se aplican las conexiones a los pesos.
6. Se repiten los pasos del 1 al 5 con todos los vectores de entrenamiento hasta que el error para todos los vectores del conjunto de entrenamiento quede reducido a un valor aceptable.

La regla delta se expresa en la siguiente ecuación

$$w_{ik}^{j+1} = w_{ik}^j + 2\mu e_k x_{ki}$$

en donde  $\mu$  es una constante positiva,  $x_{ki}$  es la  $i$ -ésima componente del  $k$ -ésimo vector de entrenamiento y

$$e_k = (t_k - y_k)$$

Se obtiene una ecuación similar cuando la red tiene más de dos capas, o cuando las funciones de salida son no lineales [Freeman/Skapura, 1994].

---

## 5.4. Algoritmo completo de propagación hacia atrás<sup>8</sup>

1. Se aplica el vector de entrada  $x_p = (x_{p1}, x_{p2}, \dots, x_{pN})^t$  a las unidades de entrada.
2. Se calculan los valores netos provenientes de las entradas para los nodos de la capa oculta

$$neta_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h$$

3. Se calculan las salidas de la capa oculta

$$i_{pj} = f_j^h(neta_{pj}^h)$$

4. Se pasa a la capa de salida. Se calculan los valores netos de las entradas para cada unidad

$$neta_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj} + \theta_k^o$$

5. Se calculan las salidas

$$o_{pk} = f_k^o(neta_{pk}^o)$$

6. Se calculan los términos de error para los nodos de salida

$$\delta_{pk}^o = (y_k - o_{pk}) f_k^{\prime o}(neta_{pk}^o)$$

---

<sup>8</sup> Tomado del libro "Redes Neuronales Algoritmos, Aplicaciones y Técnicas de Programación".

7. Se calculan los términos de error para los nodos ocultos

$$\delta_{pj}^h = f_j^{h'}(\text{net}_{pj}^h) \sum_k \delta_{pk}^o i_{pj}$$

8. Se actualizan los pesos de la capa de salida

$$w_{kj}^{o(t+1)} = w_{kj}^{o(t)} + \eta \delta_{pk}^o i_{pj}$$

9. Se actualizan los pesos de la capa oculta

$$w_{ji}^{h(t+1)} = w_{ji}^{h(t)} + \eta \delta_{pj}^h x_i$$

10. Se calcula el error

$$E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2$$

Cuando el error resulta aceptablemente pequeño para todos los pares de vectores de entrenamiento, éste puede darse por terminado.

---

## 5.5. Consideraciones

### 5.5.1. Datos de entrenamiento

Es posible utilizar todos los datos que estén disponibles para entrenar la red, aunque quizá no sea necesario utilizarlos todos. La mayoría de las veces, lo único que se necesita es un subconjunto pequeño de datos de entrenamiento de los

que se encuentran disponibles. Los datos sobrantes pueden utilizarse para probar la red.

Es necesario asegurarse que los datos de entrenamiento cubran todo el espacio de entradas esperado.

### 5.5.2. Dimensión de la red

Tres capas son suficientes la mayoría de la veces. Sin embargo, en ocasiones un problema es más fácil de resolver con más de una capa oculta. En este caso, más fácil significa que la red converge más aprisa.

### 5.5.3. Funciones de salida

Existen dos formas de función de salida

- $f_k(neta_{jk}) = neta_{jk}$
- $f_k(neta_{jk}) = (1 + e^{-neta_{jk}})^{-1}$

La primera es una función lineal de salida. La segunda función es lo que se conoce como una sigmoide, o función logística. Otra característica de la función de salida es que debe ser derivable. En el primer caso,  $f'_k = 1$ , para el segundo  $f'_k = f_k(1 - f_k) = o_{pk}(1 - o_{pk})$ .

### 5.5.4. Pesos

Los pesos deberán tener valores iniciales entre  $\pm 0.5$ , al igual que los términos de tendencia,  $\theta_i$ . Puede tratarse al

valor de tendencia como un peso más, conectado a una unidad ficticia cuya salida es siempre 1.

### 5.5.5. Parámetros de aprendizaje

La selección de un valor para el parámetro de velocidad de aprendizaje,  $\eta$ , tiene un efecto significativo en el rendimiento de la red. Normalmente,  $\eta$  debe ser un número pequeño (entre 0.05 y 0.25) para asegurar que la red llegue a converger.

Otra manera de incrementar la velocidad de convergencia consiste en hacer uso del parámetro llamado *momento*. Cuando se calcula el valor del cambio de peso,  $\Delta_p w$ , se añade una fracción del cambio anterior. Las ecuaciones de cambio de pesos de la capa de salida pasan entonces a ser

$$w_{kj}^{o(t+1)} = w_{kj}^{o(t)} + \eta \delta_{pk}^o i_{pj} + \alpha \Delta w_{kj}^{o(t-1)}$$

con una ecuación similar para la capa oculta. El parámetro  $\alpha$  suele tener un valor positivo menor que 1.

Hasta ahora hemos visto el algoritmo de propagación hacia atrás aplicado a problemas de naturaleza binaria. Para resolver problemas análogos es necesario aplicar el siguiente algoritmo<sup>9</sup> a los datos que alimentaran a la red.

---

<sup>9</sup> Tomado del trabajo "The use of microcontrollers for control and data acquisition as an alternative for chemical process modeling, using a neural network paradigm for it's programming".

***Antes de iniciar el aprendizaje o una nueva clasificación de datos:***

1. Identificar todos los valores mínimos y máximos de los datos que intervendrán en el aprendizaje.
2. Reducir de escala al valor de entrada de acuerdo a la siguiente ecuación:

$$x = (\text{valor de entrada} - \text{mínimo}) / (\text{máximo} - \text{mínimo})$$

***Después de llevar a cabo la clasificación de un nuevo caso:***

1. Para calcular la salida de una nueva clasificación de datos

$$y = (\text{valor} - 0.1) / 0.8 * (\text{máximo} - \text{mínimo}) + \text{mínimo}$$

donde valor es el resultado de la red y se encuentra entre 0 y 1.

2. Para cada valor de  $y$ :
  - Si  $y$  es menor que el valor mínimo entonces hacer el valor de entrada igual al mínimo.
  - Si  $y$  es mayor que el valor máximo entonces hacer el valor de entrada igual al máximo.

---

## **5.6. Red neuronal (Programa)**

Para implementar la red neuronal de propagación hacia atrás que pronostique los índices de contaminación, se desarrollo un programa computarizado.

### **5.6.1. Interfaz gráfica de usuario**

Fue programada en Power Builder 5.0 y permite realizar las siguientes tareas:

1. Carga de los datos de los cuales aprenderá la red.
2. Introducción de los parámetros de aprendizaje.
3. Entrenamiento de la red (aprendizaje).
4. Clasificación de nuevos casos (pronóstico).

### **5.6.2. Base de datos**

En ella se almacenaron los datos correspondientes a las variables atmosféricas y a los índices de concentración de contaminantes.

El código del programa se encuentra en el apéndice B, y la ayuda sobre el manejo del mismo en el apéndice C.

Los resultados del programa se muestran en el apéndice D.

# **VI. Alimentación de la red y generación de resultados**

---

## **6.1. Alimentación de la red**

La base de datos fue alimentada con datos del clima e índices de concentración de ozono y monóxido de carbono, ocurridos durante el periodo correspondiente a los doce meses del año 1997, y los correspondientes al primer trimestre de 1998.

### **6.1.1. Fuente de los datos**

Los índices de contaminación y las variables atmosféricas fueron proporcionados por las diez estaciones que componen la Red Automática de Monitoreo Ambiental (RAMA). También, se emplearon los registros manuales del clima proporcionados por el Servicio Meteorológico Nacional (en ausencia de registros magnéticos).

