

**APLICACION DE REDES NEURONALES ARTIFICIALES A ESTUDIOS
NEUROFISIOLÓGICOS EN NEUROPATIAS PERIFÉRICAS FOCALES**

JORGE EDUARDO GUTIERREZ G.

LYDA PEÑA PAZ

UNIVERSIDAD AUTÓNOMA DE BUCARAMANGA

INSTITUTO TECNOLÓGICO DE ESTUDIOS SUPERIORES DE MONTERREY

SANTIAGO DE CALI 2005

**APLICACION DE REDES NEURONALES ARTIFICIALES A ESTUDIOS
NEUROFISIOLÓGICOS EN NEUROPATIAS PERIFÉRICAS FOCALES**

JORGE EDUARDO GUTIERREZ G.

LYDA PEÑA PAZ

**Trabajo Realizado como
Tesis de Grado de la
Maestría en Ciencias Computacionales**

Director: Msc. Diego Fernando Almario

**UNIVERSIDAD AUTÓNOMA DE BUCARAMANGA
INSTITUTO TECNOLÓGICO DE ESTUDIOS SUPERIORES DE MONTERREY
SANTIAGO DE CALI 2005**

Nota de Aceptación:

Firma del Jurado

Firma del Jurado

TABLA DE CONTENIDO

	Pág.
SUMMARY	11
INTRODUCCION	1
1 EL PROBLEMA	3
1.1 DESCRIPCIÓN DEL PROBLEMA	3
1.2 FORMULACIÓN DEL PROBLEMA	5
1.3 OBJETIVO GENERAL	5
1.4 OBJETIVOS ESPECIFICOS	5
1.5 JUSTIFICACIÓN	6
1.6 ALCANCES Y LIMITACIONES	7
2 MARCO DE REFERENCIA	9
2.1 ANTECEDENTES DE LA INVESTIGACIÓN	9
2.2 MARCO TEÓRICO CONCEPTUAL	11
2.2.1 <i>Medicina Electrodiagnóstica</i>	12
2.2.2 <i>Inteligencia Artificial y Medicina</i>	45
2.2.3 <i>Redes Neuronales Artificiales</i>	61
2.2.4 <i>Aplicaciones de redes neuronales a Medicina</i>	94

2.2.5	<i>Aplicaciones de redes neuronales a electrodiagnóstico</i>	104
3	METODOLOGÍA	106
3.1	DATOS	106
3.1.1	<i>Salidas deseadas</i>	106
3.1.2	<i>Selección de los datos de entrada</i>	107
3.1.3	<i>Preprocesamiento de los datos de entrada</i>	109
3.1.4	<i>Datos Faltantes</i>	110
3.1.5	<i>Fuente de los datos</i>	111
3.2	ARQUITECTURA DE LA RED	113
3.2.1	<i>Tipo de red</i>	114
3.2.2	<i>Mejorar la Generalización</i>	115
3.2.3	<i>Arquitectura de la Red 1</i>	116
3.2.4	<i>Arquitectura de la Red 2</i>	121
3.3	SOFTWARE	124
3.4	HARDWARE	125
3.5	ENTRENAMIENTO	125
3.6	VALIDACIÓN DE LA RED	126
4	RESULTADOS	127
4.1	RED 1 (ESTRUCTURA DE RED GENERAL)	127
4.2	RED 2 (RED NERVIO MEDIANO)	128

4.3	RED 3 (RED NERVIO ULNAR)	130
4.4	RED 4 (RED DE GENERALIZACIÓN)	132
4.5	VALIDACIÓN DE RESULTADOS	135
5	DISCUSIÓN	137
	CONCLUSIONES	139
	RECOMENDACIONES	141
	BIBLIOGRAFIA	142
	REFERENCIAS ELECTRONICAS	150

LISTA DE TABLAS

	Pág.
Tabla 1 Clasificación de lesiones nerviosas.....	18
Tabla 2 Valores Normales de Neuroconducción Motora.....	30
Tabla 3 Valores Normales de Neuroconducción Sensitiva.....	33
Tabla 4 Parámetros de los PAUMS en diversas patologías*	37
Tabla 5 Datos de Entrada de cada nervio.....	109
Tabla 6 Valores promedio para datos de Entrada	111
Tabla 7 Distribución de Casos para Entrenamiento.....	113
Tabla 8 Codificación de las Salidas de la Red	117
Tabla 9 Salidas de la Red de Cada Nervio.....	122
Tabla 10 Entrenamiento de la Red 1	127
Tabla 11 Entrenamiento para Red 2 (Nervio Mediano).....	129
Tabla 12 Entrenamiento para Red 3 (Nervio Ulnar).....	131
Tabla 13 Entrenamiento para Red 4 (Generalización).....	133
Tabla 14 Validación de datos para la Red 2 (Nervio Mediano).....	135
Tabla 15 Validación para la Red 3 (Nervio Ulnar)	136
Tabla 16 Validación para la red 4 (Generalización)	136

LISTA DE FIGURAS

	Pag.
Figura 1 Representación esquemática de una unidad motora	13
Figura 2 Generación de un potencial de acción	14
Figura 3 Diagrama de la Unión Neuromuscular	16
Figura 4 Tipos de lesiones nerviosas	19
Figura 5 Efecto de una compresión externa (torniquete) sobre la mielina	20
Figura 6 Desmielinización Focal	22
Figura 7 Bloqueo de Conducción.....	23
Figura 8 Neuroconducción Motora	26
Figura 9 Dispersión Temporal	29
Figura 10 Potencial de Acción Sensitivo.	31
Figura 11 Potencial de Acción de Unidad Motora.....	36
Figura 12 Patrón de interferencia.....	37
Figura 13 Pseudocódigo de un algoritmo evolutivo	59
Figura 14 Neurona Artificial.....	64
Figura 15 Neurona con Umbral de Activación	65
Figura 16 Esquema del Perceptrón	69

Figura 17 Función Escalón	69
Figura 18 Gráfica del Espacio de la Función de Error.....	71
Figura 19 Gradiente Descendente	72
Figura 20 Problema Linealmente Separable	74
Figura 21 Problema No Linealmente Separable	74
Figura 22 Gráfica de la Función Sigmoidal	75
Figura 23 Esquema de una red con retropropagación.....	76
Figura 24 Gráfica con mínimos locales	77
Figura 25 Gráfica del Gradiente.....	78
Figura 26 Gráfica del Gradiente con Momento	79
Figura 27 Error de Aprendizaje y Generalización	90
Figura 28 Estructura de la Red 1	117
Figura 29 Estructura de la Red para cada Nervio	122
Figura 30 Estructura General Red 2.....	124
Figura 31 Entrenamiento de la Red 1.....	128
Figura 32 Entrenamiento de la Red 2 (Nervio Mediano).....	130
Figura 33 Entrenamiento de la Red 3 (Nervio Ulnar).....	132
Figura 34 Entrenamiento para la Red 4 (Generalización).....	134

Resultados preliminares de esta investigación, han sido presentados en los siguientes eventos internacionales:

- IV Congreso Internacional de Informática Médica de la Habana (INFORMATICA 2003) – La Habana, Cuba Marzo de 2003 (Autores de la ponencia Lyda Peña y Jorge E. Gutiérrez, presentado por Lyda Peña)
- Congreso Mundial de la Federación Internacional de Neurofisiología Clínica (IFCN) en conjunto con la Asociación Americana de Medicina Electrodiagnóstica (AAEM), San Francisco, CA, U.S.A., septiembre de 2003. Abstract publicado en la revista *Muscle and Nerve* 28(6), 2003 (Autores de la ponencia Lyda Peña y Jorge E. Gutiérrez, presentado por Lyda Pena y Jorge E. Gutierrez).
- Congreso Latinoamericano de Medicina Física y Rehabilitacion (AMLAR 2003), Ciudad de Panamá, Panamá, Noviembre de 2003 (Autores de la ponencia Lyda Peña y Jorge E. Gutiérrez, presentado por Jorge E. Gutiérrez).

SUMMARY

Introduction: Interpreting electrophysiological studies is essentially a classification task. Artificial neural networks (ANNs) are suitable tools for classification because they are based on pattern recognition techniques.

Objectives: To develop a computer system for automated detection of focal neuropathies using ANNs.

Methods: The study was based on 300 sets of nerve conduction studies (NCSs) from three different electrodiagnostic medicine laboratories. Each input data set was formed by 11 parameters, including motor and sensory latencies, amplitudes, durations, and velocities of a single nerve. The input sets were classified into 4 focal neuropathy subgroups (*distal demyelination, proximal demyelination, generalized demyelination, axon loss*) depending on the type of nerve damage plus 1 additional for normal findings. The data were presented to a backpropagation ANN with 1 hidden layer. The network structure was modified to achieve the lowest possible mean square error. The outputs from these first-level networks were presented to a second-level network in order to detect generalized neuropathies. After training the ANNs, the classification accuracy was tested using another data set that was unknown to the networks.

Results: A classification accuracy of 99% was reached for the detection of pathologic patterns. The accuracy for focal neuropathies classification was 95.2%.

Conclusions: Neural networks classify focal neuropathy subgroups with high accuracy (>95%). This method may lead to automated focal neuropathy detection.

INTRODUCCION

Para cumplir con su misión, los médicos se enfrentan diariamente con problemas de búsqueda de soluciones diagnósticas o terapéuticas dentro de un gran espacio de posibles soluciones. Actualmente existen ayudas computacionales que ofrecen medios por los cuales puede ser más fácil la obtención de un diagnóstico. Las redes neuronales artificiales son herramientas computacionales del campo de la inteligencia artificial, especialmente útiles para el reconocimiento de patrones, y por tanto, se presentan como alternativa de ayuda en los diagnósticos médicos.

El presente trabajo presenta el resultado de la investigación realizada sobre el desarrollo y aplicación de redes neuronales artificiales para el diagnóstico de neuropatías periféricas focales, incluyendo la formulación del problema y los objetivos, el reporte de una extensa investigación bibliográfica donde se tratan temas propios de la especialidad (medicina electrodiagnóstico), las redes neuronales artificiales y las aplicaciones que se han desarrollado para la medicina tanto en el campo de la inteligencia artificial como en redes neuronales artificiales específicamente, la metodología seguida para el desarrollo y entrenamiento de la red neuronal artificial, los resultados obtenidos y conclusiones y recomendaciones para trabajos futuros.

Este trabajo se presenta como tesis de grado, requisito de la Maestría en Ciencias Computacionales ofrecida por el Instituto Tecnológico de Estudios Superiores de Monterrey (México) en convenio con la Universidad Autónoma de Bucaramanga y la Universidad Autónoma de Occidente.

1 EL PROBLEMA

1.1 DESCRIPCIÓN DEL PROBLEMA

Las diferentes herramientas desarrolladas por el hombre han tendido a mejorar sus sentidos o su forma de manipular el mundo. Con el desarrollo de nuevas técnicas computacionales se espera que éstas se conviertan en una forma de apoyo para la solución de problemas por parte de los seres humanos. La medicina es un campo donde dichas herramientas se vuelven críticas para apoyo de procesos de diagnóstico y de selección de tratamientos adecuados¹. Dadas las actuales expectativas de un mejor cuidado médico y el creciente cúmulo de conocimientos especializados, es imposible que un médico pueda dedicar a cada caso individual el tiempo adecuado y abarcar todo el estado del arte del conocimiento en su campo. La mayoría de decisiones médicas deben hacerse rápido, contando únicamente con el juicio y experiencia del profesional. Hay, entonces, una gran oportunidad para que las herramientas computacionales ayuden a organizar, catalogar y recuperar el conocimiento médico, de modo que puedan sugerir diagnósticos, tratamientos y realizar pronósticos.

La Medicina Electrodiagnóstica se basa en pruebas neurofisiológicas que apoyan la toma de decisiones diagnósticas en enfermedades neuro-musculares (aunque se aplica también a otros campos como cardiología y neurología del sistema nervioso central) y que, a

¹ SZOLOVITS P. Artificial Intelligence in Medicine. Westview Press, Boluder, CO, 1982 pp 5-7.

diferencia de otros métodos diagnósticos, dependen, en gran parte, de la experiencia y del conocimiento del examinador². Dado que los estudios son tan dependientes del examinador, es deseable contar con un sistema de apoyo a la toma de decisiones diagnósticas, que ofrezca mayor confiabilidad en gran número de casos.

Los sistemas de soporte para la toma de decisiones son herramientas que han ganado importancia en los procesos de diagnóstico médico, especialmente en aquellos casos donde dichas decisiones deben ser efectivas y confiables³. Estos sistemas se basan actualmente en diversos modelos, entre ellos los árboles de decisiones y la técnica Bayesiana, los cuales, sin embargo, han mostrado poseer importantes limitaciones³.

Casi todas las decisiones médicas pueden ser vistas como búsquedas en un espacio apropiado⁴, por ejemplo, un radiólogo planeando una secuencia de dosis de radioterapia está buscando el mejor tratamiento dentro del espacio de todos los tratamientos posibles. Los espacios de búsqueda médicos son grandes y complejos, manejan gran cantidad de información para llegar a una decisión única (por ejemplo, establecer si un determinado tumor es benigno o maligno). Algunas técnicas de inteligencia artificial, por ejemplo las redes neuronales artificiales, intentan dar solución a dichos problemas pues son muy útiles formulando búsquedas en este tipo de espacios⁵.

² DUMITRI D. *Electrodiagnostic Medicine*. Williams and Wilkins, Philadelphia, 1995, pp 7-8

³ PODGOLEREC, V., KOKOL, P. Medical diagnosis using Genetic Programming. Proceedings of the 12th IEEE Symposium on Computer-based Medical Systems CBMS'99 Stamford, CN, June 1999

⁴ PEÑA-REYES C.A., SIPPER M. Evolutionary computation in medicine: an overview. *Artificial Intelligence in Medicine*. Vol 19 (1), 2000, pp 1-23

⁵ PEÑA-REYES C.A., Op. Cit

1.2 FORMULACIÓN DEL PROBLEMA

¿Cómo apoyar los estudios neurofisiológicos en el diagnóstico de neuropatías periféricas focales empleando redes neuronales artificiales?

1.3 OBJETIVO GENERAL

Construir una red neuronal artificial basada en software, que apoye los estudios neurofisiológicos en el diagnóstico de neuropatías periféricas focales.

1.4 OBJETIVOS ESPECIFICOS

- Seleccionar los patrones relevantes de anormalidad en los estudios electrodiagnósticos.
- Codificar los patrones seleccionados para ser introducidos en la red neuronal.
- Desarrollar una red neuronal artificial que determine la existencia o no de neuropatías focales a partir de los datos obtenidos en los estudios de neuroconducción.
- Recolectar datos de fuentes confiables, para realizar las etapas de entrenamiento y generalización de la red neuronal artificial.
- Entrenar la red neuronal artificial hasta optimizar su funcionamiento.
- Validar el sistema desarrollado, comparando su eficacia contra la de un grupo de neurofisiólogos expertos.

1.5 JUSTIFICACIÓN

Casi ningún tratamiento médico será adecuado si el diagnóstico es errado. En Medicina Electrodiagnóstica, en muchos casos, a pesar de contar con una adecuada historia clínica, examen físico y resultados de pruebas neurofisiológicas, no es posible llegar a una certeza diagnóstica aceptable. Los métodos de apoyo basados en redes neuronales artificiales permiten, dada la gran cantidad de información que pueden manejar, mejorar la confiabilidad del diagnóstico y, gracias a su capacidad de aprendizaje, ser más efectivos. Además, el uso de éstas herramientas de apoyo para la toma de decisiones diagnósticas, hace que éstas últimas no sean tan dependientes del conocimiento o de la memoria del médico, de igual manera, permiten diseminar, en forma económica, la mejor experticia a regiones geográficas donde no se cuenta con ella.

Contar con una aplicación basada en redes neuronales artificiales, como método de apoyo para el diagnóstico, podría permitir reducir el número de pruebas, para llegar a un diagnóstico definitivo. Lo cual redundaría en una disminución de costos para el sistema médico.

Desde un punto de vista académico, la aplicación de tecnologías de Inteligencia Artificial, tales como Redes Neuronales Artificiales, en problemas de diagnóstico en el área de medicina, es un campo relativamente inexplorado, más aún en nuestro país. Este trabajo servirá como base para el desarrollo de aplicaciones que empleen las Redes Neuronales Artificiales sobre problemas reales y extender dicho conocimiento a otros dominios.

1.6 ALCANCES Y LIMITACIONES

Dada la gran variedad de enfermedades neuro-musculares, es imposible considerar en este trabajo un modelo que las incluya a todas, por eso el estudio se limitará al análisis del proceso diagnóstico de las neuropatías periféricas focales de miembros superiores (por ejemplo, el síndrome del túnel carpiano o el síndrome de canal cubital). Estas patologías, a la vez frecuentes e importantes desde el punto de vista social y laboral, tienen un número suficiente de variables involucradas que justifican el apoyo de un método de soporte diagnóstico. Se escogieron para este trabajo las neuropatías focales de miembros superiores por su alta incidencia y el impacto económico de las mismas. La incidencia de túnel carpiano (compresión del nervio mediano en la muñeca), por ejemplo, en la población general se ha estimado entre un 7% y un 14%⁶, mientras que en algunas enfermedades metabólicas como diabetes puede presentarse hasta en un 25% de los casos y en algunas ocupaciones que involucran actividad manual repetitiva puede llegar hasta el 41%⁷. Los costos médicos y la pérdida de productividad por túnel carpiano se han estimado en Estados Unidos en U\$29,000 de promedio por trabajador afectado. El costo total por año en ese país alcanza los 20 billones de dólares⁸. No hay estudios colombianos sobre el costo de estas patologías. La incidencia de otra de las patologías más frecuentes el síndrome de canal cubital (compresión del nervio ulnar en codo) puede ser hasta de un 3% de la

⁶ NORDSTROM D. et al. Incidence of diagnosed carpal tunnel syndrome in a general population. *Epidemiology* 1998;9(3):342-5.

⁷ ZAKARIA D. et al. Work-related cumulative trauma disorders of the upper extremity: Navigating the epidemiologic literature. *Am J Ind Med* 2002; 42(3) p 258-69.

⁸ SILVERSTEIN B. et al. Claims incidence of work-related disorders of the upper extremities: Washington State, 1987 through 1995. *Am J Public Health* 1998; 88 p.1827-33.

población general y hasta el 0.8 % de trabajadores con ocupaciones que implican actividad repetitiva⁹.

En el estudio se pretende llegar hasta la implementación de una red neuronal artificial, que mejore la eficiencia y la confiabilidad de los estudios electrodiagnósticos en el problema específico de la predicción de las neuropatías periféricas, maximizando la información obtenida en cada paso del diagnóstico y minimizando la complejidad, es decir maximizar la probabilidad de un diagnóstico correcto (cerca de 1.0) y minimizar la cantidad de pasos para obtenerlo.

⁹ DESCATHA A et al. Incidence of ulnar nerve entrapment at the elbow in repetitive work. Scand J Work Environ Health. 2004 Jun;30(3) p.234-40.

2 MARCO DE REFERENCIA

2.1 ANTECEDENTES DE LA INVESTIGACIÓN

Para establecer los antecedentes de la investigación se realizó una búsqueda sistemática en las bases de datos Medline, National library of Medicine (NLM), Association for Computing Machinery (ACM) digital library, IEEE electronic library y en el buscador Google (<http://google.com>) usando los términos “artificial neural networks”, “artificial intelligence”, “medicine”, “electrodiagnosis”, “Electromyography”, “focal neuropathy”, “EMG” y “neuromuscular disorders”.

En la búsqueda, realizada en Junio de 2003, se encontraron más de 300.000 referencias sobre redes neuronales artificiales, de ellas 5234 eran relacionadas con aplicaciones médicas y 26 relacionadas específicamente con el campo de la Medicina Electrodiagnóstica. Se contactaron los autores de algunas de dichas publicaciones para solicitar mayor información y obtener bibliografía adicional.

Entre las aplicaciones médicas de las redes neuronales artificiales se destacan¹⁰:

- Diagnóstico de enfermedades cardíacas
- Clasificación de tumores

- Determinación de factores de riesgo en Diabetes
- Determinación de alteraciones del sueño
- Análisis de arritmias basándose en el electrocardiograma
- Predicción de demencias
- Imagenología médica
- Procesamiento de señales eléctricas biológicas
- Diagnóstico de enfermedades metabólicas
- Diagnóstico de Leucemias
- Diagnóstico de Epilepsia basándose en electroencefalograma.
- Diagnóstico de SIDA

Dentro de las aplicaciones de técnicas de inteligencia artificial al campo del Electrodiagnóstico se pueden destacar las siguientes:

- Diseño de diagramas de flujo para orientar los pasos diagnósticos en diferentes patologías neuromusculares¹¹.
- Uso de técnicas estadísticas Bayesianas para aumentar la sensibilidad y la especificidad de los estudios electrodiagnósticos, basadas en la incidencia y la

¹⁰PACIFIC NORTHWEST NATIONAL LABORATORY. Neural Networks: Applications in Medicine and Health. <URL: <http://www.emsl.pnl.gov:2080/proj/neuron/neural/bib/medicine.html#>>

¹¹ ZORMAN M et al. Decisión Tree's Induction Strategies on a Hard Real World Problem. Proceedings of the 13rd IEEE Symposium on Computer-based Medical Systems CBMS'00, Houston TX, June 2000

prevalencia de las patologías y en el valor predictivo de diferentes signos, síntomas y pruebas¹².

- Descomposición y análisis automático de señales electromiográficas¹³.
- Utilización de redes neuronales para clasificación de enfermedades neuromusculares basándose en análisis de datos de estudios electrodiagnósticos¹⁴.
- Uso de algoritmos genéticos en la evaluación de enfermedades neuromusculares¹⁵.

En la revisión de antecedentes, no se encontraron aplicaciones específicas de uso de redes neuronales artificiales para detección de neuropatías periféricas focales.

En el marco teórico se detallan las aplicaciones de redes neuronales en medicina y en Electrodiagnóstico descritas arriba.

2.2 MARCO TEÓRICO CONCEPTUAL

¹² NIKOVSKY Daniel. Constructing Bayesian Networks for Medical Diagnosis from Incomplete and Partially Correct Statistics. IEEE Transactions on Knowledge and Data Engineering. Vol 12 (4), 2000, pp 509-516.

¹³ ZAZULA D, KOROSIC D, SOSTARIC A. Computer-Assisted Decomposition of the Electromyograms . Proceedings of the 11th IEEE Symposium on Computer-based Medical Systems CBMS'98 Lubbock, TX, June 1998.

¹⁴ SCHIZAS C., PATTICHIS C., LIVESAY R., MIDDLETON L.: Neural Networks in Computer Aided Clinical Electromyography, Proceedings of the 13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Ed. By J.H. Nagel and W.M. Smith, 13, 1458-1459, 1991.

¹⁵ PATTICHIS CS, SCHIZAS CN. Genetics-Based Machine Learning for the assessment of certain neuromuscular disorders. IEEE Transactions on Neural Networks. Vol 7(2), 1996 pp 427-439.

2.2.1 Medicina Electrodiagnóstica

El electrodiagnóstico es definido por la Asociación Americana de Medicina Neuromuscular y Electrodiagnóstica AANEM (antigua Asociación Americana de Medicina Electrodiagnóstica AAEM) como *“los métodos científicos de registrar y analizar potenciales eléctricos biológicos de los sistemas nerviosos central, periférico, autónomo y de los músculos¹⁶”*.

Los estudios electrodiagnósticos deben ser considerados como una extensión del examen neurológico y son, ante todo, estudios funcionales que junto con la evaluación clínica y los estudios imagenológicos constituyen una tríada que se considera esencial en la evaluación de pacientes con alteraciones del sistema nervioso central o periférico. No basta con evidenciar una lesión anatómica en los estudios de imagen, hay que conocer desde el punto de vista funcional que daño esta causando a las estructuras nerviosas.

Una aplicación por excelencia de los estudios electrodiagnósticos es la detección de neuropatías focales. En estas patologías es posible determinar el sitio de la lesión, el tipo de daño del nervio, su severidad y dar un pronóstico¹⁷

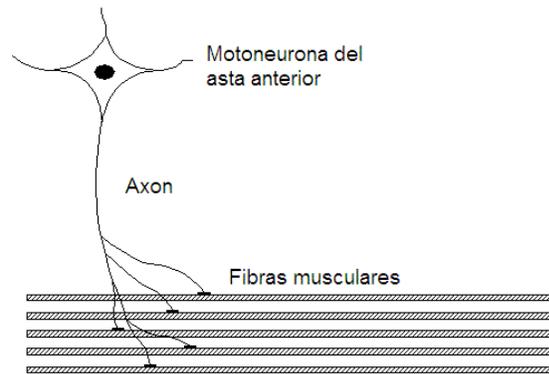
¹⁶ AAEM: Glossary of terms in Electrodiagnostic Medicine. Muscle Nerve. 2001; (suppl 10) pp S3-S12.

¹⁷ GUTIERREZ, JE. Electroagnostico. En Pedroza A (Ed): Compendio de Neurocirugia. Bogota, Quebecor Impreandes, 2002, pp 93-126.

Bases anatómicas y fisiológicas

La unidad motora: La unidad funcional de la parte motora del sistema nervioso periférico es llamada *unidad motora* y consta del soma de la motoneurona del asta anterior, su axón, la sinapsis neuromuscular y las fibras musculares que inerva. (Figura 1).

Figura 1 Representación esquemática de una unidad motora



A través de técnicas de electrofisiología clínica es posible registrar la actividad eléctrica generada en los nervios motores y sensitivos, en la placa neuromuscular y en las fibras musculares.

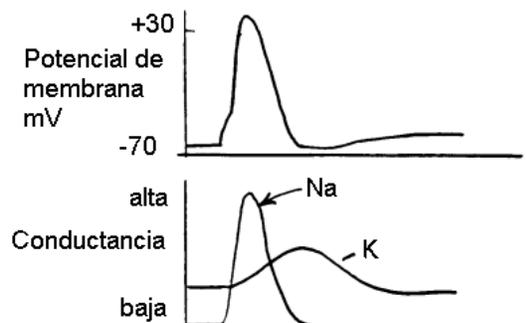
Excitabilidad del tejido nervioso y de las fibras musculares: En las neuronas al igual que en otras células hay una diferencia de voltaje entre el interior y el exterior de la célula. En reposo esta diferencia de voltaje, llamada potencial de reposo, es típicamente de -70 milivoltios (mV) en el axón de una neurona motora y de -80 mV en el músculo estriado¹⁸.

¹⁸ KANDEL ER, Schwartz JH, Jessell TM: Principles of Neural Science, 4 ed. New York, Mc Graw-Hill, 2000, pp 175-186.

El interior de la célula es negativo con respecto al exterior debido a las diferencias de concentración y permeabilidad selectiva a los iones.

Los nervios y los músculos son tejidos excitables, es decir con un estímulo consistente en una disminución del potencial de reposo hasta un punto determinado llamado umbral, se produce una despolarización de la membrana celular. Cuando el estímulo hace que alcance el umbral hay activación de canales de sodio sensibles a voltaje los cuales se abren masivamente haciendo el interior de la célula positivo con respecto al exterior con lo cual se genera un potencial de acción¹⁹. Posteriormente se cierran los canales de sodio y se activan los de potasio llevando la célula a un estado de repolarización (Figura 2).

Figura 2 Generación de un potencial de acción



El trazado superior representa las fases de la generación de un potencial de acción en una neurona motora. El trazado inferior muestra la conductancia al sodio y al potasio durante el potencial de acción.

¹⁹ Ibid.

El potencial de acción se propaga a lo largo de la célula nerviosa o muscular a medida que zonas contiguas de la membrana muscular se despolarizan. En fibras nerviosas no mielinizadas la propagación se hace a través de circuitos locales de corriente, mientras que en fibra mielinizadas la conducción se hace en forma saltatoria entre los nodos de Ranvier²⁰.

La velocidad de conducción depende del diámetro de las fibras y de la presencia o no de mielina. Las fibras grandes mielinizadas alcanzan velocidades de conducción del orden de 70 a 120 metros por segundo (m/s) mientras las fibras delgadas no mielinizadas solo alcanzan velocidades de 0.3 a 1.2 m/s. En humanos todas las fibras que conducen a más de 3 m/s son mielinizadas²¹

Debe tenerse en cuenta que la mayoría de las técnicas electrodiagnósticas de neuroconducción evalúan solamente la actividad de las fibras mielinizadas de mayor diámetro.

