

UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"

ANÁLISIS DE ESPECTROS SIMULADOS DE TIEMPO DE VIDA DE
POSITRONES MEDIANTE REDES NEURONALES ARTIFICIALES

SONIA MARÍA CORDOBA FREITEZ

Barquisimeto, Febrero de 2005

UNIVERSIDAD CENTROCCIDENTAL "LISANDRO ALVARADO"
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCIÓN: INTELIGENCIA ARTIFICIAL

ANÁLISIS DE ESPECTROS SIMULADOS DE TIEMPO DE VIDA DE
POSITRONES MEDIANTE REDES NEURONALES ARTIFICIALES

Trabajo presentado para optar al grado de
Magíster Scientiarum

Por: SONIA MARÍA CÓRDOBA FREITEZ

Barquisimeto, Febrero de 2005



UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"
DECANATO DE CIENCIAS Y TECNOLOGIA
COORDINACION DE POSTGRADO

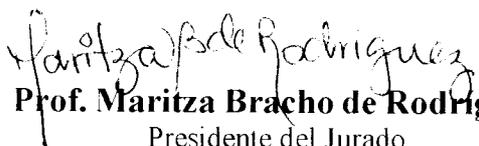


ACTA VEREDICTO TRABAJO DE GRADO

Nosotros, Miembros del Jurado Examinador del Trabajo de Grado titulado: **"Análisis de Espectros Simulados de Tiempo de Vida de Positrones Mediante Redes Neuronales Artificiales"**, presentado por la **Ing. Sonia Córdoba**, titular de la Cédula de Identidad N° 9.619.518 como requisito para optar al grado académico de **MAGISTER SCIENTIARUM EN CIENCIAS DE LA COMPUTACION**, ofrecido por el programa de Maestría en Ciencias de la Computación del Decanato de Ciencias y Tecnología de la Universidad Centroccidental "Lisandro Alvarado", hacemos constar que hoy, cuatro de Marzo del dos mil cinco (04-03-2005) a las cuatro de la tarde (4:00 p.m.) se realizó el examen Público de Defensa de Trabajo de Grado, de acuerdo a lo establecido en la Normativa sobre Trabajos de Grado de la UCLA. Una vez rendido el examen, este Jurado emite el siguiente veredicto: El Trabajo de Grado fue:

*******APROBADO CON MENCIÓN HONORIFICA*******

Dando fe de ello, levantamos la presente acta en la ciudad de Barquisimeto a los cuatro días del mes Marzo del año dos mil cinco.


Prof. Maritza Bracho de Rodriguez
Presidente del Jurado
C.I. N° 4.490.568


Prof. Freddy Torrealba
Jurado Principal
C.I. N° 7.349.103


Prof. Rafael Torrealba
Jurado Principal
C.I. N° 5.890.855

AGRADECIMIENTOS

Le doy gracias a Dios por guiar mis pasos en todo momento de mi vida, llenándola de todo lo bueno que proviene de su fuente inagotable, siendo el centro de mi existencia.

A mis padres, por ser los seres que siempre han apoyado mi formación en todos los aspectos y etapas de mi existir, con amor, esfuerzo y dedicación

A mi familia por estar siempre allí, compartiendo tantos momentos buenos y no tan buenos. Espero que mis logros sean un incentivo para ustedes.

A todos mis amigos que de alguna u otra forma me han acompañado en algún momento durante el recorrido por este camino llamado vida.

A la Ingeniero Maritza Bracho y al Dr. Freddy Torrealba, por orientarme oportunamente durante el desarrollo de este trabajo.

A la Universidad Centroccidental “Lisandro Alvarado” por ser mi casa de estudios y a los profesores que impartieron su conocimiento a lo largo de mi escolaridad en este postgrado

A todos Gracias por su apoyo.

INDICE

	Página
DEDICATORIA.....	iv
AGRADECIMIENTOS.....	v
INDICE GENERAL.....	vi
INDICE DE TABLAS.....	ix
INDICE DE FIGURAS.....	x
RESUMEN.....	xii
INTRODUCCIÓN.....	01
CAPITULO	
I. EL PROBLEMA.....	04
Planteamiento del Problema.....	04
Objetivo General.....	13
Objetivos Específicos.....	13
Justificación e Importancia.....	14
Alcance y Limitaciones.....	15
II. MARCO TEÓRICO.....	17
Antecedentes.....	17
Bases Teóricas.....	20
Caracterización de Defectos.....	20
Aniquilación Positrón/Electrón.....	21
Espectroscopia por Aniquilación de Positrones.....	23
Señal de Fondo o Background.....	25
Redes Neuronales Artificiales.....	27

Paradigma de Aprendizaje	32
Algoritmo de Retropropagación de Error	33
Método de Montecarlo	36
Glosario de Términos Básicos.....	39
III. MARCO METODOLÓGICO.....	42
Tipo de Investigación	42
Fase de Diagnóstico.....	42
Diseño de la Investigación o Procedimiento	42
Técnicas e Instrumentos de Recolección de Información	43
Técnicas de Análisis de los Datos	43
Estudio de Factibilidad	44
Factibilidad Técnica.....	44
Factibilidad Operativa.....	44
Factibilidad Económica.....	45
IV. PROPUESTA DEL ESTUDIO.....	46
Justificación.....	46
Objetivos	47
Objetivo General.....	47
Objetivos Específicos.....	47
Descripción de la Propuesta.....	47
Aspectos Generales.....	47
Procedimiento de Simulación.....	49
Diseño de la Red.....	53
Procedimientos de Entrenamiento y Prueba	58
Procesamiento del primer grupo de experimentos	58
Procesamiento del segundo grupo de experimentos	62
Experimentos Realizados	64
Resultados del primer grupo de experimentos	64

Resultados del segundo grupo de experimentos	69
Prueba de Generalización.....	74
Análisis y Comparación de Resultados.....	76
V. CONCLUSIONES Y RECOMENDACIONES.....	78
Conclusiones.....	78
Recomendaciones.....	80
 REFERENCIAS BIBLIOGRAFICAS.....	 81
ANEXOS.....	84
Carta de Aceptación del Tutor	
Listado del Programa..	

INDICE DE TABLAS

Tabla	Pag
1. Estimación de gastos.....	45
2. Rango de Variación de los parámetros de los espectros.....	49
3. Experimentos Realizados con 1500 canales y una neurona oculta. Primer grupo de experimentos.....	65
4. Experimentos Realizados con 800 canales y dos neuronas ocultas Primer grupo de experimentos.....	67
5. Experimentos Realizados con 500 canales y tres neuronas ocultas Primer grupo de experimentos.....	68
6. Experimentos Realizados con 1500 canales. Segundo grupo de experimentos.....	70
7. Experimentos Realizados con 800 canales. Segundo grupo de experimentos.....	72
8. Experimentos Realizados con 500 canales. Segundo grupo de experimentos.....	73
9. Rango de Variación de los parámetros de los espectros con ruido.....	75
10. Resultados del Procesamiento de espectros con ruido.....	75

INDICE DE FIGURAS

Figura	Pag
1. Proceso de creación de Pares.....	22
2. Aniquilación electrón-positrón.....	22
3. Espectro de tiempo de vida de positrones.....	25
4.- Distribución de probabilidad de Poisson para $\mu=10 / 6$	26
5.- Distribución de probabilidad de Gauss.	27
6. Red Neuronal Artificial Multicapa de conexión hacia delante.	29
7. Funciones de activación más comunes.	30
8. Red Neuronal Artificial con una sola capa oculta.	31
9. Red con conexiones hacia atrás.....	32
10. Ejemplos de Espectros simulados.	51
11. Algoritmo de Simulación de los Espectros.....	52
12.. Red Neuronal Artificial para 1500 canales. Primer grupo de experimentos.....	54
13.. Red Neuronal Artificial para 800 canales. Primer grupo de experimentos.....	54

14.. Red Neuronal Artificial para 500 canales. Primer grupo de experimentos.....	55
15.. Red Neuronal Artificial para 1500 canales. Segundo grupo de experimentos.....	55
16.. Red Neuronal Artificial para 800 canales. Segundo grupo de experimentos.....	56
17.. Red Neuronal Artificial para 500 canales. Segundo grupo de experimentos.....	56
18.. Algoritmo de Entrenamiento. Primer grupo de experimentos.....	59
19.. Algoritmo de Prueba. Primer grupo de experimentos.....	61
20.. Procesamiento de un espectro. Primer grupo de experimentos.....	62
21.. Algoritmo de Entrenamiento. Segundo grupo de experimentos.....	63
22.. Espectro con ruido.....	74

UNIVERSIDAD CENTROCCIDENTAL "LISANDRO ALVARADO"
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCIÓN: INTELIGENCIA ARTIFICIAL

**ANÁLISIS DE ESPECTROS SIMULADOS DE TIEMPO DE VIDA DE
POSITRONES MEDIANTE REDES NEURONALES ARTIFICIALES**

Autor: Ing. Sonia María Córdoba Freitez

Tutores: Ing. Maritza Bracho

Dr. Freddy Torrealba

RESUMEN

Esta investigación tiene como objetivo desarrollar una Red Neuronal Artificial que permita estimar los tiempos de vida medio e intensidades a partir de un espectro simulado de aniquilación de positrones. La técnica de aniquilación de positrones se utiliza para determinar el número promedio de defectos en un sólido y por ende sirve para su caracterización. En estos espectros se generaron 10^6 cuentas en un máximo de 1500 canales (tiempo) usando el método de Montecarlo, con tres componentes y sus respectivas vidas medias ($I_1, I_2, I_3, \tau_1, \tau_2, \tau_3$), los cuales están asociados a tres defectos en el sólido. En trabajos previos, Pazsit y otros (1999) y Avdic y otros (2003), lograron desarrollar una Red de retropropagación con 40 neuronas ocultas para procesar espectros de 1500 canales, obteniendo en el proceso de entrenamiento, un error cuadrático medio del 5% y 10% respectivamente. En el presente trabajo se logra desarrollar la red de retropropagación con un máximo de 10 neuronas ocultas, obteniendo un error menor al 1%, con lo cual se mejora considerablemente el rendimiento de la red, reduciendo la complejidad en su estructura y alcanzando mayor exactitud de acuerdo al mínimo error permitido. Además también se pueden procesar espectros de 800 y 500 canales. Se diseñó la red neuronal multicapa con un máximo de 1500 entradas, 6 salidas y aprendizaje supervisado basado en el algoritmo de retropropagación del error con factor de momento, la cual se entrenó y probó usando los datos simulados. El entrenamiento converge en cada caso usando periodos de tiempo no mayores a 20 minutos.

Palabras clave: espectros de tiempo de vida del positrón, redes neuronales artificiales, espectroscopia de tiempo de vida de aniquilación de positrones, método de Montecarlo.

INTRODUCCIÓN

A través de los años, la ciencia ha estudiado el comportamiento inteligente de los seres humanos con diversos fines, con lo cual se han generado métodos y técnicas que tratan de simular tal comportamiento. Al respecto Rich y Knight (1994), señalan que la Inteligencia Artificial estudia como lograr que las máquinas realicen tareas que, por el momento, son realizadas mejor por los seres humanos.

Aunque todavía se ignoran muchos aspectos de la forma en que el cerebro aprende y procesa la información, se han desarrollado modelos que tratan de mimetizar tales habilidades, denominados Redes neuronales artificiales (RNA). Estas redes poseen características que las hacen potencialmente útiles para la solución de problemas complejos, donde las técnicas computacionales convencionales fallan o son imposibles de utilizar, por lo que tienen actualmente aplicabilidad en diversas áreas.

Específicamente, el uso de RNA representa un método poderoso para tratar con problemas complejos de extracción o estimación de parámetros desde un conjunto de datos con distribución no lineal y que además contienen ruido, tal como es el caso de los espectros de tiempo de vida de aniquilación de positrones, los cuales son generados por la técnica de espectroscopia del mismo nombre (PALS: Positrón Annihilation Lifetime Spectroscopy), aplicada en la caracterización de defectos en materiales.

Los positrones son partículas que al chocar con los electrones del material estudiado sufren un proceso de aniquilación y desaparecen. Cuando estas partículas son atrapadas en áreas defectuosas, el tiempo de vida se incrementa. Un espectro de tiempo de vida de positrones es obtenido como el número de aniquilación de positrones por canal de tiempo, registrando una cantidad suficiente de eventos de aniquilación.

Es decir, que obteniendo los tiempos de vida de los positrones emitidos y sus intensidades (tasas de atrape) a partir de espectros medidos, se pueden caracterizar los defectos en el material estudiado.

Actualmente, la determinación de estos parámetros es alcanzada por procedimientos algorítmicos convencionales, que aunque dan resultados aceptables, obvian características importantes del sistema para simplificar el procesamiento matemático (no linealidad, no estacionariedad, sistemas de orden superior, presencia de fluctuaciones y ruido).

Existen programas computacionales basados en estas técnicas tradicionales, que no siempre dan un equilibrio entre la flexibilidad y la exactitud. En búsqueda de esta meta se ha abierto en el campo científico, una novel línea de investigación para la aplicación de RNA en el análisis de espectros de tiempo de vida de aniquilación de positrones. En los últimos cinco años algunos investigadores, que ya habían aplicado con éxito las RNA a problemas de control en plantas nucleares, han realizado algunos trabajos en esa línea, conociendo la capacidad de las RNA para tratar con problemas complejos.

Es así como surge la necesidad de incrementar la labor investigativa en esta área tan interesante y poco explotada, con el propósito de contribuir en la generación de técnicas, métodos, modelos, herramientas, etc., que mejoren el rendimiento de las ya existentes, con la ayuda de conocimiento interdisciplinario y de vanguardia.

La presente investigación propone diseñar una Red Neuronal Artificial que sirva como un modelo apropiado para analizar espectros de tiempo de vida de positrones, estimando los tiempos de vida e intensidades de dichos espectros, por lo que constituye una importante referencia que puede servir de base para la realización de nuevos estudios, ya sea probando con otros tipos de redes, algoritmos de entrenamiento o extendiendo su aplicación a otros procesos con comportamientos similares.

El presente trabajo está estructurado en cinco capítulos que conforman su desarrollo. El primero se refiere al planteamiento del problema, alcances y

limitaciones, justificación y objetivos de la investigación. El segundo describe el marco teórico, constituido por los antecedentes, la fundamentación teórica y la definición de términos básicos. El tercer capítulo presenta el marco metodológico, con la especificación del tipo de investigación, el diagnóstico y la factibilidad de la propuesta. El cuarto capítulo se refiere a la propuesta del estudio, especificando su justificación y objetivos, además de describir con detalle todos los aspectos relacionados con la misma. Por último, en el quinto capítulo se dan las conclusiones y recomendaciones pertinentes.

DEDICATORIA

A la Esencia Divina, fuente de todo conocimiento

A mi amado hijo Gilbertson, principal motivo de mis luchas

A la familia que me cobijó en su seno con Amor

A mis amigos que me brindan su cariño

AGRADECIMIENTOS

Le doy gracias a Dios por guiar mis pasos en todo momento de mi vida, llenándola de todo lo bueno que proviene de su fuente inagotable, siendo el centro de mi existencia.

A mis padres, por ser los seres que siempre han apoyado mi formación en todos los aspectos y etapas de mi existir, con amor, esfuerzo y dedicación

A mi familia por estar siempre allí, compartiendo tantos momentos buenos y no tan buenos. Espero que mis logros sean un incentivo para ustedes.

A todos mis amigos que de alguna u otra forma me han acompañado en algún momento durante el recorrido por este camino llamado vida.

A la Ingeniero Maritza Bracho y al Dr. Freddy Torrealba, por orientarme oportunamente durante el desarrollo de este trabajo.

A la Universidad Centroccidental “Lisandro Alvarado” por ser mi casa de estudios y a los profesores que impartieron su conocimiento a lo largo de mi escolaridad en este postgrado

A todos Gracias por su apoyo.

UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCIÓN: INTELIGENCIA ARTIFICIAL

**ANÁLISIS DE ESPECTROS SIMULADOS DE TIEMPO DE VIDA DE
POSITRONES MEDIANTE REDES NEURONALES ARTIFICIALES**

Autor: Ing. Sonia María Córdoba Freitez

Tutores: Ing. Maritza Bracho

Dr. Freddy Torrealba

RESUMEN

Esta investigación tiene como objetivo desarrollar una Red Neuronal Artificial que permita estimar los tiempos de vida medio e intensidades a partir de un espectro simulado de aniquilación de positrones. La técnica de aniquilación de positrones se utiliza para determinar el número promedio de defectos en un sólido y por ende sirve para su caracterización. En estos espectros se generaron 10^6 cuentas en un máximo de 1500 canales (tiempo) usando el método de Montecarlo, con tres componentes y sus respectivas vidas medias ($I_1, I_2, I_3, \tau_1, \tau_2, \tau_3$), los cuales están asociados a tres defectos en el sólido. En trabajos previos, Pazsit y otros (1999) y Avdic y otros (2003), lograron desarrollar una Red de retropropagación con 40 neuronas ocultas para procesar espectros de 1500 canales, obteniendo en el proceso de entrenamiento, un error cuadrático medio del 5% y 10% respectivamente,. En el presente trabajo se logra desarrollar la red de retropropagación con un máximo de 10 neuronas ocultas, obteniendo un error menor al 1%, con lo cual se mejora considerablemente el rendimiento de la red, reduciendo la complejidad en su estructura y alcanzando mayor exactitud de acuerdo al mínimo error permitido. Además también se pueden procesar espectros de 800 y 500 canales. Se diseñó la red neuronal multicapa con un máximo de 1500 entradas, 6 salidas y aprendizaje supervisado basado en el algoritmo de retropropagación del error con factor de momento, la cual se entrenó y probó usando los datos simulados. El entrenamiento converge en cada caso usando periodos de tiempo no mayores a 20 minutos.

Palabras clave: espectros de tiempo de vida del positrón, redes neuronales artificiales, espectroscopia de tiempo de vida de aniquilación de positrones, método de Montecarlo.

CAPITULO I

EL PROBLEMA

Planteamiento del Problema

La espectroscopia es una rama de la Física y la Química que se ocupa del estudio de los espectros, es decir, estudia la interacción de la luz o de cualquier radiación electromagnética, con la materia. Este estudio es muy amplio y comprende desde los diversos métodos para la obtención de espectros, su medida y aplicaciones químicas, hasta su interpretación teórica más profunda en relación con la estructura molecular de la materia.

Para la caracterización estructural de materiales, existen actualmente un amplio rango de técnicas que se diferencian entre sí por los principios físicos en los que se fundamenta, los cuales determinan los parámetros a ser medidos. Entre éstas tenemos las técnicas electroquímicas, electromagnéticas, dureza portable, réplica, ultrasónicas, de emisión acústica, termográficas y radioactivas; de las cuales, las cuatro primeras se aplican solo al estudio de materiales metálicos.

Según Shiva y Kishi (2003) las técnicas ultrasónicas son usadas para detectar fenómenos tales como dislocaciones, deformaciones, vacíos y microgrietas; en materiales metálicos, plásticos, compuestos, cerámicas, concreto, etc., a partir de cambios en las características de propagación de una onda elástica con una frecuencia determinada que es introducida dentro del material.

Las técnicas de emisión acústica se basan en detectar las ondas elásticas generadas por microgrietas o deformaciones en un sólido (metales, cerámicas, plásticos, compuestos, concreto, etc.).

Las técnicas termográficas son usadas para detectar fenómenos tales como delaminaciones, deformaciones, vacíos y microgrietas; en metales, cerámicas, compuestos, concreto, etc., usando la distribución de temperatura de la superficie.

Son de especial interés, los métodos de radiación, los cuales se basan en la propiedad de los elementos químicos de absorber la radiación en diferente forma, lo cual es usado para estudiar materiales metálicos y cerámicas. Existen técnicas radiactivas que miden la difracción de rayos X, difracción de neutrones o el efecto Mossboure en rayos gamma, para estudiar microestructuras. Específicamente, el método de análisis espectral Mossbauer, es usado para analizar el proceso de descomposición de fases en acero inempedante de fase dual. También existen dentro de estas técnicas, las que se basan en la medida de la aniquilación de positrones para detectar fenómenos tales como vacíos, deformación, microestructura, microgrietas y precipitaciones.

El uso de una técnica u otra depende del problema de caracterización a resolver y del tipo de material estudiado, pueden usarse también una combinación de éstas. Gebauer y otros (1997) afirman que frecuentemente la información detallada de los defectos no puede ser obtenida usando sólo una técnica en particular. Sin embargo, el uso de las técnicas de radiación es cada vez más común entre los científicos que trabajan en áreas interdisciplinarias tales como Física del Estado Sólido, Ciencia de Materiales, Química Analítica, Metalurgia, Ciencia de Polímeros, entre otras.

Entre éstas, una de las técnicas nucleares de vanguardia es la espectroscopia por aniquilación de positrones para el estudio y caracterización de la estructura de defectos en materiales, mediante el análisis del comportamiento de los positrones en la materia condensada. Esta presenta una serie de ventajas derivadas de la naturaleza de estas partículas.

Según PDG (2000), el positrón es una partícula elemental con propiedades idénticas al electrón pero con carga positiva. Cuando un fotón de alta energía choca con un átomo, entonces se produce una transformación generando un electrón y un positrón, proceso denominado “creación de pares”. El caso contrario es cuando un positrón se encuentra con un electrón desapareciendo ambos para formar un fotón de alta energía, proceso conocido como “aniquilación”. Los positrones que penetran en

un medio sufren inicialmente un proceso de termalización y una posterior aniquilación con los electrones del material.

En Nanopos (2004) se menciona que una característica muy significativa de los positrones, es que se puede seguir la historia de cada positrón después que ha termalizado pero antes de ser aniquilado. Es por ello que la influencia de las propiedades del material (campos internos, distribuciones de impurezas y defectos, cambios espaciales en la composición) en la difusión de positrones, puede ser medida en principio.

La espectroscopia de tiempos de vida de aniquilación de positrones (PALS: Positrón Annihilation Lifetime Spectroscopy), es una técnica que consiste, básicamente, en contar con una fuente radioactiva que emite positrones. Estas partículas ingresan en el material que se desea estudiar y se produce una aniquilación electrón/positrón y una emisión de rayos gamma. Dado que tienen carga positiva, no producen alteraciones en el material y de ese modo permiten estudiar defectos o composiciones químicas cercanas al sitio donde se produce la aniquilación.

De acuerdo a la variante del sistema que se emplee, se puede obtener información diversa acerca de los procesos que se registran en los materiales que son objeto del estudio, a escala subnanométrica, es decir, de mil millonésimas de milímetro.

Según MLUH (2004), cuando los positrones son absorbidos en áreas del material con defectos, el tiempo de vida del positrón se incrementa con respecto a las áreas libres de defectos. Es decir, si existe un defecto, aparece un componente de tiempo de vida más largo cuya fuerza o intensidad esta directamente relacionada con la concentración del defecto. Tanto el tipo como la concentración del defecto pueden ser obtenidos independientemente a través de una medida sencilla, siendo esta la mayor ventaja con respecto a otras técnicas.

La técnica PALS se basa en el hecho de que los positrones que permanecen más tiempo sin aniquilarse, y por tanto, tienen un tiempo de vida apreciable, son aquellos que están atrapados en lugares con densidad electrónica menor a la media, tales como

vacantes o huecos. En este caso, el tiempo de vida se incrementa al aumentar el tamaño de estos huecos. La medición del tiempo de vida se basa en determinar el intervalo de tiempo entre dos cuantos gamma, uno que se emite simultáneamente con la emisión del positrón en la creación de pares y otro que surge como resultado de la aniquilación de los positrones con los electrones del material.

El método de aniquilación de positrones es sensitivo a defectos microestructurales. Se puede encontrar defectos pequeñísimos en muestras de materiales y brinda la oportunidad de derivar distintas propiedades y seguir la evolución de los defectos. Estos estudios permiten ver en el nivel nanoestructural detalles imposibles de divisar con otras técnicas.

Según Morales (2004), los positrones enviados hacia el material son atrapados fácilmente por los defectos y se aniquilan con los electrones que los rodean con un tiempo de vida que es proporcional a la densidad de carga electrónica, este tiempo de vida se mide y es una huella digital del defecto, desafortunadamente muchos defectos pueden presentar tiempos de vida similares lo que hace difícil su identificación. Sin embargo la aniquilación de positrones también permite medir la energía llevada por los rayos gamma producidos en la aniquilación y esta energía contiene información de la colisión electrón-positrón, la cual ocurre entre un positrón atrapado sin energía cinética y un electrón que tiene una velocidad debida el estado cuántico de energía en que se encuentra.

Esta velocidad produce un ensanchamiento de la energía del rayo producto de la desintegración, la medida de ese ensanchamiento, trae información sobre el medio químico del defecto o la distribución atómica a su alrededor, y esta información adicional permite una mejor caracterización del defecto.

Avdic y otros (2003) afirman que este método es muy sensitivo a defectos de volumen abiertos tales como monovacancias y grupos de vacancias más grandes. Todas las técnicas experimentales de positrón proveen la tasa de atrape que es proporcional a la concentración del defecto. La mayor ventaja de la técnica es la habilidad de controlar la implantación de positrones variando la energía del rayo.

Un espectro de tiempo de vida de positrones es obtenido como el número de aniquilación de positrones por canal de tiempo. Para obtener un espectro de tiempo de vida completo deben ser registrados una cantidad suficiente de eventos de aniquilación.

Pazsit y otros (1999) afirman que en la espectroscopia de tiempo de vida de aniquilación de positrones, la cantidad medida es un espectro de tiempos de vida que consiste de una serie o suma de componentes exponenciales de la forma: $I_i \cdot \exp(-t/\tau_i)$, donde I_i es la intensidad (tasa de atrape) y τ_i es el tiempo de vida medio. La señal medida es convolucionada con la función de resolución del detector, usualmente una gaussiana. También es añadido una medida del fondo radiactivo.

Según LFN (2003), la resolución de un espectómetro es una medida de su capacidad para resolver dos picos de aniquilación que se hayan muy próximos. Por su parte, el fondo radiactivo es la tasa de cuentas que se registran en ausencia de fuentes de radiación, debido a la radiación del ambiente y la radiación cósmica, entre otras.

En un espectro, el número de componentes ($k + 1$) esta relacionado a la cantidad de defectos (k), el tiempo de vida esta relacionado a la presencia y tipo de defecto, y la intensidad da una medida de la concentración del defecto. Es decir, que obteniendo estos parámetros a partir de los espectros medidos se pueden caracterizar los defectos en el material estudiado.

Según Joran y otros (1998), los modelos de sumas exponenciales son usados frecuentemente en diversas áreas, tales como difusión de componentes químicos, series de tiempo en medicina, economía, tecnología y ciencias físicas. De allí la importancia de encontrar buenos métodos para la estimación de parámetros en sumas exponenciales.

La determinación de los tiempos de vida medio y sus distribuciones así como las amplitudes o intensidades correspondientes a partir de espectros de tiempo de vida de positrones es actualmente alcanzada por procedimientos algorítmicos que se basan en algún principio de minimización estadística.

Estos métodos clásicos de procesamiento estadístico, en su búsqueda de hacer menos complejo todo el procesamiento matemático, se basan en tres suposiciones básicas: linealidad, sistema estacionario y sistemas con estadísticas de segundo orden, enfatizando en los de tipo Gaussiano. Sin embargo, la mayoría de las señales físicas con las que se trata son generadas por procesos dinámicos que son simultáneamente no lineales y no estacionarios; mas específicamente, la distribución medida en espectroscopia de aniquilación de positrones, es una función no lineal de los parámetros involucrados y es frecuente que contenga fluctuaciones y errores de medida, por lo que se hace necesario emplear métodos mas adecuados a este tipo de problemas.

Hasta ahora, se han realizado variadas investigaciones sobre espectros de tiempo de vida en diversos materiales usando métodos algorítmicos clásicos. AVDIC y otros (2003) afirman que varios programas de computadora fueron desarrollados basándose en el procedimiento de ajuste Gauss-Newton usado para el análisis convencional de datos. En este método, se tiene que especificar el número de componentes espectrales antes del análisis. El valor para la exactitud del tiempo de resolución es también requerido para obtener una solución confiable. El análisis con el programa PATFIT da como resultado el promedio del tiempo de vida del defecto pero no la distribución de los tiempos de vida.

El programa CONTIN, basado en la técnica de inversión Laplace y el programa MELT basado en el principio de máxima entropía y el teorema de Bayes, pueden dar una distribución continua de los tiempos de vida o tasas de aniquilación. Este último tiene la ventaja que el número de componentes espectrales y el tiempo de resolución del instrumento no son requeridos como parámetros de entrada. Las aplicaciones del análisis de tiempo de vida reportados en la literatura muestran que la resolución de CONTIN es más baja que con MELT. En algunos casos, como pruebas en semiconductores, CONTIN no puede separar tiempos de vida muy cercanos.

Por otra parte, en esta misma área, se han venido realizando una serie de investigaciones aplicadas con éxito al control de plantas nucleares, las cuales usan

redes neuronales artificiales. De allí se deriva el interés de estos investigadores, de utilizar esta técnica de inteligencia artificial tan apropiada para el estudio de espectros de tiempos de vida de positrones.

Entre los trabajos aplicados al control de plantas nucleares, se destaca el de Pazsit y otros (1998), que aplican redes neuronales artificiales para determinar la posición de un eje de rotación o elevación de la barra de control de un reactor nuclear; en este trabajo se origina la red que posteriormente es usada por Pazsit y otros (1999) para el análisis de espectros de tiempos de vida de positrones. Para describir el rendimiento de la red, se calculó el error cuadrático medio (RMS) en el entrenamiento y la desviación estándar relativa para los ejemplos de prueba, la cual resultó ser más grande de lo que puede ser obtenido con algoritmos tradicionales, esto es consecuencia del gran error RMS, que fue establecido para reducir el tiempo de entrenamiento. En la presente investigación se espera mejorar estos resultados.

La creciente aplicación de las redes neuronales artificiales (RNA) en muchas áreas, se debe a que éstas poseen características que las hacen potencialmente útiles para la solución de problemas complejos, difíciles de formular matemáticamente, donde las técnicas computacionales convencionales fallan, ya que éstas aprenden o se entrenan en la tarea que ha de computarse.

Según Rich y Knight (1994), las RNA son modelos matemáticos inspirados en el funcionamiento de neuronas biológicas, que se adaptan para funcionar en computadoras convencionales. Están constituidas por un conjunto de neuronas unidas entre sí, que trabajan con una serie de patrones de entrada o ejemplos los cuales sufren una serie de transformaciones a través de toda la red para obtener una salida que esta relacionada con lo que se debe aprender.

Para lograr este aprendizaje, la estructura interna de la red pueden sufrir variaciones a medida que las neuronas van procesando los patrones de entrada, hasta que el entrenamiento en la tarea requerida sea satisfactorio. Una vez aprendida, la red debe ser capaz de dar respuestas adecuadas a patrones de entrada nunca vistos. Por el

momento, se manejan ciertos tipos de redes, para los cuales queda establecido una arquitectura y una forma de aprendizaje.

Estas redes presentan un número importante de propiedades que hacen conveniente su uso en variadas aplicaciones por lo que actualmente son ampliamente utilizadas en diversas áreas. Específicamente, el uso de una Red Neuronal Artificial representa un método poderoso para tratar con problemas complejos de extracción de parámetros a partir de datos con distribución no lineal y que además contienen ruido, tal como es el caso de los espectros de tiempo de vida del positrón planteado anteriormente.

Hasta el momento, han sido pocas las investigaciones que aplican las redes neuronales artificiales al estudio de espectros de tiempos de vida de positrones. Es así como se convierte en un área de investigación interesante y poco explotada, por lo que en este proyecto se propone diseñar una Red Neuronal Artificial que sirva como un modelo apropiado para analizar espectros de tiempo de vida de positrones, estimando los tiempos de vida e intensidades de dichos espectros, como alternativa novedosa que puede representar un avance significativo en el estudio y caracterización de defectos en materiales.

Se diseñará un modelo de red perceptrónica multicapa con entrenamiento basado en retropropagación del error, usando como patrones de entrada los espectros de tiempos de vida de positrones (número de cuentas o aniquilaciones por canal de tiempo), por lo que la capa de entrada tendrá tantas neuronas como canales tenga el espectro (en este caso un máximo de 1500). Las señales de salida a ser aprendidas o relacionadas por la red son los tiempos de vida e intensidades de cada componente del espectro, por lo que el número de neuronas de la capa de salida será el doble del número de componentes del espectro (en nuestro caso 6). Tal como es usada en la mayoría de los experimentos con redes neuronales artificiales, se tomará inicialmente el estándar de 70% de los ejemplos para el entrenamiento y el 30% restante para la validación.

Debido a que no se dispone de datos experimentales, los ejemplos de entrenamiento serán generados por un programa de simulación utilizando el método de Montecarlo, ya que los procesos involucrados son de naturaleza aleatoria.