### **6.1.2. Datos que se consideraron como relevantes en la concentración de contaminantes**

#### ***Indices de contaminación***

Ozono y Monóxido de Carbono. Se tomó como unidad al IMECA que es empleado por la Red Automática de Monitoreo Atmosférico.

#### ***Variables atmosféricas***

Punto de Rocío.	% de humedad.
Temperatura .	Grados Centígrados.
Dirección del viento.	Grados.
Velocidad del viento.	Mts/Seg.

### **6.1.3. Rangos de medición**

#### ***Indices de contaminación***

Los índices de contaminación se registran cada hora, empezando desde la 7:00 a.m. hasta las 9:00 p.m. Para fines prácticos se tomó en cuenta solamente el promedio general a nivel zona metropolitana.

#### ***Variables atmosféricas***

Se registran cada hora. Para establecer una correlación con los índices de contaminación se calculó un promedio general a nivel zona metropolitana. Ejemplo:

Mes	Hora	Punto Rocío	Temperatura	Dirección Viento	Velocidad Viento	Ozono	Monóxido de carbono
1	7	65.4	5.2	183.2	0.3	17	37
1	8	73.7	5.87	184.7	0.59	21	44
1	9	67.3	7.32	211.7	0.8	18	45
1	10	56.2	10.45	122	0.84	31	46
1	11	49.1	12.92	103.7	1.08	65	45
1	12	41.4	15.14	85.4	0.91	91	42
1	13	31.9	17.4	183.4	0.86	94	38
1	14	26.6	18.84	161.4	1.34	89	33
1	15	24.7	19.12	169.7	1.4	88	31
1	16	23.5	19.54	196.3	1.68	90	29
1	17	22.8	19.62	194.3	1.6	69	28
1	18	26	18.09	241.4	1.7	52	27
1	19	31.2	16.35	278.3	1.5	22	26
1	20	35	14.91	205.1	2.03	15	25
1	21	42	13.7	166.7	2.03	12	25

#### 6.1.4. Explicación de la elección de variables.

##### **Mes.**

Se considera que no todos los meses de año presentan las mismas condiciones climatológicas, es decir, algunos presentan mayor o menor grado de contaminación. Por ejemplo los meses que comprenden el periodo de invierno presentan las llamadas inversiones térmicas.

##### **Hora.**

En un ciclo de 24 horas existen variaciones tanto de temperatura como de actividad industrial, por lo que en las primeras horas y las últimas del día se presenta un menor grado de contaminación.

***Punto de rocío.***

Representa el porcentaje de humedad presente en el ambiente. A mayor grado de humedad menor presencia de contaminantes.

***Temperatura.***

Al aumentar la temperatura ambiente se favorece la dispersión de los gases.

***Dirección del viento.***

Dependiendo del sentido en que las corrientes de aire circulen, los contaminantes se dispersarán con mayor o menor facilidad de acuerdo con las condiciones geográficas del Valle de México.

***Velocidad del viento.***

A mayor velocidad del viento menor concentración de contaminantes (depende también de la dirección del mismo).

***Ozono.***

Por ser el contaminante más producido en la ciudad y que sobrepasa con mayor frecuencia las normas internacionales permitidas.

***Monóxido de carbono.***

Es el contaminante que ocupa el segundo lugar de producción; es generado por los vehículos automotores.

---

## 6.2. Resultados (Corrida del programa)

### 6.2.1. Carga de datos

El programa fue alimentado con datos correspondientes a todos los meses de 1997 y los correspondientes al primer trimestre de 1998.

### 6.2.2. Parámetros de aprendizaje

Se eligieron como parámetros de entrada  $\alpha = 0.2$  y  $\eta = 0.25$ , y como criterio de paro un error total de  $0.001$ .

### 6.2.3. Resultados

El aprendizaje se completó en 107 iteraciones con un error aproximado de 0.000991.

#### *Matrices de pesos*

Las matrices resultantes del aprendizaje realizado por la red se muestran a continuación:

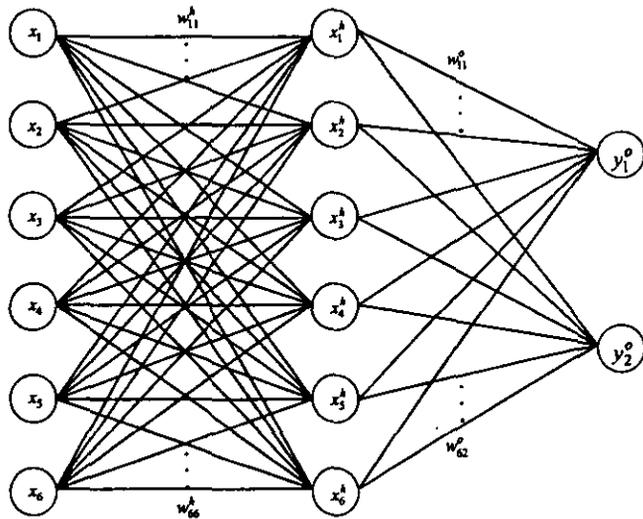
#### **Pesos entre la capa de entrada y la capa oculta**

$$W^h = \begin{bmatrix} -0.19226506 & 0.20773496 & -0.19226506 & 0.20773496 & -0.19226506 & 0.20773496 \\ 0.02764977 & 0.42765066 & 0.02764977 & 0.42765066 & 0.02764977 & 0.42765066 \\ 0.95881039 & 1.35880589 & 0.95881039 & 1.35880589 & 0.95881039 & 1.35880589 \\ -2.11235785 & -1.71235657 & -2.11235785 & -1.71235657 & -2.11235785 & -1.71235657 \\ 0.32460128 & 0.72460079 & 0.32460128 & 0.72460079 & 0.32460128 & 0.72460079 \\ 0.65739971 & 1.05739892 & 0.65739971 & 1.05739892 & 0.65739971 & 1.05739892 \end{bmatrix}$$

## Pesos entre la capa oculta y la capa de salida

$$W^o = \begin{bmatrix} -0.19228906 & 0.20773486 \\ 0.02784677 & 0.42785066 \\ 0.95981039 & 1.35880589 \\ -2.11235786 & -1.71235857 \\ 0.32480126 & 0.72480079 \\ 0.85738671 & 1.05738882 \end{bmatrix}$$

## Arquitectura de la red neuronal



## Cálculo de un nuevo valor de salida (pronóstico)

Para calcular un nuevo valor de salida deben seguirse las siguientes ecuaciones:

$$X^{(T)}W^h = X^h$$

donde

$X^{(T)}$  es el vector de valores de entrada transpuesto.

$W^h$  es la matriz de pesos entre la capa de entrada y la capa oculta.

$X^h$  es el vector de entrada que alimenta a la capa de salida.

$$X^h W^o = Y$$

donde

$W^o$  es la matriz de pesos entre la capa oculta y la capa de salida.

$Y$  es el valor esperado de salida.

Los pronósticos realizados por el programa pueden consultarse en el apéndice D.

## VII. Conclusiones

---

Los puntos que motivaron el desarrollo de éste trabajo fueron:

- 1. Los contaminantes atmosféricos incrementan o reducen su concentración dependiendo de las condiciones del clima (humedad, temperatura, dirección y velocidad del viento).*
- 2. Las propiedades de la redes neuronales para establecer relaciones entre valores de entrada y sus respectivos valores de salida.*

Basándose en los puntos anteriores se desarrolló lo siguiente:

- Una red neuronal (programa computarizado) que identifica las relaciones existentes entre el clima e índices de contaminación, es decir, pronostica los

niveles de ozono y monóxido de carbón por medio de la humedad, temperatura, dirección y velocidad del viento.