La unión neuromuscular: La unión entre el nervio y la fibra muscular esta formada por la terminación presináptica, la hendidura sináptica (200 a 500 Å) y la membrana postsináptica²². El terminal presináptico contiene abundantes mitocondrias y vesículas de

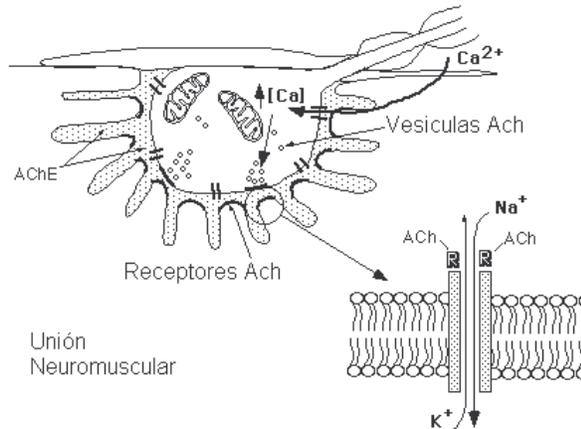
²⁰ KANDEL, Op. cit.

²¹ GUYTON AC: Textbook of Medical Physiology, 8 ed. Philadelphia, W.B. Saunders, 1991, pp 38-50.

²² ZIGMOND MJ, Bloom FE, Landis SC, eds: Fundamental Neuroscience. San Diego, Academic Press, 1998, pp 67-98.

acetilcolina (Ach), la membrana postsináptica contiene pliegues donde se ubican los receptores de Ach. (Figura 3).

Figura 3 Diagrama de la Unión Neuromuscular



Esquema de una unión neuromuscular. Se aprecian en detalle los receptores de Ach y los canales iónicos asociados a ellos.

Cuando ocurre degeneración axonal, se pierde la estructura de la placa neuromuscular y los receptores de Ach que antes sólo se encontraban en esa zona se ubican a lo largo del músculo en forma desorganizada²³. Lo anterior hace que las fibras musculares se vuelvan muy sensibles a la acción de la Ach circulante y se despolaricen espontáneamente, obteniéndose unos potenciales llamados fibrilaciones. Esta hiperexcitabilidad de la membrana muscular puede ser registrada con electrodos de aguja intramusculares.

²³ ENGEL AG, Franzini-Armstrong C, eds: Myology. New York, McGraw-Hill, 1994, pp 68-84

La fibra muscular: Un músculo esquelético está formado por miles de fibras musculares las cuales contienen haces de miofibrillas. Cada fibrilla está formada por varios filamentos y se divide en varias unidades funcionales llamadas sarcómeros. Los sarcómeros están separados por las líneas Z. Los filamentos delgados (actina, troponina y tropomiosina) se anclan a la línea Z y se extienden hasta la mitad del sarcómero, mientras que los filamentos gruesos (miosina) están en la parte media del sarcómero entremezclados con los filamentos delgados. Las diversas bandas se acortan durante la contracción.

La actividad eléctrica del músculo puede ser registrada con electrodos de aguja que evalúan parte de la unidad motora (electrodos concéntricos y monopolares), toda la unidad motora (macroelectromiografía) o la actividad de una sola fibra muscular (electromiografía de fibra única).

Lesiones de nervio periférico: El trauma de nervio periférico resulta en importante discapacidad alrededor del mundo.

Las causas de daño de las diferentes estructuras de nervio periférico son múltiples incluyendo entre otros alteraciones metabólicas (diabetes, hipertiroidismo, uremia), tóxicos, trauma abierto, trauma cerrado, heridas por arma de fuego, quemaduras eléctricas, la inyección de sustancias dentro del nervio y compresión externa en canales anatómicos.

Clasificación de las lesiones nerviosas: Las lesiones nerviosas pueden clasificarse por el mecanismo que las causa o por el tipo de daño ocasionado a las estructuras neurales.

Ambos aspectos son importantes elementos de juicio al evaluar la severidad de las lesiones, el manejo y el pronóstico de las mismas.

En cuanto al tipo de daño hay dos esquemas propuestos en la literatura: el de Seddon y el de Sunderland²⁴. (Tabla 1)

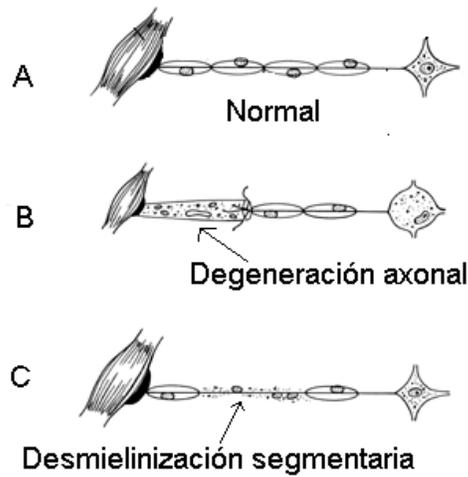
Tabla 1 Clasificación de lesiones nerviosas

SEDDON	SUNDERLAND	PATOLOGIA	PRONOSTICO
Neuropraxia	Lesión de 1 ^{er} grado	Desmielinización focal o isquemia transitoria	Excelente, recuperación completa en días o semanas.
Axonotmesis	Lesión de 2 ^{do} grado	Muerte axonal, endoneuro, perineuro y epineuro intactos	Bueno, dependiendo de la distancia al músculo
	Lesión de 3 ^{er} grado	Muerte axonal, endoneuro roto, perineuro y epineuro intactos	Regular, se pierde organización fascicular
	Lesión de 4 ^{to} grado	Muerte axonal, endoneuro y perineuro rotos, epineuro intacto	Pobre, axones crecen desorganizados, puede necesitar cirugía
Neurotmesis	Lesión de 5 ^{to} grado	Muerte axonal, endoneuro, perineuro y epineuro rotos	Malo, no hay recuperación espontánea sin cirugía.

La clasificación de Seddon es la más utilizada en la literatura. Independientemente del mecanismo de lesión de un nervio (incluyendo mecanismos no traumáticos) el daño ocurre a nivel de mielina, a nivel de los axones o a ambos niveles. (Figura 4)

²⁴ PETAJAN J: AAEM Minimonograph # 3: Motor unit recruitment. Muscle Nerve. 1991; (14) pp 489-502.

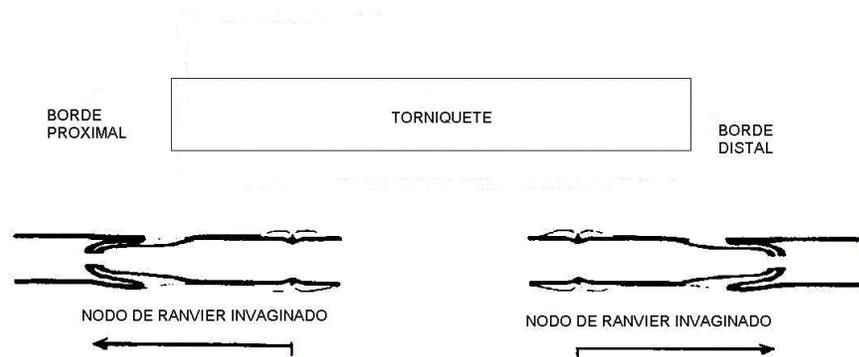
Figura 4 Tipos de lesiones nerviosas



El esquema A representa una unidad motora normal. El esquema B representa un axón que ha sufrido el proceso de degeneración Walleriana (Axonotmesis). El esquema C representa un axón que ha perdido su mielina en un segmento (Neuropráxia).

Neuropráxia: El termino Neuropráxia describe una lesión relativamente leve en la cual ocurre ya sea una isquemia reversible que no produce alteraciones estructurales o una desmielinización focalizada del nervio, sin degeneración axonal. En casos de compresión externa (por ejemplo un torniquete) la mielina se afecta pues los nodos de Ranvier de los extremos del torniquete se invaginan en los nodos vecinos. (Figura 5).

Figura 5 Efecto de una compresión externa (torniquete) sobre la mielina



Nótese que la invaginación de los nodos de Ranvier ocurre en los extremos del torniquete y no debajo de este.

Clínicamente hay un déficit motor que generalmente es incompleto y un menor compromiso de las fibras sensitivas. La recuperación de la Neuropráxia es muy rápida (horas, días, semanas o máximo unos pocos meses) y generalmente es completa.

Axonotmesis: Cuando ocurre una Axonotmesis, el axón pierde su integridad anatómica en el sitio de la lesión y dos a tres días después, el segmento distal del nervio comienza a degenerarse ocurriendo inicialmente liberación de fluido intra-axonal, edema del segmento distal, pérdida de neurofibrillas en el segmento distal y luego fragmentación del axón y de la mielina los cuales son fagocitados. (Figura 4B). En este proceso llamado degeneración Walleriana se pierde también la unión neuromuscular por lo cual el músculo queda denervado. La degeneración Walleriana puede tomar de 7 a 21 días en completarse de acuerdo a la distancia, y ocurre de proximal a distal. También hay cambios en el axón proximal (degeneración retrograda) y en el soma de la neurona.

Clínicamente en la Axonotmesis hay debilidad motora y compromiso sensitivo total o parcial dependiendo del porcentaje de axones degenerados.

Neurotmesis: En la Neurotmesis hay pérdida en la continuidad anatómica del nervio y de todas sus envolturas por lo cual el segmento distal sufre un proceso de degeneración Walleriana. Cuando los axones tratan de crecer, no encuentran el cabo distal y forman ovillos llamados neuromas. En estos casos si no se realiza cirugía no hay posibilidades de recuperación. Clínicamente hay déficit motor y sensitivo completo del nervio. Como resultado de lo anterior las fibras musculares se atrofian si no recuperan su inervación dentro de los pocos meses siguientes.

Correlación electrodiagnóstica de las lesiones nerviosas: Se debe tener en cuenta que las clasificaciones anteriores son clasificaciones histológicas y que no se deben aplicar directamente al resultado de las pruebas electrofisiológicas. Con las pruebas electrodiagnósticas se puede definir si hay desmielinización focal o generalizada, y si hay o no daño axonal el cual puede ser parcial o completo.

Lo que caracteriza a la Neuropraxia desde el punto de vista electrodiagnóstico es un *bloqueo de conducción* en el cual las zonas del nervio proximal y distal a la lesión conducen normalmente, mientras que el impulso nervioso disminuye o se pierde cuando atraviesa la zona desmielinizada pues no hay conducción saltatoria. (Figura 6).

Figura 6 Desmielinización Focal

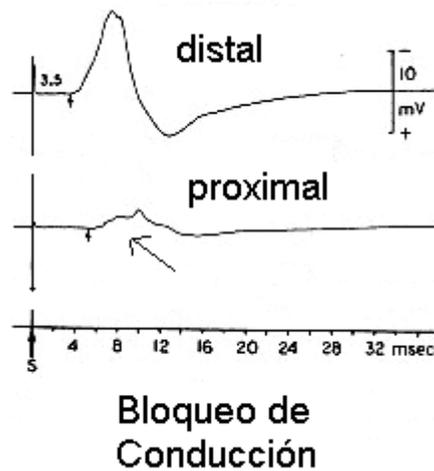


El esquema representa un axón que ha perdido su mielina entre los puntos b y c. Nótese como se pierde la conducción saltatoria produciéndose un retardo en la conducción nerviosa.

En la región internodal hay pocos canales de sodio que permitan la despolarización del nervio y además con la desmielinización se exponen canales rápidos de potasio que tienden a hiperpolarizar la membrana²⁵. Como resultado de lo anterior la velocidad del impulso nervioso en la zona desmielinizada se reduce o se bloquea completamente. Si se estimula distal a la lesión se obtiene una respuesta motora o sensitiva de amplitud normal, mientras que si se estimula proximal a la lesión, la amplitud del potencial disminuye de acuerdo al grado de bloqueo. Disminuciones de amplitud mayores de un 20 a un 30% se consideran patológicas. (Figura 7).

²⁵ DUMITRU D, Amato AA, Zwarts MJ: Electrodiagnostic Medicine 2ed. Philadelphia, Hanley & Belfus. 2001.

Figura 7 Bloqueo de Conducción



El trazado superior representa la estimulación de un nervio ulnar en la muñeca el cual tiene una respuesta de amplitud normal. Al estimular en codo (proximal a la lesión) se observa una marcada disminución de amplitud.

Lo que caracteriza a la Axonotmesis, desde el punto de vista de neuroconducciones, es la pérdida de amplitud de las respuestas motoras y sensitivas la cual es proporcional al porcentaje de axones degenerados. En lesiones axonales completas no es posible obtener registro de actividad eléctrica de las fibras sensitivas ni motoras del nervio. Durante los primeros días el segmento distal puede conducir normalmente simulando una Neuropráxia, pero después de ocurrir la degeneración Walleriana la amplitud del potencial distal disminuye o desaparece. En lesiones completas primero se pierde el potencial motor (día 9) y luego se pierde el potencial sensitivo (día 11)²⁶.

²⁶ STEWART JD: Focal Peripheral neuropathies, 3 ed. New York, Lippincott Williams & Wilkins. 1999.

En los estudios de E.M.G. de aguja se observan, a partir del día 14 a 21, signos de inestabilidad de la membrana muscular. El músculo denervado se despolariza solo y produce descargas espontáneas anormales (potenciales de fibrilación, ondas agudas positivas).

En los estudios electrodiagnósticos de la Neurotmesis hay ausencia completa de respuestas motoras y sensitivas al estimular el nervio y denervación completa del músculo en los estudios de E.M.G. de aguja. Los cambios son indistinguibles de una Axonotmesis completa y solo con el tiempo es posible diferenciarlas: en la Axonotmesis hay crecimiento axonal (1-2 milímetros por día) y reinervación subsiguiente, mientras en la Neurotmesis no hay reinervación.

Técnicas para evaluar función del nervio

Neuroconducciones: Los estudios de conducción de los nervios periféricos o craneanos representan en forma adecuada el estado fisiológico o patofisiológico de los mismos y son considerados el estándar de oro para la evaluación de las neuropatías focales y generalizadas²⁷. Hay tres clases de pruebas de neuroconducción: motoras, sensitivas y mixtas.

Neuroconducciones motoras: La conducción nerviosa de las fibras motoras de un nervio es evaluada estimulando el nervio con un estímulo único supramáximo (se estimula el

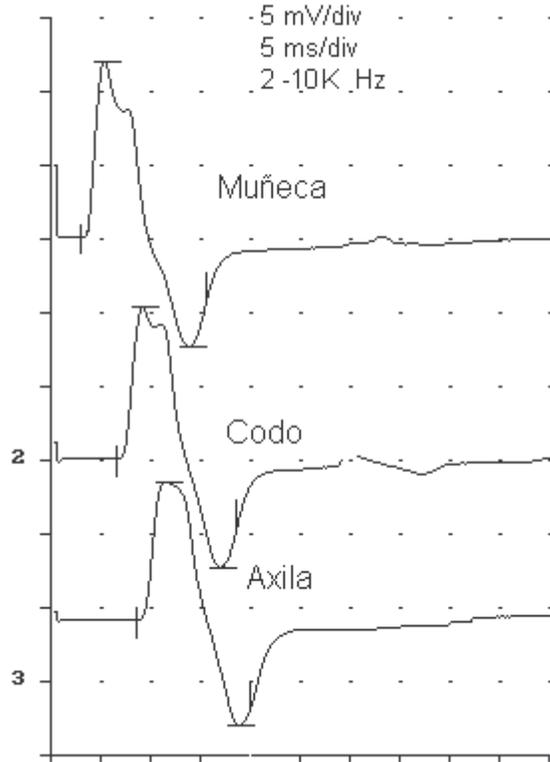
²⁷ OH, SJ: Clinical Electromyography: case studies. Baltimore, Lippincott Williams & Wilkins. 1998.

nervio hasta lograr la máxima amplitud posible y se aumenta la intensidad del estímulo en un 20% adicional para lograr normalizar la amplitud de la respuesta) en dos puntos diferentes del trayecto del nervio y registrando con electrodos de superficie sobre un músculo inervado por ese nervio. Aunque se puede registrar también con electrodos de aguja intramusculares, no se recomienda realizar las neuroconducciones motoras con este tipo de electrodos pues el área de captación de la aguja es muy pequeña y no representa la globalidad del músculo, además la morfología es muy variable incluso de estímulo a estímulo. Por lo anterior, se considera como único método válido para documentar bloqueos de conducción a los registros con electrodos de superficie²⁸.

La respuesta obtenida es llamada *Potencial de Acción Muscular Compuesto* (PAMC) u onda M pues representa la sumatoria eléctrica de la despolarización de las fibras del músculo (o en ocasiones de varios músculos) que están dentro del territorio de captación del electrodo de superficie. Como la actividad eléctrica se genera prácticamente debajo del electrodo activo, la morfología característica del PACM es bifásica con un pico negativo inicial, seguido de una deflexión positiva (figura 8).

²⁸ WILBOURN AJ: AAEM Case Report # 12: Common Peroneal Mononeuropathy at the Fibular Head. Muscle Nerve. 1986; (9) pp 825-836.

Figura 8 Neuroconducción Motora



Registro normal de un PAMC de nervio ulnar sobre el abductor del V dedo, estimulando en tres sitios diferentes: muñeca, codo y axila. Nótese la morfología bifásica y el aumento en la latencia a medida que el estímulo es más proximal.

El tiempo transcurrido entre la estimulación del punto más distal del nervio y la aparición del PAMC es llamado latencia distal o terminal.

Para calcular la velocidad de conducción motora se resta la latencia distal o terminal a la latencia obtenida al estimular proximalmente el nervio, luego se mide la distancia entre los

puntos de estímulo y se divide la distancia en milímetros sobre la resta de las latencias de acuerdo a la fórmula:

$$VCM(m/s) = \frac{Dist.(mm)}{LatP(s) - LatD(s)}$$

Donde VCM es Velocidad de Conducción Motora en metros por segundo

Dist. = distancia en milímetros

LatP= latencia proximal en segundos

LatD= latencia distal en segundos

Por ejemplo para el nervio mediano, se puede estimular a nivel del codo (respuesta proximal) y a nivel de muñeca (respuesta distal), registrando sobre el vientre del músculo abductor breve del pulgar. Si la latencia en codo es de 6.5 ms y la latencia en muñeca es de 3.2 ms, con una distancia de 18 cm entre los dos puntos de estimulación se obtendría una velocidad de conducción motora de:

$$VCM(m/s) = \frac{180mm}{6.5s - 3.2s} = \frac{180mm}{3.3s} = 54.5m/s$$

Se requiere que se pierda un 75% de las fibras motoras mielinizadas gruesas para que se altere la velocidad de conducción²⁹.

Al PAMC se le analizan varios parámetros: Latencia, amplitud, duración y área bajo la curva.

Latencia: Es el tiempo en milisegundos transcurrido entre la aplicación del estímulo y el inicio del PACM. Es una medición combinada del tiempo requerido para la despolarización del axón desde el sitio estimulado hasta la placa terminal, la transmisión neuromuscular y la despolarización de las fibras musculares. Latencias prolongadas indican generalmente alteraciones focales o generalizadas de la mielina.

Amplitud: Es el tamaño en mV del PACM (Debe medirse después de un estímulo supramaximo que estimule todas las fibras nerviosas.). Puede medirse desde la línea de base hasta el pico negativo (amplitud del pico negativo) o del pico negativo al pico positivo (amplitud pico a pico), siendo recomendable la primera. Representa en términos gruesos la cantidad de fibras musculares que se activan después de estimular el nervio y por lo tanto el número de axones viables del mismo (asumiendo una transmisión neuromuscular normal).

Cuando hay daño axonal moderado o severo hay disminución de la amplitud del PACM (sin aumento en su duración).

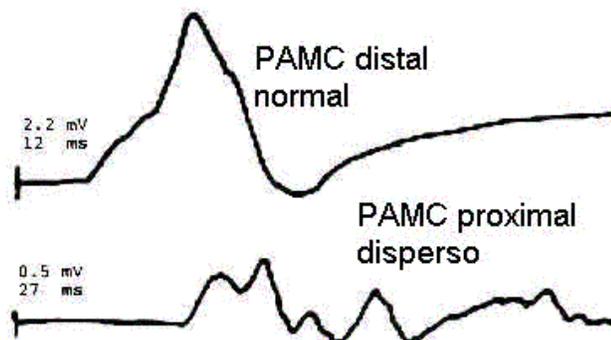
Al comparar la amplitud del PACM proximal a una lesión con la amplitud del PACM distal a la lesión se puede establecer la presencia de un bloqueo de conducción que indica desmielinización focal. En estos casos la amplitud proximal esta disminuida en un 20 a un 30% con respecto a la amplitud distal, pues el segmento del nervio distal a la lesión

²⁹ WILBOURN Op. Cit.

conduce normalmente, pero cuando el impulso eléctrico atraviesa la zona desmielinizada algunas fibras sufren retardo marcado en la conducción o se bloquean.

Duración: Representa el tiempo transcurrido entre el inicio y la finalización del PACM. Puede medirse sólo la duración del pico negativo o de todo el PACM, siendo recomendable la primera. Representa la sincronía en la descarga de las fibras musculares individuales. En situaciones normales las fibras musculares se despolarizan en forma sincrónica ante la llegada de un impulso nervioso, dando un PACM de corta duración (3 a 6 ms). Si hay fibras nerviosas que conducen a velocidad normal y otras que lo hacen más lento pues han perdido su mielina, la activación de las fibras musculares no es sincrónica y la duración del PACM aumenta causando lo que se conoce como dispersión temporal. En casos extremos de diferencias entre las velocidades de conducción de las fibras nerviosas puede verse no solo dispersión temporal sino aparición de nuevas fases del PACM. Cada fase representa una subpoblación de fibras que conduce a diferentes velocidades (Figura 9).

Figura 9 Dispersión Temporal



El trazado superior representa un PACM normal registrado estimulando distal a la lesión del nervio. El trazado inferior representa un PACM disperso registrado con un estímulo proximal a la lesión. Nótese el aumento en la duración del potencial y las diferentes fases (picos) del mismo. Área sombreada bajo la curva

Este parámetro integra la amplitud y la duración del pico negativo de PACM. Si hay daño axonal la amplitud disminuye sin aumento en la duración por lo cual hay disminución en el área. Si por el contrario hay daño de mielina, la amplitud disminuye por dispersión temporal pero la duración aumenta por lo que el área se conserva. Sus unidades son los mV/ms.

En la tabla 2 se presentan algunos valores normales de latencia, amplitud y velocidad de conducción para los nervios motores más frecuentemente examinados^{30,31}

Tabla 2 Valores Normales de Neuroconducción Motora

Nervio	Latencia distal (ms) Limite superior	Amplitud del pico negativo (mV) Limite inferior	Velocidad de Conducción (m/s) Limite inferior
Mediano	4.0	5.0	50.0
Ulnar	3.8	5.0	50.0
Radial	3.6	4.5	49.3
Femoral	4.5	4.0	---
Peroneo	5.3	3.0	40.6
Tibial	6.1	5.0	41.2

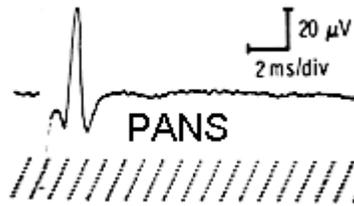
Neuroconducciones sensitivas: La conducción de las fibras sensitivas de un nervio puede ser evaluada en forma ortodrómica o antidrómica. En la forma ortodrómica se estimula la parte distal del nervio y se registra sobre una porción proximal del nervio. El registro hacerse con electrodos de superficie o con electrodos de aguja subdérmicos en la vecindad del nervio.

³⁰ AAEM: Practice Parameter for Electrodiagnostic Studies in carpal tunnel syndrome: Summary statement. Muscle Nerve. 2002; (26) pp 928-922.

³¹ WILBOURN Op. Cit.

El potencial obtenido es llamado Potencial de Acción de Nervio Sensitivo compuesto (PANS compuesto) pues representa la despolarización de todas las fibras sensitivas del nervio. El potencial es generalmente bifásico con una deflexión negativa inicial. (Figura 10).

Figura 10 Potencial de Acción Sensitivo.



Potencial de Acción Sensitivo de nervio mediano. Estimulo en la muñeca, registro en II dedo

Como no hay retardo sináptico, la velocidad de conducción de las fibras sensitivas se calcula directamente, dividiendo la distancia entre el estímulo y el electrodo de registro activo entre la latencia al inicio del PANS, según la formula:

$$VCS(m/s) = \frac{Dist.(mm)}{LatI(s)}$$

Donde VCS es Velocidad de Conducción Sensitiva en metros por segundo

Dist. = distancia en milímetros

LatI= Latencia al inicio del PANS.

Siguiendo con el ejemplo del mediano, se puede estimular a nivel de la muñeca y registrar con electrodos de anillo en el II dedo (antidrómico). Si la latencia al inicio del PANS es de 2.5 ms y la distancia es de 14 cm se obtendría una velocidad de conducción sensitiva de:

$$VCS(m/s) = \frac{140mm}{2.5s} = 56.0m/s$$

La velocidad de conducción, medida por esta técnica, representa la velocidad máxima de conducción de las fibras nerviosas sensitivas más rápidas. Si se mide la velocidad tomando la latencia al pico negativo se pueden evaluar fibras más lentas.

Al PANS compuesto se le analizan varios parámetros: Latencia, amplitud y duración.

Latencia: Es el tiempo en ms transcurrido entre la aplicación del estímulo el registro del PANS. Puede medirse la latencia al inicio del PANS o al pico negativo. El inicio del PANS no siempre es muy claro, por lo que se prefiere medir la latencia al pico negativo. Sin embargo, al medir la velocidad de conducción debe tratar de hacerse con las fibras de conducción más rápidas (latencia al inicio). Latencias prolongadas indican generalmente alteraciones focales o generalizadas de la mielina.

Amplitud: Es el tamaño en μV del PANS. Puede medirse desde la línea de base hasta el pico negativo (amplitud del pico negativo) o del pico negativo al pico positivo (amplitud pico a pico), siendo recomendable la primera. Representa en términos gruesos la cantidad de fibras sensitivas que se activan después de estimular el nervio. Cuando hay daño axonal moderado o severo hay disminución de la amplitud del PANS (sin aumento en la duración).

Duración: Representa el tiempo transcurrido entre el inicio y la finalización del PANS. Puede medirse sólo la duración del pico negativo o de todo el PANS, siendo recomendable la primera. Representa la sincronía en la descarga de las fibras sensitivas. Si hay fibras sensitivas que conducen a velocidad normal y otras que lo hacen mas lento se pierde la sincronía y la duración del PANS aumenta causando dispersión temporal. En casos extremos de diferencias entre las velocidades de conducción de las fibras nerviosa puede verse no solo dispersión temporal sino aparición de nuevas fases del PANS (especialmente en registro con electrodos de aguja cercanos al nervio). Cada fase representa una subpoblación de fibras que conduce a diferentes velocidades.

En la tabla 3 se presentan algunos valores normales de latencia, amplitud y velocidad de conducción para los nervios sensitivos más frecuentemente examinados.^{32,33} (TABLA 3)

Tabla 3 Valores Normales de Neuroconducción Sensitiva

Nervio	Latencia (ms) Limite superior	Amplitud del pico negativo (µV) Limite inferior	Velocidad de Conducción (m/s) Limite inferior
Mediano	3.5	10.0	45.6
Ulnar	3.4	10.0	45.0
Radial	3.2	10.0	44.3
Cutáneo femoral lateral	4.5	6.0	---
Sural	4.2	6.0	38.0

³² AAEM: Practice Parameter for Electrodiagnostic Studies in carpal tunnel syndrome: Summary statement. Muscle Nerve. Op. Cit.

³³ WILBOURN Op. Cit.

Neuroconducciones mixtas: Las conducciones mixtas se obtienen al estimular una porción distal de un nervio mixto (que contiene fibras motoras y sensitivas) y registrar próximalmente sobre el mismo nervio. Por ejemplo se estimula el nervio mediano en la muñeca y se registra con electrodos de superficie sobre el codo. La respuesta obtenida es llamada Potencial de Acción de Nervio Mixto compuesto y representa tanto la activación de aferentes sensitivas como de eferentes motoras).

Las mediciones, los parámetros y la interpretación clínica son similares a los del PANS.

Electromiografía de aguja: El estudio de electromiografía de aguja (E.M.G.) busca estudiar el estado fisiológico y patofisiológico de las unidades motoras por medio de una aguja insertada en el músculo estudiado. Lo que se pretende es que a partir de la actividad eléctrica registrada en el músculo se puedan inferir los cambios estructurales de la unidad motora (ejemplo pérdida de fibras musculares, aumento en el numero de fibras que dependen de una motoneurona, etc.) y por ultimo aproximarse a un diagnóstico clínico.