Según Montesinos (2001), cuando un haz de fotones interacciona con un medio material normalmente el número de procesos que tienen lugar es tan elevado que es imposible predecir los resultados de tal interacción. Esto conduce a la necesidad de realizar simulaciones numéricas del proceso de interacción. La forma más precisa de realizar estas simulaciones es utilizar el llamado Método de Montecarlo que permite resolver problemas matemáticos mediante la simulación de variables aleatorias. Este método proporciona toda la información posible a partir de la cual se pueden obtener valores medios, fluctuaciones, etc. El Método de Montecarlo para la simulación de procesos físicos es perfectamente conocido desde hace mucho tiempo, pero dado que es necesario simular un gran número de casos para conseguir resultados con significancia estadística, su uso es efectivo cuando se utiliza un computador. Su potencia se muestra en aquellos problemas que no pueden tratarse analíticamente o cuyo tratamiento es demasiado complicado.

El problema de investigación radica entonces en determinar la arquitectura del modelo de red neuronal artificial más adecuada para el análisis de espectros de vida de aniquilación de positrones simulados, específicamente para la estimación de los tiempos de vida e intensidades, verificando así su eficacia en esta tarea, lo cual constituye un aporte tanto documental como práctico a esta novedosa línea de investigación, que puede servir como base para posteriores investigaciones, incentivando la experimentación con otros modelos de red neuronal.

Objetivos

Objetivo General

Desarrollar un modelo de Red Neuronal Artificial para el análisis de espectros simulados de tiempos de vida de positrones, que permita estimar los tiempos de vida e intensidades de dichos espectros.

Objetivos Específicos

- Desarrollar un programa de simulación para generar los espectros de tiempos de vida de positrón con diferentes intensidades y tiempos de vida con el fin de cubrir las regiones de posibles valores de los mismos.
- Diseñar la Red Neuronal Artificial para el análisis de espectros simulados de tiempos de vida de positrones para establecer la arquitectura y regla de aprendizaje de la misma.
- Diseñar experimentos variando los parámetros de diseño del modelo de red y el tamaño de los conjuntos de entrenamientos, que permitan a través de los procesos de entrenamiento y validación, seleccionar el modelo de red neuronal que registre mejor rendimiento.
- Entrenar los modelos de red neuronal diseñados con parte del conjunto de entrenamiento, de acuerdo al diseño de cada experimento.
- Validar los modelos de red entrenada con ejemplos nunca procesados por la red, de acuerdo al diseño de cada experimento.
- Realizar un estudio comparativo de los resultados de cada experimento, que permitan seleccionar el modelo de red neuronal más adecuado al problema en estudio.

Justificación e Importancia

Las redes neuronales han recibido un creciente interés en diferentes áreas de la ciencia e ingeniería en problemas de reconocimiento de patrones, detección de fallas, control de reactores, y en general en problemas de solución de tareas de inversión. Las tareas de inversión consisten en la determinación de parámetros de una distribución desde valores medidos de la misma distribución. Según Pázsit y otros (1999), es muy raro que una tarea inversa pueda ser resuelta analíticamente de una manera exacta invirtiendo la relación. Aún si la inversión es única, se necesitan usar algunos criterios extras para seleccionar la verdadera solución. Además, una distribución medida frecuentemente contiene fluctuaciones y errores de medida, así que la inversión también necesita ser optimizada de alguna manera.

Las Redes Neuronales Artificiales (RNA) representan una técnica avanzada de Inteligencia Artificial muy poderosa para tratar con tales problemas. Fantoni y otros (1995) afirman que éstas encuentran su aplicación en un amplio rango de problemas, especialmente donde los fenómenos no lineales tienen que ser considerados. La razón de la difusión de estas redes es que pueden aproximar con una exactitud arbitraria algún mapeo desde un espacio de entradas a un espacio de salida

Una vez entrenadas, las RNA pueden ejecutar la inversión en una forma muy rápida, transparente y con un mínimo de esfuerzo computacional, lo cual puede ser una ventaja en el caso que muchos espectros estén siendo procesados.

Su fortaleza radica en que requieren reglas de entrenamiento que pueden ser ejecutadas sin necesidad de establecer una relación explícitamente formulada entre las variables de entrada y salida. Particularmente, son muy efectivas en la extracción de parámetros desde una distribución que es una función no lineal de los parámetros buscados, y que también contiene fluctuaciones; ya que son tolerantes al ruido en el medio ambiente y en sus componentes, además pueden determinar la esencia o características principales de un conjunto de datos. Para alcanzar esto, la red necesita

ser entrenada con un gran número de ejemplos, es decir, realizaciones de la distribución donde los parámetros buscados son también conocidos.

Entre la variedad de tipos de redes existentes, una de las más populares es el Perceptron multicapa con entrenamiento basado en retropropagación de error, ya que son capaces de aproximar cualquier función continua o modelar problemas. La ventaja de este tipo de red es que es relativamente simple de realizar y aún así es muy poderosa.

La presente investigación contribuirá aportando un modelo de red neuronal artificial, para el análisis de espectros simulados de tiempos de vida de aniquilación de positrones, que permita estimar los tiempos de vida e intensidades de dichos espectros, los cuales son parámetros físicos claves en la caracterización de defectos en materiales.

Este modelo servirá para estimar los tiempos de vida e intensidades de espectros de tiempos de vida que sean introducidos como entrada a la red. Una vez validado, dicho modelo puede fácilmente ser aplicado al procesamiento de datos experimentales. De igual forma, esta técnica de red neuronal propuesta puede ser posteriormente extendida en su aplicación al modelado de otros procesos con comportamientos parecidos

Este trabajo será una importante referencia que puede servir de base para consolidar conocimientos sobre la espectroscopia por aniquilación de positrones, métodos de simulación y redes neuronales artificiales. Así mismo, puede ayudar a la realización de nuevos estudios aplicados a la caracterización de defectos en materiales utilizando redes neuronales artificiales, ya sea probando con otros tipos de redes, algoritmos de entrenamiento u otra técnica de inteligencia artificial.

Alcance y Limitaciones

Para el entrenamiento de los modelos de red, se usará una distribución de datos de espectros de tiempo de vida de aniquilación de positrones, simulados con un máximo de tres componentes, los cuales tendrán diferentes intensidades y tiempos de vida de tal forma que cubran las regiones de posibles valores de los mismos.

Tal como es realizado en investigaciones anteriores, cada espectro registrará alrededor de 10^6 eventos de aniquilación, distribuidos en un máximo de 1500 canales de tiempo, es decir que son 1500 números enteros como máximo para cada espectro, los cuales corresponden al número de cuentas (eventos de aniquilación) registrados por cada canal de tiempo.

Los espectros serán generados por un programa de simulación, ya que no se disponen de datos reales. Igual que en investigaciones anteriores, se usará el método de Montecarlo, debido a que los procesos a simular son de naturaleza aleatoria.

Los espectros generados no serán sometidos a complejos preprocesamientos para su simplificación, ya que se busca que el modelo de red tome en cuenta todas sus características. Sin embargo, se reducirá la escala de los valores de los mismos mediante escalamiento entre los valores 0 y 1, ya que las salidas de la red neuronal arrojará, también están en ese rango.

Se diseñará un modelo de red neuronal artificial perceptrónica multicapa con entrenamiento supervisado basado en retropropagación del error, para estimar un máximo de tres tiempos de vida e intensidades. Estas salidas sirven de base para la caracterización de los defectos, sin embargo, éstas solo serán reportadas, no se someterán a procesamientos posteriores que reporten las propiedades de los mismos.

Para entrenar y probar la red, inicialmente se usaran 575 y 50 espectros respectivamente, tal como fue realizado en el trabajo de Pázsit y otros (1999), sin embargo, de acuerdo al desempeño de la misma, se podrían generar experimentos para entrenar con un número diferente de ejemplos.

Se espera que el modelo de red neuronal diseñado permita estimar los tiempos de vida e intensidades a partir de espectros de tiempos de vida de positrones, con un mínimo de error, que mejore los resultados de investigaciones anteriores.

El desarrollo de este trabajo de investigación podrá usarse para realizar nuevos estudios sobre estimación de parámetros desde espectros de aniquilación de tiempo de vida de positrones, con miras a mejorar el diseño de la red neuronal, los tiempos de entrenamiento y los errores en la estimación; lo cual llevaría a un óptimo balance entre sencillez, velocidad y exactitud de los resultados.

CAPITULO II

MARCO TEORICO

Antecedentes

El análisis de espectros de tiempo de vida de positrones para estimar los tiempos de vida e intensidades es actualmente realizado en su gran mayoría usando procedimientos algorítmicos basados en técnicas de optimización convencionales.

Por otra parte, en el área de física nuclear, se han realizado investigaciones que aplican con éxito las redes neuronales artificiales (RNA) para el control de plantas nucleares, demostrando su capacidad de aproximadores universales y su aplicabilidad a la estimación de parámetros de una distribución de datos medida. A partir de allí, se empezaron a desarrollar algunas investigaciones que aplican RNA al estudio de espectros de tiempos de vida de positrones.

A continuación se resumen algunos de los trabajos aplicados al control de plantas nucleares, y las dos investigaciones que aplican RNA al análisis de espectros de tiempos de vida de aniquilación de positrones, los cuales sirven de base para la realización del presente proyecto.

Pázsit y otros (1998), presentaron un trabajo titulado: **“Determination of PWR control rod position by core physics and neural network methods”**. En este artículo los autores aplican una red neuronal artificial (RNA) para determinar la posición de un eje de rotación o elevación de la barra de control en un reactor nuclear. En estas experiencias primarias encontraron que las redes neuronales eran muy útiles para resolver problemas de inversión o para revelar complejas e implícitas relaciones entre las variables de un problema. Particularmente, los autores encontraron que midiendo la rata del flujo de energía o de partículas a través de una superficie dada, podían determinar la posición de la barra del rotor, entrenando de manera apropiada a una RNA para aprender las relaciones necesarias para hacer la inversión, con

esfuerzos mínimos de cálculo. La red usada en este artículo es una red de tres niveles con alimentación hacia delante y retropropagación del error.

Hines y otros (1999), publicaron un trabajo denominado: **“Regularization Methods for Inferential Sensing in Nuclear Power Plants”**, donde comparan las redes neuronales artificiales con otros métodos. Investigan como la colinearidad puede afectar adversamente las técnicas de predicción de valores de variables de control y de seguridad en plantas nucleares a través del uso o aplicación de variables de la planta correlacionadas. Los métodos analizados son redes neuronales, lógica difusa y técnicas estadísticas como la regresión lineal, regresión de Ridge, mínimos cuadrados parciales, entre otros. La comparación se realiza usando datos de la Planta nuclear de Florida Power Corporation's Crystal River para predecir el rumbo en un sensor de flujo de agua.

Pázsit y otros (1999) realizaron una investigación denominada: **“Unfolding positron lifetime spectra with artificial neural networks”**, donde desarrollaron un método para extraer tiempos de vida media y amplitudes así como también distribuciones de tiempo de vida, desde espectros de tiempos de vida de positrones, basado en el uso de una Red Neuronal Artificial de retropropagación, que fue entrenada para estimar tres o cuatro tiempos de vida discretos. Se usaron datos simulados, generados por un algoritmo donde el rango de los tiempos de vida e intensidades fueron variados alrededor de los ciertos valores medios. y el número total de cuentas fue de 10 millones.

Se implementó una red de alimentación adelantada de tres capas con retropropagación del error donde una entrada consiste del espectro con 1500 canales. Existen seis nodos de salida igual al número de parámetros a determinar (3 intensidades y 3 tiempos de vida). La capa oculta fue de 40 nodos. En el entrenamiento se usaron, 575 ejemplos y 50 para probar el rendimiento de la red. El uso de este tipo de RNA se origina en trabajos anteriores de Pázsit y otros autores, quienes aplican las RNA en solución de problemas de control en plantas nucleares.

El límite permitido para el error cuadrático medio fue establecido en 5% (una precisión menos exacta que los algoritmos tradicionales usuales). Igualmente la desviación estándar relativa, en los tiempos de vida, durante la fase de prueba, es más grande de lo que puede ser obtenido de algoritmos tradicionales, esto es como consecuencia del gran error en el entrenamiento y demuestra que se requiere el diseño cuidadoso de la red, largos tiempos de entrenamiento y un método adecuado de simulación.

Avdic y otros (2003), presentaron su trabajo **“Analysis of the experimental positron lifetime spectra by neural networks”**. Esta publicación trata del análisis de espectros de tiempo de vida de positrón experimental en materiales de polímeros usando varios algoritmos de redes neuronales. En este trabajo, la aplicabilidad de las redes neuronales al análisis de espectros de positrón experimentales ha sido verificado en el caso de espectros de materiales de polímeros con tres componentes.

En este artículo, el autor hace referencia a varias investigaciones donde se han usado programas de análisis convencional de espectros de tiempo de vida de positrones (PATFIT, CONTIN, MELT), realizando una comparación entre estos.

Se demostró que la red neuronal de retropropagación puede determinar los parámetros espectrales con alta exactitud y ejecutar la descomposición de los tiempos de vida las cuales difieren por 10% o más, sin embargo, no ha sido apropiada para la identificación al mismo tiempo, del número de componentes espectrales.

Por ello, se diseñó un módulo de red neuronal separado basado en mapas auto-organizativos. También fueron probados algoritmos de cuantización del vector de aprendizaje, con los cuales se obtuvo mejor rendimiento y confiabilidad.

De la revisión de los trabajos anteriormente mencionados se desprende la necesidad de realizar investigaciones más profundas en el área de estudio, experimentando con diferentes técnicas alternativas tales como las redes neuronales artificiales, de las cuales existe una clasificación variada de acuerdo a diferentes criterios de diseño. De esta forma se podría establecer la efectividad de uno u otro

diseño escogido, comparando con resultados de investigaciones similares y concluyendo de acuerdo a dicha comparación.

En la presente investigación se pretende experimentar con un diseño de red neuronal artificial de tipo retropropagación, aplicada a datos simulados de espectros de tiempos de vida de aniquilación de positrones para comparar y evaluar el rendimiento de dicho diseño. En esta comparación se tomará como punto de referencia la investigación de Pázsit y otros (1999), igualmente como la de Avdic y otros (2003). Además, en conjunto con los restantes trabajos mencionados, sirven de apoyo para el desarrollo teórico en que se basa el trabajo desarrollado.

Bases Teóricas

Caracterización de Defectos

Para la caracterización estructural de materiales, existen actualmente un amplio rango de técnicas que se diferencian entre sí por los principios físicos en los que se fundamenta, los cuales determinan los parámetros a ser medidos. Entre estas tenemos las técnicas electroquímicas, electromagnéticas, dureza portable, réplica, ultrasónicas, de emisión acústica, termográficas y radioactivas; de las cuales, las cuatro primeras se aplican solo al estudio de materiales metálicos.

Las técnicas ultrasónicas son usadas para detectar fenómenos tales como dislocaciones, deformaciones, vacíos y microgrietas; en materiales metálicos, plásticos, compuestos, cerámicas, concreto, etc., a partir de cambios en las características de propagación de una onda elástica con una frecuencia determinada que es introducida dentro del material.

Las técnicas de emisión acústica se basan en detectar las ondas elásticas generadas por microgrietas o deformaciones en un sólido (metales, cerámicas, plásticos, compuestos, concreto, etc.).

Las técnicas termográficas son usadas para detectar fenómenos tales como delaminaciones, deformaciones, vacíos y microgrietas; en metales, cerámicas, compuestos, concreto, etc., usando la distribución de temperatura de la superficie.

Las técnicas de radiación se basan en la propiedad de los elementos químicos de absorber la radiación en diferente forma, lo cual es usado para estudiar materiales metálicos y cerámicas. Existen técnicas radiactivas que miden la difracción de rayos X, difracción de neutrones o el efecto Mossboure en rayos gamma, para estudiar microestructuras. Específicamente, el método de análisis espectral Mossbauer, es usado para analizar el proceso de descomposición de fases en acero inempedante de fase dual. También existen dentro de estas técnicas, las que se basan en la medida de la aniquilación de positrones para detectar fenómenos tales como vacíos, deformación, microestructura, microgrietas y precipitaciones.

Aniquilación Positrón/Electrón

Según PDG (2000), un positrón es una partícula elemental de masa y espín (momento angular intrínseco de una partícula subatómica) iguales a los del electrón y cuya carga eléctrica es también igual a la de éste pero de signo contrario, es decir, positiva. El positrón, por lo tanto, se considera como antipartícula del electrón. Prevista su existencia por Dirac en 1930, Anderson consiguió detectarlo en los rayos cósmicos en el año 1932. Según Eisberg y otros (2002), esta radiación consiste en un flujo de fotones con gran contenido energético y partículas cargadas que inciden sobre la tierra provenientes de fuentes extraterrestres.