- Se analizaron los resultados arrojados por el programa, mostrando que el algoritmo de propagación hacia atrás encontró la relación existente entre condiciones ambientales y agentes contaminantes. Se observa que existe un comportamiento similar entre los índices reales de contaminación y los pronosticados. Tanto en los datos reales como en sus respectivos pronósticos, la cantidad del contaminante:

- *Registra una mayor concentración de las 12:00 a 17:00 hrs. para el ozono, y de las 10:00 a 15:00 hrs. para el monóxido de carbón.*

- *La concentración de los contaminates disminuye al principio y término del día.*

Las características de las redes neuronales de adaptarse, aprender, auto-organizarse, generalizar, ser entrenadas, proveer de memoria asociativa y distributiva y de ser resistentes a fallas, hacen que los resultados sean cercanos a los datos reales.

Lograr estos resultados implicó superar las siguientes dificultades:

1. *La gran cantidad de datos que tuvieron que ser promediados antes de alimentar el programa: 26,848 para el clima y 13, 424 entre ozono y monóxido de carbón<sup>10</sup>. El promedio se realizó, en base a las 10 estaciones de monitoreo que compone la RAMA.*
2. *En ocasiones la falta de registros del clima en medios magnéticos. Esto provocó que se recurriera a la captura manual de los mismos<sup>11</sup>. Alrededor de 3 meses (aproximadamente 5, 040 datos).*
3. *Relacionar cada promedio del clima con su respectivo cálculo de concentración de ozono y monóxido de carbono. Se realizó en base al día y hora<sup>12</sup>.*
4. *La necesidad de probar diferentes valores de momento  $\alpha$  y del parámetro de aprendizaje  $\eta$  hasta llegar a un resultado óptimo.*
5. *La red necesita realizar un gran número de iteraciones para llegar a un resultado favorable que minimice el error entre datos reales y sus aproximaciones.*

---

<sup>10</sup> 6,712 por cada variable ( humedad, temperatura, velocidad, dirección del viento, ozono y monóxido de carbón).

<sup>11</sup> Datos proporcionados por el Servicio Meteorológico Nacional

<sup>12</sup> Clima y contaminación se registran por separado.

Recomendaciones para mejorar los resultados de la red:

- Incluir más casos de entrenamiento, esto no implica que haya que desechar el aprendizaje anterior, sólo es necesario alimentar la red con nuevos datos y tomar como valores iniciales los pesos obtenidos con anterioridad.
- En el caso del ozono se pueden incluir como valor de entrada a la radiación solar<sup>13</sup>. Con la finalidad, de caracterizar de una manera más precisa su concentración.

**Conclusión final:** *si es posible el empleo de la redes neuronales para pronosticar los niveles de contaminación en el Valle de México a través de las condiciones atmosféricas.*

---

<sup>13</sup> El ozono es producto de procesos fotoquímicos entre los óxidos de nitrógeno (Nox) e hidrocarburos (HC).

# Apéndice A

---

## Índice Metropolitano de Calidad del Aire (IMECA)

Los IMECAS representan una serie de valores numéricos que relacionan el grado de contaminación atmosférica con posibles daños a la salud de las personas que habitan la Cd. de México.

El IMECA considera solamente cinco tipos de contaminantes (NO<sub>2</sub>, SO<sub>2</sub>, CO, O<sub>3</sub>, PST), de tal forma que el producto de dos de ellos representa el efecto de ambos en la atmósfera.

Este índice se basa en valores obtenidos a partir de las normas mexicanas de calidad del aire y de niveles de contaminación de los cuales se tiene conocimiento que dañan la salud. De esta forma a los primeros se les asigna un valor de 100, y a los segundos un valor de 500. Lo que da como resultado un intervalo de 0 a 500, con la excepción del producto SO<sub>2</sub> x PST que tiene una escala más amplia.

IMECAS	DESCRIPCION
100	Norma de calidad para el contaminante
200	Nivel considerado como alerta, ya que pueden presentarse daños al medio ambiente
300	Nivel de alerta se incrementa la posibilidad de daños al ambiente
400	Nivel de alarma existe mayor riesgo de que se presenten daños al ambiente
500	Se tiene la certeza de que existen daños significativos a la salud.

Como podemos ver en la tabla anterior, 100 puntos IMECA corresponde al valor de la norma para cada uno de los contaminantes. A continuación se brinda el procedimientos de conversión de concentraciones en partes por millón y microgramos por metro cúbico a puntos IMECA, y viceversa.

#### CONVERSION DE PARTES POR MILLON O MICROGRAMOS POR METRO CUBICO A PUNTOS IMECA

CONTAMINANTE	CUANDO LA CONCENTRACION DEL CONTAMINANTE SE ENCUENTRE EN EL INTERVALO	APLIQUE LA ECUACION
PARTICULAS SUSPENDIDAS TOTALES (PST)	DE 0 A 260 $\mu\text{g} / \text{m}^3$	IMECA PST=0.384615*CONCENTRACION DE PST
	DE 260 A 1000 $\mu\text{g} / \text{m}^3$	IMECA PST=0.540540*CONCENTRACION DE PST-40.540541
PARTICULAS FRACCION RESPIRABLE (PM-10)	DE 0 A 50 $\mu\text{g} / \text{m}^3$	IMECA PM10=CONCENTRACION DE PM10
	DE 51 A 350 $\mu\text{g} / \text{m}^3$	IMECA PM10=(1/2)*CONCENTRACION DE PM10+25
	DE 351 A 420 $\mu\text{g} / \text{m}^3$	IMECA PM10=(10/7)*CONCENTRACION DE PM10-300
	DE 421 A 500 $\mu\text{g} / \text{m}^3$	IMECA PM10=(10/8)*CONCENTRACION DE PM10-225

**CONVERSION DE PARTES POR MILLON O MICROGRAMOS  
POR METRO CUBICO A PUNTOS IMECA**

CONTAMINANTE	CUANDO LA CONCENTRACION DEL CONTAMINANTE SE ENCUENTRE EN EL INTERVALO	APLIQUE LA ECUACION
	DE 501 A 600 $\mu\text{g} / \text{m}^3$	IMECA PM10=CONCENTRACION DE PM10-100
BIOXIDO DE AZUFRE (SO2)	DE 0 A 0.13 PPM	IMECA SO2=769.230769*CONCENTRACION DE SO2
	DE 0.13 A 1.00 PPM	IMECA SO2=459.770114*CONCENTRACION DE SO2+40.22989
MONOXIDO DE CARBONO (CO)	DE 0 A 11 PPM	IMECA CO=9.090909*CONCENTRACION DE CO
	DE 11 A 50 PPM	IMECA CO=10.256410*CONCENTRACION DE CO-12.820512
BIOXIDO DE NITROGENO (NO2)	DE 0 A 0.21 PPM	IMECA NO2=476.190476*CONCENTRACION DE NO2
	DE 0.21 A 2.00 PPM	IMECA NO2=223.463687*CONCENTRACION DE NO2+53.07264
OZONO (O3)	0 A 0.11 PPM	IMECA O3=909.090909*CONCENTRACION DE O3
	DE 0.11 A 0.6 PPM	IMECA O3=816.32653*CONCENTRACION DE O3+10.20409