El electrodo de aguja, que puede ser concéntrico, bipolar o monopolar, tiene según su diseño, un determinado territorio de captación y dentro de ese territorio registra la actividad del músculo en reposo, en contracción voluntaria mínima y en contracción voluntaria máxima. Se pueden hacer registros con cualquiera de estos electrodos aunque los valores normales cambian para cada uno de ellos.

Un músculo sano en reposo generalmente no presenta actividad eléctrica espontánea, es decir presenta un silencio eléctrico visualizado en la pantalla del osciloscopio como una línea de base plana.

Durante el estudio en reposo se evalúa la actividad eléctrica generada al insertar o movilizar la aguja y la presencia de descargas espontáneas.

Los potenciales espontáneos normales se registran cuando la aguja esta cerca de la placa terminal y son: el ruido de placa terminal y los potenciales de placa terminal.

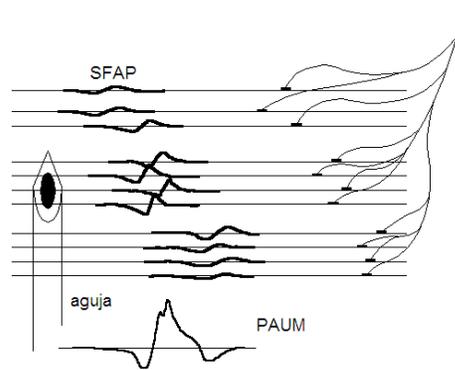
La actividad espontánea anormal puede ser de 6 tipos: Fibrilaciones, ondas agudas positivas, descargas repetitivas complejas, fasciculaciones, descargas miotónicas y descargas miokimicas.

Las fibrilaciones, las ondas agudas positivas y las descargas repetitivas complejas tienen el mismo significado: inestabilidad de la membrana muscular. Se gradúan de + a +++++, según la cantidad presente. Se originan en cualquier situación patológica que altere el potencial de reposo de la membrana muscular, por ejemplo cuando el músculo se denerva, aparecen receptores de Ach ectópicos haciendo que la fibra muscular se despolarice por efecto de la Ach circulante. Alteraciones de la estructura de la fibra muscular, por ejemplo de tipo inflamatorio (miositis), o alteraciones de los canales iónicos también causan inestabilidad de la membrana muscular la cual se despolariza en forma espontánea dando origen a la actividad anormal descrita.

Al pedirle al individuo que haga una contracción voluntaria mínima del músculo se activan las unidades motoras de umbral más bajo (1er orden). Con un electrodo de aguja dentro de

las fibras musculares es posible registrar la actividad eléctrica generada por esas unidades motoras. Cada fibra muscular al despolarizarse genera un potencial de acción de fibra muscular. La resultante de la sumatoria eléctrica de los potenciales de fibra muscular que están dentro del territorio de captación de la aguja (generalmente 5 a 12 fibras) es el llamado *Potencial de Acción de Unidad Motora*³⁴ (PAUM). (Figura 11).

Figura 11 Potencial de Acción de Unidad Motora

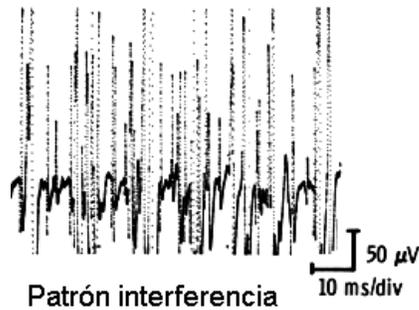


Registro de un PAUM. La aguja concéntrica registra los potenciales de acción de cada fibra muscular individual (SFAP). La sumatoria eléctrica de los potenciales de fibra única que detecta la aguja es el PAUM.

A medida que se reclutan nuevas unidades motoras la pantalla del osciloscopio comienza a llenarse de PAUMs, hasta un punto donde la línea de base se pierde completamente. A esta actividad máxima se le llama *patrón de interferencia*. (Figura 12).

³⁴ AAEM: Glossary of terms in Electrodiagnostic Medicine. Op. Cit.

Figura 12 Patrón de interferencia



Registro de actividad voluntaria en un músculo sano. Nótese la pérdida de la línea de base y la dificultad para observar PAUMs individuales

La tabla 4 resume los hallazgos característicos en la morfología de los PAUMs y en el patrón de interferencia en diversas situaciones clínicas.^{35,36}

Tabla 4 Parámetros de los PAUMS en diversas patologías*

Enfermedad	Amplitud del PAUM (mV)	Duración del PAUM (ms)	PAUMs polifásicos	Patrón de Interferencia
Normal	0.3 -2.5	6-15 ms	Amplitud y duración normal. Menos del 20% de los PAUMs son polifásicos	Completo
Miopatía	< 0.3	< 6	↓ amplitud y duración	Reclutamiento temprano
Denervación	> 2.5	> 15	↑ en amplitud y duración. Más del 20% de los PAUMs son polifásicos	Patrón ↓ con unidades motoras disparando rápido
Enfermedad Motoneurona	> 8.0	> 20	↑↑ en amplitud y duración	Patrón ↓ con unidades motoras “gigantes” disparando rápido

* Parámetros aplicables sólo a estudio con aguja concéntrica y a músculos de las extremidades.

³⁵ AMINOFF MJ: Electromyography in clinical practice, 3ed. New York, Churchill Livingstone. 1998; pp 87-112

Neuropatías periféricas

Neuropatías focales compresivas: Las neuropatías focales compresivas, llamadas también síndromes de atrapamiento cuando ocurren en canales anatómicos, son problemas muy frecuentes en la práctica electrodiagnóstica. Inicialmente son lesiones desmielinizantes, pero si la compresión es muy severa o muy prolongada puede ocurrir también daño axonal³⁷.

Lo característico de este tipo de lesión es la disminución focal de la velocidad de conducción, la presencia de bloqueo de conducción y la dispersión temporal en el sitio de la lesión. Siempre que sea posible se debe estimular antes y después de la lesión para buscar esas anormalidades.

Compresión aguda: La compresión aguda es causada por un período de presión sobre el nervio. Por ejemplo en la parálisis del sábado en la noche, la presión sobre el nervio radial produce mano caída. En este tipo de compromiso, generalmente hay un déficit incompleto, con mayor compromiso motor que sensitivo. Los estudios animales han demostrado la presencia de desmielinización focal de fibras gruesas mielinizadas.³⁸ Las anormalidades eléctricas son compatibles con desmielinización focal: bloqueo de conducción y puede haber además dispersión temporal y disminución focal de la velocidad de conducción. El pronóstico es bueno y no se requiere manejo quirúrgico.

³⁶ WILBOURN Op. Cit.

³⁷ STEWART Op. Cit.

Otros ejemplos de compresión aguda son: parálisis por torniquete, parálisis de la pierna cruzada, parálisis peri-operatoria del ulnar.

Compresión crónica (atrapamiento): El termino atrapamiento se refiere a una neuropatía focal que ocurre por compromiso de un nervio dentro de un canal anatómico estrecho. El mejor ejemplo es la compresión del nervio mediano en el canal del carpo.

Lo característico de este tipo de lesiones es³⁹:

- El mayor compromiso es sensitivo y se expresa en forma de dolor o especialmente parestesias. En casos severos puede haber compromiso motor (debilidad y atrofia).
- Puede haber dolor retrogrado hacia los segmento proximales, simulando el dolor de una radiculopatía.
- Es frecuente encontrar un signo de Tinel en el sitio afectado
- Se caracterizan inicialmente por desmielinización paranodal en los extremos de la compresión. Luego puede haber desmielinización segmentaria y en etapas tardías daño axonal.

Por lo crónico de la compresión es difícil encontrar siempre el bloqueo de conducción o la dispersión temporal, por eso el parámetro más sensible es la disminución focal de la

³⁸ WILBOURN Op. Cit

³⁹ WILBOURN Op. Cit

velocidad de conducción. En estas patologías es importante estudiar segmentos cortos para localizar la lesión por ejemplo estimulando el nervio cada centímetro y buscando cambios en la amplitud o la latencia (técnica de “**inching**”).

A continuación se ofrecen como ejemplo las tres neuropatías focales más frecuentes, los conceptos discutidos son aplicables al estudio de otras neuropatías focales:

Túnel carpiano: El atrapamiento del nervio mediano en la muñeca causa parestesias de predominio nocturno en los 3.5 dedos que inerva, puede haber Tinel y Phalen y en casos severos atrofia tenar.

La Asociación Americana de Medicina Electrodiagnóstica (AAEM) propone el siguiente protocolo para la evaluación del síndrome de túnel carpiano el cual está basado en una revisión amplia de la literatura⁴⁰. Los resultados de cada prueba están expresados en términos de sensibilidad y especificidad. Sensibilidad es la probabilidad de clasificar correctamente a un individuo enfermo, es decir, la probabilidad de que para un sujeto enfermo se obtenga en la prueba un resultado positivo. La sensibilidad es, por lo tanto, la capacidad del test para detectar la enfermedad. La especificidad es la probabilidad de clasificar correctamente a un individuo sano, es decir, la probabilidad de que para un sujeto sano se obtenga un resultado negativo. En otras palabras, se puede definir la especificidad como la capacidad para detectar a los sanos.

Las pruebas recomendadas son:

- Conducción sensitiva del nervio mediano entre muñeca e índice (sensibilidad 0.65, especificidad 0.98)
- Si la conducción sensitiva es normal realizar alguno de los siguientes estudios:
- Conducción mixta comparativa entre mediano y ulnar estimulando cada nervio en la palma y registrado en muñeca (sensibilidad 0.71, especificidad 0.97)
- Conducción comparativa del mediano y ulnar al IV dedo o del mediano y el radial al I dedo (sensibilidad para el IV dedo: 0.85, especificidad 0.97)
- Comparación de la velocidad de conducción sensitiva o mixta a través del carpo con la velocidad de conducción en antebrazo o en la porción distal de la palma
- Conducción motora del mediano estimulando en muñeca (sensibilidad 0.63, especificidad 0.98).

Conducciones complementarias:

- Conducción motora comparativa de mediano y ulnar al segundo lumbrical y segundo Interóseo palmar respectivamente.
- Conducción motora entre muñeca y palma
- Conducción a través del carpo en segmentos cortos (“**inching**”)
- Electromiografía incluyendo músculo del tenar y de raíces C5 a T1.

Lo característico del túnel carpiano es la prolongación de las latencias sensitivas inicialmente y luego de las latencias motoras indicando desmielinización. En casos

⁴⁰ AAEM: Practice Parameter for Electrodiagnostic Studies in carpal tunnel syndrome: Summary statement.

avanzados hay disminución de la amplitud distal a la lesión, indicando daño axonal. Según la severidad del compromiso se considera en leve (sólo compromiso de latencia sensitiva), moderado (latencia motora prolongada pero menor de 5ms) y severo (latencia motora mayor de 5 ms). En caso de daño axonal hay indicación de liberación quirúrgica.

Canal cubital: La neuropatía focal del ulnar en el codo es la segunda mononeuropatía más frecuente. La compresión puede ocurrir en el canal retroepicondilar, en la arcada humeroulnar (canal cubital) o al salir del flexor ulnar del carpo. Los síntomas son parestesias en IV y V dedos, debilidad de intrínsecos de la mano inervados por ulnar y presencia de signo de Tinel en codo.

La AAEM propone la siguiente guía de práctica clínica para la evaluación del síndrome de canal cubital la cual esta basada en una revisión amplia de la literatura⁴¹ :

Conducción motora al abductor del V dedo estimulando en muñeca, por encima y por debajo de codo. La distancia entre los puntos de estimulación en codo es de 10 cm y la flexión del codo de 70-90 grados. Las anomalías posibles en orden de importancia son:

- Velocidad de conducción motora a través del codo menor de 50 m/s (sensibilidad 0.75, especificidad 0.96)
- Disminución relativa de velocidad de conducción motora al comparar la conducción a través del codo con la conducción en antebrazo de más de 10 m/s.

Muscle Nerve. 2002; (26) pp 928-922.

⁴¹ AAEM: Practice Parameter for Electrodiagnostic Studies in Ulnar Neuropathy at the elbow: Summary statement. Muscle Nerve. 1999; (suppl 8) pp S171-S174.

- Disminución de amplitud del PAMC en codo de un 20% o más al compararlo con el de muñeca.
- Cambio en la morfología del PAMC de codo al compararlo con el de muñeca.

Si el estudio anterior es normal se puede realizar una de las siguientes pruebas:

- Registro de PACM en el 1er interóseo dorsal (evalúa fascículos diferentes)
- Estudio de segmentos cortos a través del codo (“**inching**”)
- Conducción motora a los músculos del antebrazo
- Electromiografía de aguja incluyendo el primer interóseo dorsal.

Todas las pruebas descritas buscan demostrar el compromiso focal a nivel del codo. Las lesiones son inicialmente desmielinizantes pero puede haber luego daño axonal en cuya presencia pueden convertirse en quirúrgicas.

Neuropatía focal de peroneo en cabeza de la fíbula: Es la mononeuropatía más frecuente de miembros inferiores. Ocurre asociada a trauma directo, compresión externa (por ejemplo una férula o un yeso), a posiciones que afectan la microcirculación del nervio (posición de cuclillas sostenida), tumores del nervio etc. Generalmente se afecta el peroneo común aunque puede haber compromiso selectivo de las ramas superficial o profunda. Clínicamente hay debilidad de dorsiflexores y evertores, con poco compromiso sensitivo.

Según los hallazgos electrodiagnósticos han sido descritos varios patrones de lesión⁴²:

- Tipo 1: Pérdida axonal
- Tipo 2: Bloqueo de conducción parcial o completo
- Tipo 3: Lesión mixta axonal y desmielinizante
- Tipo 4: Doble lesión: radiculopatía L5 y neuropatía del peroneo

Cuando hay compromiso sólo de mielina el pronóstico de recuperación es bueno. En el otro extremo cuando hay daño axonal severo y ausencia de PAMC el pronóstico de recuperación es reservado.

Polineuropatías: Al evaluar pacientes con sospecha de neuropatía periférica el estudio electrodiagnóstico tiene dos fines primordiales: afirmar o descartar el diagnóstico y establecer el tipo de daño y su distribución. Las etiologías de las polineuropatías son múltiples y aunque no es posible determinarlas exactamente son el estudio electrodiagnóstico, si es posible disminuir el número de etiologías posibles según el tipo de daño (axonal, desmielinizante, mixto) y su distribución (Predominio motor o sensitivo, simétrico o asimétrico, distal o proximal etc.).

La mayoría de pacientes con neuropatía tienen compromiso de fibras gruesas mielinizadas por lo que los estudios de neuroconducción motora y sensitiva, las respuestas tardías y la

⁴² WILBOURN Op. Cit.

E.M.G. pueden mostrar anormalidades^{43,44}. Si el compromiso es de fibras pequeñas los estudios de neuroconducción y E.M.G. son normales. En caso de sospecha de compromiso de fibras autónomas hay pruebas especiales para evaluar el sistema simpático (respuesta simpática de la piel) y parasimpático (variabilidad del intervalo R-R con maniobras de inclinación y de Valsalva).

Los cambios patofisiológicos de las polineuropatías son compatibles ya sea con daño axonal, con desmielinización o mixtos. El daño axonal es de predominio distal, posiblemente por falta de transporte axonal a los segmentos mas distales del nervio. La desmielinización puede ser generalizada o multifocal. En etapas iniciales puede haber disminuciones en las velocidades de conducción sin daño anatómico demostrable posiblemente por alteraciones metabólicas (por ejemplo hiperglicemia).

2.2.2 Inteligencia Artificial y Medicina

Antes de discutir los aspectos específicos de la aplicación de Inteligencia artificial en Medicina vale la pena revisar una definición general de Inteligencia Artificial (IA):

“La inteligencia artificial es el estudio de la ideas que permiten a los computadores hacer las cosas que hacen que la gente parezca inteligente....La idea central de la Inteligencia

⁴³ STEWART Op. Cit.

⁴⁴DONOFRIO PD, Albers JW: Polyneuropathy: Classification by nerve conduction studies and electromyography. Muscle Nerve. 1990; (13) pp 889-903

Artificial es hacer computadores mas útiles y entender los principios que hacen posible la inteligencia”⁴⁵

Szolovits⁴⁶ (1982) definió el campo de la Inteligencia Artificial en Medicina (IAM) como el resultado de la combinación de modelos complejos y técnicas computacionales con la mirada de médicos expertos, para construir herramientas que mejoren el cuidado de la salud.

Los investigadores que trabajan en IAM usan métodos de razonamiento en los programas que utilizan y están en un punto intermedio entre encontrar una equivalencia humano/computador o desarrollar las mejores técnicas para entender y capturar la experticia de los médicos.

En teoría no es indispensable depender del experto humano para desarrollar un sistema computacional aplicado a medicina. Un conocimiento profundo de aspectos fisiológicos y fisiopatológicos podrían ser suficientes para realizar diagnósticos y proponer terapias. Sin embargo en ningún área de la medicina se ha avanzado tanto en el conocimiento como para no depender del juicio del médico experto. Por lo anterior la mayoría de técnicas de IAM tratan de reproducir las técnicas usadas por los médicos.

Otras ventajas de basarse en el conocimiento de los expertos humanos para construir sistemas computacionales expertos son: Poder explicar las decisiones y recomendaciones

⁴⁵ WINSTON, P. W., *Artificial Intelligence*. Addison-Wesley, Reading, Mass., (1977).

⁴⁶ SZOLOVITS Op. Cit

de un programa a sus usuarios finales en términos que les sean familiares, posibilidad de comparar directamente los resultados dados por el programa con la conducta del experto.

Las principales metas de un sistema de IAM son:

- Desarrollar sistemas computacionales expertos para uso clínico, haciendo posible la diseminación a diferentes regiones geográficas o grupos de practicantes que no poseen dicha experticia.
- Formalizar la experiencia médica, permitiendo a los médicos entender mejor su propio conocimiento y dándoles una estructura sistemática para enseñar su experticia a los estudiantes de medicina.
- Probar y utilizar el sistema de Inteligencia artificial en situaciones del “mundo real”.

En los años 70, los investigadores de IAM trabajaron en aspectos como la representación del conocimiento y organización del conocimiento en métodos para resolver problemas y en el manejo de la incertidumbre. En la siguiente década se hizo énfasis en aplicaciones. En los 90, se buscó, ante todo, el desarrollo de sistemas de soporte para la toma de decisiones. A pesar de todo el desarrollo logrado, dichos sistemas aún no hacen parte de las herramientas rutinarias de los médicos.

Antes del desarrollo de técnicas de IAM se usaron otros procedimientos computacionales para analizar los datos obtenidos de las historias clínicas o de los ensayos clínicos. Dichas técnicas son los diagramas de flujo, la búsqueda en bases de datos y técnicas basadas en la teoría de decisión. Estas tres técnicas contribuyeron de algún modo al desarrollo de la IAM⁴⁷ y se resumen a continuación:

Diagramas de flujo. Es la herramienta de toma de decisiones más simple. Codifican las secuencias de acciones que el clínico tendría que realizar para llegar a un diagnóstico o tratamiento. Tienen en cuenta, por ejemplo, los síntomas, los signos, los resultados de laboratorio, los procedimientos realizados y la terapéutica aplicada en la mayoría de situaciones para proponer un flujograma de manejo. En cada paso hay una toma de decisión dependiendo de un hallazgo o resultado. Su mayor inconveniente radica en que al ser utilizadas en problemas complejos su tamaño debería ser enorme para considerar todas las posibilidades. Además, por su misma construcción, cada decisión parece ser independiente de las otras y no hay registro de los sitios donde se utilizó cada información. Cuando hay datos faltantes el diagrama de flujo se detiene pues no sabe que camino seguir.

Reconocimiento de patrones en bases de datos. Con la creación de grandes bases de datos a partir de las historias clínicas de los pacientes, se ha recogido información de casos clínicos “típicos” útiles para usos clínicos y de investigación. Las grandes bases de datos para diversas enfermedades ayudan a clarificar la verdadera incidencia de las mismas, los

⁴⁷ SZOLOVITS Op. cit.

factores demográficos y la respuesta terapéutica a medicaciones o procedimientos. Con estas herramientas se busca saber si un determinado caso a mano corresponde o no al conjunto de una determinada patología apoyándose en algunas herramientas estadísticas. Hay que realizar un trabajo intenso para identificar las generalizaciones importantes entre los miles de datos registrados. Las dificultades radican en los costos y el tiempo requerido para crear y mantener actualizadas las bases de datos. Además, los casos antiguos pueden contener información diferente de los casos nuevos por los avances en medicina. También, en casos de enfermedades “raras” o presentaciones atípicas de las mismas, estas herramientas son poco eficientes. Por lo anterior solo se justificaría crear bases de datos en los problemas clínicos más frecuentes y severos. Otro problema de las bases de datos médicas es que no permiten recuperar la experiencia de los médicos usando sólo herramientas estadísticas.

Teoría de la decisión. Es una teoría matemática de la toma de decisiones bajo incertidumbre. Asume que es posible cuantificar las probabilidades a priori y condicionales de los estados existentes y sus manifestaciones y que se puede de igual forma realizar una evaluación (utilidad) de todos los desenlaces posibles. Dados estos datos, la teoría de la decisión ofrece una teoría normativa de la toma óptima de decisiones en problemas médicos. La teoría trata de hacer explícitas las bases para la toma de decisiones.

La principal desventaja de esta teoría radica en la dificultad de obtener estimaciones razonables de las probabilidades y los desenlaces de un caso particular. Existe el riesgo de hacer simplificaciones artificiales. Esta teoría no ha podido ser usada con éxito en dominios médicos complejos o en enfermedades coexistentes.

Un enfoque importante en las técnicas de IA es la de entender las ideas que permiten a los computadores hacer las cosas que hacen parecer inteligente a los humanos. Aunque hay mucha controversia al respecto, los investigadores están de acuerdo en que la solución de problemas es un buen ejemplo de las tareas que deben realizar los programas de IA. Tradicionalmente la habilidad para resolver problemas se basa en dos componentes: Conocimiento y habilidad para razonar. Este último componente ha llevado al desarrollo de programas de razonamiento complejos aplicados a bases de datos. La mayoría de estos programas trabajan descomponiendo una meta en metas parciales y estableciendo nuevas metas basados en la diferencia entre el estado actual y el deseado. Varias técnicas y algoritmos se utilizan destacándose el uso del método de contradicción y las técnicas de inferencia.

Es importante además saber como esta representado el conocimiento, entre mas estructurado esté el mismo, mas fácil es razonar sobre él. Se trata de estructurar tanto los objetos como los procesos.

La mayoría de programas de AIM usan tanto herramientas de razonamiento poderosas como mecanismos de representación bastante estructurados. En algunos programas el énfasis esta hecho en el razonamiento mientras que en otros se hace énfasis en la estructura de los datos.

Los primeros programas desarrollados en el campo de IAM se concentraban en un aspecto particular de un problema médico diagnóstico o terapéutico para tratar de compensar las deficiencias de los métodos tradicionales de la toma de decisiones en medicina. A continuación se revisan brevemente algunos de ellos:

CASNET desarrollado por la Universidad de Rutgers es una herramienta para diagnóstico y manejo de glaucoma y otras enfermedades oculares relacionadas. Se basa en el concepto de causalidad, según el cual para cada fenómeno anormal siempre hay un factor etiológico que puede ser rastreado. Por el contrario si no se encuentra ninguna ruta hacia un proceso o agente patógeno, se puede descartar la presencia de enfermedad. Tiene como limitante la dificultad de establecer en muchos casos la relación de causalidad.

El sistema MYCIN desarrollado por la Universidad de Stanford para el diagnóstico de enfermedades infecciosas se basa en el principio de condensar cientos de páginas de un flujograma en unas pocas reglas concisas aplicadas a un sencillo algoritmo recursivo. Sólo las reglas que se necesitan para llegar a una conclusión en determinado momento del análisis son utilizadas. Por ejemplo si se está en la etapa de identificación de un germen, solo se usan las reglas que aporten a dicho proceso. Una ventaja de este sistema modular es que cada grupo de expertos pueden trabajar sobre algunas reglas en particular. Además no solo se trabaja sobre el conocimiento sino sobre las reglas usadas para la toma de decisiones.

Un programa desarrollado por el M.I.T. para recomendar el uso del medicamento Digital a pacientes con enfermedad cardiaca se basa en incluir además de los datos del paciente individual tomados de la historia clínica, las metas del tratamiento y el resultado de la aplicación del medicamento a otros pacientes. Se basa más en análisis procedimientos que en reglas.

El sistema INTERNIST desarrollado por la Universidad de Pittsburg para diagnóstico en Medicina Interna en general, usa un enfoque heurístico en la formulación de problemas para escoger entre todas las posibles enfermedades el conjunto de las que deben considerarse de acuerdo a la información que se tengan a mano (historia clínica, examen físico, exámenes etc.). La formulación del problema diagnóstico se hace en una forma muy estructurada a la cual se aplican métodos convencionales para hacer el diagnóstico diferencial.

También en el M.I.T. se desarrolló el Present Illnes Program (PIP) un programa que usa un formalismo con una representación causal de las enfermedades que permite considerar en gran detalle cada parte del proceso diagnóstico. Se usó inicialmente en enfermedades renales.

Todos los programas anteriores demostraron ser semejantes al experto médico en caso sencillo, pero su desempeño era muy inferior en casos complejos o con múltiples enfermedades.

La nueva generación de programas de IAM tiene resultados muy superiores y muchos de ellos han sido juzgados como indistinguibles del experto médico usando pruebas de validación objetivas. Se puede decir de ellos que han logrado capturar, en un determinado dominio médico, la esencia de lo que significa ser un experto en dicho campo. Como todos los programas de computación funcionan muy bien en un conjunto de aplicaciones anticipadas pero su desempeño cae notablemente si se apartan de su dominio. Esto ocurre especialmente al evaluar casos atípicos, casos poco frecuentes o con múltiples interacciones o enfermedades.

La principal falla de los programas de IAM radica en la dificultad de reconocer la presencia de un problema o patología determinados en algunos casos y en ofrecer un análisis mas detallado de los mecanismos envueltos. La mayoría del conocimiento representado en IAM está basado especialmente en la experiencia o en la creencia del médico y no tanto en los mecanismos que producen las enfermedades.

El gran reto de las técnicas de IAM es tratar de codificar el conocimiento médico experto. La dificultad radica en entender cómo un médico razona, cómo infiere lo que sabe y estructurar ese vasto conocimiento en un modelo computacional.

Cualquier herramienta de IAM es tan buena como su poder de razonamiento y su habilidad para representar el conocimiento.

El uso generalizado de técnicas de IAM sólo se dará cuando el practicante de la medicina considere que su trabajo es más exitoso usando dichas herramientas que sin ellas.

Sistemas de Soporte a la Toma de Decisiones Diagnósticas

Los sistemas de apoyo a la toma de decisiones diagnósticas en Medicina se basan en considerar dicho proceso como un proceso de diagnóstico diferencial⁴⁸. Este término se refiere a un tipo particular de tarea analítica donde el médico confronta un conjunto de alternativas diagnósticas. Su trabajo es determinar si la información a mano es suficiente o requiere pruebas adicionales para llegar a una aproximación diagnóstica.

Un creciente número de procedimientos, diagramas de flujo y algoritmos han sido diseñados para ayudar a realizar el diagnóstico diferencial. En algunos casos, los algoritmos procuran capturar la lógica envuelta en el proceso diagnóstico de una enfermedad particular. En otros, se basan en procedimientos genéricos que se modifican para una aplicación particular. La idea detrás de procedimientos es ofrecer al clínico un espectro amplio de posibles diagnósticos diferenciales y ayudarlo a la toma de decisiones.

Algunas de las técnicas usadas como apoyo a la toma de decisiones en medicina son:

⁴⁸ SZOLOVITS Op. cit.

Árboles de decisión. Una secuencia de decisiones se estructura en forma de un árbol. Cada nodo representa una pregunta y la respuesta determina que rama seguir para llegar a la siguiente. El resultado final se obtiene al llegar a una hoja del árbol.

Técnicas Bayesianas. Es un método estadístico que utiliza el teorema de Bayes para calcular el valor predictivo positivo o negativo de cada prueba o síntoma. Requiere conocer la incidencia y prevalencia de las diferentes enfermedades. La prevalencia (P) cuantifica la proporción de individuos de una población que padecen una enfermedad en un momento o periodo de tiempo determinado. Su cálculo se estima mediante la expresión:

$$P = \frac{\text{Nº de casos con la enfermedad en un momento dado}}{\text{Total de población en ese momento}}$$

La incidencia (I) se define como el número de casos nuevos de una enfermedad que se desarrollan en una población durante un período de tiempo determinado.

$$IA = \frac{\text{Nº de casos nuevos de una enfermedad durante el seguimiento}}{\text{Total de población en riesgo al inicio del seguimiento}}$$

Teoría del análisis de decisión. Se asocia cada pieza de información con el costo de obtenerla y se mide el beneficio que se deriva de poseer dicha información. Se continúa recogiendo información hasta llegar a un punto en el cual el costo de obtener mas información supera el beneficio de hacerlo.