El positrón se crea por transformación de un fotón en un par electrón-positrón y a pesar de su relativa estabilidad, se aniquila por interacción con los electrones. Cuando un fotón de alta energía choca con un átomo, entonces se produce un electrón (e^-) y un positrón (e^+). Esto se denomina creación de pares (Ver Figura 1). En este proceso, tanto el electrón como el positrón se crean a partir de energía y ninguno de ellos es parte de un residuo nuclear.

La energía se conserva en este proceso; es decir, la energía cinética del fotón incidente se convierte en las masas y energía cinéticas del par. La carga también se conserva, dado que cada vez que se crea un positrón, positivamente cargado; un electrón, de carga negativa, se crea con él.

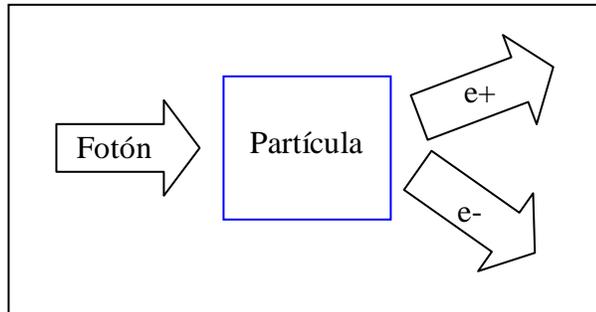


Figura 1. Proceso de creación de Pares. Geocities (2004)

Si el positrón creado se encuentra con un electrón, el electrón y el positrón desaparecen y en su lugar encontramos fotones de alta energía. Este proceso se denomina aniquilación y representa la situación inversa a la creación (Ver Figura 2).

Cuando una partícula y su antipartícula se encuentran, su energía total, incluyendo la energía de sus masas, se convierte en fotones y las partículas originales dejan de existir.

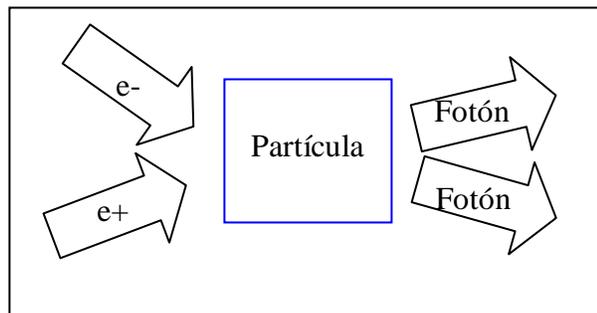


Figura 2. Aniquilación electrón-positrón. Geocities (2004)

La medición del tiempo de vida del positrón se basa en determinar el intervalo de tiempo entre dos cuantos gamma, uno que se emite simultáneamente con la emisión del positrón, y otro que surge como resultado de la aniquilación de los positrones con los electrones del material en estudio.

Según MLUH (2004), el tiempo de vida es una función de la densidad del electrón en el sitio de aniquilación..

Espectroscopia por Aniquilación de Positrones

La espectroscopia es una rama de la Física y la Química que estudia los espectros, es decir, la interacción de la luz o de cualquier radiación electromagnética, con la materia. Específicamente, para el estudio y caracterización de la estructura de defectos en materiales existen actualmente varias técnicas nucleares entre las que se encuentra la espectroscopia de aniquilación de positrones.

En dicha técnica de espectroscopia se analiza el comportamiento de los positrones cuando se aniquilan con los electrones de la materia condensada.

La técnica de espectroscopia de aniquilación de tiempos de vida de positrones (PALS). Consiste, en contar con una fuente radioactiva que emite positrones. Estos ingresan en el material estudiado y son absorbidos en áreas del material con defectos en un tiempo de vida mayor al normal, permitiendo estudiar los defectos cercanos al sitio donde se produce la aniquilación. La intensidad de este componente esta directamente relacionada con la concentración del defecto.

La técnica PALS se basa en el hecho de que los positrones que permanecen más tiempo sin aniquilarse, y por tanto, tienen un tiempo de vida apreciable, son aquellos que están atrapados en lugares con densidad electrónica menor a la media, tales como vacantes o huecos. En este caso, el tiempo de vida se incrementa al aumentar el tamaño de estos huecos.

El tiempo de vida se mide determinando el intervalo entre un cuanto gamma que se emite simultáneamente con la emisión del positrón y otro cuanto gamma que surge en la aniquilación de los positrones con los electrones del material.

Un espectro de tiempos de vida consiste de tres o más exponenciales donde están reflejados una cantidad específica de eventos de aniquilación.

Pázsit y otros (1999) señalan que un componente de un espectro medido $y(t)$ puede ser escrito en la forma:

$$y(t) = N_0 R(t) * \int_0^{\infty} \frac{f(\tau)}{\tau} \exp\left(-\frac{t}{\tau}\right) d\tau + B \quad (1)$$

Donde $R(t)$ es la función de resolución del detector, asumida usualmente como Gaussiana, con la cual la señal medida es convolucionada, lo que se indica por el asterisco. B es una señal de fondo aleatoria y $f(\tau)$ es el perfil de tiempo de vida del positrón.

Si hay solo líneas discretas en el perfil de tiempos de vida, este puede ser representado como:

$$f(\tau) = \sum_{\tau=1}^3 I_i \delta(\tau - \tau_i) \quad (2)$$

Donde I son las intensidades y τ los tiempos de vida. En casos mas generales, algunos de los tres tiempos de vida pueden exhibir una variación alrededor del valor medio, por ejemplo, una distribución Gaussiana estrecha en vez de la forma dada en la ecuación 2, es decir:

$$f(\tau) = \sum_{i=1}^3 \frac{I_i}{\sqrt{2\pi\sigma_i}} \exp\left[-\frac{(\tau - \tau_i)^2}{2\sigma_i^2}\right] \quad (3)$$

En la siguiente figura se puede ver un espectro de tiempo de vida de aniquilación de positrones.

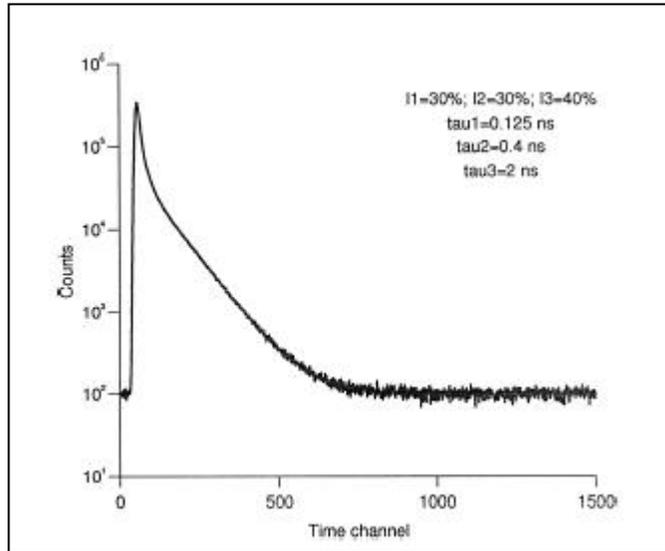


Figura 3. Espectro de tiempo de vida de positrones. Paszit y otros (1998)

Señal de Fondo o Background

El fondo radiactivo (“background”) es la tasa de cuentas que se registra en ausencia de fuentes de radiación y se debe fundamentalmente a:

- La radiación γ del ambiente y radiación cósmica
- Las partículas β de impurezas del material del detector
- Las descargas espúreas

El número de cuentas que se registra en un detector debido al background sigue una distribución de Poisson, que viene dada por la expresión:

$$P_p(x, \mu) = \frac{\mu^x}{x!} e^{-\mu} \quad (4)$$

en la que $P_p(x, \mu)$ da la probabilidad de observar x sucesos en un intervalo de tiempo t , si μ es el número medio de sucesos observados en el mismo intervalo de tiempo.

En esta distribución la desviación estándar es $\sigma = \sqrt{\mu}$. Se trata de una distribución discreta, es decir, x sólo toma valores enteros. En la figura 4 se muestra la distribución de probabilidad de Poisson para $\mu = 10/6$.

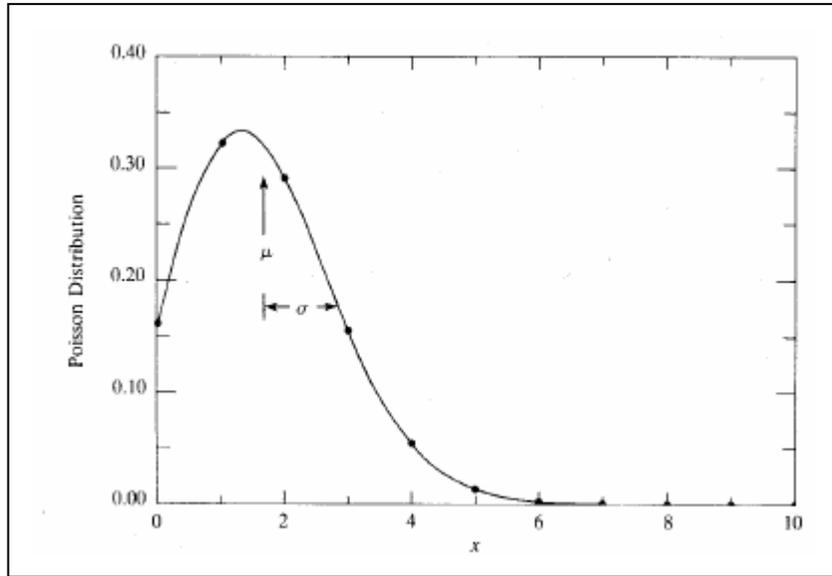


Figura 4.- Distribución de probabilidad de Poisson para $\mu = 10/6$. Medenhall y Sincich (1997)

Los valores de la distribución son los puntos; la curva sirve para guiar el ojo. Si el valor de $\mu > 15$, la distribución de Poisson se convierte en una distribución de Gauss o Normal cuya expresión es:

$$P_G(x, \mu, \sigma = \sqrt{\mu}) = \frac{1}{\sqrt{2\pi\mu}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\mu}} \quad (5)$$

Ésta es una función que describe la probabilidad de que el valor de una observación aleatoria sea x , de una distribución con una media μ y una desviación estándar $\sigma = \sqrt{\mu}$.

Puesto que se trata de una distribución “continua”, debemos definir un intervalo que debe contener el valor de x ; en este caso, la función de probabilidad se define

correctamente como la probabilidad de que el valor de una observación aleatoria se encuentre dentro de un intervalo dx alrededor de x . En la siguiente figura se muestra la distribución de probabilidad de Gauss..

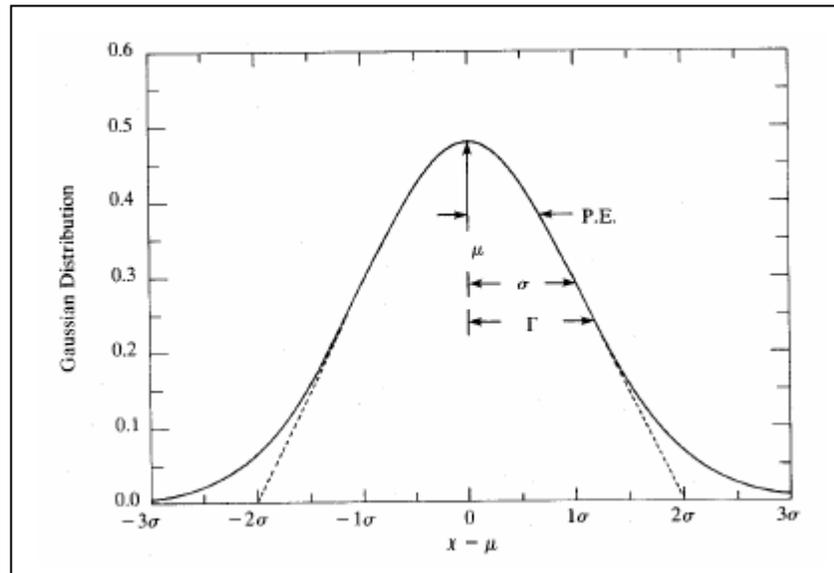


Figura 5.- Distribución de probabilidad de Gauss.. Mendenhall y Sincich (1997)

Redes Neuronales Artificiales

Las Redes Neuronales Artificiales han sido aplicadas en gran número de problemas cuya solución resulta difícil a través de tecnologías convencionales. Según Rich y Knight (1994), los modelos de redes neuronales se basan en una imagen cerebral de la computación.

Esta técnica trata de imitar ciertas características de la fisiología del cerebro como base para su funcionamiento, siendo además sistemas que se optimizan al ir aprendiendo de la experiencia.

Las Redes Neuronales Artificiales (R.N.A.) son modelos matemáticos inspirados en sistemas biológicos, adaptados y simulados en computadoras convencionales. Los elementos que las conforman se asemejan a las neuronas biológicas.

Al respecto, Nilsson (2001) señala que las RNAs son circuitos, algoritmos computacionales o representaciones matemáticas de un conjunto de neuronas conectadas en forma masiva.

Una RNA es un conjunto de neuronas unidas entre sí por medio de conexiones. Estas pueden disponer de conjuntos de neuronas agrupadas en uno o varios niveles o capas.

Cada neurona trabaja con una serie de patrones de entrada y una serie de pesos para cada uno de ellos, con la opción de añadir o no un peso para una entrada de umbral (bias) de valor constante unitario. Incluir este término puede ayudar a la convergencia de los pesos proporcionando un grado de libertad adicional.

Hilera y Martinez (1995) señala que la topología consiste en como se organizan y disponen las neuronas en la red. De acuerdo a su topología las redes neuronales pueden agruparse en: (a) Redes con conexiones hacia delante (feedforward) y (b) Redes con conexiones hacia atrás (feedback).

Normalmente, todas las neuronas de una capa reciben señales de entrada de otra capa anterior cercana a la entrada de la red, y envían las señales de salida a una capa posterior cercana a la salida de la red, por lo que las comunicaciones van en una sola dirección de las neuronas de entrada hacia las neuronas de salida.

A este tipo de redes se les denomina de conexiones hacia delante y según Muller y Reinhardt (1990), esta topología no admite caminos cerrados.

En la figura número 6 puede verse la estructura general de una red multicapa de este tipo con conexión hacia delante, también llamada de alimentación adelantada, la cual funciona de la siguiente forma:

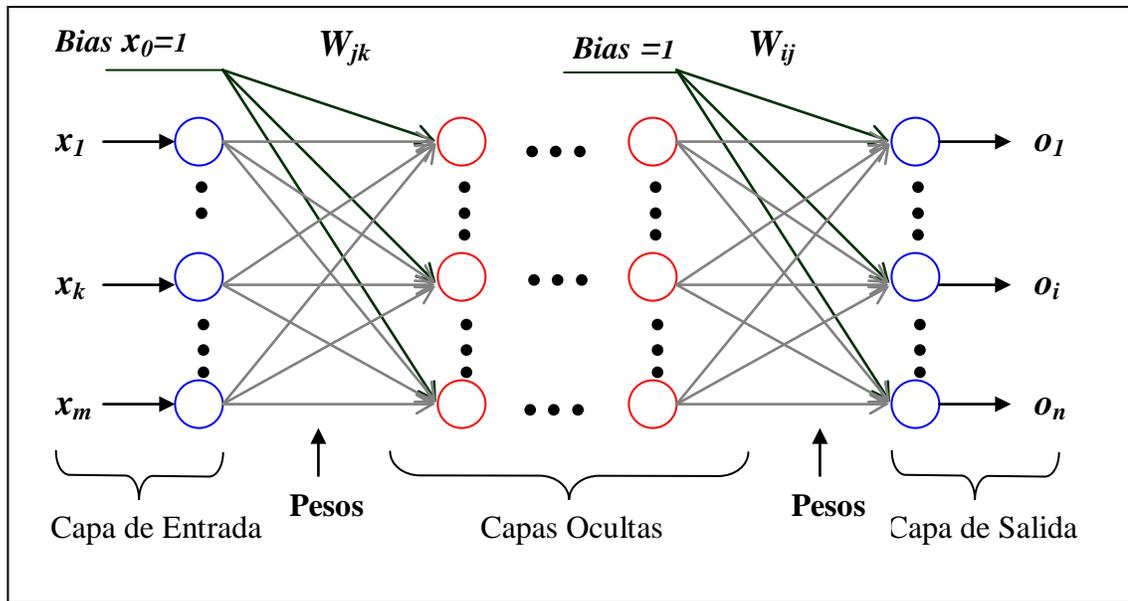


Figura 6. Red Neuronal Artificial Multicapa de conexión hacia adelante. Hilera y Martínez (1995)

1. El funcionamiento de este tipo de redes comienza con una identificación de los patrones a procesar. Cada patrón consiste de un vector de m entradas ($x_1, x_2, \dots, x_k, \dots, x_m$). Luego se presenta a la capa de entrada de la Red la información de cada patrón.
2. En la primera capa oculta cada neurona combina las entradas en un solo valor. Al respecto Muller y Reinhardt (1990) señalan que el nuevo estado v_j de la j -ésima neurona, esta determinado por la influencia de todas las otras neuronas de las que recibe señales y esta expresada como una combinación lineal de sus valores de salida. Esta combinación es realizada de la siguiente forma:

$$h = \sum w_{kj} x_k$$

3. Este resultado se filtra por medio de una función, llamada función de activación $f(h)$. Esta función puede ser lineal con o sin umbral, estocástica o no lineal. Las más comunes son las siguientes:

Lineal: $v_j = c * h$ donde c es un valor constante.