**CONVERSION DE PUNTOS IMECA A PARTES POR MILLON  
O MICROGRAMOS POR METRO CUBICO**

CONTAMINANTE	CUANDO LA CONCENTRACION DEL CONTAMINANTE SE ENCUENTRE EN EL INTERVALO	APLIQUE LA ECUACION
PARTICULAS SUSPENDIDAS TOTALES (PST)	DE 0 A 100 IMECA	CONCENTRACION DE PST=IMECA PST/0.384645
	DE 100 A 500 IMECA	CONCENTRACION DE PST=(IMECA PST+40.540541)/0.540540
PARTICULAS FRACCION RESPIRABLE (PM-10)	DE 0 A 50 IMECA	CONCENTRACION DE PM10=IMECA PM10
	DE 51 A 350 IMECA	CONCENTRACION DE PM10=(IMECA PM10-25)*1/4
	DE 351 A 420 IMECA	CONCENTRACION DE PM10=(IMECA PM10+300)/710
	DE 421 A 500 IMECA	CONCENTRACION DE PM10=(IMECA PM10+225)*8/10

### CONVERSION DE PARTES POR MILLON O MICROGRAMOS POR METRO CUBICO A PUNTOS IMECA

CONTAMINANTE	CUANDO LA CONCENTRACION DEL CONTAMINANTE SE ENCUENTRE EN EL INTERVALO	APLIQUE LA ECUACION
	DE 501 A 600 IMECA	CONCENTRACION DE PM10=(IMECA PM10+100)
BIOXIDO DE AZUFRE (SO2)	DE 0 A 100 IMECA	CONCENTRACION DE SO2=(IMECA SO2/769.230769)
	DE 100 A 500 IMECA	CONCENTRACION DE SO2=(IMECA SO2-40.22969)/459.770114
MONOXIDO DE CARBONO (CO)	DE 0 A 100 IMECA	CONCENTRACION DE CO=(IMECA CO/9.090909)
	DE 100 A 500 IMECA	CONCENTRACION DE CO=(IMECA CO+12.820541)/10.258410
BIOXIDO DE NITROGENO (NO2)	DE 0 A 100 IMECA	CONCENTRACION DE NO2=(IMECA NO2/476.190476)
	DE 100 A 500 IMECA	CONCENTRACION DE NO2=(IMECA NO2-53.07264)/223.463687
OZONO (O3)	DE 0 A 100 IMECA	CONCENTRACION DE O3=(IMECA O3/909.090909)
	DE 100 A 500 IMECA	CONCENTRACION DE O3=(IMECA O3-10.20409)/816.32653

### NORMAS DE CALIDAD DEL AIRE VIGENTES A PARTIR DE ENERO DE 1995

CONTAMINANTE	CONCENTRACION PARTES POR MILLON (PPM)	CONCENTRACION MICROGRAMOS POR METRO CUBICO ( $\mu\text{g} / \text{m}^3$ )	PROMEDIO ANUAL MICROGRAMOS POR METRO CUBICO ( $\mu\text{g} / \text{m}^3$ )
MONOXIDO DE CARBONO (CO)	11 EN 8 HORAS	12.595 EN 8 HORAS	
OZONO (O3)	0.11 EN 1 HORA	216 EN 1 HORA	
BIOXIDO DE NITROGENO (NO2)	0.21 EN 1 HORA	395 EN 1 HORA	
BIOXIDO DE AZUFRE (SO2)	0.13 EN 24 HORAS	341 EN 24 HORAS	79 MEDIA ARITMETICA ANUAL
PARTICULAS SUSPENDIDAS TOTALES (PST)		260 EN 24 HORAS	75 MEDIA ARITMETICA ANUAL
PARTICULAS SUSPENDIDAS FRACCION		150 EN 24 HORAS	50 MEDIA ARITMETICA ANUAL

**NORMAS DE CALIDAD DEL AIRE VIGENTES A PARTIR DE  
ENERO DE 1995**

CONTAMINANTE	CONCENTRACION PARTES POR MILLON (PPM)	CONCENTRACION MICROGRAMOS POR METRO CUBICO ( $\mu\text{g} / \text{m}^3$ )	PROMEDIO ANUAL MICROGRAMOS POR METRO CUBICO ( $\mu\text{g} / \text{m}^3$ )
RESPIRABLE			
PLOMO (PB)		1.5 PROMEDIO TRIMESTRAL	

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

## Apéndice B

---

### Código fuente de las funciones de propagación hacia atrás

#### f\_propagacion\_hacia\_atrás

Se determina el valor del momento, la tasa de aprendizaje, el número de nodos ocultos.

```
randomize(0)
learn = real(w_aprendizaje.sle_learn.text)
mo = real(w_aprendizaje.sle_momento.text)
superior = 0.5
inferior = -0.5
v_acelera = 1
hidden = f_calcula_nodos_ocultos(entradas, salidas)
```

Generación de los valores iniciales para comenzar el aprendizaje

```
For j = 1 To hidden
  For i = 1 To entradas
    wf[j,i] = f_inicializa_pesos_y_deltas(superior, inferior)
    dw[j,i] = 0
  Next
Next

For j = 1 To hidden
  tf[j] = f_inicializa_pesos_y_deltas(superior, inferior)
  dt[j] = 0
Next

For j = 1 To hidden
  For k = 1 To salidas
```

```

    wh[k,j] = f_inicializa_pesos_y_deltas(superior,inferior)
    dwh[k,j] = 0
  Next
Next

For j = 1 To salidas
  th[j] = f_inicializa_pesos_y_deltas(superior,inferior)
  dh[j] = 0
Next

itera = 0

```

Comienza el algoritmo de propagación hacia atrás

```

Do
  itera = itera + 1
  w_aprendizaje.st_itera.text = " Iteración No. " + string(itera)
  For pat = 1 To caso

```

Se calculan los valores de entrada para los nodos de la capa oculta

```

  For j = 1 To hidden
    sum1 = 0
    For i = 1 To entradas
      sum1 = sum1 + w[j,i] * entra[pat,i]
    Next
    sum1 = sum1 + tt[j]
    xh[j] = f_calcula_sigmoide(sum1)
  Next

```

Cálculo de los valores de entrada para los nodos de la capa de salida

```

  For k = 1 To salidas
    sum2 = 0
    For j = 1 To hidden
      sum2 = sum2 + wh[k,j] * xh[j]
    Next
    sum2 = sum2 + th[k]
    d[k] = f_calcula_sigmoide(sum2)
  Next

```

Cálculo de los términos de error para los nodos de salida

```

  errt = 0
  For k = 1 To salidas

```

```

err[k] = f_derivada_sigmoide(sum2) *
(sale[pat,k] - d[k])
errt = errt + err[k] ^ 2
Next
For j = 1 To salidas
dh[j] = learn * err[j] * d[j] + mo * dh[j]
th[j] = th[j] + dh[j]
Next

```

### Cálculo de los términos de error para los nodos ocultos

```

For j = 1 To hidden
sum = 0
For k = 1 To salidas
sum = sum + err[k] * wh[k,j]
Next
erro[j] = xb[j] * (1 - xb[j]) * sum
Next
For j = 1 To hidden
dt[j] = learn * erro[j] * xb[j] + mo * dt[j]
tt[j] = tt[j] + dt[j]
Next

```

### Actualización de los pesos en la capa de salida

```

For k = 1 To salidas
For j = 1 To hidden
dwh[k,j] = learn * err[k] * v_acelera *
xb[j] + mo * dwh[k,j]
wh[k,j] = wh[k,j] + dwh[k,j]
Next
Next

```

### Actualización de los pesos en la capa oculta

```

For i = 1 To entradas
For j = 1 To hidden
dw[j,i] = learn * erro[j] * v_acelera *
entra[pat,i] + mo * dw[j,i]
w[j,i] = w[j,i] + dw[j,i]
Next
Next

```

### Cálculo del error del caso

```

ep[pat] = 0.5 * errt
w_aprendizaje.sle_pat.text = string(pat)
w_aprendizaje.sle_ep.text = string(ep[pat])
w_aprendizaje.sle_co.text = string(f_calcula_rescala(d[2], min[8],
max[8]))