Estas técnicas tienen importantes limitaciones tales como la inadecuada representación del conocimiento, falta de flexibilidad y la dependencia de grandes cantidades de información previa sobre probabilidades y condicionantes.

Los nuevos sistemas de soporte tienden a basarse en construcción de grandes y bien documentados modelos jerárquicos de las enfermedades y de sus síntomas asociados, con los cuales se puede estimar formalmente la probabilidad de una patología en particular. Otro método se centra en el uso de modelos sofisticados de procesos fisiopatológicos, donde se puede aislar un sistema (como el sistema nervioso periférico en el caso de las neuropatías) para lograr un mejor entendimiento.

Según Kokol, las características más deseables de los sistemas expertos son⁴⁹: tener conductas orientadas hacia metas, aprender de la experiencia, usar grandes cantidades de conocimiento, tolerar errores y responder en tiempo real.

Técnicas avanzadas de IA

Además de los sistemas expertos y de las técnicas de minería en bases de datos descritos arriba hay otros métodos de cómputo blando que son la lógica difusa (“fuzzy logic”), la programación evolutiva (algoritmos genéticos, programación genética, estrategias evolutivas etc.) y las redes neuronales artificiales. Estos sistemas pueden usarse solos o combinados (sistemas híbridos). En esta sección se revisan brevemente las dos primeras técnicas y más adelante se revisan en detalle las redes neuronales artificiales y su aplicación a la medicina en general y al electrodiagnóstico en particular.

Lógica difusa: La lógica difusa es un paradigma computacional que provee una herramienta matemática para tratar con la incertidumbre y con la imprecisión típica del razonamiento humano⁵⁰. Una característica primordial de la lógica difusa es la de poder expresar el conocimiento en una forma lingüística usando palabras comunes. Otra característica es que las reglas de cómputo son simples y conocidas.

Un sistema de lógica difusa es un sistema basado en reglas que en lugar de usar lógica Booleana usa lógica difusa para razonar acerca de los datos. El conjunto de datos se divide en subconjuntos de acuerdo a reglas previamente establecidas y cada dato tiene un determinado grado de pertenencia a los diferentes subconjuntos establecidos. En otras palabras los subconjuntos no son excluyentes.

Los principales componentes de un sistema difuso son: 1) un difusor que convierte los datos reales de las entradas en valores difusos 2) una maquina de inferencia que aplica mecanismos de razonamiento difuso para obtener una salida difusa 3) un de-difusor que convierte la salida difusa en valores reales y 4) una base de conocimiento que contiene reglas de lógica difusa y una base de datos que define las funciones de pertenencia usadas en lógica difusa⁵¹.

⁴⁹ KOKOL P et al. Some Ideas About Intelligent Medical System Design. Proceedings of the 12th IEEE Symposium on Computer-based Medical Systems CBMS'99 Stamford, CN, June 1999

⁵⁰ YAGER RR, Zadeh LA. Fuzzy Sets, Neural Networks, and Soft Computing. New York: Van Nostrand Reinhold, 1994.

⁵¹ PEÑA-REYES CA, Sipper M. A fuzzy-genetic approach to breast cancer diagnosis. Artif Intell Med

El diseño de un sistema de lógica difusa es una tarea de optimización donde se busca definir en la mejor forma posible los parámetros y reglas necesarias para obtener el comportamiento deseado.

Programación Evolutiva: La programación evolutiva es la aplicación de técnicas computacionales basadas en los principios de la evolución natural. Dichas técnicas son básicamente métodos de búsqueda o de optimización de la mejor solución posible a un problema⁵².

Los mecanismos naturales en los cuales se basa la programación evolutiva son reproducción, mutación y selección. La metáfora utilizada es que cada problema es un ambiente donde vive un determinado número de individuos, cada uno representando una posible solución al problema. Al igual que la evolución natural, los algoritmos evolucionarios se encargan de mantener individuos más fuertes (mejores soluciones) y descartar los débiles. Nuevo material genético puede ser introducido a través de mutaciones para mejorar las soluciones.

Hay varios tipos de algoritmos evolucionarios: algoritmos genéticos, programación genética, estrategias evolutivas y programación evolutiva. Aunque difieren en su

1999;17(2):131–5.

⁵² MICHALEWIS Z. Genetic Algorithms + Data Structures = Evolution programs, 3 ed. Berlin, Springer, 1996.

funcionamiento específico, el proceso general es el mismo⁵³: Una población inicial de individuos $P(0)$ es generada al azar. En cada paso evolucionario t conocido como una generación, los individuos que conforman la población actual $P(t)$ son decodificados y evaluados de acuerdo a un criterio de calidad llamado función de ajuste (“fitness”). Luego una subpoblación $P'(t)$ es seleccionada para reproducirse con los individuos más aptos según el criterio de ajuste. De acuerdo a lo anterior, solo los individuos más aptos se reproducen, mientras los otros tienden a desaparecer. Para ampliar el espacio de búsqueda se introducen cruces y mutaciones para producir una nueva población $P''(t)$ que se introduce en la siguiente generación $P(t+1)$. El algoritmo se detiene después de alcanzar un número predeterminado de generaciones o al alcanzar un nivel de ajuste aceptable.

La figura 13 muestra la estructura de un algoritmo evolutivo en pseudocódigo

Figura 13 Pseudocódigo de un algoritmo evolutivo⁵⁴

```
begin EC  
  t:=0  
  Initialize population  $P(t)$   
  while not done do  
    Evaluate  $P(t)$   
     $P'(t) := \text{Select}[P(t)]$   
     $P''(t) := \text{ApplyGeneticOperators}[P'(t)]$   
     $P(t+1) := \text{Introduce}[P''(t), P(t)]$   
    t:=t+1  
  end while  
end EC
```

⁵³ PEÑA-REYES CA, Sipper M. Evolutionary computation in medicine: an overview. Artificial Intelligence in Medicine 2000: 19, 1-23.

⁵⁴ PEÑA-REYES. Op. Cit.

Las principales ventajas de las técnicas de programación evolutiva son: 1) Se requiere escaso conocimiento inicial 2) Se extiende la búsqueda evitando caer en soluciones óptimas locales y 3) se intensifica la búsqueda de soluciones en las regiones más promisorias.

En el área de diagnóstico médico la programación evolutiva se ha utilizado para diagnóstico de tumores, detección de membrana hialina en prematuros, detección de cáncer de mama, clasificación de tumores cerebrales, evaluación de riesgo de enfermedades cardiovasculares, detección de prolapso de válvula mitral en ecografías, determinación de perfiles de riesgo en diabetes, detección de enfermedades a partir de estudios diagnósticos (electrocardiograma, electroencefalograma, mamografía etc.)⁵⁵. En el área de pronóstico médico se ha usado la programación evolutiva para predecir la posibilidad de recurrencia de cáncer de mama, compatibilidad de donantes, resultados de tratamientos en úlceras duodenales, desenlace de pacientes en unidad de cuidados intensivos, predicción de depresión después de manía, estimación de supervivencia en diferentes tipos de cáncer y clasificación de enfermedades neuromusculares basándose en hallazgos electromiográficos. En el campo de imágenes médicas y de procesamiento de señales electrofisiológicas se ha utilizado la programación evolutiva para interpretación de radiografías de tórax, angiografías retinianas y cardíacas, tomografías computarizadas, mamografías, resonancias magnéticas nucleares, magnetoencefalografías, ecografías, electroencefalografía, electrocardiografía etc. El otro grupo de aplicaciones médicas donde se ha usado la programación evolutiva es el de planeación y programación (“**scheduling**”)

⁵⁵ PEÑA-REYES CA, SIPPER M. Op Cit

el cual incluye aplicaciones como distribución de recurso humano hospitalario, planeación tridimensional de radioterapia, dosimetría de implantes de radioterapia para tratamiento de cáncer de próstata, programación de citas medicas, planeación de cirugía estereotáxica, etc.

2.2.3 Redes Neuronales Artificiales

Las redes neuronales artificiales son una técnica de inteligencia artificial que intenta imitar el comportamiento de las redes neuronales biológicas, las cuales constituyen el componente principal del cerebro humano y lo proveen de habilidades como la memoria, el razonamiento o la capacidad de relacionar eventos actuales con eventos pasados.

Una definición bastante amplia de lo que es una red neuronal artificial indica que: *“es una estructura matemática flexible capaz de identificar complejas relaciones no lineales entre un conjunto de entradas y de salidas; es por ello que han resultado muy útiles y eficientes, en problemas cuyo procesamiento es difícil de describir a partir de ecuaciones físicas. Las redes neuronales tienen grandes capacidades para la inferencia y generalización, por ello, una Red Neuronal que ha sido entrenada con un número representativo de ejemplos de un proceso determinado, puede extrapolar estados no presentados en el conjunto de datos de ejemplo”*⁵⁶.

⁵⁶ HECHT-NIELSEN R., 1991, Neurocomputing. Addison-Wesley

La Neurona Artificial: Al igual que en las redes neuronales biológicas, en las redes neuronales artificiales, el principal elemento constitutivo es la neurona.

Cada neurona en una red es una unidad de procesamiento independiente, su trabajo consiste en tomar las entradas proporcionadas por otras neuronas o por el medio y con base en ellas determinar la salida correspondiente, la cual a su vez es entregada a otras neuronas de la red o regresada al medio.

Cada neurona se enlaza a las neuronas de sus capas vecinas mediante una conexión que tiene asociado un peso, el peso de la entrada i en la neurona j se indica como w_{ij} .

La señal de entrada real a la neurona (s_j) se entiende como la sumatoria de cada valor de entrada (x_i) multiplicado por el peso correspondiente de la conexión, a esta regla se le conoce como *regla de propagación*:

$$s_j = \sum_i^n x_i w_{ij}$$

Cada neurona dentro de la red se encuentra en un estado determinado en cada momento, el conjunto de estados en los que puede encontrarse la neurona se denominan *estados de activación*, la continuidad o discreción de este conjunto al igual que su rango se determinan de acuerdo al problema que se trate.

El nuevo estado de activación de la neurona (en el tiempo $t+1$), se determina por medio de la *función de activación*, la cual considera el valor de su entrada y el estado de activación

actual de la neurona (en el tiempo t), la función de activación se define de la siguiente forma:

$$a_j(t+1) = F(a_j(t), s_j)$$

La función de activación debe ser capaz de relacionar todos los patrones posibles y no solamente aquellos para los cuales la red fue entrenada.

En la mayoría de los casos, la función de activación es la función identidad por lo que el estado anterior de la neurona es despreciable, de forma que el estado siguiente de la neurona es equivalente a su entrada real (s_j).

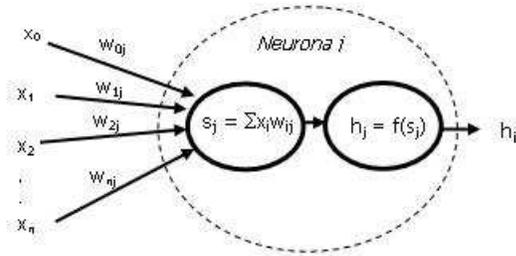
Para establecer la salida de la neurona, el nuevo estado de activación (s_j) se pasa por una función, denominada *función de transferencia*, la cual calcula el valor de salida de la neurona (h_j) a partir de sus entradas:

$$h_j(t+1) = f\left(\sum_i^n w_{ij} x_i(t)\right)$$

Se han establecido diferentes funciones de transferencia que se aplican dependiendo del tipo y objetivo de la red, se encuentran entre otras: función escalón, lineal y mixta, sigmoideal y gaussiana.

La figura 14 resume el funcionamiento de la neurona artificial en término de sus entradas y salidas:

Figura 14 Neurona Artificial

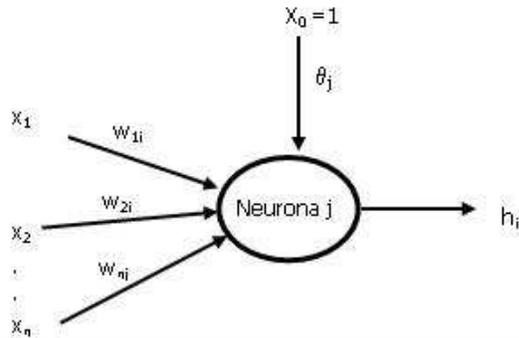


La función de activación generalmente no se centra en el origen del eje que representa la entrada, sino que tiene un cierto desplazamiento, el cual depende de las características de cada neurona, a este desplazamiento se le denomina sesgo de activación y se representa con θ . Incluyendo el sesgo, la función de activación general queda definida de la siguiente forma:

$$h_j(t+1) = f\left(\sum_i^n w_{ij} x_i(t) - \theta_j\right)$$

La representación del sesgo de activación generalmente se maneja como una neurona ficticia, con valor 1 en su entrada y con una conexión de peso θ_j para la neurona j , tal como se observa en la siguiente figura.

Figura 15 Neurona con Umbral de Activación



Aprendizaje: Las redes neuronales artificiales no pueden ser programadas para realizar labores específicas, ya que han sido diseñadas para aprender mediante los ejemplos que se le presenten. La red descubre las relaciones existentes en la información presentada como ejemplo, logrando establecer generalizaciones, que le permiten interpolar y extrapolar de una manera sensible ante elementos que no pertenecían al conjunto de entrenamiento⁵⁷.

Considerando que los ejemplos presentados son los que le entregarán el conocimiento a la red, es muy importante seleccionarlos de forma adecuada, de otra manera se gastaría mucho tiempo en el entrenamiento o en el peor de los casos la red aprendería de manera incorrecta.⁵⁸

⁵⁷ INSIGNARES. ., Víctor Manuel. Redes Neuronales. Principios y aplicaciones en el campo biológico. pg. 19

⁵⁸ SIGAMOS Dimitrios. Why neural networks? <URL: (http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol1/ds12/article.html)>

Las redes neuronales aprenden variando los pesos de las conexiones entre sus neuronas, de forma que cada vez se mejoren las respuestas obtenidas. Se entiende que el proceso de aprendizaje ha terminado cuando los pesos de las conexiones permanecen estables, es decir:

$$\frac{\partial w_{ij}}{\partial t} = 0$$

Considerando que no siempre se logra el estado ideal, donde el error es cero, normalmente se considera que el problema ha sido solucionado, si se logra una combinación de pesos en la red, para la cual, la *función de error* definida ha alcanzado un mínimo aceptable, el cual depende de la clase de aplicación que se trate.

Los criterios que sigue la red, para la variación de los pesos, determinan lo que se conoce como regla de aprendizaje de la red. Cada modelo de red neuronal artificial emplea una clase particular de *regla de aprendizaje*.

De acuerdo a la forma como se realiza el aprendizaje las redes pueden ser catalogadas de las siguientes formas:

- **Aprendizaje en Funcionamiento:** Una red con aprendizaje ON-LINE, puede aprender mientras está en funcionamiento, una red con aprendizaje OFF-LINE, tiene dos formas de funcionamiento, un modo de aprendizaje o entrenamiento y un modo de operación. Estas últimas dejan sus pesos fijos una vez que ha terminado la fase de aprendizaje, lo cual permite mayor estabilidad en el funcionamiento de la red.
- **Aprendizaje Supervisado o no Supervisado:** En el aprendizaje supervisado se supone la existencia de un agente externo que evalúa las salidas de la red a partir de un conjunto

de entradas establecidas y determina los ajustes necesarios en los pesos de las conexiones. Se consideran formas de aprendizaje supervisado: Aprendizaje por corrección de error, Aprendizaje por refuerzo y Aprendizaje estocástico.

El aprendizaje no supervisado, también llamado aprendizaje autosupervisado no tiene influencia externa para la evaluación de salidas y ajuste de los pesos, en cambio, la red misma busca correlaciones o categorías en los patrones de entrada. El tipo de respuestas que puede generar una red de este tipo está más relacionado con los patrones de sus entradas que con un patrón generado a partir de ella. Entre los algoritmos para aprendizaje no supervisado se encuentra: Aprendizaje hebbiano y Aprendizaje competitivo y cooperativo.

La red Neuronal Artificial: Las redes neuronales artificiales se forman de capas de neuronas interconectadas, cada capa es un conjunto de neuronas cuyas entradas provienen de la misma fuente, sea del exterior de la red o de otra capa de neuronas, y cuyas salidas se dirigen al mismo destino, puede ser el exterior de la red u otra capa de neuronas.

Una red neuronal artificial puede tener tres clases de capas diferentes: *capa de entrada*, que se encarga de recibir el patrón de entrada suministrado a la red, *capa de salida*, que entrega un resultado como respuesta al patrón de entrada y opcionalmente una o más *capas ocultas*, que se encargan de realizar el procesamiento interno.

Las redes neuronales artificiales pueden poseer una sola capa de neuronas, en cuyo caso se denominan redes *monocapa* o pueden poseer dos o más capas de neuronas, denominándose redes *multicapa*.

Cuando todas las salidas de un nivel en la red, sirven como entrada a cada una de las neuronas del siguiente nivel, se dice que la red se encuentra *totalmente conectada*, de otra forma se diría que la red está *parcialmente conectada*. De acuerdo a la forma como se realiza la conexión entre las neuronas de una red, estas se clasifican así:

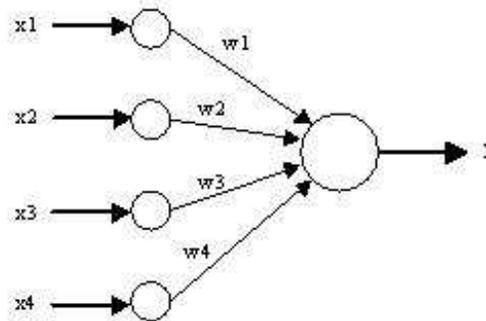
- **FeedForward o Propagación hacia delante:** Cuando ninguna salida de una neurona sirve como entrada a neuronas de la misma capa o de capas precedentes. Algunas redes de esta clase son: Perceptron, Adaline, Madaline, Backpropagation, Liner Adaptive Memory.
- **FeedBack o Propagación hacia atrás:** Cuando las salidas de una neurona pueden servir como entrada a cualquier neurona e incluso a ella misma. Las redes ART (**Adaptive Resonance Theory**) y BAM (**Bidirectional Associative Memory**) pertenecen a esta clase de redes.

El Perceptron: El Perceptrón es uno de los modelos más simples de redes neuronales artificiales, fue propuesto por Rosenblatt en 1959 para el reconocimiento de patrones.

Un perceptrón es una red multicapa de propagación hacia adelante, compuesto únicamente por dos capas: una capa de entrada, que puede contener cualquier número de neuronas y

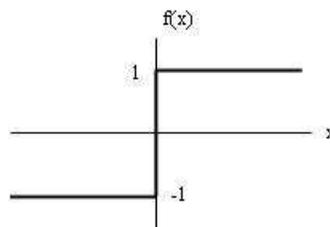
una capa de salida, con una sola neurona. La figura 16 presenta el esquema general de un perceptrón.

Figura 16 Esquema del Perceptrón



El perceptrón emplea la función escalón (figura 17) como función de activación, ya que esta limita la cantidad de salidas: +1 para un patrón de clase A y -1 (ó 0 en algunos casos) cuando el patrón es de clase B; permitiendo establecer patrones claros que generarían una de dos salidas.

Figura 17 Función Escalón



Aprendizaje del Perceptrón: Debido a que su estructura es simple, el aprendizaje del perceptron también se torna simple. La regla de aprendizaje empleada en este caso se conoce como *regla delta*, Regla Widrow – Hoff o LMS (Least Mean Squares), y se resume de la siguiente forma:

1. El resultado en la neurona de salida, se calcula a partir de las entradas, sus pesos y el sesgo, así:

$$y = f\left(\sum_i^n w_i x_i - \theta\right)$$

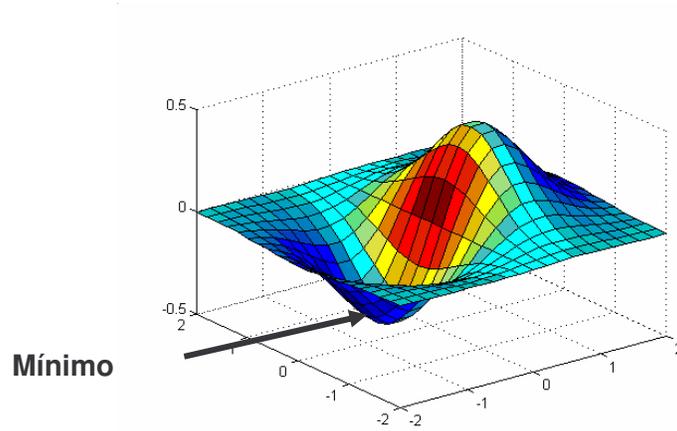
La salida resultante será 1 si el valor de la sumatoria es no negativo, y -1 (o 0) en otro caso.

2. Para cada patrón presentado a la red, se calcula el error cuadrado medio para la red, a partir de los errores resultantes en la salida de cada patrón presentado. Considerando que o_k es la salida obtenida para el patrón k , y t_k es la salida real para el patrón k , el error cuadrado promedio para la red se calcula así:

$$E = \sum_k E_k = \frac{1}{2} \sum_k (t_k - o_k)^2$$

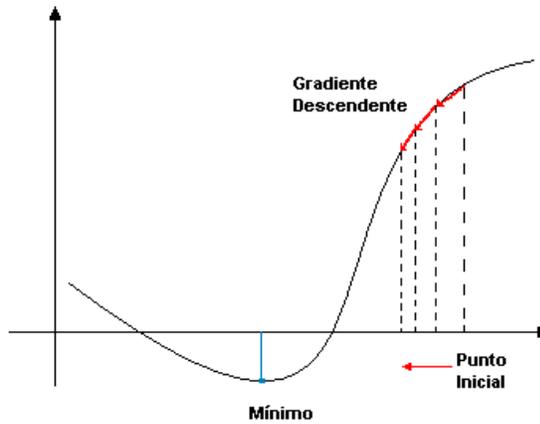
La idea del aprendizaje es minimizar esta función, lo cual se logra por variación de los pesos, ya que los otros valores (salidas reales y deseadas de la red) son constantes o están en función de estos. La forma de minimizar una función es alcanzar el mínimo global del espacio multidimensional creado por ella, como se presenta en la figura 18.

Figura 18 Gráfica del Espacio de la Función de Error



3. Para alcanzar el mínimo deseado, los pesos en la red son variados proporcionalmente al gradiente descendente negativo de la función de error, lo cual permite descender por el espacio de la función hasta alcanzar el mínimo de esta, como se presenta en la figura 19.

Figura 19 Gradiente Descendente



La actualización de los pesos se realiza en una proporción α del gradiente descendente, lo cual establece el tamaño del paso para el movimiento a través de la función de error; a este parámetro se le conoce como tasa de aprendizaje de la red; de esta manera, el cambio para la conexión con peso w_j cuando se analiza el patrón k , se establece con la ecuación:

$$\Delta_k w_j = (-\alpha) \frac{\partial E_k}{\partial w_j}$$

Aplicando técnicas de cálculo se puede reducir esta expresión a:

$$\Delta_k w_j = \alpha(t_k - o_k)x_j$$

4. Los pasos 1 a 3 se repiten hasta que se alcanza el mínimo global en el espacio generado por la función de error o se alcanza un error aceptable, establecido para cada caso de forma particular.

Esta regla permite que el perceptrón sea entrenado. Sin embargo, Minsky y Papert probaron que el perceptrón solamente puede llegar a soluciones en problemas que sean linealmente separables, lo cual significa que se pueda trazar una línea en el espacio de las entradas de forma que las entradas con salidas comunes queden en el mismo lado, como prueba de su tesis, Minsky y Papert presentaron el caso del o - exclusivo (XOR), cuyo espacio de solución no es linealmente separable⁵⁹ (las figuras 20 y 21 ejemplifican estos casos).

La forma de resolver el problema presentado para el perceptrón fue incluir capas ocultas en la red neuronal y variar las funciones de activación presentadas; a esta nueva red se le conoce como perceptrón multicapa, ejemplo de estas son las redes ADALINE, MADALINE y BackPropagation.

⁵⁹ MINSKY, Marvin and PAPERT, Seymour. Perceptron: An Introduction to Computational Geometry. 1969. MIT Press

Figura 20 Problema Linealmente Separable

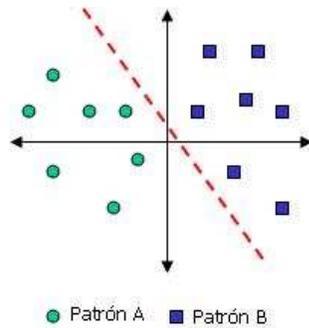
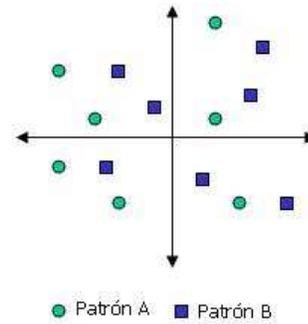


Figura 21 Problema No Linealmente Separable



BackPropagation: Esta clase de red fue propuesta por Rumelhart, Hunton y Williams en 1986, el objetivo primordial de su trabajo fue desarrollar una red que aprendiera las asociaciones entre las entradas y los estados de activación de la red, solucionando los problemas que presentaba el perceptrón.

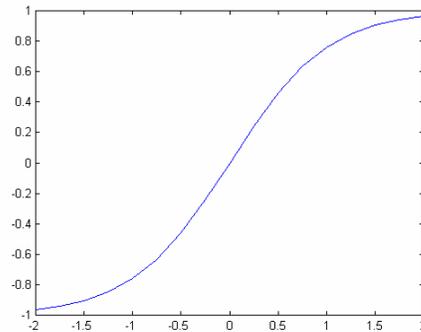
Una red BackPropagation es una clase particular de red de alimentación hacia adelante (FeedForward), que posee una capa de entrada, una de salida y al menos una capa oculta, donde todas las neuronas de una capa (excepto en la capa de salida) están totalmente conectadas a la capa siguiente, el aprendizaje es supervisado y de tipo OFF-LINE.

Generalmente, las redes BackPropagation emplean la función sigmoideal como función de activación, ya que esta función permite obtener una salida continua entre 0 y 1. La función sigmoideal está definida de la forma:

$$y = \frac{1}{1 + e^{-ax}}$$

y su gráfica se presenta en la figura 22:

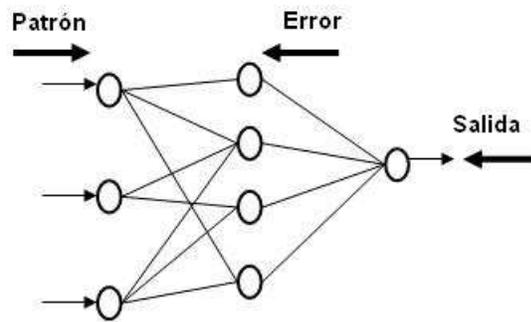
Figura 22 Gráfica de la Función Sigmoidal



Aprendizaje por Retropropagación: El proceso de aprendizaje en una red BackPropagation sigue el mismo principio que el del perceptrón: se propone alcanzar el mínimo global del espacio formado por la función de error, empleando el método del gradiente descendente para ajustar los pesos en las conexiones entre los nodos.

La regla delta original (usada en el Perceptrón) fue extendida para redes con capas intermedias, conexiones hacia adelante y neuronas con funciones de activación continua, originando lo que se conoce como algoritmo de retropropagación (backpropagation) o regla delta generalizada, cuyo esquema general se presenta en la figura 23:

Figura 23 Esquema de una red con retropropagación



El aprendizaje en una red Backpropagation se realiza en dos etapas:

1. Cada patrón de entrada es presentado a la red y propagado hacia delante, hasta obtener la salida correspondiente. Con la salida obtenida se calcula el error para cada neurona de la capa de salida, de la siguiente forma:

$$E = \frac{1}{2} \sum_k (t_k - o_k)^2$$

2. Con el error hallado, se ajustan los pesos de las conexiones que llegan a la capa de salida:

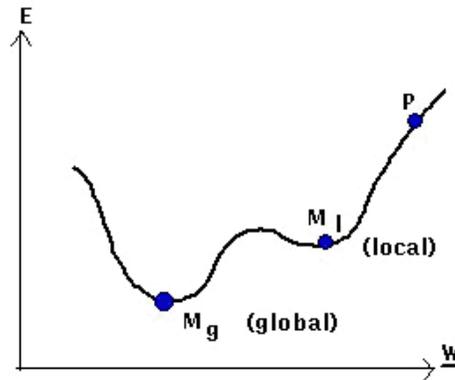
$$\Delta_k w_j = (-\alpha) \frac{\partial E_k}{\partial w_j}$$

Se calcula el error para las neuronas de las capas intermedias (capas ocultas) y se propaga el error hacia atrás, ajustando los pesos de cada neurona en los niveles anteriores, de forma proporcional a su participación en la salida obtenida.