Escalón: $v_j = 1$ si $h > c$

$v_j = 0$ si $h \leq c$

donde c es un valor constante, generalmente cero (0)

Sigmoide: $v_j = 1/(1 + \exp(-h))$

En la figura número 7 puede verse gráficamente cada una de las anteriores funciones de activación.

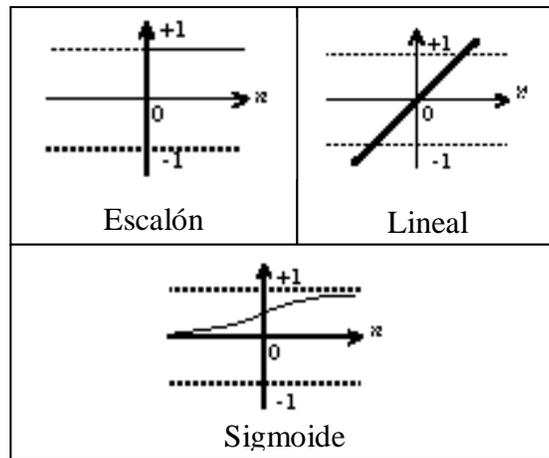


Figura 7. Funciones de activación más comunes. MathWorks, Inc. (1998)

4. Las salidas de la primera capa oculta ($v_1, v_2, \dots, v_j, \dots, v_l$) entran a la segunda capa oculta repitiéndose el proceso hasta obtener las salidas de la red ($o_1, o_2, \dots, o_i, \dots, o_n$).

La mayoría de las aplicaciones en Redes Neuronales Artificiales utilizan una topología de una sola capa oculta. En la figura 8 se puede observar la estructura de este tipo de red.

También existe la posibilidad de conectar las salidas de las neuronas de capas posteriores a las entradas de las capas anteriores (feedback), tal como se muestra en la figura 9. Un ejemplo de estas se puede ver en las redes recurrentes y las redes de Hopfield.

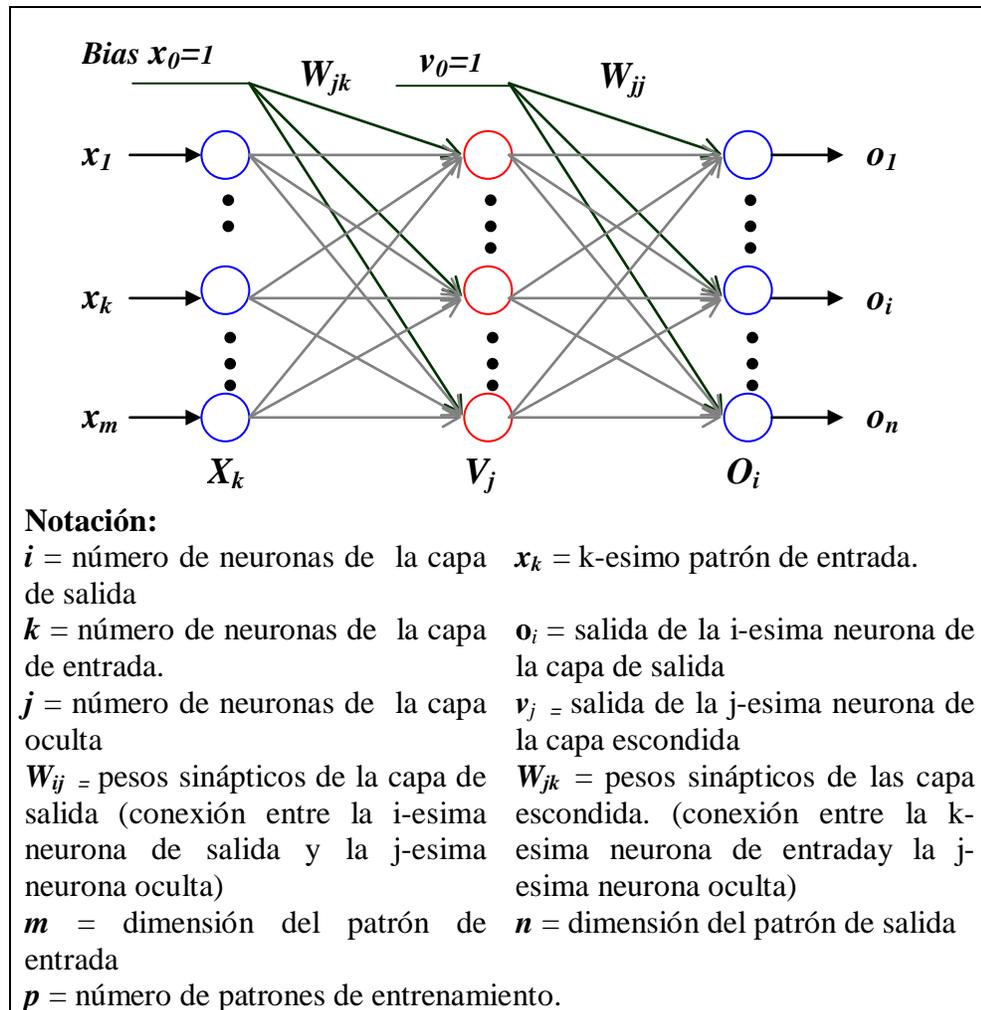


Figura 8. Red Neuronal Artificial con una sola capa oculta. Hilera y Martínez (1995)

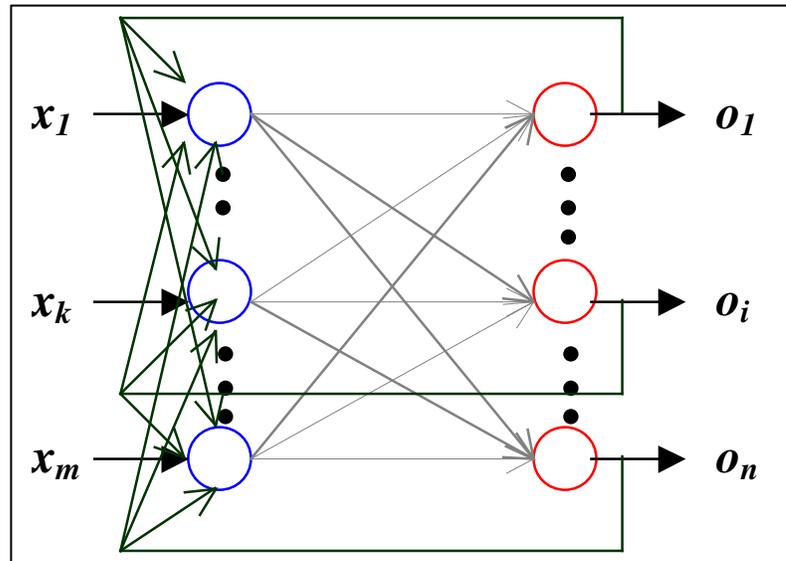


Figura 9. Red con conexiones hacia atrás. Hilera y Martínez (1995)

Paradigma de Aprendizaje

Según Rich y Knight (1994), el aprendizaje en un sistema, implica cambios que le permiten realizar una misma tarea en las mismas condiciones de forma más eficiente y eficaz. En una RNA el conocimiento se almacena en la configuración y parámetros internos de la misma los cuales serán ajustados o adaptados mediante un algoritmo que usa algún criterio para realizar estos ajustes (regla de aprendizaje). Generalmente la adaptación se realiza en las conexiones o pesos sinápticos de la red durante una fase llamada entrenamiento..

En esta fase, se presentaran repetidamente un conjunto de patrones de entradas a la red con el fin de entrenarla. Con los patrones anteriores, se ajustarán los factores de los pesos y a través de estos ajustes, la RNA “aprenderá” el correcto comportamiento. En el desarrollo de una RNA, esta fase es la que por lo general toma más tiempo.

Posteriormente, en la fase de validación, se somete a la RNA un extenso arreglo de patrones de entrada no vistos en el entrenamiento.

En la fase de generalización, se proporcionan a la red patrones de entrada que no han sido utilizados o vistos anteriormente, pero cuyas salidas son conocidas previamente, monitoreando así el desempeño del sistema.

Los métodos de aprendizaje pueden ser agrupados en dos categorías: a) Métodos de aprendizaje supervisados y b) Métodos de aprendizaje no supervisados.

En el aprendizaje supervisado, se guía a la red indicándole la salida correcta o esperada. Al respecto, Nilsson (2001) señala que los ejemplos consistirán en parejas formadas por un posible valor de las variables de entrada junto con la acción que se debe tomar para dicho valor, es decir la salida deseada. Por otra parte, Muller y Reinhardt (1990), afirman que en cada paso, la red es ajustada comparando la salida actual con la salida deseada. El aprendizaje supervisado se puede dividir en aprendizaje estructural, cuyo propósito es encontrar la mejor relación entrada-salida; y aprendizaje temporal relacionado con la captura de una serie secuencial de patrones que se necesitan para alcanzar una meta. Entre estos tipos de algoritmos tenemos los de corrección del error (e.g. Perceptron, Retropropagación, Adaline, Madaline), aprendizaje por reforzamiento y aprendizaje Estocástico (Boltzmann Machine, Cauchy Machine).

El entrenamiento no supervisado consiste en un proceso auto-organizativo que se basa en información local, por lo que no se necesita la guía externa.. Entre estos podemos mencionar el entrenamiento hebbiano (e.g. Red de Hopfield), entrenamiento competitivo y cooperativo (e.g. Red ART).

Algoritmo de Retropropagación de Error.

Según Muller y Reinhardt (1990), este algoritmo de aprendizaje esta basado en un simple pero muy efectivo principio: los pesos sinápticos son modificados iterativamente de forma que la señal de salida difiera tan poco como sea posible de la señal de entrada. Esto es alcanzado por aplicación del método de gradiente, el cual ofrece las modificaciones requeridas para los pesos.

Este algoritmo emplea un ciclo *propagación-adaptación* de dos fases: primero se aplica un patrón de entrada para la primera capa de las neuronas de la red, se va propagando a través de todas las capas superiores hasta generar una salida, se compara el resultado obtenido con la salida que se desea obtener y se calcula un valor del error para cada neurona de salida. Luego esos errores se transmiten hacia atrás partiendo de la capa de salida, hacia todas las neuronas de la capa intermedia recibiendo el porcentaje de error aproximado a la participación de la neurona en la salida. Este proceso se repite para cada capa oculta. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que la próxima vez que se presente el mismo patrón, el error disminuya.

De esta forma, esta red tiene la capacidad de autoadaptar los pesos para aprender la relación entre un conjunto de patrones de entrenamiento y luego poder aplicar esa misma relación a nuevos vectores de entrada, dando una salida activa si la nueva entrada es parecida a las presentadas durante el aprendizaje.

Este es el algoritmo de entrenamiento más popular también llamado ***regla delta generalizada*** por la forma en que se derivan las ecuaciones de actualización de los pesos. La finalidad es minimizar el error existente entre la salida de la red O_i y la salida esperada Y_i para un vector de entrada.

La función de calidad a optimizar comúnmente es el error cuadrático medio:

$$E(\vec{w}) = \frac{1}{2} \sum_{\mu} \sum_i (y^{\mu} - O_i^{\mu})^2 \quad (6)$$

Como la función de salida de la red es continua y derivable con respecto a los pesos sinápticos, esta función de costo puede optimizarse usando un método de descenso de gradiente.

Para las unidades de salida, tenemos:

$$\begin{aligned}\Delta W_{ij} &= -\eta \frac{\partial E}{\partial W_{ij}} = \eta \sum_{\mu} (y_i^{\mu} - O_i^{\mu}) f'(\sum_j W_{ij} V_j) V_j^{\mu} \\ &= \eta \sum_{\mu} \delta_i^{\mu} V_i^{\mu}\end{aligned}\quad (7)$$

Donde

$$\delta_i^{\mu} = f'(\sum_j W_{ij} V_j) (y_i^{\mu} - O_i^{\mu}) \quad (8)$$

Para las unidades de la capa escondida tenemos:

$$\begin{aligned}\Delta w_{jk} &= -\eta \frac{\partial E}{\partial w_{jk}} = \eta \sum_k \frac{\partial E}{\partial V_j^{\mu}} \frac{\partial V_j^{\mu}}{\partial w_{jk}} \\ &= \eta \sum_{\mu} (y_i^{\mu} - O_i^{\mu}) f'(\sum_j W_{ij} V_j) W_{ij} f''(\sum_k w_{jk} x_k^{\mu}) x_k^{\mu} \\ &= \eta \sum_{\mu} \delta_i^{\mu} W_{ij} f'(\sum_j w_{jk} x_k^{\mu}) x_k^{\mu} \\ &= \eta \sum_{\mu} \delta_j^{\mu} x_k^{\mu}\end{aligned}\quad (9)$$

donde

$$\delta_j^{\mu} = f'(\sum_k w_{jk} x_k^{\mu}) \sum_i \delta_i^{\mu} W_{ij} \quad (10)$$

Según Mitchell (1997), el algoritmo retropropagación es la técnica de aprendizaje más comúnmente usada y es apropiada para problemas con las siguientes características: (a) Las instancias son representadas por muchos pares atributo-valor; (b) La salida de la función destino pueden ser valores discretos, reales o un vector que combina ambos tipos; (c) Los ejemplos de entrenamiento pueden contener errores; (d) Son aceptables largos tiempos de entrenamiento; (e) Puede ser requerida la evaluación rápida de la función aprendida; (f) La habilidad de los humanos para entender la función aprendida no es importante.

Método de Montecarlo

La simulación es la técnica que imita el comportamiento de un sistema del mundo real cuando evoluciona en el tiempo. El modelo de simulación es el conjunto de hipótesis acerca del funcionamiento del sistema expresado como relaciones matemáticas y/o lógicas entre los elementos del sistema. Un proceso de simulación es la ejecución del modelo a través del tiempo en un computador para generar muestras representativas del comportamiento del sistema. La simulación se basa en un muestreo aleatorio, es decir, la salida de la simulación está sujeta a variaciones aleatorias y debe ser examinada utilizando pruebas de inferencia estadística. Los elementos de una simulación son los siguientes:

- Salidas: objetivos del estudio expresadas mediante valores numéricos.
- Entradas: valores numéricos que permiten iniciar la simulación y obtener las salidas:
- Condiciones iniciales: valores que expresan el estado del sistema al principio de la simulación
- Datos determinísticos: valores conocidos necesarios para realizar los cálculos que producen las salidas
- Datos probabilísticos: cantidades cuyos valores son inciertos pero necesarios para obtener las salidas de la simulación. Los valores específicos de estos datos deben conocerse a través de una distribución de probabilidad.

El método de Montecarlo es un modelo de simulación estática ya que representa a un sistema en un instante de tiempo determinado y que se emplea para resolver problemas empleando muestras aleatorias. Según Sobol (1976) su algoritmo es muy sencillo, elaborando un programa para la realización de una prueba aleatoria, la misma se repite N veces de forma independiente y luego se toma la media de los resultados.

Este método permite simular procesos que dependen de factores aleatorios, de lo contrario, se puede modelar en forma probabilística dicho proceso para resolverlo. Un mismo problema puede ser resuelto aplicando distintas variantes del método de Montecarlo.

El mismo autor señala que las variables aleatorias pueden ser obtenidas a través de: a) tablas de cifras aleatorias, donde cada número de la cifra es tomado al azar de entre 10 posibles valores (0 al 9) teniendo cada uno una probabilidad de 0.1 de ser seleccionado; b) generadores de números aleatorios, donde las cifras son generadas a partir de números binarios en los que el valor de los bits depende de una condición booleana; c) números pseudoaleatorios, generados por diversos métodos y utilizados en su inmensa mayoría en los cálculos del método de Montecarlo.