```

```

        w_aprendizaje.sle_ozono.text = string(f_calcula_rescala(d[1],
        min{7}, max{7}))
    Next

    egt = 0
    For i = 1 To caso
        egt = egt + cp[i]
    Next

    w_aprendizaje.sle_egt.text = string(egt)
    Loop While (egt > 0.001)

```

### **f\_calcula\_nodos\_ocultos**

```

return round(2 * sqrt(numero_entradas + numero_salidas) + 0.5, 0)

```

### **f\_calcula\_sigmoide**

```

return 1 / (1 + exp(-producto_interno))

```

### **f\_derivada\_sigmoide**

```

return 1 * (exp(-sigmoide) / ((1 + exp(-sigmoide)) ^ 2))

```

### **f\_inicializa\_pesos\_y\_deltas**

```

return rand(1) * (limite_superior - limite_inferior) + limite_inferior

```

# Apéndice C

---

## Manual de operación

### ¿Qué es el propagador neuronal?

#### *Descripción*

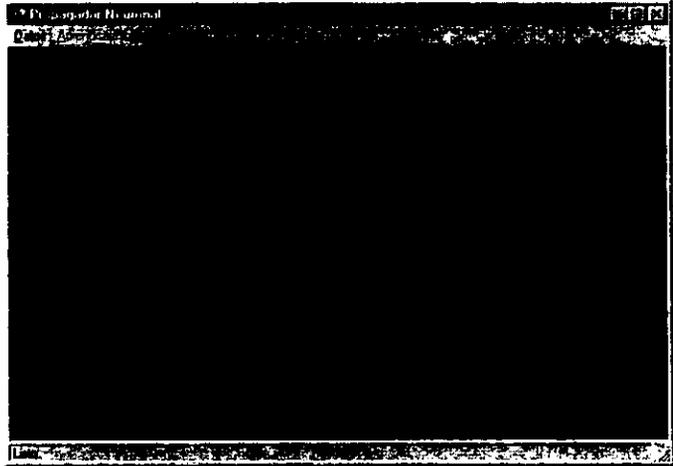
Es una red neuronal basada en la arquitectura de propagación hacia atrás. Se desarrolló exclusivamente para dar una aproximación de los índices de contaminación que se registran en el valle de México.

### ¿Cómo funciona?

Basado en un conjunto de datos de entrenamiento (casos o mediciones reales) compuestos de las condiciones ambientales y los índices de contaminación que corresponden a las mismas; el propagador neuronal trata de establecer una relación entre los datos de entrada y los datos de salida, es decir, lleva a cabo un aprendizaje que nos indique que a una determinada entrada le corresponde una salida determinada la cual tendrá un error mínimo con respecto al valor real de salida.

## Menú Principal

El menú principal consta de la opciones: Datos, Aprendizaje Salir y Ayuda.



## Menú Datos

Esta opción permite acceder al conjunto de casos de entrenamiento así con generar una nueva clasificación de datos de entrada después de realizado el aprendizaje.

### *Importar*

- Permite cargar los datos (casos de aprendizaje) almacenados en el servidor de datos SQL Anywhere. En la figura se muestra como aparece en el menú principal del propagador neuronal.



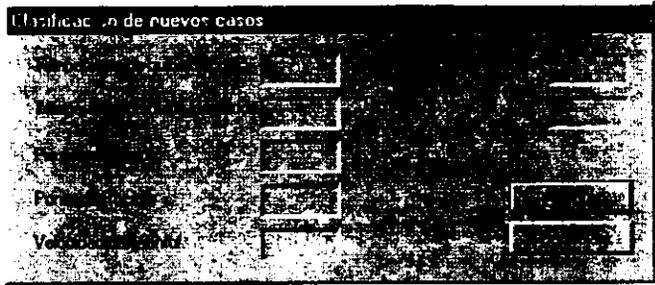
Una vez que los datos han sido cargados en memoria aparece la ventana que se muestra a continuación:



Esto nos indica que la red esta lista para comenzar el aprendizaje.

### **Clasificar**

- Realiza la clasificación de nuevos casos de entrada y obtiene como resultado un nuevo valor de salida. Esta opción solo esta disponible hasta que el aprendizaje a tomado lugar.

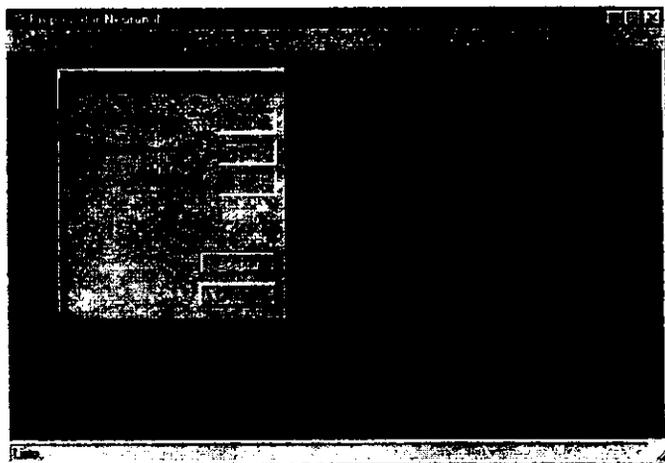


## Menú Aprendizaje

Es aquí donde se lleva a cabo el aprendizaje de la red neuronal mediante la implementación del algoritmo de propagación hacia atrás. Solo esta disponible cuando se ha realizado la importación de los casos de entrenamiento.

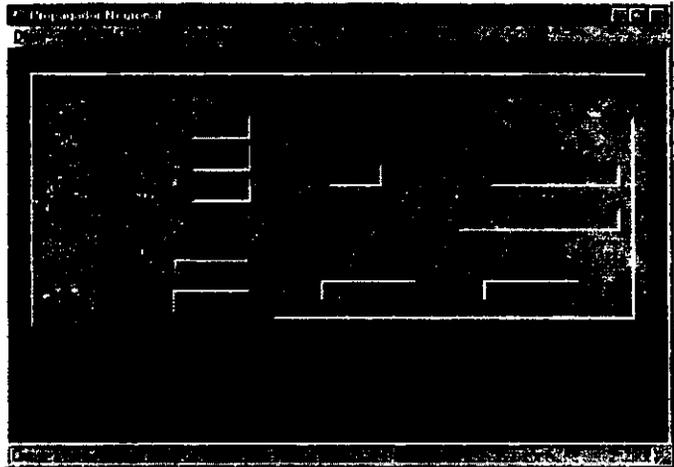
### ***Establecer parámetros***

Permite establecer el valor del momento y la tasa de aprendizaje.



### ***Iniciar Aprendizaje***

Comienza el proceso de aprendizaje una vez que se han establecido los parámetros de aprendizaje.