Considerando que la regla delta busca hallar el mínimo de la función, sin importar si este es un mínimo local o global, puede ocurrir que la función se estabilice en un mínimo que no es

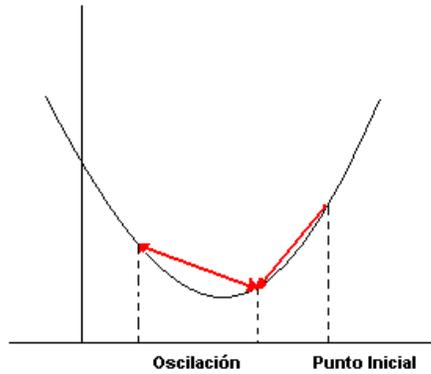
el global de la función, tal como se presenta en la figura 24, sin embargo la validez de esta respuesta será dictada por el error aceptado de acuerdo a los criterios establecidos para el problema.

Figura 24 Gráfica con mínimos locales



Adición del Momento: La cantidad de cálculos necesaria para el ajuste de los pesos en una red backpropagation es considerable. La tasa de aprendizaje (α) indica la amplitud del paso de cambio en los pesos manejados, una tasa de aprendizaje alta permitirá un ajuste mayor de los pesos, lo cual incrementa la velocidad para llegar al mínimo, pero también puede ocasionar oscilaciones alrededor de este sin lograr alcanzarlo, tal como se presenta en la figura 25:

Figura 25 Gráfica del Gradiente

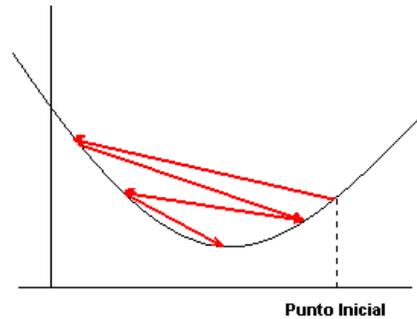


Rumelhart, Hinton y Williams, sugirieron filtrar las oscilaciones agregando un momento (β) en la expresión que calcula el ajuste de los pesos, quedando de la siguiente forma:

$$w_{ji}(t+1) = w_{ji}(t) + \alpha \delta_{pj} y_{pj} + \beta (w_{ji}(t) - w_{ji}(t-1)) = \Delta w_{ji}(t+1) = \alpha \delta_{pj} y_{pj} + \beta \Delta w_{ji}(t)$$

La adición del momento, permite disminuir el número de iteraciones necesarias para lograr la convergencia de la red al mínimo global, ya que el momento considera los valores de los pesos en cada instante, permitiendo ajustes acorde a la cercanía del mínimo buscado, como se muestra en la figura 26:

Figura 26 Gráfica del Gradiente con Momento



Las redes BackPropagation han sido definidas como "cajas negras", debido a que una vez se ha terminado la etapa de aprendizaje, no es posible conocer con exactitud la función generada, al respecto se han desarrollado investigaciones que emplean otras técnicas para inferir la función contenida en la red⁶⁰.

Aplicaciones: Las redes neuronales artificiales han sido aplicadas en gran variedad de proyectos. Aunque se pueden emplear sistemas de redes neuronales (artificiales) para interpretación, predicción, diagnóstico, planeación, monitoreo, depuración, reparación, instrucción y control, las aplicación mas exitosas de las redes neuronales (artificiales) son en la categorización y reconocimiento de patrones⁶¹.

⁶⁰ BRAUSE,R., FRIEDRICH, F., A Neuro-Fuzzy Approach as Medical Diagnostic Interface. European Symposium on Artificial Neural networks. ESANN 2000.

⁶¹ KLERFORS, Daniel. Artificial Neural Networks. Saint Louis University.

Estructura de la Red: El primer paso para la construcción de la red neuronal es determinar el tipo de red a emplear, considerando las características propias del problema; una vez determinado el tipo de red, se habrán definido varios aspectos de su estructura, sin embargo, hay otros aspectos que el diseñador debe establecer, tales como el número de capas que tendrá la red, la cantidad de neuronas que se dispondrá en cada capa, la forma como serán presentados los datos a la red y la forma como esta entregará el resultado; considerando que no existen resultados teóricos que indiquen la arquitectura óptima de la red para un problema dado, cobran especial importancia la intuición y la experiencia del diseñador.

Datos de Entrada: La forma como se presenten los datos a la red, depende de la clase de problema que se este trabajando y de la estructura interpretativa de los datos mismos; en aplicaciones médicas los datos que alimentan a la red pueden ser de diversos tipos: datos clínicos, registros del paciente, resultados de laboratorio, etc.

La selección de los datos que conformarán las entradas de la red, es una de las primeras decisiones que debe tomarse en el diseño; la forma más sencilla para establecer los datos de entrada, es considerar como entrada todos los datos disponibles, sin embargo es conveniente realizar una selección de los datos, siguiendo alguno de estos métodos:

- *Intervención de un experto:* Un experto en el campo de trabajo, puede ayudar a determinar, con su experiencia, cuales datos de entrada no son relevantes para el diagnóstico que se persigue.

- *Significancia estadística:* Se pueden emplear métodos de análisis de varianza para determinar los valores indicativos de la significancia de cada variable en el proceso de discriminación, sin embargo, puede ocurrir que las variables que se han descartado por su poca significancia, al ser combinadas con otras variables contribuyan para la obtención del resultado.
- *Reducción del Umbral:* El algoritmo de aprendizaje puede rechazar nodos cuyo peso cae por debajo de un umbral definido; esta aproximación, puede resultar inconveniente debido a que ciertos pesos caen después de un número definido de iteraciones para posteriormente elevar su peso.
- *Correlaciones:* Para disminuir el número de entradas, pueden eliminarse aquellos nodos que se encuentran correlacionados con otros o que resultan de la combinación de otros nodos.

Debido a que las redes neuronales trabajan con entradas numéricas para realizar sus cálculos, aquellos datos que se refieran a características o datos cualitativos, deben ser codificados en un formato numérico, para lo cual se emplea una de las siguientes técnicas:

- *Datos nominales:* Permiten representar categorías por medio de valores numéricos asociados, los cuales de ninguna manera representan ordenamiento. Por ejemplo, 0 - normal, 1- daño axonal, 2 - desmielinización.
- *Datos ordinales:* A diferencia de los nominales, esta categorización de alguna manera implica ordenamiento, sin embargo la distancia entre los valores asignados a las categorías no se puede asumir como un valor válido. Por ejemplo: Se pueden determinar las siguientes categorías para la severidad del dolor: 0 - sin dolor, 1 - dolor leve, 2 -

dolor moderado, 3 - dolor agudo, pero no es correcto afirmar que un valor ubicado entre 1 y 3 corresponde a un dolor moderado.

- *Intervalos:* En este nivel se ubican aquellos datos, para los que efectivamente la distancia entre los valores asignados es representativa. Por ejemplo: la edad de una persona.

La forma como se presentan los datos a la red, incide de manera directa en el desempeño de la misma, ya que si las entradas son complejas la red requerirá mayor procesamiento interno (mayor número de neuronas ocultas), mientras que si las entradas son de sencilla interpretación para la red, se puede disminuir de forma considerable el número de neuronas internas.

Existen varios métodos para codificar las entradas que se presentan a la red, de forma que se simplifique su procesamiento, estas técnicas pueden incrementar el número de neuronas en la capa de entrada, pero disminuyen la necesidad de procesamiento de la red:

- *Codificación 1-de-C:* En esta codificación, una variable con C categorías diferentes se representa a través de C entradas booleanas, donde cada entrada se activa para representar una categoría particular.
- *Codificación 1-de-(C-1):* Esta codificación es una variación del anterior, en la cual, variables ordinales o de intervalo se representan con C-1 entradas booleanas, de forma que la n-ésima entrada se pone en 1 para representar la n-ésima categoría, y la última categoría se presentará cuando todas las entradas estén en cero (o en -1), sin embargo, esta codificación requiere que se disponga de umbral de activación (bias).

- *Codificación de Termómetro*: Similar a la codificación 1-de-C conservando el orden de las categorías.

Datos de Salida: Gran parte de las aplicaciones de redes neuronales, especialmente para diagnóstico médico involucran la toma de decisiones, razón por la cual, los datos resultantes son variables nominales u ordinales. Estos datos de salida se pueden presentar de diversas formas:

- *Codificación 1-de-n*: donde n nodos de salida, representan n posibles categorías, de tal manera que si el nodo k tiene una salida positiva (1) mientras los otros nodos permanecen en 0, se establece la categoría k como salida. Al emplear este esquema de codificación, las salidas deben ser interpretadas como probabilidades, es decir que la salida con valor más alto, corresponde a la más probable respuesta ante la entrada presentada.
- *Codificación z-de-n*: Es una derivación de la codificación 1-de-n, en la cual, se asigna un número a cada salida posible, y se representan los números (en binario) por variables ficticias (dummy); aunque esta codificación permite disminuir el número de neuronas en la capa de salida, puede requerir mayor grado de procesamiento.

Capas Ocultas: Diversos estudios⁶² han demostrado que en una red BackPropagation, una capa oculta es suficiente para aproximar cualquier función no lineal, de esta forma, la

⁶² BISHOP, Christopher M. Neural Network for Pattern Recognition. Pag. 130

decisión en esta capa se reduce a determinar el número de unidades (neuronas) que debe contener.

A diferencia de las capas de entrada y salida, cuyo número de neuronas está definido por el problema mismo, el número de unidades en la capa oculta, reviste una complejidad mayor: cuando se incrementa el número de neuronas ocultas, se está incrementando la complejidad de la red y por tanto, el tiempo necesario para su entrenamiento así como la cantidad de casos requeridos para una buena generalización. Si el número de neuronas ocultas no es suficiente, la red no podrá alcanzar el grado de generalización esperado, pero, si el número de neuronas ocultas es excesivo, puede ocasionarse un sobreajuste (**overfitting**) de la red.

La cantidad de neuronas óptima para la capa oculta, depende de varios aspectos:

- El número de neuronas de entrada y salida.
- El número de casos de entrenamiento disponibles.
- La cantidad de ruido de los datos.
- La complejidad de la función de clasificación que debe ser aprendida por la red.
- La arquitectura.
- El tipo de función de activación de las unidades ocultas.
- El algoritmo de entrenamiento.

En la mayoría de los casos, no hay forma de determinar el mejor número de neuronas ocultas sin realizar entrenamiento de la red y hallar el error de generalización. Algunos

autores han propuesto reglas empíricas⁶³ para seleccionar la arquitectura, sin embargo, la mayoría de ellas, se basan en uno o dos de los aspectos mencionados anteriormente, omitiendo los restantes.

La recomendación general, es ensayar con diferentes arquitecturas para la red, estimar el error de generalización para cada una y seleccionar la red que alcance el menor error estimado.

Entrenamiento: En general, la capacidad de generalización de una red neuronal artificial está determinada por (1) la arquitectura de la red, (2) el número de patrones de entrenamiento y (3) la complejidad del problema⁶⁴, aunque los tres elementos están íntimamente relacionados, de tal forma que si un problema es complejo, se requerirá una red mas grande para representarlo, lo que obliga a tener mas patrones de entrenamiento.

Patrones de Entrenamiento: Se ha demostrado⁶⁵ que una red de n entradas y h neuronas ocultas con un total de pesos w , requiere un número de patrones del orden de w/ϵ (donde ϵ es el error de generalización). Para un error de 0.1, el número de patrones requeridos será $p=10w$.

⁶³ El concepto exacto esta dado por el término “Rules of Thumbs”

⁶⁴ HAYKIN S. (Neural Networks. A comprehensive Foundation) citado por MARTIN DEL BRIO y Sanz Molina en Redes Neuronales y Sistemas Difusos (2ª Ed)

⁶⁵ [BAUM 89] citado por MARTIN DEL BRIO y Sanz Molina en Redes Neuronales y Sistemas Difusos (2ª Ed)

En general, el número de patrones disponibles para el entrenamiento de la red suele ser reducido respecto al ideal, por lo que se emplean técnicas de reducción de la arquitectura tales como compartir pesos, podado de red o decaimiento de pesos que disminuyen el número de pesos efectivos en la red, de manera que el número de patrones de entrenamiento requeridos también disminuye.

Sin embargo debe asegurarse para los patrones empleados para el entrenamiento, que:

- La colección de datos es una representación bastante aproximada del espacio de problemas que se va a desarrollar
- No aparezcan inconsistencias en los datos, que la red no será capaz de resolver

Datos de entrada faltantes: Cuando se determina el conjunto de datos de entrada para el entrenamiento de la red, pueden obtenerse conjuntos incompletos de datos, se proponen varias formas de resolver este problema:

- 1) Incluir una entrada adicional que se active solamente cuando falte un determinado dato.
- 2) Establecer la entrada booleana en 0.5 para el dato faltante, como un valor promedio.
- 3) Establecer el valor de los datos faltantes a través de un algoritmo de "Maximización de Esperanza", el cual emplea modelos probabilísticos para calcular un valor apropiado, sin embargo, este modelo requiere de un alto nivel de cálculo, por lo cual se recomienda su empleo solamente si se puede realizar de forma automática.

Patrones faltantes: Cuando se trata de diagnóstico médico, es común que se encuentren gran cantidad de datos sobre una patología particular, mientras otros resultados tienen una

presencia bastante baja, lo que podría llevar a que la red neuronal realice una mala clasificación de estos casos; para resolver esta situación, se sugieren varias técnicas:

- 1) Emplear el criterio de “mínimo error cuadrado” durante el entrenamiento, lo cual es equivalente a replicar los patrones de las clases no representadas.
- 2) Manipular el error, agregando una salida p_i que representa el predominio de una categoría, para asegurar que la red construya una función discriminante para cada clase, de acuerdo a su influencia.
- 3) Si se conoce el costo de la clasificación equivocada de una clase, se puede utilizar este valor para el entrenamiento de la red, de manera que la red intente minimizar estas malas clasificaciones.

Proceso de Entrenamiento: Normalmente para el entrenamiento de una red neuronal artificial, deben ensayarse diferentes topologías con diferentes parámetros determinando para cada una de ellas el error de aprendizaje, hasta alcanzar los niveles establecidos dependiendo de la magnitud del problema. Si no se alcanza el nivel de error esperado, deberán revisarse nuevamente los patrones elegidos, los rangos de datos establecidos y el modelo de red neuronal seleccionado.

Los parámetros que se modifican durante el entrenamiento son:

- ***Pesos iniciales:*** existen varios métodos para iniciar los pesos en la red, aunque el más común es emplear pesos aleatorios fijados dentro de un rango aceptable.
- ***Tasa de aprendizaje:*** Debe recordarse que la tasa de aprendizaje establece el tamaño de los pasos en el movimiento hacia el mínimo global de la función de aprendizaje, una

tasa demasiado pequeña disminuye la velocidad de convergencia y puede ocasionar que la función se quede atrapada en un mínimo local; pero una tasa de aprendizaje alta puede ocasionar inestabilidad de la red, por lo cual es aconsejable seleccionar una tasa de aprendizaje que pueda variar según la necesidad del proceso.

- **Número de neuronas ocultas:** Ya que no se puede determinar con seguridad el número de neuronas ocultas requeridas para una red, es necesario ajustarlas durante el entrenamiento.
- **Criterio de Parada:** El momento en que el entrenamiento puede detenerse está guiado por el nivel de error establecido, se entiende que cuando el proceso de entrenamiento alcance este error, la red funciona de la manera esperada, sin importar si ha alcanzado o no el mínimo global de la función.

Generalización: Uno de los aspectos fundamentales del uso de Redes Neuronales Artificiales es su capacidad de generalización, es decir, la capacidad de responder de forma adecuada a casos que no fueron presentados durante el entrenamiento.

Sin embargo, las Redes Neuronales Artificiales, tal como otros métodos de estimación no lineales, pueden sufrir de *sobreajuste* (**overfitting**) o *subajuste* (**underfitting**). Una red que no es lo suficientemente compleja, puede fallar para detectar los casos presentados (subajuste), pero una red muy compleja puede incluir el ruido de los datos de aprendizaje como parte de los mismos (sobreajuste), ocasionando predicciones incorrectas.

La mejor forma de evitar el sobreajuste es emplear gran cantidad de datos de entrenamiento, si se tiene al menos 30 veces mas casos de entrenamiento que pesos en la red, es poco probable tener sobreajuste, sin embargo, si se puede asegurar que los datos están libres de ruido, bastaría tener 5 veces mas.

El sobreajuste y el subajuste también pueden ocurrir cuando:

- El número de variables de entrada es grande respecto al número de casos de entrenamiento.
- Las variables de entrada están altamente correlacionadas unas con otras.

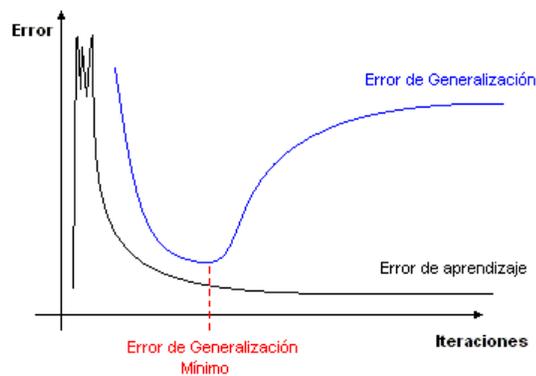
Validación Cruzada: Durante el entrenamiento se considera el *error de aprendizaje*, calculado como el error cuadrático medio de los resultados entregados por la red para el conjunto de patrones presentados. Pero, existe un error de generalización, que se puede medir empleando un conjunto representativo de patrones diferentes a los usados para el entrenamiento. Considerando estos dos errores, se puede entrenar la red usando un conjunto de patrones de aprendizaje hasta alcanzar el error de aprendizaje esperado y luego comprobar la eficiencia real de la misma mediante un conjunto de prueba que verifica el *error de generalización*.

De manera ideal, la red debería entrenarse hasta alcanzar un error de generalización mínimo (que idealmente es igual al error de aprendizaje), ya que si se considera solamente el error de aprendizaje y se intenta lograr un error muy pequeño, puede llegarse a que la red

memorice los patrones presentados pero no sea capaz de generalizar a patrones nuevos (sobreaprendizaje).

La figura 27 muestra la relación entre el error de aprendizaje y el error de generalización, como se observa existe un punto óptimo donde debería detenerse el entrenamiento, cuando el error de generalización se hace mínimo.

Figura 27 Error de Aprendizaje y Generalización



El procedimiento de entrenar y validar a la vez para alcanzar el error de generalización mínimo se denomina *validación cruzada* (**cross validation**), el cual es ampliamente usado para el desarrollo de redes neuronales artificiales supervisadas como los perceptrón multi-capas.

Parada Temprana (Early Stopping): Para asegurar la generalización de la red se pueden emplear diversos métodos, uno de los más usados es el de *parada temprana*, debido a su

facilidad de implementación y a que ha sido reportado como superior ante otros métodos de regularización⁶⁶.

La parada temprana agrega un conjunto de patrones para validación y permite que el proceso de entrenamiento realizado durante la validación cruzada finalice en el momento en que el error del conjunto de validación empieza a incrementarse.

El proceso de *parada temprana* incluye⁶⁷:

- Dividir los datos disponibles en dos conjuntos (entrenamiento y validación).
- Usar una gran cantidad de neuronas ocultas.
- Usar valores iniciales muy pequeños
- Usar una tasa de aprendizaje lenta.
- Calcular el error de validación periódicamente durante el entrenamiento.
- Detener el entrenamiento cuando el error de validación empiece a incrementar.

Este método es rápido y puede aplicarse a redes en las que el número de pesos supera la cantidad de patrones disponibles, no obstante, presenta algunos inconvenientes como determinar con exactitud cuantos y cuales patrones incluir en cada conjunto, y cómo verificar que el error de validación efectivamente ha empezado a incrementarse⁶⁸.

⁶⁶ PREHCELT, Lutz. Early Stopping – but when?

⁶⁷ What is early stopping? (<http://www.faqs.org/ai-faq/neural-nets/part3/section-5.html>)

⁶⁸ What is early stopping? Op. Cit

El conjunto de datos que se empleen para validación debe incluir casos no presentados a la red durante el entrenamiento e todas las posibles salidas de la red; usualmente, del conjunto de patrones disponibles se emplea el 80% para entrenamiento y se reserva el 20% para validación⁶⁹.

Evaluación de la Red: La valoración de una red neuronal artificial como sistema para el soporte a la toma de decisiones, incluye⁷⁰:

- La verificación de la integridad de los datos de entrenamiento.
- El desempeño técnico del algoritmo de aprendizaje.
- El desempeño del algoritmo de aprendizaje en conjunto con los datos de entrenamiento.

Verificación de los datos: La verificación de los datos para entrenamiento debe realizarse desde dos puntos de vista: su origen y su precisión.

Los datos para entrenamiento de sistemas para soporte de decisión en medicina pueden tener diversas fuentes: revisión de historias clínicas, bases de datos o estudios prospectivos; cada una de las cuales presenta ventajas y desventajas. Es recomendable emplear estudios prospectivos, sin embargo, la elaboración de estos resulta bastante dispendiosa y puede llegar a tomar mucho tiempo, por ello, es común que se empleen historias clínicas disponibles, aunque no siempre incluyen los datos esperados.

⁶⁹ HAYKIN S. (Neural Networks. A comprehensive Foundation) citado por MARTIN DEL BRIO y Sanz Molina en Redes Neuronales y Sistemas Difusos (2ª Ed)

Respecto a la pertinencia de los datos, debe examinarse si se posee una cantidad de datos suficiente por cada categoría disponible en la red, si estos son consistentes y están libres de error (al menos, tanto como es posible).

Validación del algoritmo de aprendizaje: Las herramientas adquiridas, en general, han sido ampliamente probadas, por lo cual no es necesario realizar nuevas pruebas de validación sobre los algoritmos; en contraposición, si el algoritmo es desarrollado debería validarse de manera exhaustiva.

Al evaluar algoritmos de aprendizaje deben considerarse los aspectos teóricos involucrados en su desarrollo, algunos aspectos prácticos como el tiempo de ejecución o la capacidad de procesamiento y la comparación con otros algoritmos conocidos.

Evaluación del desempeño: Una vez que se han validado de manera independiente los datos y el algoritmo de entrenamiento, deben integrarse para verificar su funcionalidad, para lo cual se empleará un conjunto de datos diferente a los usados hasta el momento.

El método de evaluación primario, es el desempeño de la red sobre el conjunto de prueba; se aconseja emplear medidas como la sensibilidad y la especificidad de la precisión del conjunto de prueba para cada categoría y para el total de los datos presentados. En general, para aplicaciones médicas se considera un éxito el equilibrio en las dos partes, aunque,

⁷⁰ HUDSON, Donna. COHEN, Maurice. Neural Networks and Artificial Intelligence for Biomedical Engineering.

existen casos en los cuales es preferible obtener falsos negativos que falsos positivos o viceversa.

Otra forma de establecer el desempeño es comparar los resultados con otro método, por ejemplo métodos estadísticos tradicionales. Los resultados también pueden ser comparados con datos históricos encontrados en literatura especializada o con datos proporcionados por expertos.

2.2.4 Aplicaciones de redes neuronales a Medicina

El uso de redes neuronales artificiales (RNAs) en medicina se concentra especialmente en tareas de clasificación⁷¹. Las redes neuronales artificiales se han usado principalmente en cinco campos médicos⁷²:

- Modelamiento de funcionamiento de órganos
- Procesamiento de señales eléctricas biológicas
- Diagnostico medico usando hallazgos clínicos e instrumentales
- Pronostico Medico basado en análisis de parámetros
- Control de equipos médicos basándose en parámetros biológicos

A continuación se listan algunas aplicaciones de las redes neuronales artificiales en medicina organizadas según los campos médicos descritos:

⁷¹ JAKUS V: The concept and applications of artificial neural networks in Medicine. Bratisl Lek Listy 1999: 110(11), 625-637.

Modelamiento de funcionamiento de órganos

Una red neuronal artificial ha sido usada para modelar la regulación de la frecuencia cardiaca⁷³.

Guigo y colaboradores, estudiaron y modelaron el proceso de aprendizaje de la corteza cerebral prefrontal⁷⁴.

Una red neuronal artificial fue diseñada para modelar el sistema auditivo⁷⁵.

Burstein realizo modelamiento del proceso de embriogénesis usando redes neuronales artificiales⁷⁶.

Para clasificar proteínas y determinar secuencias de aminoácidos han sido diseñadas varias redes neuronales artificiales⁷⁷.

⁷² PAPIK K et al: Application of neural networks in medicine-a review. Med Sci Monit 1988; 4(3), 538-546.

⁷³ HOYER D, SCHMIDT K, ZWIENER U: Principles and experiences for modeling chaotic attractors of heart rate fluctuations with artificial neural networks. Biomed Tech Berl, 1995; 40(7-8): 190-4

⁷⁴ GUIGON E et al. Neural correlates of learning in the prefrontal cortex of the monkey: a predictive model. Cereb Cortex, 1995; 5(2): 135-47

⁷⁵ CHITTAJALLU SK, WONG D: Connectionist networks in auditory system modeling. Comput Biol Med, 1994; 24(6): 431-9

⁷⁶ BURSTEIN Z: A network model of developmental gene hierarchy. J Theor Biol 1995; 174(1): 1-11

⁷⁷ WU C, WHITSON G, MCLARTY J et al: Protein classification artificial neural system. Protein Sci, 1992; 1(5): 667-77

Procesamiento de señales eléctricas biológicas

Varias redes artificiales han sido desarrolladas para interpretación de electrocardiogramas⁷⁸.

La evaluación de monitoria con electrocardiograma de 24 horas (Holter) ha sido automatizada usando redes neuronales artificiales con una sensibilidad del 99.9%⁷⁹.

Para interpretación de señales electroencefalográficas (EEG) también se ha desarrollado varias redes neuronales artificiales. Una red backpropagation se diseño para detectar automáticamente complejos punta-onda de alto voltaje en ratas⁸⁰. Gabor y colaboradores diseñaron una red neuronal que detecta alteraciones epileptiformes en pacientes con convulsiones⁸¹. Varias redes neuronales han sido utilizadas para detectar anomalías durante estudios de EEG durante el sueño⁸².

Diagnostico medico

En cardiología, se ha desarrollado una red neuronal artificial para diagnostico de infarto agudo de miocardio (IAM) basándose en niveles enzimáticos⁸³. Otra red diseñada analiza niveles de enzimas y hallazgos electrocardiográficos para diagnostico de infarto de

⁷⁸ PAPIK K et al: Op Cit.

⁷⁹ MACFARLANE PW: Recent developments in computer analysis of ECGs. Clin Physiol 1992; 12(3): 313-7

⁸⁰ JAND. G, SIEGEL RM, HORV.th Z et al: Pattern recognition of the electroencephalogram by artificial neural networks. Electroencephalogr Clin Neurophysiol, 1993; 86(2): 100-9

⁸¹ GABOR AJ, SEYAL M. Automated interictal EEG spike detection using artificial neural networks. Electroencephalogr Clin Neurophysiol,1992; 83(5): 271-80

⁸² PFURTSCHELLER G, FLOTZINGER D, MATUSCHIK K. Sleep classification in infants based on artificial neural networks. Biomed Tech Berlin, 1992; 37(6): 122-30

⁸³ BAXT WG. Use of an artificial neural network for the diagnosis of myocardial infarction. Ann Intern Med. 1991; 115(11): 843-8

miocardio⁸⁴. También se han usado redes neuronales para diagnóstico de falla cardíaca⁸⁵, detección de arritmias cardíacas, diagnóstico de hipertrofia ventricular, clasificación de anomalías del segmento S-T y análisis de complejos QRS de morfología compleja⁸⁶.

En unidades de cuidado intensivo y durante anestesia, se han utilizado redes neuronales artificiales para detectar signos de alarma y alertar tempranamente al médico⁸⁷.

Basándose en hallazgos radiológicos, una red neuronal artificial fue desarrollada para clasificar nódulos pulmonares solitarios⁸⁸. Otra red neuronal fue más efectiva que su contraparte humana para diagnosticar embolismo pulmonar en 1064 pacientes⁸⁹.