Uno de los métodos para generar números pseudoaleatorios es utilizar números binarios de 42 ordenes donde la posición cero es el signo de la mantisa, de la uno a la 35 representa la mantisa, la posición 36 es el signo del orden y las restantes representan el orden. Cada número se obtiene del número anterior multiplicándolo por 10^{17} , el resultado se desplaza en 7 ordenes a la izquierda y se toma el valor absoluto del número obtenido.

Se pueden sortear los valores de una variable aleatoria ξ mediante la transformación de uno o varios valores de la variable aleatoria γ uniformemente distribuida en el intervalo (0,1). Si ξ es una variable discreta con la siguiente distribución:

$$\left(\begin{array}{c} x_1, x_2, \dots, x_n, \\ p_1, p_2, \dots, p_n \end{array} \right)$$

tomamos $\xi = x_i$ cuando $p_1 + p_2 + \dots + p_{i-1} < \gamma < p_1 + p_2 + \dots + p_i$

Si ξ es una variable aleatoria continua distribuida con densidad $p(x)$ en el intervalo (a,b), los valores de ξ pueden determinarse de la ecuación: $\int_b^a p(x)dx = \gamma$ (11)

Si la variable esta uniformemente distribuida en el intervalo $[0,1]$ su densidad es constante e igual a: $1/(b-a)$, por lo que la integral se calcula fácilmente quedando que $\xi = a + \gamma(b - a)$

Cuando no se puede resolver fácilmente la integral y la densidad esta acotada, es decir, $p(x) \leq M_o$, los valores de ξ se pueden sortear escogiendo dos valores γ' y γ'' , considerando el punto aleatorio $\Gamma(\eta', \eta'')$ donde $\eta' = a + \gamma'(b - a)$ y $\eta'' = \gamma''M_o$. Si el punto Γ esta por debajo de la curva $y = p(x)$ entonces $\xi = \eta'$, de lo contrario escogemos un nuevo par de valores para γ' y γ'' .

Para variables aleatorias normales ζ los procedimientos anteriores pueden ser poco efectivos, por lo que existen tablas de valores normales ya creadas. Por ejemplo, para valores normales con Esperanza matemática $M\zeta = 0$ y varianza $D\zeta = 1$, los valores de ζ se obtienen con la fórmula $\alpha + \sigma\zeta$ donde $M\zeta = \alpha$ y $D\zeta = \sigma^2$.

Glosario de Términos Básicos.

Antipartícula: Cada tipo de partícula material tiene una antipartícula correspondiente. Cuando una partícula material choca con su antipartícula se aniquilan ambas, quedando sólo energía

Átomo: Unidad básica de la materia ordinaria, compuesta de un núcleo diminuto (consistente en protones y neutrones) rodeado por electrones que giran alrededor de él. (Daintith, 1984)

Ciencia de Materiales: Tiene como objetivo estudiar y entender los mecanismos microscópicos estructurales y dinámicos responsables de las propiedades macroscópicas de los sólidos, así como la influencia de los procesamientos de los materiales sobre su estructura y por lo tanto sobre sus propiedades. Todo esto sirve por una parte para entender el comportamiento de la materia sólida como parte de la comprensión de nuestro ambiente y por otra, para poder manipular la materia y obtener de ella propiedades específicas que sirvan a uno o varios fines.

Ciencia de Polímeros: Ciencia que estudia los polímeros. Un polímero es una molécula grande que consiste en muchas subunidades similares idénticas

Cuanto: Unidad indivisible, en la que las ondas pueden ser emitidas o absorbidas. Wikipedia. (2004)

Detector: Dispositivo que responde a un efecto físico, utilizado para indicar la presencia de una señal o para medir ésta. (Daintith, 1984)

Espectro: Separación de, por ejemplo, una onda electromagnética en sus frecuencias componentes. (Daintith, 1984)

Espectroscopio: Instrumento adecuado para descomponer la luz en su espectro, por medio de un prisma o de un retículo de difracción.

Espin: Propiedad interna de las partículas elementales, relacionadas con el concepto ordinario de giro. Wikipedia. (2004)

Evento: Lo que ocurre cuando dos partículas colisionan, o cuando una partícula decae. Las teorías de partículas predicen la probabilidad de que ocurran varios acontecimientos posibles, cuando se estudian muchas colisiones o decaimientos similares. No se puede predecir el resultado para un evento en particular.

Fermión: Cualquier partícula que tiene momento angular intrínseco (spin) impar semientero ($1/2$, $3/2$, ...), medido en unidades de h -barra. Como una consecuencia de este momento angular peculiar, los fermiones obedecen una regla llamada el Principio de Exclusión, que establece que dos fermiones no pueden existir en el mismo estado al mismo tiempo. Muchas de las propiedades de la materia ordinaria surgen como consecuencia de esta regla. Los electrones, protones, y neutrones son todos fermiones, lo mismo que todas las partículas de materia fundamentales, tanto quarks como leptones. Wikipedia. (2004)

Fotón: Del griego φως, φωτος, luz. También denominado gamma o γ , es la partícula mediadora de la interacción electromagnética. Los fotones están asociados a cualquier radiación electromagnética. Wikipedia. (2004)

Inteligencia Artificial: Es la simulación de los procesos de la inteligencia humana por medio de máquinas, especialmente sistemas computarizados. Esos procesos incluyen aprendizaje, razonamiento y auto-corrección.

Leptón: Un fermión fundamental que no participa en las interacciones fuertes. Los leptones eléctricamente cargados son: los electrones, los muones, las partículas tau, y sus antipartículas. Los eléctricamente neutros son llamados neutrinos. Wikipedia. (2004)

Masa: Propiedad característica de un objeto, que establece su inercia, es decir, la resistencia del objeto a iniciar o cambiar su movimiento. La masa se mide en kilos. No es igual al peso. (Daintith, 1984)

Parámetro: Es un término utilizado muy frecuentemente en el lenguaje común y, no necesariamente, con demasiada propiedad. Parámetro es condición variable a la que se asignan unos valores determinados y fijos. En informática puede ser cualquier condición para el desarrollo de un programa, que modifica o escinde su forma de funcionar.

Partículas materiales: El Modelo Standard establece que la mayoría de las partículas de las cuales tenemos conocimiento están compuestas en realidad de partículas más fundamentales llamadas quarks. Hay otra clase de partículas fundamentales llamadas leptones (el electrón es un ejemplo).

Partículas portadoras de fuerza:

Cada tipo de fuerza fundamental es "transportada" por una partícula portadora de fuerza (el fotón es un ejemplo).

Rayos gamma: Radiación correspondiente a la parte más energética (o de mayor frecuencia) de todo el espectro electromagnético. Los rayos gamma son producidos en los procesos de desintegración nuclear junto con los rayos alfa y beta. La radiación gamma puede originarse también por aniquilación de la materia con la antimateria. Wikipedia. (2004)

Traza: El registro de la trayectoria de una partícula cuando atraviesa un detector. (Daintith, 1984)

CAPITULO III

MARCO METODOLOGICO

Tipo de Investigación

De acuerdo con lo establecido por el Manual para la presentación de Trabajos de Especialización, Maestría y Doctorado (UCLA, 2002), la modalidad de investigación en la que está enmarcado el presente trabajo es la de proyecto factible.

El proyecto se apoya en la investigación documental realizada para la revisión y recopilación de toda la información referida a los dominios de conocimiento involucrados, tanto en el área de estudio de materiales como en la de Redes Neuronales Artificiales.

Este trabajo se enmarca dentro de la Línea de Investigación de Inteligencia Artificial de la Maestría en Ciencia de la Computación, y más específicamente en el área de Aprendizaje Automático a través de Redes Neuronales Artificiales.

Para lograr la culminación del proyecto, se seguirá una serie de fases contempladas dentro del esquema establecido por las normas de presentación de trabajo de grado. Las fases de diagnóstico y factibilidad se detallan a continuación.

Fase de Diagnóstico

Diseño de la investigación o Procedimiento

1. Primeramente, se desarrollará un programa de simulación basado en el método de Montecarlo, para generar los datos de los espectros de tiempo de vida de

aniquilación de positrones, que servirán de ejemplos para el entrenamiento y validación del modelo de red neuronal artificial a diseñar.

2. Con este programa, se generarán los patrones de entrenamiento que posteriormente serán procesados por la red.
3. Una vez obtenido los espectros simulados, se procederá a diseñar y programar el modelo de Red Neuronal Artificial (RNA) seleccionando una determinada topología y algoritmo de entrenamiento
4. Luego de culminar esta fase, se procederá a ejecutar una serie de experimentos, donde se varía la arquitectura y parámetros de la RNA. Se entrenará y se validará cada modelo con los espectros simulados.
5. Se registrarán los resultados obtenidos por cada modelo, para posteriormente seleccionar el modelo de red que presente mejor rendimiento.

Técnicas e Instrumentos de Recolección de Información

Para la recolección de información acerca de las características propias de los esquemas de redes neuronales existentes, se llevará a cabo la revisión documental, a través de consultas en la bibliografía de referencia, búsquedas en Internet, y en la biblioteca de la UCLA. Además, se realizarán consultas a académicos y expertos en materias afines a la investigación. La recolección de los datos requeridos para el entrenamiento de la red, se realizará de forma automática con el computador. Se utilizarán herramientas de software para desarrollar el programa necesario para que esto pueda realizarse adecuadamente.

Técnicas de análisis de los datos.

Los datos de los espectros generados por el programa de simulación, serán analizados según el método de simulación utilizado, para asegurar que los mismos

representen lo mejor posible al comportamiento del sistema, abarcando las regiones de posibles valores de los parámetros.

La información obtenida acerca de los diversos tipos de red neuronal con aprendizaje supervisado se analizará de acuerdo a las aplicaciones que han tenido en diversas áreas, a fin de establecer cual es el diseño más apropiado para el problema que se esta modelando.

Los resultados obtenidos de los procesos de entrenamiento, validación y prueba de la red neuronal serán analizados basándose en el error cuadrático medio, el tiempo de duración de cada proceso y la complejidad de la red diseñada

Estudio de Factibilidad

Factibilidad Técnica

Este proyecto se puede considerar técnicamente factible ya que los recursos de hardware y software requeridos para el desarrollo del mismo están disponibles en el ámbito nacional. Para el basamento teórico-práctico del mismo, se dispone de bibliografía adecuada con respecto a los dominios de conocimiento involucrados en la investigación, además de contar con el acceso a Internet.

Por otra parte se cuenta con recurso humano especializado para el desarrollo del mismo, así como también con el asesoramiento de expertos en el área de física e inteligencia artificial.

Factibilidad Operativa

Este proyecto se considera factible operativamente, ya que se desarrollará una red neuronal artificial que una vez entrenada puede operar con tiempos de

identificación relativamente cortos. La interfaz del programa será diseñada de tal forma que pueda ser fácilmente operado por los usuarios.

Factibilidad Económica

El desarrollo de esta investigación se considera factible económicamente ya que se cuenta con los recursos de hardware y software necesarios, es decir, se cuenta con un computador con la configuración acorde a los requerimientos de desarrollo; igualmente se cuenta con el software de lenguaje en el que se realizará la programación. Por otra parte, los tutores que se encargan de asesorar este proyecto se encuentran establecidos en esta región por lo que no se realizaran gastos de traslado a otras regiones del país.. Así mismo, el material bibliográfico necesario se puede encontrar en este mismo estado. Igualmente, se cuenta con acceso gratuito a Internet por medio de la conexión en la Universidad Centroccidental “Lisandro Alvarado”.

En resumen, la inversión económica a realizar estaría representada por materiales y suministros varios además de los aranceles, entre otros gastos. El costo estimado en este aspecto, se muestra en la tabla 2, el cual se encuentra dentro de un rango admisible por parte del encargado de financiar y llevar a cabo el proyecto.

Tabla 1.

Estimación de gastos

Descripción	Costo Estimado (Bs.)
Materiales y suministros	900.000.00
Aranceles y otros gastos	800.000.00
TOTAL	1.700.000.00

Fuente: El Autor.

CAPÍTULO IV

PROPUESTA DEL ESTUDIO

Justificación

La Red Neuronal Artificial para el análisis de espectros de vida de aniquilación de positrones, propuesta como resultado de la investigación realizada, es una herramienta útil para el estudio de materiales, ya que sirve de base para determinar los tiempos de vida e intensidades de aniquilación de los positrones que interactúan con dicho material, a partir del espectro de tiempos de vida medido como producto de dichas interacciones.

Igualmente, representa un aporte significativo en la novedosa línea de investigación de aplicación de redes neuronales artificiales para el análisis de espectros de tiempos de vida de aniquilación de positrones, que podría incentivar a realizar nuevas investigaciones en esta área, ya sea por parte de la misma autora o de otros investigadores interesados en la misma, aplicando otros diseños o tipos de red neuronal artificial.

Esta propuesta mejora los resultados obtenidos en investigaciones anteriores, tanto en exactitud, como en complejidad y tiempos de entrenamiento, estableciendo un buen balance entre dichos criterios.

Objetivos

Objetivo General

Analizar espectros simulados de tiempos de vida de aniquilación de positrones

Objetivos Específicos

- Generar espectros de tiempos de vida de aniquilación de positrones
- Entrenar la red para que aprenda los espectros simulados
- Probar el rendimiento de la red con espectros no usados en el entrenamiento.
- Determinar las intensidades y tiempos de vida de un espectro dado
- Graficar un espectro dado.

Descripción de la Propuesta

Aspectos Generales

En la presente investigación se diseñó y desarrolló una Red Neuronal Artificial para el análisis de espectros simulados de tiempos de vida de positrones, los cuales son partículas elementales que se aniquilan al chocar con los electrones de algún material.

La Red Neuronal Artificial para el análisis de espectros simulados de tiempos de vida de aniquilación de positrones (POSIREN 1.0), cuenta con procedimientos para generar y graficar los espectros simulados, e igualmente con procedimientos para el entrenamiento y prueba de la red neuronal.

Los espectros son simulados basándose en el método de Montecarlo, ya que las variables involucradas son de naturaleza aleatoria. Los datos de los espectros son preparados para ser usados como patrones de entrenamiento de la red.

Un patrón de entrada consiste de una cantidad de valores de entrada igual al número de canales del espectro. Estos valores corresponden a la cantidad de cuentas por canal, es decir, la cantidad de eventos de aniquilación ocurridos dentro del lapso de tiempo representado en el canal. Igualmente, el patrón contiene los seis valores para las salidas deseadas, es decir, las tres intensidades y los tres tiempos de vida del espectro simulado

Estos espectros son procesados por la red neuronal para aprender la relación existente entre las entradas y salidas deseadas. Esto permite que la red, una vez entrenada, pueda determinar estos valores implícitos dentro del espectro, funcionando así, como un aproximador de la función asociada a la generación de espectros.

Las intensidades y tiempos de vida de aniquilación de positrones son relevantes para el estudio de defectos en materiales, ya que mediante estos, se puede determinar si un material estudiado tiene defectos, al presentar tiempos de vida más largos que los normales; por otro lado, las intensidades dan una idea de la magnitud o concentración del defecto. El número de componentes (en este caso tres),. esta relacionado a la cantidad de defectos

La red neuronal implementada es de tipo perceptrónica multicapa con entrenamiento basado en retropropagación del error. Para acelerar el proceso de aprendizaje se añadió un término de momento, además de someter a los patrones de entrenamiento a un preprocesamiento para llevar los datos a una escala común..

Para la programación y el procesamiento de los datos se utilizó un computador personal con procesador AMD Athlon. 1,00 GHz y 120 MB de memoria RAM, disco duro de 20 GB, Monitor SVGA. La codificación de los programas fue desarrollada en lenguaje de programación ANSI C, permitiendo la portabilidad entre diferentes plataformas de sistema operativo.

POSIREN 1.0 permite generar espectros simulados de 1500, 800 y 500 canales. También es posible graficar en pantalla un espectro generado. De igual forma, se puede entrenar y validar la red neuronal con los datos simulados, mostrando como resultado el error cuadrático medio obtenido en ambos procesos. Además, permite

procesar un espectro para determinar tres intensidades y tres tiempos de vida de tres componentes, mostrando como resultado las salidas deseadas, las salidas obtenidas por la red y el error cometido en cada uno de los seis parámetros calculados.

Procedimiento de Simulación

Para la generación de los patrones de entrenamiento de la red neuronal se desarrolló un programa de simulación de espectros de tiempos de vida de aniquilación de positrones, donde se generan los datos con diferentes intensidades y tiempos de vida que oscilan entre determinados rangos de valores, los cuales se muestran en la tabla 2.