# Apéndice D

---

## **Pronósticos calculados por el programa computarizado**

Los resultados del programa se muestran junto con las observaciones reales, para el periodo comprendido entre los días 20 y 31 de Marzo, se muestran en la siguiente tabla:

Fecha	Hora	Punto de Ruido	Temperatura	Direccion del Viento	Velocidad del Viento	Uzono (real)	CO (real)	Ozono (promedio)	CO (promedio)	Ozono (diferencial)	CO (diferencial)
20/03/98	7	49.56	10	203.64	1.22		27		25		2
20/03/98	8	50.96	10.51	197.43	1.06		29		25		4
20/03/98	9	44.23	14.02	152.97	0.92		34		28		6
20/03/98	10	38.69	17.03	141.44	0.73		37		31		6
20/03/98	11	29.4	17.72	94.7	1.03		39		33		6
20/03/98	12	27.57	22.59	114.38	1.44		41		34		7
20/03/98	13	22.29	25.22	123.01	1.6		42		35		7
20/03/98	14	18.83	27.78	158.23	1.4		43		37		6
20/03/98	15	16.1	28.98	194.7	1.73		41		36		5
20/03/98	16	17.2	28.77	232.4	3.16		37		30		7
20/03/98	17	18.2	27.57	227.4	4.48		33		26		7
20/03/98	18	18.8	26.58	226.1	4.35		29		26		3
20/03/98	19	21.6	24.5	218.7	3.71		28		25		3
20/03/98	20	24.5	22.1	200	2.73		26		26		0
20/03/98	21	25.1	20.73	199.5	2.05		27		26		1
21/03/98	7	38	12.7	225.9	0.81		46		28		18
21/03/98	8	38.3	13.33	205	0.99		47		28		19
21/03/98	9	36	15.56	184.6	0.92		48		29		19
21/03/98	10	30.7	19.03	151	1.2		50		32		18
21/03/98	11	23.7	22.84	116.3	1.61		49		35		14
21/03/98	12	21.3	25.95	185.8	1.68		48		35		13
21/03/98	13	20.9	27	204.3	2.62		46		32		14
21/03/98	14	19.7	27.76	247.6	3.09		45		30		15
21/03/98	15	20.6	27.83	265.4	4.58		41		26		15
21/03/98	16	23.7	26.36	272.9	5.02		37		24		13
21/03/98	17	26.3	25.42	271.1	4.83		33		23		10
21/03/98	18	25.6	24.31	267.8	4.39		30		23		7
21/03/98	19	25	22.96	261.3	3.61		28		24		4
21/03/98	20	27.3	21.37	254.4	2.83		28		24		4
21/03/98	21	28.7	20.05	249.5	2.62		29		23		6

Fecha	Hora	Punto de Rocío	Temperatura	Dirección del Viento	Velocidad del Viento	Ozono (real)	CO (real)	Ozono (pronóstico)	CO (pronóstico)	Ozono (diferencia)	CO (diferencia)
22/03/98	7	51.5	10.49	209.5	1.41	19	33		25	14	8
22/03/98	8	56.2	10.22	108.1	2.22	19	32	5	24	14	8
22/03/98	9	55.2	11.78	223.7	2.81	22	31	5	22	17	9
22/03/98	10	51	14.24	123.5	2.25	33	30	5	25	28	5
22/03/98	11	35.8	16.83	88.6	2.48	50	30	16	28	34	2
22/03/98	12	25.5	19.06	92.5	2.78	57	29	13	30	28	-1
22/03/98	13	24	20.55	113.4	2.76	57	28	15	30	31	-2
22/03/98	14	23.8	21.08	117.4	2.97	55	28	17	29	38	-1
22/03/98	15	22.5	21.72	156.8	3	55	28	15	28	40	0
22/03/98	16	21.5	22.2	123.7	2.94	79	26	43	29	36	-3
22/03/98	17	20.7	22.12	189.7	3.32	76	26	27	27	49	-1
22/03/98	18	19.8	21.51	152.2	3.61	68	26	23	26	43	0
22/03/98	19	20	19.74	86.09	4.4	55	27	14	25	41	2
22/03/98	20	18.9	17.78	84.7	3.71	41	27	12	26	29	1
22/03/98	21	22.8	16.33	120	3.25	32	27	5	24	27	3
23/03/98	7	33.9	5.21	177.8	1.08	11	28	5	26	6	2
23/03/98	8	33.3	6.8	176.4	1.05	20	32	5	27	15	5
23/03/98	9	27.4	11.4	81	0.71	30	36	21	32	9	4
23/03/98	10	23.1	15.16	102.4	0.77	44	41	43	34	1	7
23/03/98	11	19.3	17.62	90	1.58	68	43	53	34	15	9
23/03/98	12	17.4	19.24	128	2.23	82	44	48	32	34	12
23/03/98	13	15.3	20.93	195.9	1.85	99	46	56	32	43	14
23/03/98	14	13.8	22.9	187.4	1.23	111	46	84	34	27	12
23/03/98	15	11.2	24.24	157	1.45	105	46	69	35	3	11
23/03/98	16	9.95	24.82	120.5	1.72	90	42	103	36	-16	6
23/03/98	17	9.1	24.98	174.3	2.18	81	38	80	33	-5	5
23/03/98	18	8.8	24.69	196.9	2.07	74	33	79	32	-5	1
23/03/98	19	9.8	23.15	238	2.29	59	34	52	29	7	5
23/03/98	20	16.1	20.41	171.8	3.16	47	35	23	26	24	9
23/03/98	21	24	18.01	212.6	2.53	20	35	7	24	13	11

1

Fecha	Hora	Punto de Rocío	Temperatura	Dirección del Viento	Velocidad del Viento	Ozono (real)	CO (real)	Ozono (promedio)	CO (promedio)	Ozono (diferencia)	CO (diferencia)
24/03/98	7	41.9	8.02	187.3	0.86	37	37	26	11	11	11
24/03/98	8	39.5	9.3	172.4	0.74	40	40	27	14	13	13
24/03/98	9	32.6	13.17	98.7	1.01	44	44	31	1	13	13
24/03/98	10	27.9	16.72	115	1.07	48	48	32	4	16	16
24/03/98	11	22.4	20.44	146.8	1.28	51	51	34	27	17	17
24/03/98	12	16	23.67	97.2	1.34	51	51	38	6	13	13
24/03/98	13	13.1	24.93	26.9	2.42	52	52	38	12	14	14
24/03/98	14	11.8	25.79	62	2.42	52	52	37	8	15	15
24/03/98	15	10.6	26.6	93.3	2.46	51	51	36	6	15	15
24/03/98	16	9.9	26.83	92.8	2.37	45	45	37	-18	8	8
24/03/98	17	9.5	26.64	164.3	3.08	39	39	32	7	7	7
24/03/98	18	9.8	25.98	94.8	3.3	35	35	33	-1	2	2
24/03/98	19	10.8	24.33	92.7	3.26	32	32	31	0	1	1
24/03/98	20	13.2	22.07	92.5	2.58	33	33	31	-17	2	2
24/03/98	21	19.6	19.9	53	2.97	33	33	28	-1	5	5
25/03/98	7	34.1	9.69	195	0.98	39	39	28	8	11	11
25/03/98	8	34.2	10.73	192.6	0.76	42	42	28	9	14	14
25/03/98	9	26.9	14.67	107.6	0.54	48	48	33	-23	15	15
25/03/98	10	23.7	18.06	85.7	0.92	53	53	35	-22	18	18
25/03/98	11	21.3	20.95	99.5	1.23	56	56	36	6	20	20
25/03/98	12	20	24.24	55.9	1.05	57	57	39	1	18	18
25/03/98	13	15.5	26.27	212.5	1.47	57	57	35	25	22	22
25/03/98	14	13.8	26.83	93.6	1.85	55	55	38	-2	17	17
25/03/98	15	13.2	27.3	124.3	1.41	53	53	38	1	15	15
25/03/98	16	12	27.77	182.8	1.42	47	47	36	9	11	11
25/03/98	17	11.2	27.72	231.9	1.52	40	40	34	12	6	6
25/03/98	18	11	27.21	303.8	1.85	33	33	31	28	2	2
25/03/98	19	13	25.48	211.6	2.66	29	29	29	20	0	0
25/03/98	20	16.4	22.85	50.1	3.98	29	29	28	18	1	1
25/03/98	21	16.6	21.05	76.7	2.91	26	26	29	-14	1	1