Varias redes neuronales han sido descritas para uso en oncología. Una red desarrollada por Fogel reconoce adecuadamente cáncer de mama usando información de mamográfica, citológica y epidemiológica⁹⁰. Baker diseñó una red que diagnostica cáncer de mama basándose en el sistema de codificación radiológico BI-RADS⁹¹.

⁸⁴ BAXT WG. Analysis of the clinical variables driving decision in an artificial neural network trained to identify the presence of myocardial infarction [see comments]. *Ann Emerg Med*, 1992; 21(12): 1439-44

⁸⁵ ORTIZ J, SABBATINI RM, GHEFTER CG, Silva CE: Use of artificial neural networks in survival evaluation in heart failure. *Arq Bras Cardiol*, 1995; 64(1): 87-90

⁸⁶ ABREU-LIMA C., de Sa J.P.: Automatic classifiers for the interpretation of electrocardiograms. *Rev. Port. Cardiol.*, 17, 1998, è. 5, s. 415-428.

⁸⁷ MYLREA KC, ORR JA, WESTENSKOW DR. Integration of monitoring intelligent alarm anesthesia: neural networks - can they help? *J Clin Monit*, 1993; 9(1): 31-7

⁸⁸ GURNEY JW, SWENSEN SJ. Solitary pulmonary nodules: determining the likelihood of malignancy with neural network analysis. *Radiology*, 1995; 196(3): 823-9

⁸⁹ TOURASSI GD et al. Artificial neural network for diagnosis of acute pulmonary embolism: effect of case and observer selection. *Radiology*, 1995; 194(3): 889-93

⁹⁰ FOGEL DB, Wasson EC 3rd, Boughton EM: Evolving neural networks for detecting breast cancer. *Cancer Lett*, 1995; 96(1): 49-53

En neurología las redes neuronales artificiales han sido utilizadas para diferenciar entre enfermedad de Alzheimer y demencia vascular basándose en información de imágenes de escanografías de emisión de positrones (SPECTs) cerebrales⁹².

Maeda y colaboradores usaron redes neuronales para interpretación de videoqueratografías en pacientes con patología de cornea⁹³. Otra red neuronal llamada OPTONET ha sido desarrollada para interpretación automática de campimetría visual⁹⁴. Otra aplicación ha sido desarrollada para clasificar escotomas usando información de campimetrías⁹⁵.

El campo de la radiología es uno de los que mas ha sido explotado para el uso de redes neuronales artificiales. Entre las aplicaciones radiológicas se incluyen⁹⁶: detección de lesiones frías en SPECTs cerebrales, detección de lesiones coronarias en SPECTs cardiacos, detección de microcalcificaciones en mamografía⁹⁷ y clasificación de ecografías hepáticas y de próstata.

⁹¹ BAKER JA et al. Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon. *Radiology*, 1995; 196(3): 817-22

⁹² DEFIGUEIREDO RJ et al. Neural-network-based classification of cognitively normal, demented, Alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain. *Proc Natl Acad Sci USA*, 1995; 92(12): 5530-4

⁹³ MAEDA N, KLYCE SD, SMOLEK MK: Neural network classification of corneal topography. Preliminary demonstration. *Invest Ophthalmol Vis Sci*, 1995; 36(7): 1327-35

⁹⁴ ACCORNERO N, CAPOZZA M: OPTONET: neural network for visual field diagnosis. *Med Biol Eng Comput*, 1995; 33(2): 223-6

⁹⁵ JAKUS V: Op. Cit

⁹⁶ PAPIK K et al: Op Cit.

⁹⁷ PAPAPOULOS A, FOTIADIS DI, LIKAS A: A automatic microcalcification detection system based on a hybrid neural network classifier. *Artificial Intelligence in Medicine*, 2002; 25(2), 149-167

En Patología, las redes neuronales artificiales han sido usadas para estudio automatizado de microscopia en cáncer de mama⁹⁸, para diferenciar entre carcinoma tubular y adenosis esclerosante⁹⁹, para gradación de astrocitomas basándose en inmunohistología¹⁰⁰, determinación de extensión de cáncer de próstata¹⁰¹, clasificación automática de citologías vaginales (PAPNET)¹⁰², clasificación de biopsias gástricas¹⁰³ etc.

Usando un mapa cromosómico como base, se diseñó una red neuronal artificial que clasifica alteraciones genéticas¹⁰⁴.

Varias redes artificiales neuronales han sido creadas para análisis automático de pruebas bioquímicas de laboratorio^{105, 106}.

Forina y colaboradores desarrollaron una red neuronal para clasificación de nuevos medicamentos cardiotónicos¹⁰⁷.

⁹⁸ DAWSON AE, AUSTIN RE Jr, WEINBERG DS: Nuclear grading of breast carcinoma by image analysis. Classification by multivariate and neural network analysis. *Am J Clin Pathol*, 1991; 95(4Suppl1): S29-37

⁹⁹ O'LEARY TJ, MIKEL UV, BECKER RL: Computer-assisted image interpretation: use of a neural network to differentiate tubular carcinoma from sclerosing adenosis. *Mod Pathol*, 1992; 5(4): 402-5

¹⁰⁰ KOLLES H et al. Automated grading of astrocytomas based on histomorphometric analysis. Classification results of neuronal networks and discriminant analysis. *Anal Cell Pathol*, 1995; 8(2): 101-16

¹⁰¹ STOTZKA R, et al. A hybrid neural and statistical classifier system for histopathologic grading of prostatic lesions. *Anal Quant Cytol Histol*, 1995; 17(3): 204-18

¹⁰² BOON ME et al. Neural network processing of cervical smears can lead to a decrease in diagnostic variability and an increase in screening efficiency: a study of 63 false-negative smears. *Mod Pathol*, 1994; 7(9): 957-61

¹⁰³ MOLNAR B, et al. Application of multivariate, fuzzy set and neural network analysis in quantitative cytological examinations. *Anal Cell Pathol*, 1993; 5: 161-164

¹⁰⁴ ERRINGTON PA, GRAHAM J: Application of artificial neural networks to chromosome classification. *Cytometry*, 1993; 14(6): 627-39

¹⁰⁵ BURTIS CA.; Technological trends in clinical laboratory science. *Clin Biochem*, 1995; 28(3): 213-9

¹⁰⁶ EKLUND P, FORSSTROM JJ: Computational intelligence for laboratory information systems. *Scand J Clin Lab Invest Suppl*, 1995; 222: 21-30

Binder describió una red neuronal para diagnóstico de lesiones pigmentadas de piel a partir de sus imágenes¹⁰⁸.

Choi y colaboradores reportaron la aplicación de una red neuronal artificial para diagnóstico de carcinoma de vejiga a partir de imágenes¹⁰⁹.

Bounds y colaboradores describieron una red neuronal para diferenciar entre dolor lumbar simple, dolor radicular, dolor espinal y simuladores. La red fue superior a los médicos especialistas para detectar casos complejos de dolor lumbar mientras que los humanos detectaron mejor los casos de dolor lumbar simple¹¹⁰.

Juhola y colaboradores usaron redes neuronales artificiales para clasificar en forma exitosa diversas clases de vértigo basándose en parámetros clínicos relevantes¹¹¹.

Otras patologías en las cuales se han usado exitosamente redes neuronales artificiales para diagnóstico médico, son^{112,113}: arteritis temporal, cirrosis hepática, trastornos hepato-

¹⁰⁷ FORINA M. et al. Zupan.s descriptors in QSAR applied to the study of a new class of cardiotonic agents. *Farmaco*, 52, 1997, è. 6.7, s. 411.419.

¹⁰⁸ BINDER M. et al. Epiluminiscence microscopy-based classification of pigmented skin lesions using computerized image analysis and a n artificial neural network. *Melanoma. Res.*, 8, 1998, è. 3, s. 261.266.

¹⁰⁹ JAKUS V: Op. Cit

¹¹⁰ BOUNDS D., LLOYD P.J.: A multi-layer perceptron network for the diagnosis of low back pain. *Proc. IEEE Int. Conf. on Neural Networks*, (1988); Vol.II, 481-489

¹¹¹ JUHOLA M et al: Neural network recognition of otoneurological vertigo diseases. In Horn W.... (eds): *Artificial Intelligence in Medicine*, Springer, Berlin, 1999, pp 217-226.

¹¹² Jakusv: Op cit.

biliares, abscesos intra-abdominales, nefropatías, clasificación de enfermedades psiquiátricas, diagnóstico de depresión, síndrome de estrés postraumático, dermatomiositis, ovario poliquístico, retinopatías diabéticas, autismo, tratamiento de migrañas, diagnóstico de apendicitis aguda, toxoplasmosis, etc.

Pronóstico médico

Se han usado redes neuronales para determinar el pronóstico de pacientes después de hepatectomías por tumores de hígado, el porcentaje de acierto fue de 100% en uno de los estudios¹¹⁴. Otros tumores en los cuales se ha establecido pronóstico usando redes neuronales son cáncer de mama, carcinomas pulmonares, cáncer de próstata y cáncer de ovario¹¹⁵.

Una red neuronal detectó correctamente la posibilidad de recurrencia de tumores de mama en 960 de 1008 casos, usando como información de entrada el tamaño del tumor, el número de nodos linfáticos palpables y el estado hormonal de la paciente¹¹⁶.

En obstetricia las redes neuronales artificiales han sido usadas para predecir la teratogenicidad de medicamentos administrados perinatalmente¹¹⁷. Otra red ha sido usada para predecir el peso al nacimiento basándose en varios parámetros perinatales¹¹⁸.

¹¹³ DYBOWSKI R, GANT V (Eds): Clinical applications of neural networks. Cambridge University Press, United Kingdom, 2001.

¹¹⁴ HAMAMOTO I et al. Prediction of the early prognosis of the hepatectomized patient with hepatocellular carcinoma with a neural network. *Comput Biol Med*, 1995; 25(1): 49-59

¹¹⁵ JAKUS V: Op. Cit.

Benigni desarrollo una red neuronal para predecir la toxicidad química de varias sustancias¹¹⁹.

Domine y colaboradores usaron una red neuronal para predecir la aparición de efectos secundarios después de administración de varios medicamentos¹²⁰.

Brause describió una red neuronal para establecer pronóstico de muerte o sobrevida en pacientes con choque séptico tomado como entradas varios parámetros clínicos y de laboratorio.

Una red neuronal demostró ser más útil para predecir sobrevida después de trauma severo que escalas tradicionales como TRISS y ASCOT¹²¹.

Otras aplicaciones de pronóstico medico incluyen¹²²: prediccion de nefrotoxicidad de gentamicina, predicción de desenlace después de neumonías, predicción de desenlace en

¹¹⁶ RAVDIN PM, CLARK GM, HILSENBECK SG et al.: A demonstration that breast cancer recurrence can be predicted by neural network analysis. *Breast Cancer Res Treat*, 1992; 21(1): 47-53

¹¹⁷ BENESOVA O et al.[Perinatal pharmacotherapy and the risk of functional teratogenic defects]. *Cesk Fysiol*, 1995; 44(1): 11-4

¹¹⁸ LAPEER RJ et al. Application of neural networks to the ranking of perinatal variables influencing birth weight. *Scand J Clin Lab Invest Suppl*, 1995; 222: 83-93

¹¹⁹ BENIGNI R., RICHARD A.M.: Quantitative structure-based modeling applied to characterization and prediction of chemical toxicity. *Methods*, 14, 1998, è. 3, s. 264.276

¹²⁰ DOMINE D et al. Nonlinear neural mapping analysis of the adverse effects of drugs. *SAR QSAR Environ. Res.*, 8, 1998, è. 1.2, s. 109.120.

¹²¹ Mcgonigal M. : A New Technique for Survival Prediction in Trauma Care Using a Neural Network. *Proc. World Conference on Neural Networks*, 1994 pp.3495-3498

teratomas testiculares, determinación de riesgo de isquemia cardiaca según niveles de lípidos, predicción de hepatitis según fosfatas alcalinas, predicción de nivel de anticoagulación con warfarina, predicción de complicaciones en unidades de cuidados intensivo neonatales,

Control de equipos médicos

Se han usado RNAs para controlar desfibriladores cardiacos basándose en la señal electrocardiográfica¹²³.

El ajuste adecuado de audífonos usados por personas con deficiencias auditivas y la filtración de ruido han sido logrados usando redes neuronales artificiales¹²⁴.

Redes neuronales han sido desarrolladas para controlar prótesis de miembros superiores después de amputación, basándose en actividad electromiográfica de superficie de músculos por encima del nivel de amputación¹²⁵.

Otra aplicación usa redes neuronales artificiales para controlar equipos de estimulación eléctrica funcional (FES) que ayudan a mejorar la marcha en pacientes con trauma raquimedular¹²⁶.

¹²² DYBOWSKI R, Gant V (Eds): Op. Cit.

¹²³ FARRUGGIA S, YEE H, NICKOLLS P. Implantable cardiverter defibrillator electrogram recognition with a multilayer perceptron. PACE Pacing Clin Electrophysiol, 1993; 161(1Pt2): 228-34

¹²⁴ WANG RJ, JABRI MA. Artificial neural network-based channel selection and loudness mapping. Ann Otol Rhinol Laryngol Suppl, 1995; 166:381-4

¹²⁵ LUNDBORG G et al: Motor control of tomorrow's artificial hand: Based on the combined use of artificial neural networks and a data glove. ASSH 56th annual Meeting, 2001, paper #1.

Una red neuronal artificial ha sido diseñada para controlar los niveles intraoperatorios de diversos anestésicos¹²⁷.

2.2.5 Aplicaciones de redes neuronales a electrodiagnóstico

Las redes neuronales han sido utilizadas para diferenciar entre patologías miopáticas y neuropáticas basándose en los parámetros del potencial de acción de unidad motora (PAUM) en la electromiografía de aguja¹²⁸.

Giacometti y colaboradores diseñaron un sistema híbrido llamado SHADE que usa redes neuronales combinadas con sistemas expertos para diagnóstico de enfermedades neuromusculares usando parámetros obtenidos de estudios de neuroconducción¹²⁹.

Cristodolou describió una red neuronal que realiza descomposición del patrón de interferencia en PAUMs individuales y clasifica enfermedades neuromusculares¹³⁰.

¹²⁶ TONG, K, GRANAT MH: Artificial Neural Network control on functional electrical stimulation assisted gait for persons with spinal cord injury. In Lisboa PJG...(eds): Artificial Neural networks in Biomedicine. Springer, London, 2001, pp 181-193.

¹²⁷ KANGAS LJ, KELLER, PE: Neurometric assessment of adequacy of intraoperative anesthetic. In Lisboa PJG...(eds): Artificial Neural networks in Biomedicine. Springer, London, 2001, pp 81-91

¹²⁸ PATTICHIS CS, SCHIZAS CN, MIDDLETON L: Neural Networks models in EMG diagnosis. IEEE transactions on biomedical Engineering, 1995, 42(5), 486-496

¹²⁹ GIACOMETTI A et al: A Hybrid approach to computer-aided diagnosis in electromyography. Proceedings of the Annual International conference of the IEEE Engineering in Medicine and Biology Society, vol. 14, 1992.

¹³⁰ CRISTODOULOU CU, PATTICHIS CS: A new technique for the classification and decomposition of EMG signals. Proceedings of the IEEE international conference on Neural Networks, 1995, vol. 5, pp 2303-2308

Zazula ha utilizado redes neuronales artificiales para descomposición automática de electromiografía de superficie (SEMG)¹³¹.

Jervis diseño una red neuronal artificial para interpretación de potenciales evocados cognitivos¹³².

Schizas y colaboradores describieron una red neuronal artificial para clasificación de señales de macroelectromiografía (MacroEMG)¹³³.

¹³¹ ZAZULA D, KOROSSEC, D, SOSTARIC A: Computer-assisted decomposition of the electromyograms. Proceedings of the 11th IEEE symposium on Computer-Based Medical Systems (CBMS), 1998, Lubbock, TX, USA, pp 26-31.

¹³² JERVIS BW: The application of Neural Networks to interpret Evoked Potential waveforms. In Lisboa PJG... (eds): Artificial Neural networks in Biomedicine. Springer, London, 2001, pp 195-210.

¹³³ SCHIZAS C et al: Unsupervised learning in computer aided macro Electromyography. In Computer-based Medical Systems, IEEE Computer Society Press, Los Alamitos, CA, USA, 1991, pp 305-312.

3 METODOLOGÍA

3.1 DATOS

Como se ha mencionado anteriormente, el éxito del funcionamiento de una red neuronal depende en su mayor parte de la calidad de los datos que se le presenten para entrenamiento, ya que de estos depende la generalización que la red pueda realizar para tratar datos que no conoció con anterioridad, tratándose de datos médicos la calidad de los datos toma una relevancia aún mayor.

3.1.1 Salidas deseadas

La primera definición establecida corresponde a las salidas que se quieren obtener de la red, lo cual se encuentra definido por los objetivos del estudio. En el presente caso, se pretende lograr una red neuronal artificial que pueda contribuir al diagnóstico de neuropatías periféricas focales, para lo cual se determinaron los siguientes valores por categorías (además del diagnóstico Normal):

Extensión del Daño:

- Mononeuropatía
- Mononeuropatía múltiple
- Polineuropatía

- Mononeuropatía + Polineuripatía

Localización del daño:

- Distal
- Proximal
- Difusa
- Multifocal

Tipo de Daño:

- Desmielinizante
- Axonal
- Mixto

De manera que para un diagnóstico efectivo es preciso establecer la extensión, localización y tipo del daño, además del nervio afectado y la severidad de la lesión.

3.1.2 Selección de los datos de entrada

La cantidad de datos que se presenten a la red para su aprendizaje determinará el número de neuronas que contenga la capa de entrada y por ende la complejidad de la red, y en este caso particular, los datos de entrada deben reflejar los datos sobre los cuales el médico basa su diagnóstico.

Considerando las salidas establecidas, inicialmente se definieron 36 datos de entrada que contemplaban los distintos parámetros de las conducciones motoras y sensitivas para los tres nervios principales de miembro superior (ulnar, mediano y radial). Los datos de la historia clínica y del examen físico (signos y síntomas) no fueron incluidos por considerarse distractores para el proceso de entrenamiento de la red y por la baja sensibilidad encontrada en los ensayos clínicos¹³⁴.

Los datos definidos fueron los siguientes:

Conducción Motora:

- Latencia distal
- Latencia proximal
- Amplitud distal
- Amplitud proximal
- Variación de amplitud
- Velocidad de conducción segmento distal
- Velocidad de conducción segmento proximal
- Dispersión temporal

Conducción Sensitiva:

- Latencia distal
- Amplitud Distal
- Velocidad de conducción segmento distal
- Dispersión temporal

¹³⁴ AAEM: Practice Parameter for Electrodiagnostic Studies in carpal tunnel syndrome: Summary statement. Muscle Nerve. 2002; (26) pp 928-922.

3.1.3 Preprocesamiento de los datos de entrada

Al analizar la correlación entre los datos de entrada, se encontró que la *variación de amplitud* es un valor calculado a partir de las amplitudes distal y proximal, por lo cual se eliminó este dato de entrada.

Para el procesamiento de la red neuronal artificial es necesario asegurar que todas las entradas se presenten de forma numérica, por cuanto, el aprendizaje de la red se realiza mediante una función numérica; considerando que los datos definidos para la entrada son obtenidos del electromiógrafo, estos son datos numéricos, a excepción de la dispersión temporal, la cual el médico debe definir como presente o no, de tal forma que este parámetro se codificó como 0 para cuando no exista dispersión temporal y 1 para cuando se presente la dispersión temporal.

La tabla 5, presenta los datos de entrada a la red, indicando para cada uno de ellos la unidad de medida y su rango.

Tabla 5 Datos de Entrada de cada nervio

Dato	Unidad	Rango
<i>Conducción Motora</i>		
Latencia distal	ms	(0 – 30]
Latencia proximal	ms	[2 – 40]
Amplitud distal	mV	[0 – 30]
Amplitud proximal	mV	[0 – 30]
Velocidad de conducción segmento distal	m / s	(0 – 120]
Velocidad de conducción segmento proximal	m / s	(0 – 120]

Dispersión temporal	No (0) , Si (1)	
<i>Conducción Sensitiva</i>		
Latencia distal	ms	(0 – 20]
Amplitud Distal	μV	[0 – 99]
Velocidad de conducción segmento distal	m / s	(0 – 120]
Dispersión temporal	No (0) , Si (1)	

3.1.4 Datos Faltantes

Una vez revisados los casos obtenidos se observa que en menos del 10% de los mismos se ha realizado el examen de 3 nervios, la mayoría de los exámenes considera solamente 2 nervios (mediano y ulnar) para el diagnóstico por ser las neuropatías focales más frecuentes de miembros superiores y por hacer parte de todos los protocolos desarrollados para estudio de estas patologías, tomando en cuenta este aspecto se decide trabajar solamente con dos nervios (ulnar y mediano) para realizar el diagnóstico con la red neuronal artificial.

Los casos que tuvieran menos del 80% de los datos definidos fueron descartados y para los casos donde se encontraron datos faltantes se decidió tomar el valor promedio de la población normal según lo publicado en la literatura¹³⁵, tal como se observa en la tabla 6. Debe anotarse que los límites de normalidad se definen usando el valor promedio más o menos dos desviaciones estándar según el caso.

¹³⁵ OH, SJ. Op. Cit.

Tabla 6 Valores promedio para datos de Entrada

Dato	Valor por defecto
<i>Conducción Motora Nervio Mediano</i>	
Latencia distal	2.8
Latencia proximal	6.5
Amplitud distal	6.5
Amplitud proximal	6.5
Velocidad de conducción segmento distal	58.7
Velocidad de conducción segmento proximal	65.7
Dispersión temporal	0 (No)
<i>Conducción Sensitiva Nervio Mediano</i>	
Latencia distal	2.5
Amplitud Distal	19
Velocidad de conducción segmento distal	41.8
Dispersión temporal	0 (No)
<i>Conducción Motora Nervio Ulnar</i>	
Latencia distal	2.1
Latencia proximal	5.9
Amplitud distal	6.1
Amplitud proximal	6.1
Velocidad de conducción segmento distal	61.1
Velocidad de conducción segmento proximal	52.3
Dispersión temporal	0 (No)
<i>Conducción Sensitiva Nervio Ulnar</i>	
Latencia distal	2.03
Amplitud Distal	16
Velocidad de conducción segmento distal	47.4
Dispersión temporal	0 (No)

3.1.5 Fuente de los datos

Se revisaron en forma prospectiva 5860 estudios electrodiagnósticos realizados desde Enero de 2003 a Diciembre de 2004 en tres laboratorios diferentes: Centro de Atención de Salud Ocupacional del Instituto de Seguro Social – Cali (Electromiógrafo de dos canales Neurónica, fabricado por el Centro de Neurociencias, Cuba), Hospital Universitario del

Valle (Electromiógrafo de 4 canales Keypoint, fabricado por Dantec Inc., Dinamarca), Centro de Electrodiagnóstico y Rehabilitación POTENCIALES (electromiógrafo de dos canales Neuropack 2, fabricado por Nihon Khoden Inc., Japon). De los 5860 estudios 2152 correspondieron a neuropatías focales y de ellos 953 a neuropatías focales de miembros superiores. De los 953 estudios correspondientes a neuropatías focales de miembros superiores se seleccionó una muestra de 354 casos a todos los cuales se les había realizado neuroconducciones de mediano y ulnar y que cumplieron además con los criterios establecidos de inclusión y exclusión, listados a continuación.

Criterios de inclusión

- Estudios electrodiagnósticos realizados para evaluar posible neuropatía focal de miembros superiores.
- Descripción detallada de los sitios de registro y estímulo para cada nervio evaluado.
- Registro explícito sobre la normalidad o anormalidad del resultado del estudio y sobre la severidad de la neuropatía focal.
- Valores de referencia del laboratorio claramente establecidos.

Criterios de exclusión

- Patologías asociadas a la neuropatía focal tales como radiculopatías, plexopatías, miopatías, alteración de la unión neuromuscular, enfermedades del asta anterior, etc.
- Cirugía previa para tratamiento de neuropatía focal.
- Diagnóstico electrofisiológico no concluyente.
- Patologías muy severas donde fue imposible obtener las neuroconducciones motoras y sensitivas.
- Estudios donde faltaban más del 20% de los parámetros de entrada.

Finalmente se obtuvo un conjunto de 300 casos que incluían los diversos diagnósticos esperados, como se observa en la tabla 7.

Tabla 7 Distribución de Casos para Entrenamiento

Extensión del daño	Porcentaje
Normal	7.0 %
Mononeuropatía	55.0 %
Mononeuropatía Múltiple	5.0 %
Polineuropatía	13.7 %
Mononeuropatía + Polineuropatía	19.3 %

3.2 ARQUITECTURA DE LA RED

Considerando que la red neuronal en construcción trabajaría sobre dos nervios que tenían los mismos parámetros de entrada y que en algún momento podrían llegar a interferirse, se discutió la opción de construir un esquema de tres redes, donde dos redes iniciales se entrenarían para diagnosticar de forma particular un nervio cada una y las salidas de estas redes serían presentadas a una tercera red, la cual se encargaría de presentar un diagnóstico general.

Finalmente, se decidió trabajar con dos esquemas para luego comparar los resultados obtenidos. El primer esquema, corresponde a una red normal, que recibe todos los parámetros de entrada y clasifica de acuerdo a las salidas establecidas. El segundo esquema, es el de tres redes, dos iniciales para cada nervio y una tercera para generalizar.

3.2.1 Tipo de red

Aproximar la función que define un diagnóstico médico no es sencillo, debido principalmente a que los datos médicos son diversos y en ocasiones, altamente complejos, por esta razón es importante determinar el tipo de red apropiado, de manera tal que se logre una aproximación muy exacta de la función objetivo, para ello, se definieron las siguientes características deseadas para el tipo de red a emplear:

- Alta aplicabilidad para casos de clasificación y reconocimiento de patrones.
- Emplee aprendizaje supervisado, específicamente por corrección de error, de tal forma que el proceso de aprendizaje sea controlado y los pesos se ajusten como resultado de la comparación de las salidas obtenidas con las salidas reales, llevando a que la red determine patrones a partir de los casos presentados (la experiencia) y no por correlación entre los datos de entrada, lo cual podría generar soluciones improbables o inexistentes.
- Realice aprendizaje OFF – LINE, de manera que se pueda calcular la eficiencia de la red que entrará en funcionamiento y se eviten procesos de sobreaprendizaje o memorización.
- Red Heteroasociativa, especializada para computar funciones que relacionan las entradas con las salidas y que en general, no se podrían establecer de forma analítica.
- Capacidad de generalizar para casos no presentados durante el proceso de aprendizaje.
- Poca complejidad computacional, de forma que se logre llegar a una solución apropiada sin requerir excesivos recursos.

Analizadas las condiciones mencionadas se encontró que las redes Backpropagation (Red de alimentación hacia delante con algoritmos de retropropagación) cubrían las especificaciones descritas y considerando que diferentes estudios han demostrado que este tipo de red es especialmente apropiada para tratar casos de diagnóstico médico ^{136,137,138}, se seleccionó la estructura Backpropagation para el desarrollo del trabajo.

3.2.2 Mejorar la Generalización

Para mejorar la generalización que realice la red neuronal artificial, se empleó validación cruzada y parada temprana, para lo cual el conjunto de 300 casos seleccionados se dividió en tres grupos organizados en forma aleatoria:

- *Grupo de entrenamiento:* el cual permitirá entrenar la red hasta lograr el error mínimo deseado, en este grupo se incluyeron 150 casos. Se verificó que este grupo incluyera casos que proveyeran todas las posibles salidas de la red.
- *Conjunto de validación:* este conjunto de datos permite verificar que la red no haya caído en un sobreentrenamiento. Se destinaron 75 casos en forma aleatoria a este conjunto.
- *Conjunto de prueba:* permite comparar las diferentes estructuras de la red que se ensayen. Se destinaron 75 casos en forma aleatoria a este conjunto.

¹³⁶ FU, L.M. Neural networks in computer intelligence. McGrawHill, Singapore (1994)

¹³⁷ SWINGLER, K. Applying neural networks. Academic Press. London (1996)

¹³⁸ PESONEN, E., Eskelinen, M., Juhola, M. Comparison of different neural networks algorithm in the diagnosis of acute appendicitis. Int. J Bio-Med. Comput. 40 (1996) 227-233.