De igual forma, como han sido generados en investigaciones anteriores, se usará el método de Montecarlo, debido a que los procesos a simular son de naturaleza aleatoria.

Tabla 2

Rango de Variación de los parámetros de los espectros

		Límite Inferior	Límite Superior
Primer Componente	Intensidad	10	40
	Tiempo de Vida	0.1	0.16
Segundo Componente	Intensidad	30	60
	Tiempo de Vida	0.25	0.44
Tercer Componente	Intensidad	20	45
	Tiempo de Vida	1.6	4.0

Los datos simulados son generados bajo las siguientes características preestablecidas:

- Número de componentes: 3
- Canal de tiempo cero: 39.0 ps
- Ancho de Canal: 25.0 ps

El número de cuentas que serán registrados en los espectros se pide como dato de entrada, el cual debe ser suministrado por el usuario de la aplicación.

Primero se generan aleatoriamente los valores de las intensidades y tiempos de vida para el espectro a generar, tomando en cuenta los rangos de variación de dichos valores. Se verifica que la suma de las intensidades alcance el valor de 100.

Para normalizar estos valores se convierten los tiempos de vida de nanosegundos a picosegundos, se calcula la relación del tiempo de vida de cada componente con la anchura de canal

Luego se calcula la fracción de eventos relacionados al tiempo de vida de cada componente basándose en las fórmula de la derivada del espectro de decaimiento del positrón, que según Bondarenko (2003) se expresa de la siguiente forma:

$$N(T) = \sum_{1 < i < k+1} \frac{I_i}{\tau_i} \exp\left(\frac{t-t_0}{\tau_i}\right) \quad (12)$$

Con este resultado se calcula el número de cuentas por componente de acuerdo a la fracción de eventos y a las intensidades de cada uno. Con los valores mencionados anteriormente se inicia la simulación

Para cada evento de cada componente (positron emitido), se genera un número aleatorio γ diferente de cero. El tiempo recorrido τ es una variable aleatoria que puede tomar cualquier valor positivo con una distribución exponencial de la forma:

$p(x) = \alpha e^{-\alpha x}$ con esperanza matemática de $1/\alpha$. Por lo tanto, la fórmula correspondiente al sorteo de τ es:

$$\int_b^a \alpha e^{-\alpha x} dx = \gamma \quad (13)$$

Calculando la integral del primer miembro, obtenemos la relación:

$$1 - \alpha e^{-\alpha x} = \gamma$$

de donde resulta: $\tau = 1/\alpha \ln(1-\gamma)$. Como la variable $1-\gamma$ tiene la misma distribución de γ , se puede usar la fórmula:

$$\tau = 1/\alpha \ln(\gamma). \quad (14)$$

En este caso $1/\alpha$ corresponde al valor esperado para el tiempo de vida del componente.

Luego de calcular el tiempo de vida para el evento de aniquilación se convoluciona con la función de resolución Gaussiana y se determina en que canal de tiempo cae dicho evento, incrementándose de esta forma el número de cuentas en ese canal. Después de procesar todas las cuentas se agrega a los canales un background o fondo aleatorio, siguiendo también una distribución Gaussiana.

En la figura 10 se pueden observar ejemplos de espectros simulados con tres intensidades y tiempos de vida, con canal de tiempo cero centrado en 25.0 ps. En la figura 11 se muestra gráficamente el algoritmo seguido para la simulación.

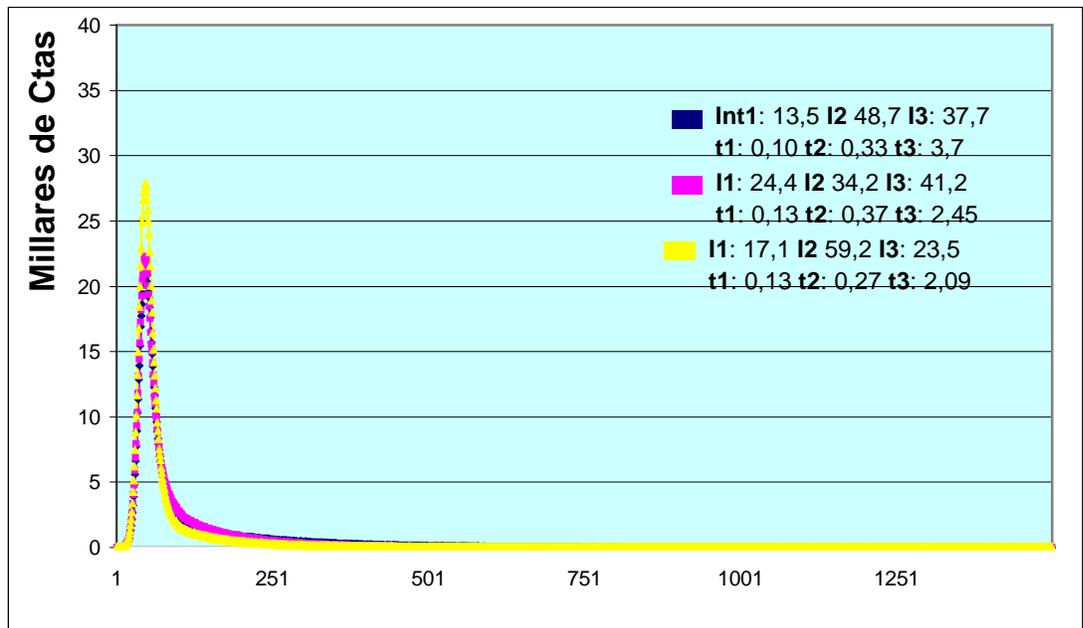


Figura 10. Ejemplos de espectros simulados
Córdoba(2005)

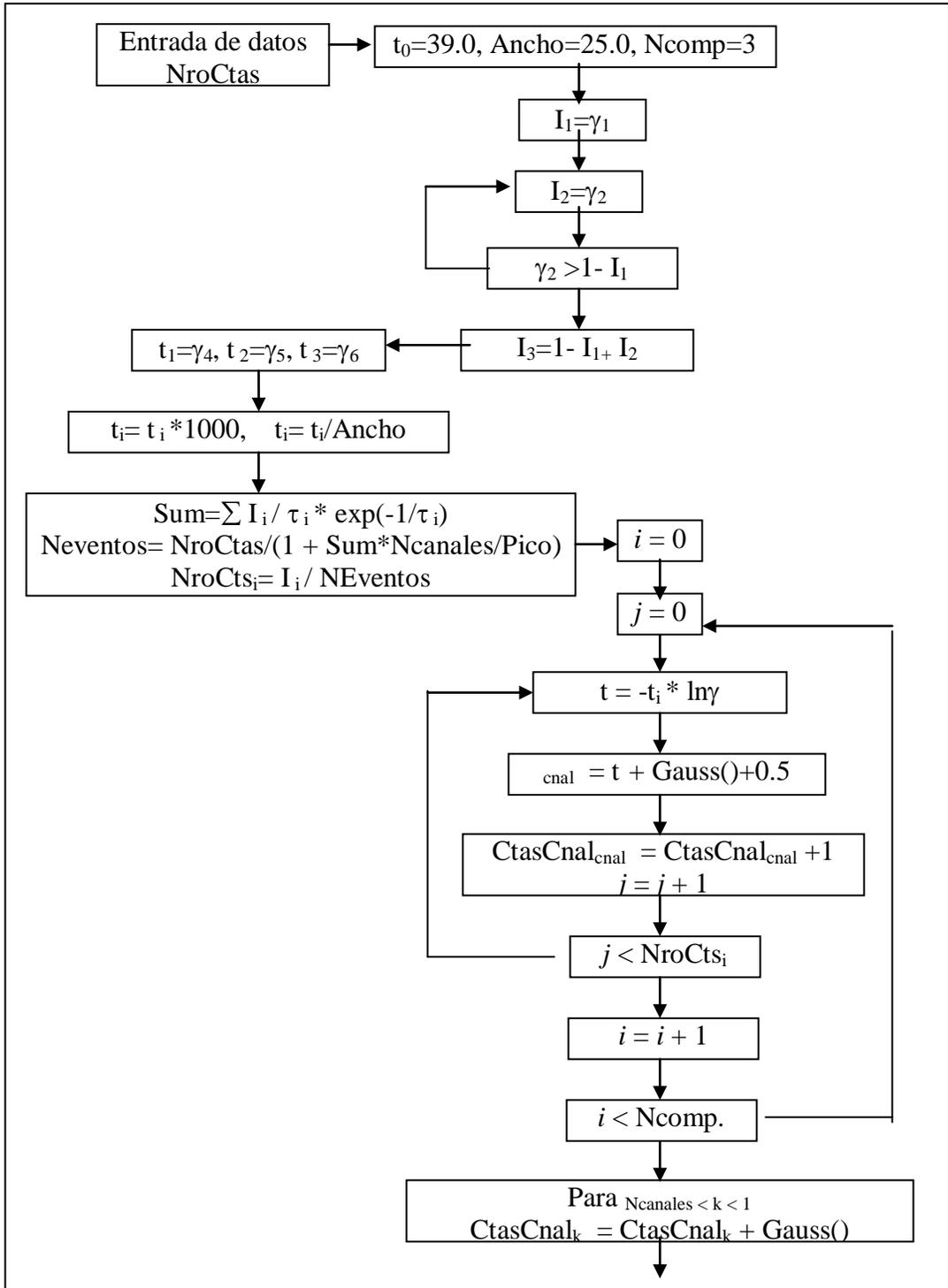


Figura 11. Algoritmo de Simulación de los Espectros
Córdoba(2005)

Diseño de la Red

La red neuronal implementada es de tipo perceptrónica multicapa con entrenamiento basado en retropropagación del error. La activación de las neuronas se realiza a través de la función sigmoide. La capa de entrada tiene tantas neuronas como número de canales se simulen, en este caso, se experimentó la red con 1500, 800 y 500 canales. La capa de salida tiene seis neuronas que corresponden a los seis parámetros que serán aprendidos por la red, es decir las tres intensidades y los tres tiempos de vida. El número de neuronas de la capa oculta fue establecido en base a los experimentos realizados y a los recursos de memoria del equipo utilizado.

El diseño general de la red para el primer grupo de experimentos tiene un máximo de 3 neuronas en la capa oculta. De acuerdo a los resultados de los experimentos para espectros de 1500 canales, la capa oculta fue establecida con una neurona, para espectros de 800 canales la capa oculta tiene dos neuronas y para los espectros de 500 canales, se estableció una capa oculta de tres neuronas. El valor para la tasa de aprendizaje, que regula el tamaño del paso en la superficie del error, fue establecido en 0.5. Para acelerar aun más este proceso se añadió un término de momento, el cual se colocó en 0.9.

En el segundo grupo de experimentos, la red fue diseñada con un máximo de 9 neuronas ocultas. Con los resultados obtenidos quedó establecido para espectros de 1500 canales, una capa oculta con tres neurona, para espectros de 800 canales la capa oculta tiene cuatro neuronas y para los espectros de 500 canales, se estableció una capa oculta de ocho neuronas. El valor para la tasa de aprendizaje usada es de 0.5 para el primer caso y 0.25 para el segundo y tercer caso. Se mantiene un término de momento de 0.9 en todos los experimentos. Las arquitecturas de la red para los tres casos estudiados (1500, 800 y 500 canales), del primer grupo de experimentos, se muestran en las figuras 12, 13, y 14, respectivamente. En las figuras 15, 16 y 17 se observan las arquitecturas para el segundo grupo de experimentos. Estos diseños, ya

permiten observar una reducción de la complejidad de la red, lo cual puede mejorar los tiempos de entrenamiento.

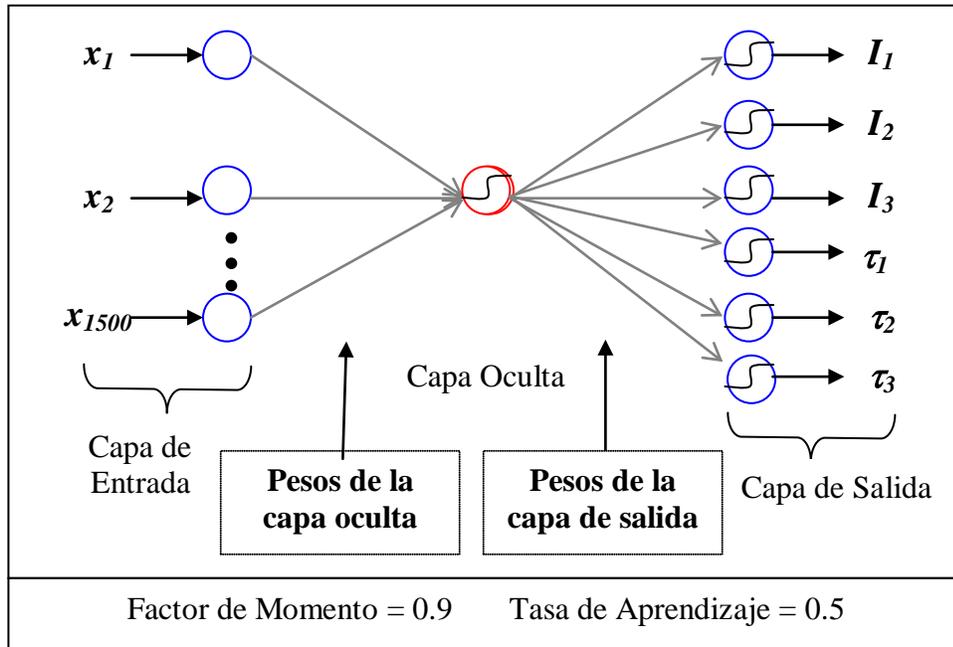


Figura 12. RNA, 1500 Canales. Primer grupo de experimentos. Córdoba(2005)

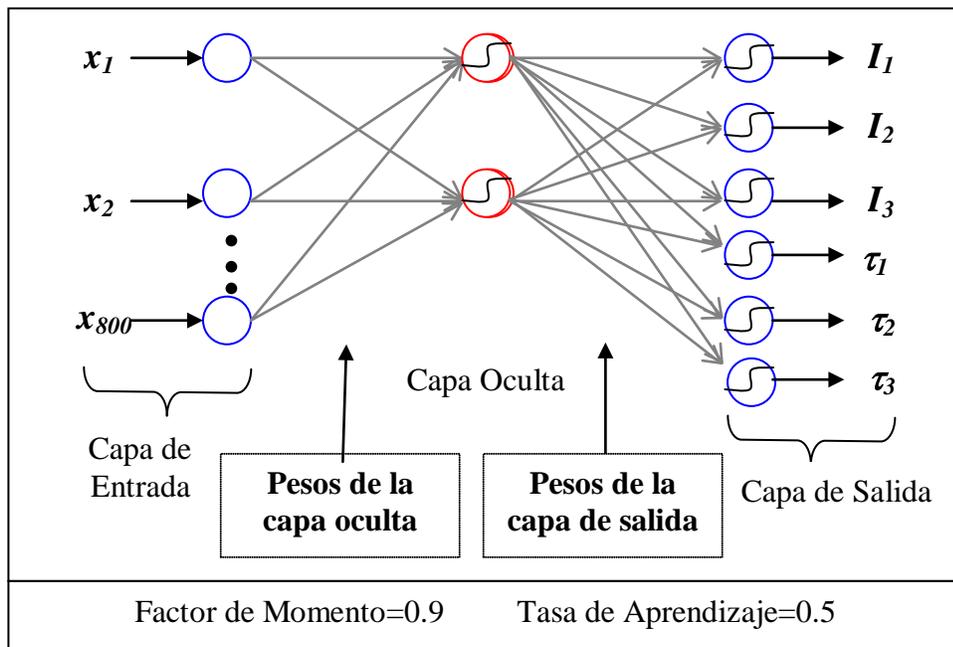


Figura 13. RNA, 800 Canales. Primer grupo de experimentos. Córdoba(2005)