Día	Hora	Punto de Rocío	Temperatura	Dirección del Viento	Velocidad del Viento	Ozono (real)	CO (real)	Ozono (pronóstico)	CO (pronóstico)	Ozono (diferencia)	CO (diferencia)
26/03/98	7	31.5	10.5	177.7	0.89	17	34	5	29	12	5
26/03/98	8	31.8	11.5	185.8	0.8	18	38	8	29	10	9
26/03/98	9	26.1	15.15	118.5	1.1	26	43	34	32	-8	11
26/03/98	10	22.3	18.27	46.3	1.47	43	49	66	35	-23	14
26/03/98	11	19.1	21.49	72.7	1.02	80	53	98	38	-18	15
26/03/98	12	16.4	24.39	81.7	1.02	133	55	121	39	12	16
26/03/98	13	12.8	26.69	92.7	1.18	136	55	136	40	0	15
26/03/98	14	10.7	27.64	94.2	1.97	128	54	130	39	-2	15
26/03/98	15	10.1	27.87	121.3	1.31	134	53	136	39	-2	14
26/03/98	16	9.3	28.37	139.2	1.35	120	47	135	39	-15	8
26/03/98	17	8.9	28.2	150	1.77	110	41	123	37	-13	4
26/03/98	18	10.6	26.97	205.4	2.19	92	36	90	33	2	3
26/03/98	19	16	24.55	145.8	2.56	68	32	64	30	4	2
26/03/98	20	22.2	22.23	198.9	2.67	55	32	28	26	27	6
26/03/98	21	24.3	20.98	133.4	2.13	32	33	32	27	0	6
27/03/98	7	38.9	12.17	116.7	0.75	18	34	12	30	6	4
27/03/98	8	37.9	13.11	201.7	0.74	14	38	8	28	6	10
27/03/98	9	31.5	16.67	154.4	0.72	20	43	36	32	-16	11
27/03/98	10	29.9	19.55	118.7	0.89	42	46	60	34	-18	12
27/03/98	11	27.1	22.42	100	1.04	87	48	86	36	1	12
27/03/98	12	23.7	24.63	56.3	1.19	123	50	112	38	11	12
27/03/98	13	21.5	26.08	105	1.3	139	50	113	37	26	13
27/03/98	14	20.5	26.91	100.4	1.5	152	50	116	37	36	13
27/03/98	15	19.6	27.66	135	1.4	144	49	116	37	28	12
27/03/98	16	19	27.84	81.3	1.58	126	43	122	37	4	6
27/03/98	17	18.6	27.76	83.9	2	108	39	112	36	-4	3
27/03/98	18	18.4	27.17	76.4	2.51	94	35	97	34	-3	1
27/03/98	19	18.3	25.64	31.1	2.8	75	33	86	33	-11	0
27/03/98	20	22.4	23.56	54.2	2.09	52	32	69	32	-17	0
27/03/98	21	30.1	21.48	157.9	1.77	37	32	30	27	7	5

Da	Hora	Punto de Roca	Temperatura	Dirección del Viento	Velocidad del Viento	Ozono (real)	CO (real)	Ozono (pronóstico)	CO (pronóstico)	Ozono (diferencia)	CO (diferencia)
28/03/98	7	38.6	11.95	172.6	0.85	26	37	29	14	8	
28/03/98	8	37.3	12.52	187.3	0.85	26	39	28	19	11	
28/03/98	9	31.9	16.23	133.1	0.72	25	41	32	-11	9	
28/03/98	10	26.2	19.88	78.8	0.87	25	43	36	-25	7	
28/03/98	11	22.7	22.77	68.4	0.91	27	44	105	38	-13	6
28/03/98	12	19.4	24.86	56.7	1.03	114	44	124	40	-10	4
28/03/98	13	16.8	26.76	105.4	1.06	124	43	130	39	-6	4
28/03/98	14	16	27.66	135.3	0.99	117	42	157	39	10	3
28/03/98	15	14.9	28.4	94.3	1.08	105	40	141	40	22	0
28/03/98	16	13.9	28.77	109.7	1.2	171	35	185	39	33	-4
28/03/98	17	14.7	28.46	120.8	1.12	171	33	192	38	43	-5
28/03/98	18	21.4	27	152.2	1.92	157	31	87	33	45	-2
28/03/98	19	29.7	24.63	162.8	2.82	70	29	37	27	33	2
28/03/98	20	31.9	22.83	201.1	2.39	55	29	24	25	39	4
28/03/98	21	33.1	21.67	193.5	2.37	56	29	17	25	39	4
29/03/98	7	46.8	13.06	169.3	0.91	31	36	5	28	18	8
29/03/98	8	44.6	13.71	181.3	0.86	24	36	8	28	18	8
29/03/98	9	37.4	17.34	119.4	0.56	23	35	40	32	-7	3
29/03/98	10	33.1	20.52	98.9	0.84	68	34	67	34	1	0
29/03/98	11	28.8	23.38	109.4	0.88	102	34	92	36	11	-2
29/03/98	12	26.1	25.36	59.2	1.25	111	34	112	38	9	-4
29/03/98	13	24.2	26.65	112.9	1.11	192	32	115	37	27	-5
29/03/98	14	23.2	27.99	115.2	0.99	153	31	126	38	27	-7
29/03/98	15	22.4	28.66	121.9	1.24	155	30	124	37	31	-7
29/03/98	16	22.1	29.04	117	1.43	132	28	122	37	10	-9
29/03/98	17	24.7	28.38	133.4	2.3	109	26	91	33	18	-7
29/03/98	18	28.4	26.75	150	3.19	83	24	51	28	32	-4
29/03/98	19	31.5	24.83	184.7	3.17	69	24	29	26	40	-2
29/03/98	20	33.7	22.93	196	2.72	61	24	19	25	42	-1
29/03/98	21	33.5	21.64	199	2.16	54	24	18	25	36	-1

Día	Hora	Punto de Rocio	Temperatura	Dirección del Viento	Velocidad del Viento	Ozono (real)	CO (real)	Ozono (promedio)	CO (promedio)	Ozono (diferencia)	CO (diferencia)
30/03/98	7	44.7	13.82	195	0.73	28	28		28	13	0
30/03/98	8	42.3	14.83	206.4	0.79	18	32	10	28	8	4
30/03/98	9	34.3	18.56	150.1	0.94	31	36	42	32	-11	4
30/03/98	10	28.8	21.58	142.2	1.67	60	38	59	33	1	5
30/03/98	11	25.3	24.1	130.9	1.87	91	39	81	34	10	5
30/03/98	12	23.1	26.09	199	2.56	91	40	122	32	21	8
30/03/98	13	21	27.96	196	2.47	91	40	89	33	2	7
30/03/98	14	16.6	26.04	198.4	1.85	94	39	87	34	3	5
30/03/98	15	15.9	26.2	199.6	2.8	80	36	79	31	-12	5
30/03/98	16	15.7	25.87	191.3	3.17	78	30	63	30	15	0
30/03/98	17	16.3	25.22	190.6	3.6	75	26	58	28	27	-2
30/03/98	18	17.7	23.92	192.1	3.98	61	23	54	26	38	-3
30/03/98	19	20.5	21.84	187.9	3.55	65	23	62	25	44	-2
30/03/98	20	22.6	19.84	185.1	2.39	72	23	69	26	34	-3
30/03/98	21	23.1	18.62	153.8	1.63	41	24	41	27	20	-3
31/03/98	7	32.6	9.92	170.8	0.92	27	27	27	29	18	-2
31/03/98	8	30.9	11.08	171.4	0.61	17	31	10	30	9	1
31/03/98	9	24.9	14.82	132.6	0.6	70	35	34	33	19	2
31/03/98	10	21.4	17.59	147	0.89	60	39	54	34	6	5
31/03/98	11	16.7	20.98	129.3	1.01	119	42	43	36	10	6
31/03/98	12	12.4	23.41	94.4	1.25	166	44	54	39	11	5
31/03/98	13	10.16	28.12	73.1	1.6	163	45	53	41	17	4
31/03/98	14	10.1	30.12	95.4	1.66	133	46	50	41	18	5
31/03/98	15	10.02	30.93	113.4	1.82	98	43	59	40	41	3
31/03/98	16	10.82	30.64	139.5	3.32	74	38	51	35	21	3
31/03/98	17	11.42	29.78	158.9	4.09	69	33	66	32	12	1
31/03/98	18	11.82	28.61	191.4	4.09	63	29	68	29	6	0
31/03/98	19	14.5	26.27	203.5	3.77	63	27	48	28	15	-1
31/03/98	20	17.5	23.87	174.8	2.59	51	27	49	28	2	-1
31/03/98	21	19.2	22.36	186.2	1.67	39	28	47	29	-8	-1

# Bibliografía

---

*"An introduction to Neural Nets"*. Chemical Engineering Progress, agosto, 1990.