3.2.3 Arquitectura de la Red 1

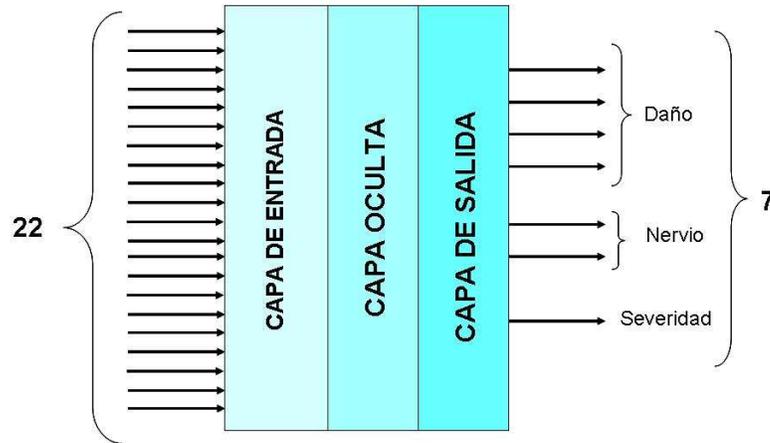
Las redes Backpropagation tienen al menos tres capas: una capa de entrada, una capa de salida y al menos una capa oculta. Tomando en cuenta los estudios presentados¹³⁹, en los cuales se demuestra que en una red Backpropagation es suficiente una capa oculta para lograr un buen nivel de generalización, se determina mantener una sola capa oculta.

El número de neuronas para la capa de entrada está dado por el número de datos de entrada, en este caso 22 (11 por cada nervio).

Para disminuir la complejidad de la red (dada por el número de enlaces que contiene) se decidió trabajar con la combinación de salidas más comunes y codificarlas mediante números binarios, de manera que se requieran menos neuronas en la capa de salida, así cuatro neuronas (que representan bits) indican el tipo de lesión que sufre el nervio, dos neuronas establecen el o los nervio(s) lesionado(s) y otra neurona indica la severidad. La estructura general de la red, se presenta en la figura 28:

¹³⁹ BISHOP, op.cit

Figura 28 Estructura de la Red 1



La codificación definida para las diferentes clases de lesiones se presenta en la tabla 8.

Tabla 8 Codificación de las Salidas de la Red

Parámetro	Valor	Codificación
<i>Daño</i>		
Normal	0	0000
Mononeuropatía distal desmielinizante	1	0001
Mononeuropatía proximal desmielinizante	2	0010
Mononeuropatía Múltiple multifocal desmielinizante	3	0011
Polineuropatía difusa desmielinizante	4	0100
Mononeuropatía + Polineuropatía difusa desmielinizante	5	0101
Mononeuropatía distal axonal	6	0110
Mononeuropatía proximal axonal	7	0111
Mononeuropatía difusa axonal	8	1000
Mononeuropatía Múltiple multifocal axonal	9	1001
Polineuropatía difusa axonal	10	1010
Mononeuropatía Múltiple multifocal mixta	11	1011
Polineuropatía difusa mixta	12	1100
Mononeuropatía + Polineuropatía difusa mixta	13	1101
<i>Nervio Afectado</i>		
Nervio Mediano	1	01
Nervio Ulnar	2	10
Ambos Nervios	3	11
<i>Severidad</i>		
Leve a Moderado	0	0
Severo	1	1

Como punto de partida para el entrenamiento de la red 1, se definieron las siguientes características:

1. *Número de Capas:* Las redes neuronales de tipo Backpropagation, emplean al menos tres capas: una de entrada, una de salida y una oculta, como se mencionó anteriormente, varios estudios han demostrado que una capa oculta es suficiente para lograr un buen nivel de generalización, por lo cual se decidió emplear inicialmente 3 capas para la red.
2. *Capa de Entrada:* El número de neuronas contenidas en la capa de entrada esta dado por el número de características de entrada seleccionadas para la red, en este caso la capa de entrada tiene 22 neuronas, cada una de las cuales codifica una característica de un nervio (ver tabla 6).
3. *Capa de Salida:* La capa de salida posee 7 neuronas, estas neuronas permiten entregar un diagnóstico codificado de forma que las 4 primeras neuronas identifican el tipo de daño, las 2 siguientes el nervio sobre el cual ocurre y la última la severidad del daño (ver tabla 8).
4. *Capa Oculta:* Como se mencionó anteriormente, no existe una fórmula para establecer con total precisión el número de neuronas ideal para la capa oculta, sin embargo se puede obtener un buen punto de partida aplicando la fórmula propuesta por Livingstone y Manallack¹⁴⁰: $D = \frac{mo}{w}$, donde m es el número de casos de entrenamiento, o es el número de salidas de la red, y w es el número total de pesos en la red (conexiones de la capa de entrada a la capa oculta y de la capa oculta a la capa de salida). Se considera que un valor superior a 3 para D asegura la generalización y evita la memorización durante el entrenamiento.

¹⁴⁰ PENNY, Hill, FROST, Davis. Neural Networks in Clinical Medicine.

Considerando los aspectos definidos para la red, se tiene:

$$4 = \frac{150 * 7}{w} \quad \rightarrow \quad w = \frac{150 * 7}{4} \quad \rightarrow \quad w = 263$$

Ya que la red posee 22 entradas y 7 salidas, y que la red backpropagation es totalmente conectada, para lograr una cantidad aproximada de 263 pesos en la red, se establecieron 10 neuronas en la capa oculta como parámetro inicial.

5. *Función de Transferencia en la capa oculta:* Debido a que las redes Backpropagation requieren una función derivable como función de transferencia, se decidió emplear la función sigmoideal, ya que esta, además de ser derivable, permite a la red aprender funciones tanto lineales como no lineales para relacionar las entradas y salidas de la misma. Específicamente se empleó la función tansig la cual se acota entre -1 y 1.
6. *Función de Transferencia en la capa de salida:* Debido a que se quiere restringir las salidas de la red al rango [0,1] y considerando lo dicho en el apartado anterior, se empleó la función de transferencia sigmoideal (logsig) para las neuronas de la capa de salida.
7. *Iniciación de pesos y sesgos:* El punto de inicio de la búsqueda en el espacio de la función de error no es tan relevante al momento de realizar el entrenamiento, ya que este se ha diseñado de forma tal que siempre se aproxime hacia un mínimo local o global de la función, por esta razón se decidió realizar la iniciación de los pesos y sesgos con una función aleatoria, de manera que se asegure un valor inicial diferente a cero.
8. *Tasa de Aprendizaje:* Como se mencionó anteriormente, se decidió utilizar validación cruzada y parada temprana para mejorar el proceso de generalización de la red, considerando que estos métodos exigen una tasa de aprendizaje pequeña, se fijó 0.001 como parámetro inicial para la red.
9. *Función de aprendizaje:* Se seleccionó la función de aprendizaje en lotes con gradiente descendente y momento (learngdm), ya que este tipo de función permite alcanzar de

manera más rápida y efectiva los mínimos dentro de una función, evitando que el proceso se quede estancado.

10. *Forma de entrenamiento:* Para realizar el entrenamiento se seleccionó el aprendizaje por lotes (batch), en el cual, se presentan a la red todos los casos de entrenamiento y solamente después de esto se modifican los pesos basados en el error acumulado; este tipo de entrenamiento permite una mejor aproximación de la función logrando que el tiempo de procesamiento se disminuya sin afectar el desempeño de la red.
11. *Algoritmo de entrenamiento:* Como algoritmo de entrenamiento se seleccionó el algoritmo Levenberg-Marquardt, el cual fue diseñado para mejorar la velocidad de entrenamiento sin necesidad de calcular la matriz Hessiana, debido a que ha sido reconocido como el método más rápido de entrenamiento para redes neuronales *feedforward*¹⁴¹, ideal cuando la red contiene pocos pesos, ya que en redes con muchos pesos puede llegar a requerir gran cantidad de recursos.
12. *Error:* Se determinó un error esperado de 0.01 para el entrenamiento, de manera que en condiciones ideales, se obtenga un 99% de confianza en la generalización lograda por la red; este error se determinó considerando que en las guías para estudios electrodiagnósticos se ha encontrado que la sensibilidad promedio de dichas pruebas es alrededor del 85 – 87%^{142,143}.
13. *Número de pasadas:* El número de pasadas para entrenamiento inicial se fijó en 100, es decir que la red terminará su entrenamiento cuando alcance el error fijado o haya realizado este número de pasadas de los casos de entrenamiento.

¹⁴¹ Matlab. BACKPROPAGATION. Speed and Memory comparison. Pg 5-35, 5.36.

¹⁴² AAEM: Practice Parameter for Electrodiagnostic Studies in carpal tunnel syndrome: Summary statement. Muscle Nerve. 2002; (26) pp 928-922.

¹⁴³ AAEM: Practice Parameter for Electrodiagnostic Studies in Ulnar Neuropathy at the elbow: Summary statement. Muscle Nerve. 1999; (suppl 8) pp S171-S174.

3.2.4 Arquitectura de la Red 2

Esta arquitectura realmente esta compuesta por tres redes. Las dos primeras, que analizarán cada uno de los nervios tendrán una estructura similar, la última, que corresponde a la red de generalización, tiene una estructura diferente.

Redes para cada nervio

Para definir la red que analizaría cada uno de los nervios, era necesario empezar por definir sus posibles salidas, las cuales se establecieron de la siguiente forma:

- Estudio Normal
- Desmielinización focal distal
- Desmielinización focal proximal
- Desmielinización generalizada
- Daño Axonal
- Severidad

Siguiendo la idea de disminuir el número de neuronas en la capa de salida, se propuso codificar las salidas que genera la red, tal como se presenta en la tabla 9. En este caso se consideraron las patologías mas frecuentes.

La figura 29 presenta la estructura las redes que analizaran cada nervio, en este estudio se considerarán el nervio mediano y ulnar, sin embargo, la misma estructura puede emplearse para otros nervios.

Figura 29 Estructura de la Red para cada Nervio

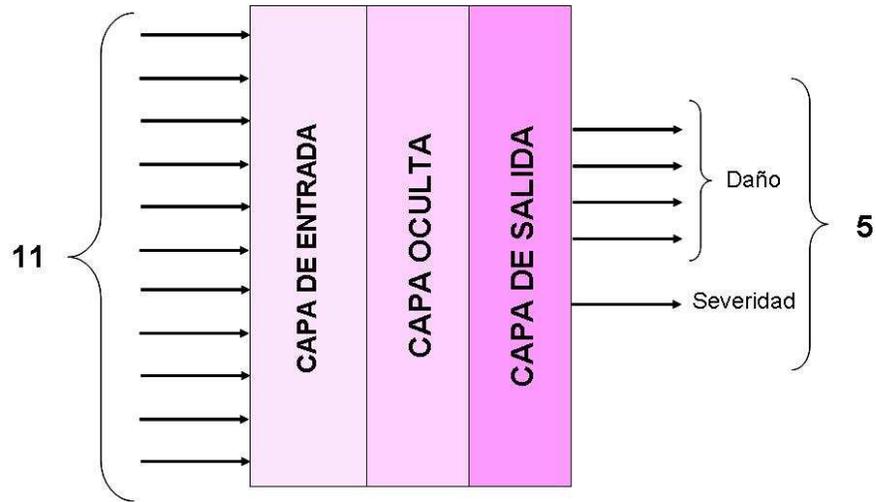


Tabla 9 Salidas de la Red de Cada Nervio

Tipo de Daño	Valor	Codificación
Estudio Normal	0	0000
Neuropatía Focal Distal	1	0001
Neuropatía Focal Proximal	2	0010
Neuropatía Generalizada	3	0011
Daño Axonal	4	0100
Neuropatía Focal Distal + Generalizada	5	0101
Neuropatía Focal Distal + Daño Axonal	6	0110
Neuropatía Focal Proximal + Generalizada	7	0111
Neuropatía Focal Proximal + Daño Axonal	8	1000
Neuropatía Generalizada + Daño Axonal	9	1001
Severidad	0 – leve a moderado 1 - Severo	

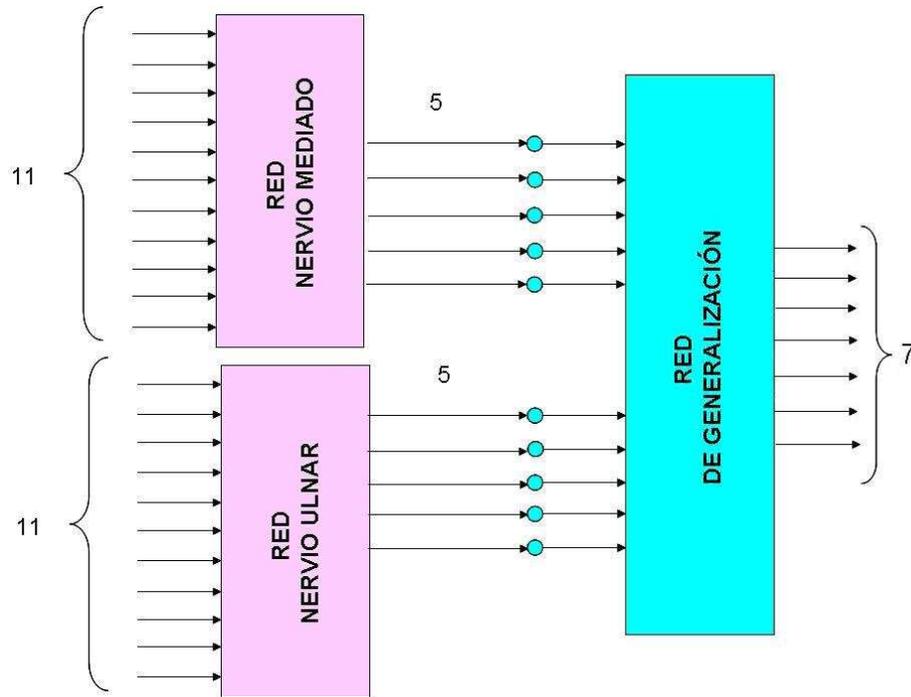
Las características iniciales para esta red son las mismas definidas para la red 1, variando únicamente la cantidad de neuronas en cada capa, de la siguiente forma:

1. *Capa de Entrada:* La capa de entrada para la red que analiza cada nervio tendrá 11 neuronas.
2. *Capa de Salida:* La capa de salida posee 5 neuronas, de acuerdo a la codificación presentada.
3. *Capa Oculta:* Empleando la ecuación propuesta, se obtiene que la red para cada nervio debe poseer aproximadamente 12 neuronas en su capa oculta.

Red de Generalización

Esta red recibe como entradas las salidas provenientes de las dos redes de cada nervio y sus salidas serán las mismas propuestas para la red 1. La figura 30 muestra el esquema general de la red de generalización.

Figura 30 Estructura General Red 2



Las características iniciales definidas para la red de generalización son las mismas definidas para las otras redes, variando solamente en el número de neuronas en cada capa, de la siguiente forma:

1. *Capa de Entrada:* La capa de entrada para la red tendrá 10 neuronas correspondientes a las salidas de cada una de las redes de los nervios.
2. *Capa de Salida:* La capa de salida tendrá 7 neuronas de acuerdo a la codificación propuesta inicialmente y empleada para la red 1.
3. *Capa Oculta:* Empleando la ecuación propuesta, se obtiene para esta red una cantidad de 263 pesos, lo que lleva a tener 16 neuronas en la capa oculta.

3.3 SOFTWARE

Para el desarrollo de la red neuronal artificial y su proceso de entrenamiento se empleó MATLAB 6.5 fabricado por MathWorks Inc. considerando la trayectoria de este software para el manejo de aplicaciones matemáticas y la calidad de su módulo de redes neuronales Neural nets Tool 4.0. que incluye los algoritmos básicos para el manejo de redes neuronales de diferente clase.

3.4 HARDWARE

Para el proceso de entrenamiento y verificación de la red, se empleó un computador con procesador INTEL Pentium IV a 2.4GHz, con 512 Mbytes de Memoria RAM, Sistema Operativo Windows XP Service Pack 2 producido por Microsoft Corporation.

3.5 ENTRENAMIENTO

El entrenamiento de la red se realizó inicialmente con los parámetros establecidos, proporcionando a la red el conjunto de datos de entrenamiento (150), el conjunto de datos de validación (75) y el conjunto de datos de prueba (75) aunque este último conjunto solamente fue considerado para establecer el error de generalización, pero en ningún momento para el entrenamiento.

Se realizaron variaciones en:

- Número de Neuronas en la capa oculta, incrementando y disminuyendo a partir del valor definido inicialmente.

- Tasa de aprendizaje.
- Valores iniciales de los pesos.
- Número de capas ocultas, se incrementó una capa oculta y se realizaron algunos entrenamientos con diferente combinación de neuronas en cada capa.

El proceso de entrenamiento de las redes cubrió diferentes cambios a pesar de haber alcanzado un error aceptable, para verificar si a pesar de esto era posible lograr un mejor desempeño de la red y poder comparar los dos esquemas propuestos.

3.6 VALIDACIÓN DE LA RED

El conjunto de prueba constituido por 75 casos fue presentado a dos neurofisiólogos expertos en forma ciega para que a partir de los datos de las neuroconducciones establecieran un diagnóstico electrofisiológico. Los resultados fueron comparados con el desempeño de las redes propuestas.

4 RESULTADOS

4.1 RED 1 (ESTRUCTURA DE RED GENERAL)

Durante el entrenamiento de la Red General, propuesta inicialmente, se obtuvieron los siguientes resultados (tablas 10)

Tabla 10 Entrenamiento de la Red 1

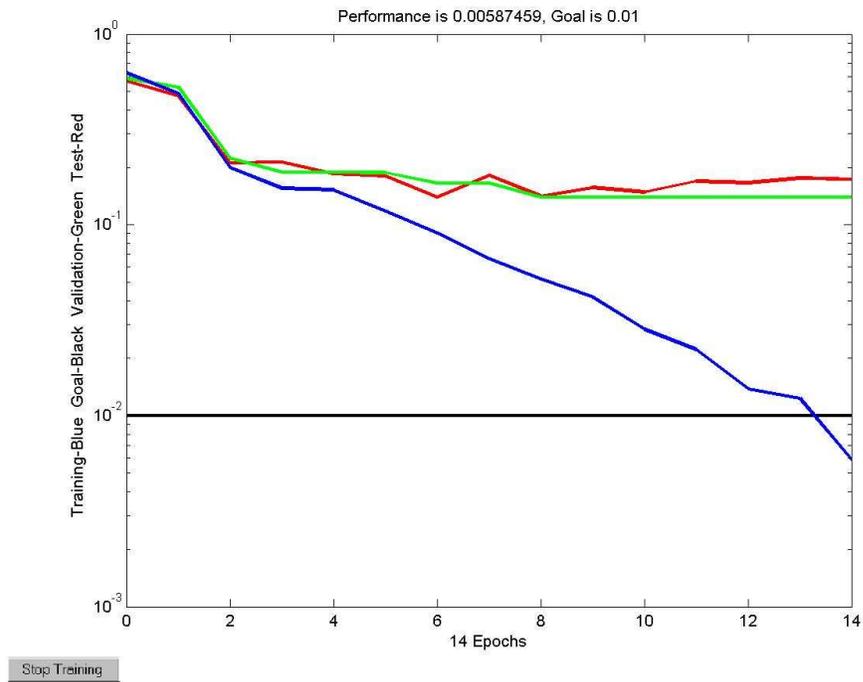
Neuronas de Entradas	Neuronas de Salidas	Neuronas Capa Oculta 1	Neuronas Capa Oculta 2	Tasa de Error	Tasa de Aprendizaje	Error Obtenido	Iteraciones Realizadas
22	7	10	NI	0.01	0.001	0.0962804	37
22	7	12	NI	0.01	0.001	0.0446425	19
22	7	14	NI	0.01	0.001	0.0409499	14
22	7	20	NI	0.01	0.001	0.0132245	16
22	7	25	NI	0.01	0.001	0.00587459	14
22	7	30	NI	0.01	0.001	0.0158104	13
22	7	32	NI	0.01	0.001	0.0396736	25
22	7	27	NI	0.01	0.001	0.0241254	11
22	7	25*	NI	0.01	0.001	0.0167966	15
22	7	25*	NI	0.01	0.001	0.0137648	17
22	7	25	NI	0.01	0.01	0.00587459	14
22	7	25	NI	0.01	0.1	0.00587459	14
22	7	10	10	0.01	0.001	0.0124389	17
22	7	12	12	0.01	0.001	0.01562	19
22	7	14	14	0.01	0.001	0.00781096	13

* Se variaron los pesos iniciales
NI = No implementada

El menor error de entrenamiento conseguido fue 0.00587459, con una estructura 22-25-7.

El error de generalización para este caso se calculó en 0.1417, es decir una efectividad de 85.83%. La figura 31 muestra el resultado de entrenamiento de la red.

Figura 31 Entrenamiento de la Red 1



4.2 RED 2 (RED NERVIO MEDIANO)

Para la red 2, dedicada al entrenamiento de los datos para el nervio mediano, se obtuvieron los siguientes datos (tabla 11).

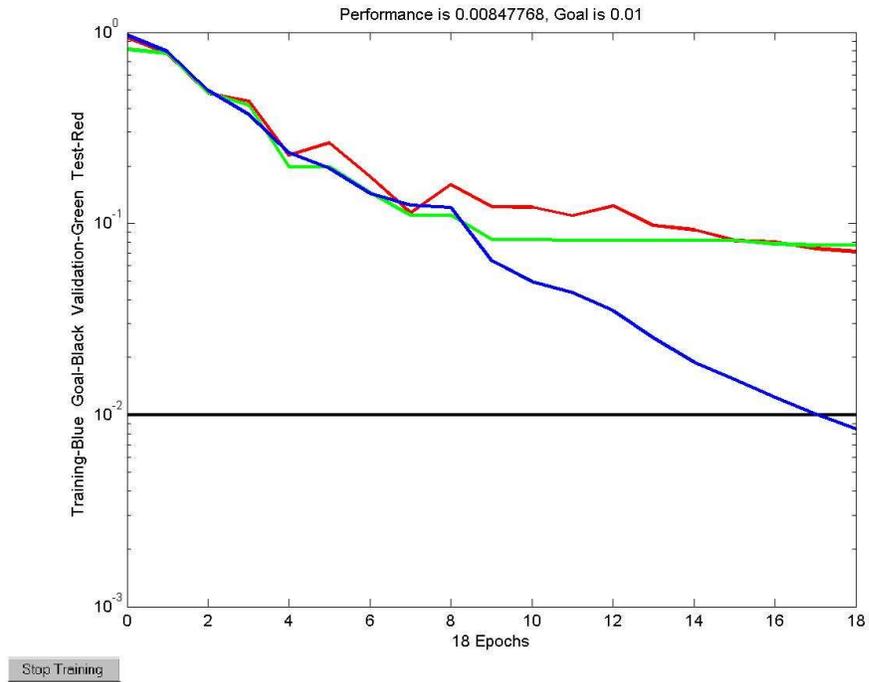
Tabla 11 Entrenamiento para Red 2 (Nervio Mediano)

Neuronas de Entradas	Neuronas de Salidas	Neuronas Capa Oculta 1	Neuronas Capa Oculta 2	Tasa de Error	Tasa de Aprendizaje	Error Obtenido	Iteraciones Realizadas
11	5	12	NI	0.01	0.001	0.11318	32
11	5	14	NI	0.01	0.001	0.0168348	17
11	5	16	NI	0.01	0.001	0.0196215	17
11	5	18	NI	0.01	0.001	0.0151185	19
11	5	20	NI	0.01	0.001	0.0226181	14
11	5	22	NI	0.01	0.001	0.033063	20
11	5	24	NI	0.01	0.001	0.0157561	14
11	5	10	NI	0.01	0.001	0.043623	19
11	5	8	NI	0.01	0.001	0.0360276	15
11	5	25	NI	0.01	0.001	0.0126002	16
11	5	30	NI	0.01	0.001	0.00984231	18
11	5	35	NI	0.01	0.001	0.0235687	14
11	5	32	NI	0.01	0.001	0.0259881	18
11	5	30	NI	0.01	0.01	0.00984231	18
11	5	30	NI	0.01	0.1	0.00984231	18
11	5	30*	NI	0.01	0.001	0.00974239	18
11	5	30*	NI	0.01	0.001	0.78664	29
11	5	10	10	0.01	0.001	0.0110845	17
11	5	12	12	0.01	0.001	0.00960978	14
11	5	14	14	0.01	0.001	0.00847768	18
11	5	10	14	0.01	0.001	0.0130054	17
11	5	14	16	0.01	0.001	0.00887554	12

* Se variaron los pesos iniciales
 NI = No implementada

El menor error de entrenamiento conseguido fue 0.00984231, con una estructura 11-30-5 cuando se dispuso una sola capa oculta; con dos capas ocultas se mejoró el desempeño de la red, obteniendo un error de 0.00847768 con una estructura de 11-14-14-5. El error de generalización para este caso se calculó en 0.074, es decir una efectividad de 92.6%. La figura 32 muestra el resultado de entrenamiento de la red.

Figura 32 Entrenamiento de la Red 2 (Nervio Mediano)



4.3 RED 3 (RED NERVIO ULNAR)

Para la red 3, dedicada al entrenamiento de los datos para el nervio ulnar, se obtuvieron los siguientes datos (tabla 12).

Tabla 12 Entrenamiento para Red 3 (Nervio Ulnar)

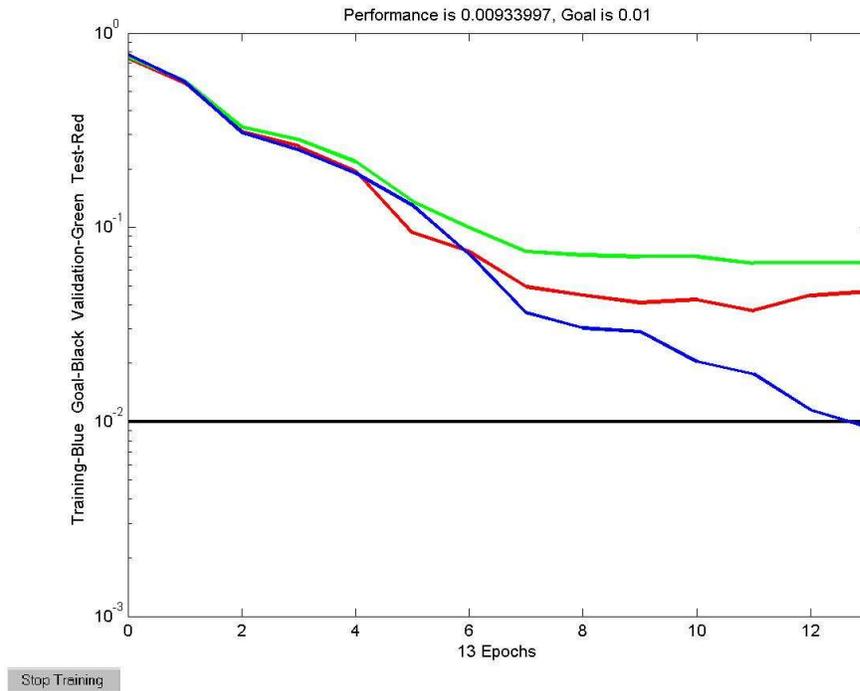
Neuronas de Entradas	Neuronas de Salidas	Neuronas Capa Oculta 1	Neuronas Capa Oculta 2	Tasa de Error	Tasa de Aprendizaje	Error Obtenido	Iteraciones Realizadas
11	5	12	NI	0.01	0.001	0.187378	57
11	5	14	NI	0.01	0.001	0.0143109	21
11	5	16	NI	0.01	0.001	0.0229125	25
11	5	18	NI	0.01	0.001	0.0149204	15
11	5	20	NI	0.01	0.001	0.0191713	15
11	5	22	NI	0.01	0.001	0.00926019	21
11	5	24	NI	0.01	0.001	0.00808746	18
11	5	10	NI	0.01	0.001	0.280441	28
11	5	8	NI	0.01	0.001	0.0199048	25
11	5	25	NI	0.01	0.001	0.0142937	14
11	5	30	NI	0.01	0.001	0.0119402	21
11	5	32	NI	0.01	0.001	0.00989668	16
11	5	35	NI	0.01	0.001	0.00953449	13
11	5	40	NI	0.01	0.001	0.0121785	18
11	5	38	NI	0.01	0.001	0.00933997	13
11	5	10	10	0.01	0.001	0.0164272	32
11	5	12	12	0.01	0.001	0.00984463	15
11	5	14	14	0.01	0.001	0.00916231	12
11	5	10	14	0.01	0.001	0.00867688	12
11	5	14	16	0.01	0.001	0.00738405	13
11	5	38	NI	0.01	0.01	0.00933997	13
11	5	38	NI	0.01	0.1	0.00933997	13
11	5	38*	NI	0.01	0.001	0.0942831	13
11	5	38*	NI	0.01	0.001	0.0947074	16

* Se variaron los pesos iniciales
 NI = No implementada

El menor error de entrenamiento conseguido fue 0.00933997, con una estructura 11-38-5.