*"Contaminación ambiental, causas, efectos y prevención"*. Desarrollo y medio ambiente. Vol. 2. núm 1, 1989.

*"Información general"*. SEDUE, 1988.

Dr. Dobb's, *"Neural Nets for predicting behavior"*. 1993.

*"¿Qué es el IMECA?"*. SEDUE, 1988.

ANDERSON, James A. *"An introduction to Neural Networks"*. The MIT Press, 1995.

CHANLETT. *"La protección del medio ambiente"*. Limusa, 1989.

FILLOY García, Enrique R. *"Análisis de comportamientos dinámicos y propiedades computacionales en redes neuronales"*. Tesis. UNAM, 1992.

MORAN, Joseph M. *"Meteorology"*. 3a. ed. Macmillan. USA, 1991.

SLOANE, Christine S. *"Atmospheric Chemistry. Models and Predictions for Climate and Air Quality"*. Lewis Publishers, 1991.

SLOANE, Christine S. *"Air Pollution. Mathematical models"*. Lewis Publishers, 1991.

STOKER. *"Química ambiental"*. Ed. Prentice Hall, 1990.

VENKATRAM, A. *"A Statistical Model to Estimate Long-Term Concentrations of Pollutants Associated with Long-Range Transport"*. Atmospheric Environment, 1982.

CIHAN H. DAGLI. *"Artificial Neural Networks for Intelligent Manufacturing"*. Edit. Chapman & Hall, 1994, 469 p.p.

VÁZQUEZ ZAMORA, EDUARDO. *"¿Qué son las redes neuronales?"*. Revista de Tópicos de Investigación y Posgrado. Vol. II. Núm. 3., U.N.A.M., 1992.

VÁZQUEZ ZAMORA, EDUARDO. *"The use of microcontrollers for control and data acquisition as an alternative for chemical processes modeling, using a neural network paradigm for its programming"*. U.N.A.M., 1997.

JACEK M. ZURADA. *"Introduction to Artificial Neural Systems"*. Edit. West Publishing Company, 1992.

JAMES A. FREDMAN Y DAVID M. SKAPURA. *“Redes Neuronales Algoritmos, Aplicaciones y Técnicas de Programación”*. Edit. Addison-Wesley/Diaz Santos, 1994, 431 p.p.

JOSE R. HILERA. *“Redes Neuronales Artificiales”*. Edit. Addison Wesley Iberoamericana.

# Glosario de Términos

## **ADALINE**

Red neuronal que consta de un único elemento de procesamiento.

## **Dióxido de Azufre (SO<sub>2</sub>)**

Los óxidos de azufre (SO<sub>x</sub>) son producidos por el proceso de combustión de los combustibles (gasóleo, diesel) con contenido de azufre (carbón y petróleo), durante el proceso de combustión son oxidados, formando como producto principal dióxido de azufre (SO<sub>2</sub>).

## **Estabilidad Atmosférica**

Es una característica de la atmósfera que determina la intensidad del movimiento vertical del aire. Si la atmósfera se encuentra estable el aire permanece inmóvil o tiende a descender. Si la atmósfera está inestable, se favorece el mezclado vertical y la dispersión de los contaminantes.

### **Fluido intersticial**

Es el líquido que ocupa el espacio existente entre las neuronas.

### **IMECAS**

Representan una serie de valores numéricos que relacionan el grado de contaminación atmosférica con posibles daños a la salud de las personas.

### **Indicadores ambientales**

Representan una señal, síntoma o manifestación de una variable ambiental que permite describirla o caracterizarla.

### **Indices ambientales**

Agrupación de uno o más indicadores ambientales.

### **Inversión térmica**

Aire caliente que se extiende sobre el aire más frío como una tapadera que impide el movimiento ascendente de las corrientes del aire y retiene los contaminantes contra el suelo.

### **Iones**

Átomos o grupos de átomos cargados eléctricamente.

### **Microgramos por Metro Cubico ( $\mu\text{g}/\text{m}^3$ )**

Unidad utilizada para expresar la concentración de Partículas Suspendidas en un volumen de Aire. La concentración de partículas se determina por la diferencia de peso de un filtro, después de la extracción de un volumen de aire conocido.

### **Monóxido de Carbono (CO)**

Es un gas inodoro e incoloro generado por la combustión incompleta del carbón, petróleo y otros combustibles.

### **Neurodo**

Elemento de procesamiento dentro de una red neuronal.

### **Neurona**

Célula altamente especializada que generalmente tienen una forma estrellada, con un gran número de ramificaciones cortas (dendritas) y una sola más larga llamada axón.

### **Neurotransmisores**

Sustancias que al ser liberadas permiten la comunicación entre las neuronas.

### **Oxidos de Nitrógeno ( $\text{NO}_x$ )**

Los óxidos de nitrógeno en la atmósfera se emiten en forma de  $\text{N}_2\text{O}$ ,  $\text{NO}$ ,  $\text{NO}_2$ ,  $\text{N}_2\text{O}_3$ ,  $\text{N}_2\text{O}_4$ ,  $\text{N}_2\text{O}_5$ ,  $\text{NO}$  y  $\text{NO}_2$ . Los últimos dos compuestos son contaminantes generados

cuando el carbón o petróleo contenidos en los combustibles son quemados.

### **Ozono (O<sub>3</sub>)**

El ozono es un oxidante fotoquímico, producido por la reacción entre los hidrocarburos reactivos, óxidos de nitrógeno y la intensidad de la radiación solar. En la Ciudad de México es el principal contaminante atmosférico, su control es el más difícil debido al alto consumo de combustibles y por su complejo proceso de formación.

### **Partes por Millón (PPM)**

Unidad utilizada para expresar la concentración de contaminantes gaseosos, especificada por el Sistema Americano de Ingeniería (A.E.I.) adoptado por los E.U.A., el cual es una modificación del Sistema Británico.

### **Partículas Suspendidas Totales (PST)**

Las partículas suspendidas se constituyen por cenizas, humos, polvos, metales, alquitrán, neblinas y smog, generados por los procesos de combustión, calentamiento, producción, transporte y manipulación de materiales pulverizados.

### **Redes neuronales**

Colección de procesadores paralelos conectados unos a otros formando un grafo dirigido, organizado de tal manera que la red resultante sea adecuada para resolver el tipo de problema que se a planteado de antemano.

**Regla delta**

Aprendizaje de la diferencia entre el valor real y el valor estimado de las salidas.

**Sinapsis**

Unión existente entre dos neuronas.

**Sistemas neuronales artificiales**

Modelos computarizados que reproducen la funcionalidad cerebral.

**Tasa de declinación adiabática**

Porción de aire que pasa de un nivel bajo de altitud a un nivel de altitud mucho más elevado, y no intercambia calor con el aire de su entorno expandiéndose conforme la presión decrece y se enfría.