El error de generalización para este caso se calculó en 0.0373, es decir una efectividad de 96.27%. La figura 33 muestra el resultado de entrenamiento de la red.

Figura 33 Entrenamiento de la Red 3 (Nervio Ulnar)



4.4 RED 4 (RED DE GENERALIZACIÓN)

Para la red 4, que genera un diagnóstico a partir de las salidas entregadas por las redes de cada nervio, se obtuvieron los siguientes datos (tabla 13).

Tabla 13 Entrenamiento para Red 4 (Generalización)

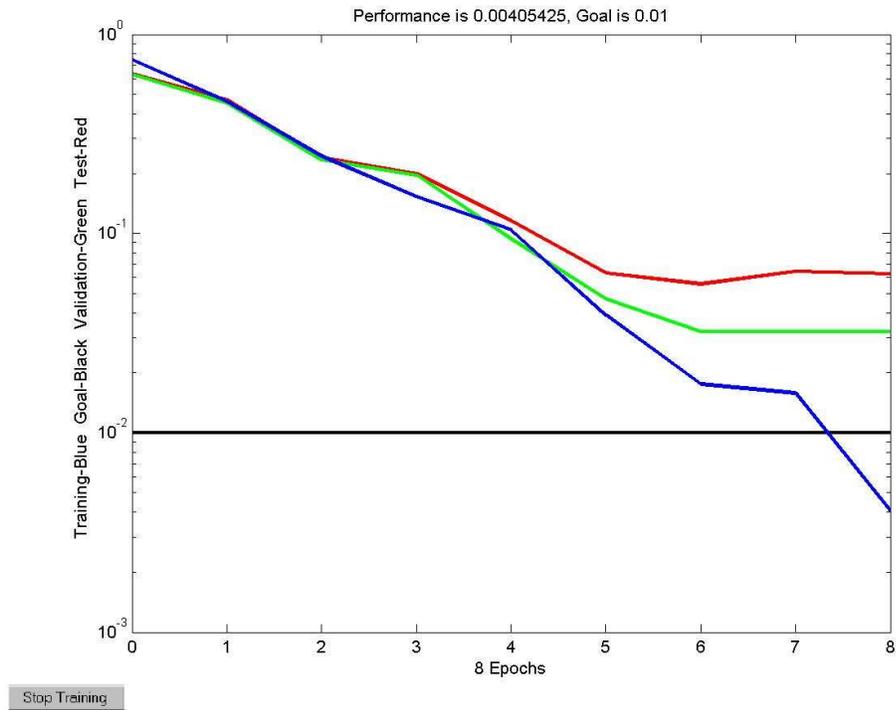
Neuronas de Entradas	Neuronas de Salidas	Neuronas Capa Oculta 1	Neuronas Capa Oculta 2	Tasa de Error	Tasa de Aprendizaje	Error Obtenido	Iteraciones Realizadas
10	7	16	NI	0.01	0.001	0.0071389	15
10	7	18	NI	0.01	0.001	0.50381	57
10	7	20	NI	0.01	0.001	0.87933	22
10	7	17	NI	0.01	0.001	0.00948825	10
10	7	14	NI	0.01	0.001	0.00791259	8
10	7	15	NI	0.01	0.001	0.00988104	7
10	7	16	NI	0.01	0.01	0.0071389	15
10	7	16	NI	0.01	0.1	0.0071389	15
10	7	16*	NI	0.01	0.001	0.0416514	52
10	7	16*	NI	0.01	0.001	0.257833	49
10	7	16*	NI	0.01	0.001	0.00405425	8
10	7	8	8	0.01	0.001	0.976745	48
10	7	10	10	0.01	0.001	0.721905	13
10	7	12	12	0.01	0.001	0.618745	23
10	7	16	16	0.01	0.001	0.75098	32

* Se variaron los pesos iniciales
 NI = No implementada

El menor error de entrenamiento conseguido fue 0.00405425, con una estructura 10-16-7.

El error de generalización para este caso se calculó en 0.0558, es decir una efectividad de 94.42%. La figura 34 muestra el resultado de entrenamiento de la red.

Figura 34 Entrenamiento para la Red 4 (Generalización)



Calculo de Error

Considerando que la segunda arquitectura propuesta involucra la conexión de tres redes neuronales, cada una con un porcentaje de error calculado, el error total de la estructura debe calcularse para efectos reales.

Ya que cada una de las redes que analizan los nervios, aportan igual cantidad de información a la tercera red, el error total de la estructura se calculó de la siguiente forma:

$$Error_{total} = 1 - ((aciertos_{red2} * 0.5 + aciertos_{red3} * 0.5) * aciertos_{red4})$$

Para este caso, se denominó aciertos al porcentaje de efectividad dado por la red (1 – error de generalización). La red 2 corresponde a la red del nervio mediano, la red 3 la red del nervio ulnar y la red 4 es la red de generalización.

El resultado obtenido fue:

$$Error_{total} = 1 - ((0.926 * 0.5 + 0.9627 * 0.5) * 0.9442)$$

$$Error_{total} = 1 - 0.89165527$$

$$Error_{total} = 0.10834473$$

4.5 VALIDACIÓN DE RESULTADOS

En la tabla 14 se presentan los resultados comparativos entre el mejor desempeño alcanzado por las redes y los expertos en neurofisiología consultados.

Tabla 14 Validación de datos para la Red 2 (Nervio Mediano)

Patrón evaluado	Precisión red neuronal artificial	Precisión Neurofisiólogo #1	Precisión Neurofisiólogo #2
Estudio Normal	99.07%	98%	96%
Neuropatía Focal Distal	96.3%	96%	95%
Neuropatía Focal Proximal	85.4%	82%	80%
Neuropatía Generalizada	89.0%	87%	88%
Daño Axonal	80.2%	82%	80%
Neuropatía Focal Distal + Generalizada	81.5%	72.3%	65%
Neuropatía Focal Distal + Daño Axonal	84.23%	83%	78%
Neuropatía Focal Proximal + Generalizada	81.27%	71%	59%
Neuropatía Focal Proximal + Daño Axonal	80.12%	65%	53%
Neuropatía Generalizada + Daño Axonal	82.02%	74%	62%

Tabla 15 Validación para la Red 3 (Nervio Ulnar)

Patrón evaluado	Precisión red neuronal artificial	Precisión Neurofisiólogo #1	Precisión Neurofisiólogo #2
Estudio Normal	99.5%	99%	97%
Neuropatía Focal Distal	85.7%	83%	79%
Neuropatía Focal Proximal	97.3%	95%	93%
Neuropatía Generalizada	89.0%	87%	88%
Daño Axonal	82.2%	81%	79%
Neuropatía Focal Distal + Generalizada	79.0%	71%	63%
Neuropatía Focal Distal + Daño Axonal	83.7%	81%	75%
Neuropatía Focal Proximal + Generalizada	82.0%	72%	55%
Neuropatía Focal Proximal + Daño Axonal	81.3%	72%	60%
Neuropatía Generalizada + Daño Axonal	84.2%	78%	67%

Tabla 16 Validación para la red 4 (Generalización)

Patrón Evaluado	Precisión red neuronal artificial	Precisión Neurofisiólogo #1	Precisión Neurofisiólogo #2
Normal	95.3%	98%	96%
Mononeuropatía distal desmielinizante	94.5%	94%	91%
Mononeuropatía proximal desmielinizante	91.8%	94%	90%
Mononeuropatía Múltiple multifocal desmielinizante	87.5%	85%	73%
Polineuropatía difusa desmielinizante	85.4%	87%	88%
Mononeuropatía + Polineuropatía difusa desmielinizante	92.0%	70%	64%
Mononeuropatía distal axonal	89.2%	82%	79%
Mononeuropatía proximal axonal	81.4%	79%	68%
Mononeuropatía difusa axonal	81.4%	84%	79%
Mononeuropatía Múltiple multifocal axonal	82.7%	67%	55%
Polineuropatía difusa axonal	80.5%	82%	75%
Mononeuropatía Múltiple multifocal mixta	78.2%	61%	49%
Polineuropatía difusa mixta	83.0%	78%	69%
Mononeuropatía + Polineuropatía difusa mixta	75.7%	63%	55%

5 DISCUSIÓN

Al analizar las modificaciones hechas mientras se buscaba la arquitectura óptima de las redes se observó que los cambios en la tasa de aprendizaje de las redes no afectan el desempeño de las mismas pero sí incrementan el tiempo de computo, agregar una capa de neuronas ocultas puede llevar a una pequeña mejora del desempeño de la red, el cual no es significativo frente al incremento ocasionado en la complejidad y el tiempo de cálculo, aumentar el número de neuronas en la capa oculta mejora el desempeño de la red hasta cierto punto, después del cual la red se torna muy compleja y disminuye su nivel de desempeño, emplear el método de parada temprana permite que se obtenga un resultado mas cercano a la realidad, por cuanto maneja un conjunto de datos para validación.

Se observó que las redes definidas para los nervios (mediano y ulnar) aunque reciben los mismos parámetros de entrada y se consideró la misma arquitectura presentaron comportamientos diferentes durante el proceso de entrenamiento, lo cual puede estar relacionado con el número de patrones presentados para cada una de las salidas definidas por la prevalencia diferente de patologías para cada uno de los nervios, por ejemplo, en el nervio mediano son mucho más frecuentes las neuropatías distales (síndrome de túnel carpiano), mientras que en el nervio ulnar son mas frecuentes las neuropatías proximales (síndrome de canal cubital).

La red 4 o red de generalización que recibe los datos obtenidos por las redes de cada nervio, logró un muy buen nivel de desempeño en poco tiempo, debido probablemente a la depuración en los datos de entrada (diagnóstico de cada nervio).

El desempeño logrado en la arquitectura 2, que incluía una red para cada nervio y posteriormente una red de generalización fue verificado considerando los errores propagados de una red a otra, a pesar de lo cual, este esquema resultó mas eficiente que trabajar con una sola red.

La estructura de la red seleccionada mostró ser tan efectiva como su contraparte humana para diferenciar entre estudios normales y anormales con una precisión igual o superior a 95%. El rango de clasificación adecuada de los distintos patrones estuvo entre 75 y 97% mientras que para los expertos estuvo entre 55 y 96%. Al clasificar los patrones de enfermedad se observó que el desempeño de la red es muy semejante al de los expertos cuando se trata de patologías comunes y no combinadas. La red mostró un desempeño muy superior al clasificar patologías poco comunes y sobretodo al encontrar combinaciones de patologías que el experto humano no reconoció fácilmente.

CONCLUSIONES

En este trabajo de investigación se realizó el diseño, el entrenamiento y la aplicación de una red neuronal artificial para la detección de neuropatías periféricas focales cumpliendo con los objetivos propuestos. Los datos obtenidos de estudios de neuroconducción motoras y sensitivos fueron ofrecidos a redes neuronales usando diferentes arquitecturas.

El modelo mas eficiente fue el de una red inicial para el diagnóstico y localización de neuropatías focales de cada nervio individual y una segunda red para determinar la presencia de patologías multifocales o generalizadas, debido probablemente a la especialidad que se logra con cada red, separando el diagnóstico en etapas lógicas que permiten una mejor aproximación.

La arquitectura de red diseñada puede adaptarse fácilmente para generalizar su aplicación a cualquier combinación de nervios tanto de miembros superiores o inferiores.

Una vez entrenada la red mostró ser tan efectiva como expertos humanos para detectar anomalía en los estudios de neuroconducción (mas del 95% de casos normales correctos) y sobre todos para detectar y clasificar los distintos patrones de neuropatías focales y generalizadas. En los casos de patrones que combinan diversa patologías la red es superior que su contraparte humana pues considera todas las posibles combinaciones con las que fue entrenada y no solo las más frecuentes en la práctica clínica.

Dados los excelentes resultados obtenidos no se considera necesario usar sistemas híbridos que combinen redes artificiales neuronales con otras técnica de inteligencia artificial para diagnostico de estas patologías.

La red neuronal artificial desarrollada puede integrarse fácilmente en equipos de Electrodiagnóstico basados en PC, para detección automática de neuropatías focales. Esto sería muy útil como segunda opinión para el neurofisiólogo, independientemente de su nivel de experticia.

RECOMENDACIONES

El presente trabajo es un avance en la incorporación de herramientas de inteligencia artificial a la labor médica, es importante en la medida en que permite probar la eficacia de estas técnicas como apoyo a la toma de decisiones médicas, no obstante, puede entenderse como un punto de partida para el desarrollo de posibles proyectos en otras áreas de ciencias de la salud, tanto en lo referente a la aplicación de redes neuronales como de otras técnicas de inteligencia artificial.

A partir de los resultados obtenidos con esta investigación se pueden derivar propuestas diversas para proyectos futuros que involucren, la incorporación de otros nervios en la estructura para diagnóstico, aplicación de estructuras similares para diagnóstico de otras patologías, por ejemplo en miembros inferiores y la construcción de un software que implemente la red y facilite la interacción de este con el médico.

BIBLIOGRAFIA

- AAEM: Glossary of terms in Electrodiagnostic Medicine. En: Muscle Nerve. No 10. 2001.
- _____ Practice Parameter for Electrodiagnostic Studies in carpal tunnel syndrome: Summary statement. En Muscle Nerve. No 26. 2002.
- _____ Practice Parameter for Electrodiagnostic Studies in Ulnar Neuropathy at the elbow: Summary statement. En Muscle Nerve. No 8. 1999.
- ABREU-LIMA C. and DE SA, J.P. Automatic classifiers for the interpretation of electrocardiograms. Rev. Port. Cardiol. 17. 1998.
- ACCORNERO N and CAPOZZA M OPTONET: neural network for visual field diagnosis. Med Biol Eng Comput, Vol 2. No 33. 1995.
- AMINOFF MJ Electromyography in clinical practice, 3ed. New York, Churchill Livingstone. 1998.
- BAKER JA et al Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon. Radiology, 1995.
- BAXT WG Analysis of the clinical variables driving decision in an artificial neural network trained to identify the presence of myocardial infarction. Ann Emerg Med, 1992.
- _____ Use of an artificial neural network for the diagnosis of myocardial infarction. Ann Intern Med. 1991.
- BENESOVA O et al Perinatal pharmacotherapy and the risk of functional teratogenic defect]. Cesk Fysiol, 1995.
- BENIGNI R. and RICHARD A.M. Quantitative structure-based modeling applied to characterization and prediction of chemical toxicity. Methods, 14, 1998.
- BINDER M. et al. Epiluminiscence microscopy-based classification of pigmented skin lesions using computerized image analysis and an artificial neural network. Melanoma. 1998.
- BISHOP, Christopher M. Neural Network for Pattern Recognition. 1995.

BOON M.E. et al. Neural network processing of cervical smears can lead to a decrease in diagnostic variability and an increase in screening efficiency: a study of 63 false-negative smears. *Mod Pathol*, 1994.

BOUNDS D. and LLOYD P.J. A multi-layer perceptron network for the diagnosis of low back pain. *Proc. IEEE Int. Conf. on Neural Networks*. Vol.II. 1988. p. 481-489

BRAUSE,R. and FRIEDRICH, F. A Neuro-Fuzzy Approach as Medical Diagnostic Interface. *European Symposium on Artificial Neural networks*. ESANN 2000.

BURSTEIN Z. A network model of developmental gene hierarchy. *J Theor Biol* 1995.

BURTIS CA. Technological trends in clinical laboratory science. *Clin Biochem*, 1995.

CHITTAJALLU SK and WONG D. Connectionist networks in auditory system modeling. *Comput Biol Med*, 1994.

CRISTODOULOU CU and PATTICHIS CS. A new technique for the classification and decomposition of EMG signals. *Proceedings of the IEEE international conference on Neural Networks*. Vol 5. 1995.

DAWSON AE; AUSTIN RE Jr and WEINBERG DS. Nuclear grading of breast carcinoma by image analysis. Classification by multivariate and neural network analysis. *Am J Clin Pathol*, 1991.

DEFIGUEIREDO R et al. Neural-network-based classification of cognitively normal, demented, Alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain. *Proc Natl Acad Sci USA*, 1995.

DESCATHA A et al. Incidence of ulnar nerve entrapment at the elbow in repetitive work. *Scand J Work Environ Health*. 2004. p.234-40.

DOMINE D. et al. Nonlinear neural mapping analysis of the adverse effects of drugs. *SAR QSAR Environ. Res.*, 8, 1998.

DONOFRIO PD and ALBERS JW. Polyneuropathy: Classification by nerve conduction studies and electromyography. *En Muscle Nerve*. No 13. 1990.

DUMITRI D. *Electrodiagnostic Medicine*. Williams and Wilkins, Philadelphia, 1995.

DUMITRU D.; AMATO AA and ZWARTS MJ: *Electrodiagnostic Medicine* 2ed. Philadelphia, Hanley & Belfus. 2001.

DYBOWSKI R, Gant V (Eds): *Clinical applications of neural networks*. Cambridge University Press, United Kingdom, 2001.

- EKLUND P and FORSSTROM JJ. Computational intelligence for laboratory information systems. *Scand J Clin Lab Invest Suppl*, 1995. p.222
- ENGEL AG and FRANZINI-ARMSTRONG C, eds. *Myology*. New York, McGraw-Hill. 1994.
- ERRINGTON PA and GRAHAM J. Application of artificial neural networks to chromosome classification. *Cytometry*. 1993.
- FARRUGGIA S; YEE H and NICKOLLS P. Implantable cardioverter defibrillator electrogram recognition with a multilayer perceptron. *PACE Pacing Clin Electrophysiol*. 1993.
- FOGEL DB; WASSON EC 3rd and BOUGHTON EM. Evolving neural networks for detecting breast cancer. *Cancer Lett*. 1995.
- FORINA M. et al. Zupan.s descriptors in QSAR applied to the study of a new class of cardiotoxic agents. *Farmaco*, 52, 1997.
- FU, L.M. *Neural networks in computer intelligence*. McGrawHill, Singapore. 1994.
- GABOR AJ and SEYAL M: Automated interictal EEG spike detection using artificial neural networks. *Electroencephalogr Clin Neurophysiol*. 1992.
- GIACOMETTI A et al. A Hybrid approach to computer-aided diagnosis in electromyography. *Proceedings of the Annual International conference of the IEEE Engineering in Medicine and Biology Society*. vol. 14. 1992.
- GUIGON E et al. Neural correlates of learning in the prefrontal cortex of the monkey: a predictive model. *Cereb Cortex*, 1995.
- GURNEY JW and SWENSEN SJ. Solitary pulmonary nodules: determining the likelihood of malignancy with neural network analysis. *Radiology*. 1995.
- GUTIERREZ, JE. *Electroagnóstico en PEDROZA A (Ed). Compendio de Neurocirugía*. Bogota, Quebecor Impreandes, 2002.
- GUYTON AC: *Textbook of Medical Physiology*, 8 ed. Philadelphia, W.B. Saunders, 1991.
- HAMAMOTO I. et al. Prediction of the early prognosis of the hepatectomized patient with hepatocellular carcinoma with a neural network. *ComputBiol Med*. 1995.
- HECHT-NIELSEN R. *Neurocomputing*. Addison-Wesley. 1991.

HOYER D; SCHMIDT K and ZWIENER U: Principles and experiences for modeling chaotic attractors of heart rate fluctuations with artificial neural networks. Biomed Tech Berl, 1995.

HUDSON, Donna and COHEN, Maurice. Neural Networks and Artificial Intelligence for Biomedical Engineering. 2000.

INSIGNARES, Víctor Manuel. Redes Neuronales. Principios y aplicaciones en el campo biológico

JAKUS V. The concept and applications of artificial neural networks in Medicine. Bratisl Lek Listy. 1999.

JAND. G et al. Pattern recognition of the electroencephalogram by artificial neural networks. Electroencephalogr Clin Neurophysiol. 1993.

JERVIS BW. The application of Neural Networks to interpret Evoked Potential waveforms. In LISBOA PJG (eds): Artificial Neural networks in Biomedicine. Springer, London, 2001.

JUHOLA M et al. Neural network recognition of otoneurological vertigo diseases. In HORN W (eds): Artificial Intelligence in Medicine, Springer, Berlin, 1999.

KANDEL ER; SCHWARTZ JH and JESSELL TM: Principles of Neural Science, 4 ed. New York, Mc Graw-Hill, 2000.

KANGAS LJ and KELLER, PE. Neurometric assessment of adequacy of intraoperative anesthetic. In LISBOA PJG...(eds): Artificial Neural networks in Biomedicine. Springer, London, 2001.

KOKOL P et al. Some Ideas About Intelligent Medical System Design. Proceedings of the 12th IEEE Symposium on Computer-based Medical Systems CBMS'99 Stamford, CN, June 1999

KOLLES H et al. Automated grading of astrocytomas based on histomorphometric analysis. Classification results of neuronal networks and discriminant analysis. Anal Cell Pathol, 1995.

LAPEER RJ et al. Application of neural networks to the ranking of perinatal variables influencing birth weight. Scand J Clin Lab Invest Suppl.

LUNDBORG G et al. Motor control of tomorrow's artificial hand: Based on the combined use of artificial neural networks and a data glove. ASSH 56th annual Meeting, 2001, paper #1.

MACFARLANE PW. Recent developments in computer analysis of ECGs. Clin Physiol 1992.

MAEDA N; KLYCE SD and SMOLEK MK. Neural network classification of corneal topography. Preliminary demonstration. Invest Ophthalmol Vis Sci, 1995.

MARTIN DEL BRIO, Bonifacio y SANZ MOLINA, Alfredo. Redes Neuronales y Sistemas Difusos (2ª Ed). 2002.

MATWORKS. Documentación de MatLab.

McGONIGAL M. A New Technique for Survival Prediction in Trauma Care Using a Neural Network. Proc. World Conference on Neural Networks. 1994.

MICHALEWIS Z. Genetic Algorithms + Data Structures = Evolution programs, 3 ed. Berlin, Springer, 1996.

MINSKY, Marvin and PAPER, Seymour. Perceptron: An Introduction to Computational Geometry. MIT Press. 1969.

MOLNAR B et al. Application of multivariate, fuzzy set and neural network analysis in quantitative cytological examinations. Anal Cell Pathol. 1993.

MYLREA KC; ORR JA and WESTENSKOW DR: Integration of monitoring intelligent alarm anesthesia: neural networks - can they help? J Clin Monit. 1993.

NIKOVSKY Daniel. Constructing Bayesian Networks for Medical Diagnosis from Incomplete and Partially Correct Statistics. IEEE Transactions on Knowledge and Data Engineering. Vol 12 (4), 2000.

NORDSTROM D. et al. Incidence of diagnosed carpal tunnel syndrome in a general population. Epidemiology 1998..

OH, SJ. Clinical Electromyography: case studies. Baltimore, Lippincott Williams & Wilkins. 1998.

O'LEARY TJ; MIKEL UV and BECKER RL: Computer-assisted image interpretation: use of a neural network to differentiate tubular carcinoma from sclerosing adenosis. Mod Pathol, 1992.

ORTIZ J. et al. Use of artificial neural networks in survival evaluation in heart failure. Arq Bras Cardiol, 1995.

PAPADOPOULOS A; FOTIADIS DI and LIKAS A: A automatic microcalcification detection system based on a hybrid neural network classifier. Artificial Intelligence in Medicine, 2002.

PAPIK K et al: Application of neural networks in medicine-a review. Med Sci Monit 1988.

PATTICHIS CS; SCHIZAS CN and MIDDLETON L. Neural Networks models in EMG diagnosis. IEEE transactions on biomedical Engineering, 1995.

_____ Genetics-Based Machine Learning for the assessment of certain neuro-muscular disorders. IEE Transactions on Neural Networks. Vol 7(2), 1996.

_____ Genetics-Based Machine Learning for the assessment of certain neuro-muscular disorders. IEE Transactions on Neural Networks. Vol 7(2), 1996.

PENNY, Hill and FROST, Davis. Neural Networks in Clinical Medicine.

PEÑA-REYES C.A. and SIPPER M. Evolutionary computation in medicine: an overview. Artificial Intelligence in Medicine. Vol 19 (1), 2000.

_____ A fuzzy-genetic approach to breast cancer diagnosis. Artif Intell Med.

PESONEN, E.; ESKELINEN, M. and JUHOLA, M. Comparison of different neural networks algorithm in the diagnosis of acute appendicitis. Int. J Bio-Med. Comput. 40. 1996.

PETAJAN J: AAEM Minimonograph # 3: Motor unit recruitment. Muscle Nerve. No 14. 1991.

PFURTSCHELLER G; FLOTZINGER D and MATUSCHIK K: Sleep classification in infants based on artificial neural networks. Biomed Tech Berlin. 1992.

PODGOLEREC, V. and KOKOL, P. Medical diagnosis using Genetic Programming. Proceedings of the 12th IEEE Symposium on Computer-based Medical Systems CBMS'99 Stamford, CN, June 1999.

PREHCELT, Lutz. Early Stopping – but when?

RAVDIN PM et al.: A demonstration that breast cancer recurrence can be predicted by neural network analysis. Breast Cancer Res Treat. 1992.

SCHIZAS C et al: Unsupervised learning in computer aided macro Electromyography. In Computer-based Medical Systems, IEEE Computer Society Press, Los Alamitos, CA, USA. 1991.

_____. Neural Networks in Computer Aided Clinical Electromyography, Proceedings of the 13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Ed. By J.H. Nagel and W.M. Smith, 13. 1991. p. 1458-1459

SILVERSTEIN B et al. Claims incidence of work-related disorders of the upper extremities: Washington State, 1987 through 1995. Am J Public Health 1998. p 1827-33.

STEWART JD. Focal Peripheral neuropathies, 3 ed. New York, Lippincott Williams & Wilkins. 1999.

STOTZKA R. et al. A hybrid neural and statistical classifier system for histopathologic grading of prostatic lesions. Anal Quant Cytol Histol, 1995.

SWINGLER, K. Applying neural networks. Academic Press. London. 1996.

SZOLOVITS P. Artificial Intelligence in Medicine. Westview Press, Boulder, CO, 1982.

TONG, K and GRANAT MH. Artificial Neural Network control on functional electrical stimulation assisted gait for persons with spinal cord injury. In Lisboa PJG...(eds): Artificial Neural networks in Biomedicine. Springer, London. 2001.

TOURASSI GD et al. Artificial neural network for diagnosis of acute pulmonary embolism: effect of case and observer selection. Radiology, 1995.

WANG RJ and JABRI MA. Artificial neural network-based channel selection and loudness mapping. Ann Otol Rhinol Laryngol Suppl, 1995.

WILBOURN AJ: AAEM Case Report # 12: Common Peroneal Mononeuropathy at the Fibular Head. Muscle Nerve. 1986.

WINSTON, P. W. Artificial Intelligence. Addison-Wesley, Reading, Mass. 1977.

WU C et al: Protein classification artificial neural system. Protein Sci, 1992.

YAGER RR and ZADEH LA. Fuzzy Sets, Neural Networks, and Soft Computing. New York: Van Nostrand Reinhold, 1994.

ZAKARIA D. et al. Work-related cumulative trauma disorders of the upper extremity: Navigating the epidemiologic literature. Am J Ind Med. 2002.

ZAZULA D; KOROSEC D and SOSTARIC A. Computer-Assisted Decomposition of the Electromyograms . Proceedings of the 11th IEEE Symposium on Computer-based Medical Systems CBMS'98 Lubbock, TX. 1998.

ZIGMOND MJ; BLOOM FE and LANDIS SC, eds: Fundamental Neuroscience. San Diego, Academic Press. 1998.

ZORMAN M et al. Decision Tree's Induction Strategies on a Hard Real World Problem. Proceedings of the 13rd IEEE Symposium on Computer-based Medical Systems CBMS'00, Houston TX. 2000.

REFERENCIAS ELECTRONICAS

INTERNET FAQ CONSORTIUM. What is early stopping? <URL: (http://www.faqs.org/ai-faq/neural-nets/part3/section-5.html)>

KLERFORS, Daniel. Artificial Neural Networks. Saint Louis University. <URL : (http://neurona.tuke.sk/~hudecm/d/Artificial%2520Neuronal%2520Networks.htm)>

PACIFIC NORTHWEST NATIONAL LABORATORY. "Neural Networks: Applications in Medicine and Health". <URL: http://www.emsl.pnl.gov:2080/proj/neuron/neural/bib/medicine.html#>

SIGAMOS Dimitrios. Why neural networks? <URL: http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol1/ds12/article.html>