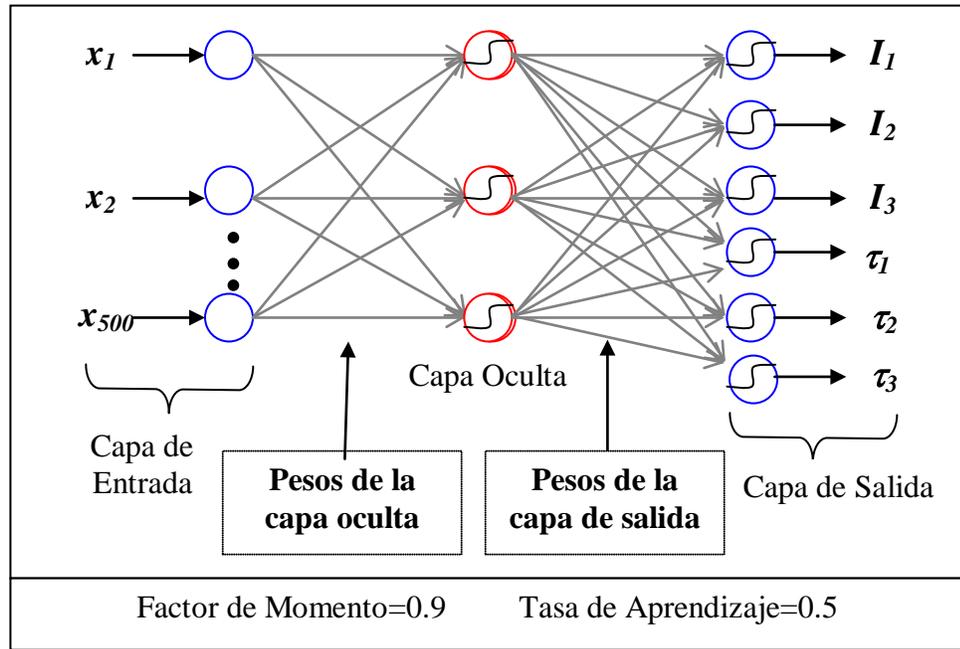


Figura 14. RNA, 500 Canales. Primer grupo de experimentos. Córdoba(2005)

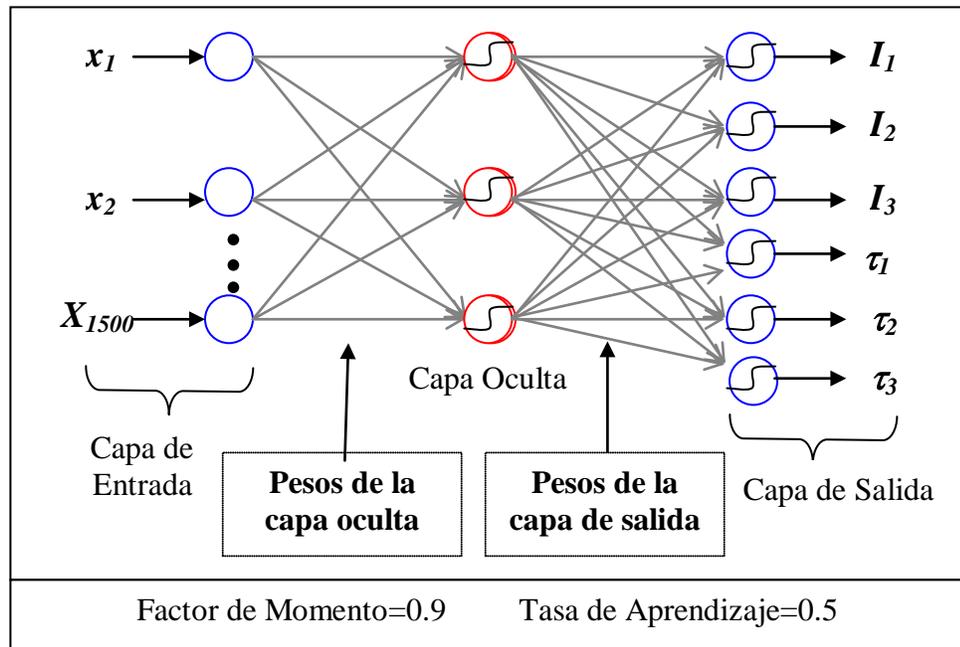


Figura 15. RNA, 1500 Canales. Segundo grupo de experimentos. Córdoba(2005)

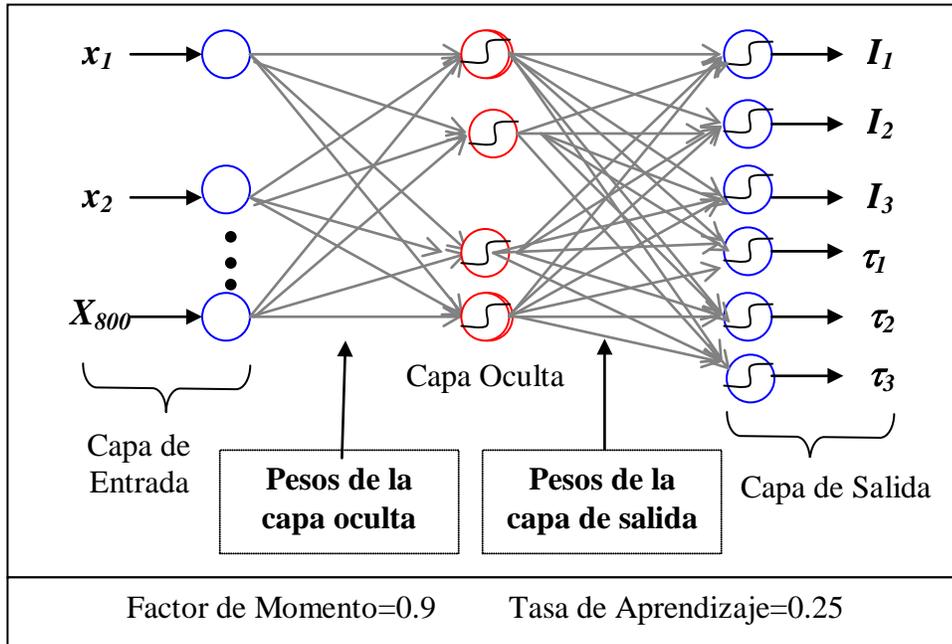


Figura 16. RNA, 800 Canales. Segundo grupo de experimentos. Córdoba(2005)

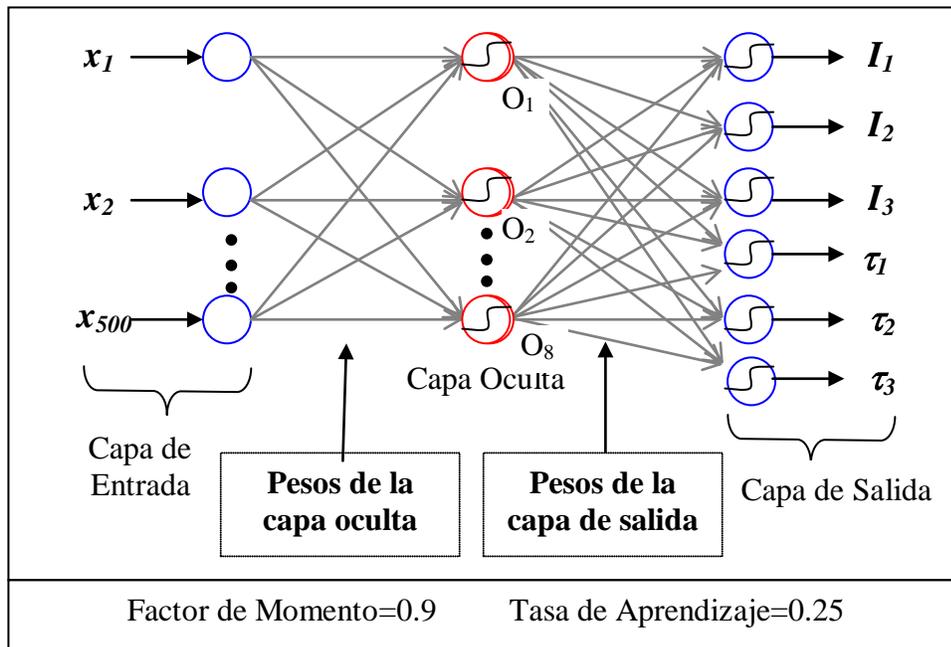


Figura 17. Red Neuronal Artificial, Caso 500 Canales. Segundo grupo de experimentos. Córdoba(2005)

Se utilizaron estructuras de datos acordes a los requerimientos del problema.

Entre estas tenemos:

- Un arreglo unidimensional para almacenar el patrón de entrada en proceso, donde las tres primeras posiciones corresponden a las intensidades, en las tres siguientes se almacenan los tiempos de vida y las restantes son para almacenar el número de cuentas por canal de tiempo, ya sea para 1500, 800 o 500 canales según la opción seleccionada.
- Tres arreglos estáticos de seis posiciones para almacenar las salidas de la red, las salidas deseadas y los errores de las unidades de salida.
- Dos arreglos unidimensionales estáticos para almacenar las salidas de las unidades ocultas y los errores cometidos por las mismas. Estos tendrán un máximo de 3, 4 u 8 posiciones según el caso (1500, 800 y 500 canales)
- Dos arreglos dinámicos unidimensionales, uno para almacenar los pesos de las conexiones entre las unidades de la capa de entrada y la capa oculta, el otro para almacenar las variaciones de los errores para dichas conexiones. Estos tendrán una dimensión correspondiente a la multiplicación del número de unidades de entrada (número de canales) por el número de unidades ocultas
- Dos arreglos estáticos bidimensionales, uno para almacenar los pesos de las conexiones entre las unidades de la capa oculta y la capa de salida, el otro para almacenar las variaciones de los errores para dichas conexiones. Estos tendrán una dimensión correspondiente a la multiplicación del número de unidades ocultas por el número de unidades de salida.

Procesos de Entrenamiento y Prueba

Procesamiento del Primer Grupo de Experimentos

En esta etapa se realizaron varios grupos de experimentos con cada uno de los casos planteados. Primeramente se desarrolló un programa en donde el entrenamiento y prueba se realizó con espectros generados en el mismo momento de ser procesados. En las figura 18 se puede observar gráficamente el algoritmo de entrenamiento.

El proceso de entrenamiento comienza con la presentación a la red de un patrón de entrada. Como se mencionó anteriormente un patrón de entrada contiene seis valores iniciales, que corresponden a las salidas deseadas (las tres intensidades y tres tiempos de vida del espectro generado); los valores restantes son las cuentas por cada canal de tiempo del espectro.

Cada patrón de entrenamiento es generado en el mismo momento que va a ser procesado, con el fin de ahorrar recursos de memoria, que son necesarios para almacenar las estructuras dinámicas para los pesos de la capa oculta. Los datos de cada patrón procesado son posteriormente almacenados en un archivo de texto. Para entrenar la red se realizaron una serie de experimentos variando los valores de los parámetros claves en la red. Entre estos se encuentran:

- Número de canales
- Número de unidades en la capa oculta
- Cantidad de épocas de entrenamiento
- Cantidad de espectros de entrenamiento
- Tasa de Aprendizaje

Antes de ser procesado, el espectro generado es sometido a un proceso de escalamiento entre los valores 0 y 1, permitiendo reducir los valores de cada espectro a una escala común. Para este proceso se usó la siguiente fórmula:

$$\text{Valor escalado} = \frac{(\text{Valor Original} - \text{Menor Valor})}{(\text{Mayor Valor} - \text{Menor Valor})} \quad (15)$$

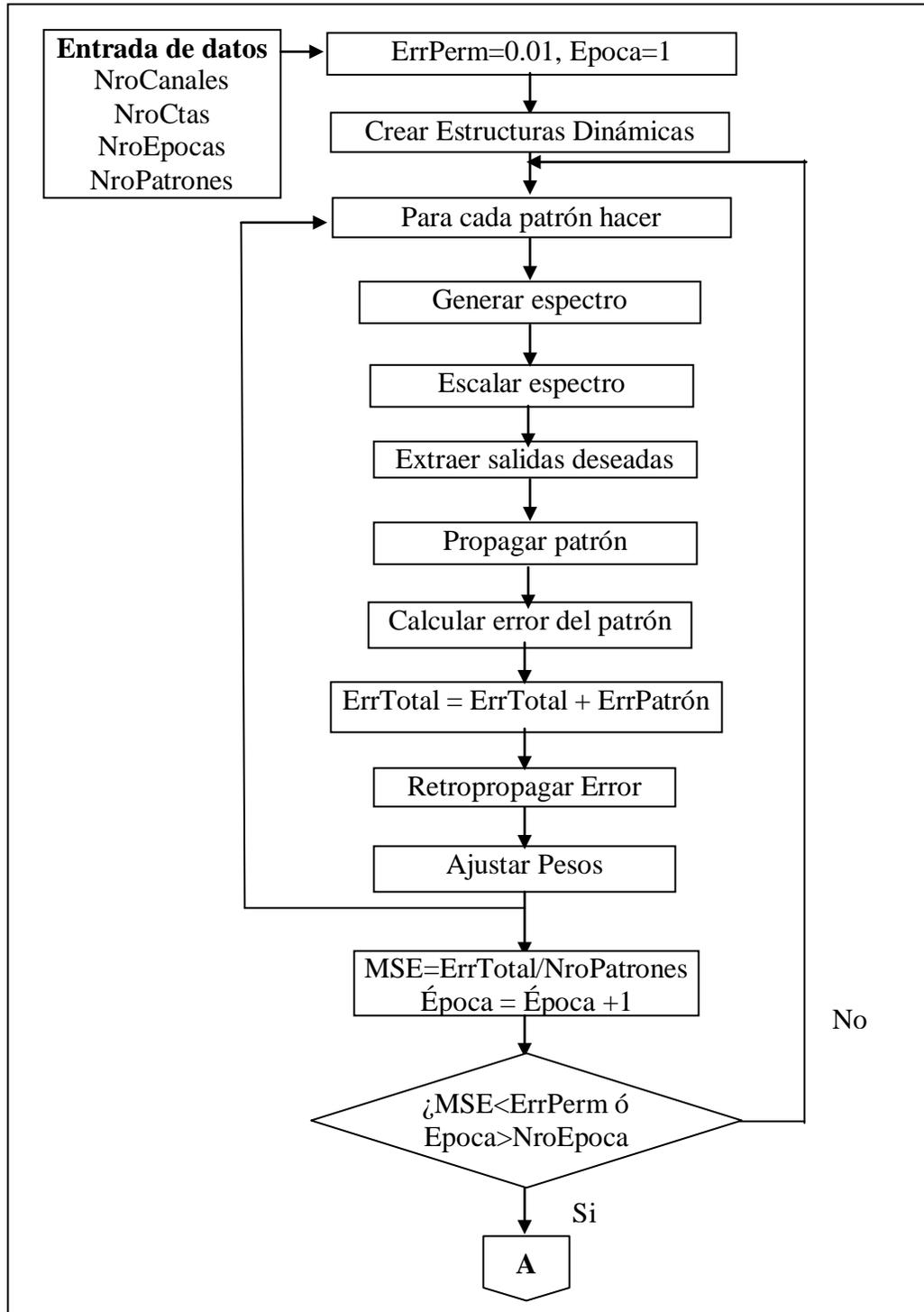


Figura 18. Algoritmo de Entrenamiento. Primer grupo de experimentos Córdoba(2005)

Del espectro generado se extraen las salidas deseadas para luego propagar las entradas a través de la red y así obtener las salidas de la misma. Para verificar el aprendizaje de la red es necesario calcular el error cuadrático medio (MSE) para todos los patrones procesados (época), a partir de las salidas obtenidas del procesamiento de cada patrón.

El error cometido en el patrón se obtiene como la sumatoria de las diferencias entre las salida deseada y obtenida en cada neurona elevadas al cuadrado y dividida este resultado entre dos. Este error se acumula para posteriormente calcular el error cometido por la red en cada época (MSE), dividiendo el error acumulado entre el número de patrones procesados.

El error mínimo permitido para el MSE se tomó, como es estándar, en 10^{-3} . El entrenamiento culmina cuando el error desciende hasta éste valor o cuando se alcanza el máximo número de épocas. El error MSE obtenido es reportado en pantalla y los pesos resultantes del aprendizaje son almacenados en un archivo de texto, para su uso posterior en la prueba de la red.

Se somete la red entrenada al proceso de prueba, con espectros no usados por la red en la fase de entrenamiento. Igual que en el proceso de entrenamiento, estos espectros son generados en el mismo momento de ser procesados. Como único dato de entrada para el proceso de prueba, solo se pide el número de patrones de prueba que se procesarán.

En el proceso de prueba se propagan los patrones usando los pesos obtenidos en el proceso de entrenamiento y calculando igualmente el error MSE, el cual es también reportado en pantalla. Los datos de los espectros usados en el proceso de prueba, son igualmente almacenados en un archivo de texto. En la figura 19 se muestra gráficamente el algoritmo del proceso de prueba.

Una vez culminado el proceso de prueba, es posible procesar un espectro con la red entrenada para verificar como se comporta, en este caso, son reportadas en pantalla las salidas deseadas, las salidas obtenidas y el error cometido en el cálculo de las intensidades y tiempos de vida de cada componente. Este proceso puede verse gráficamente en la figura 20.

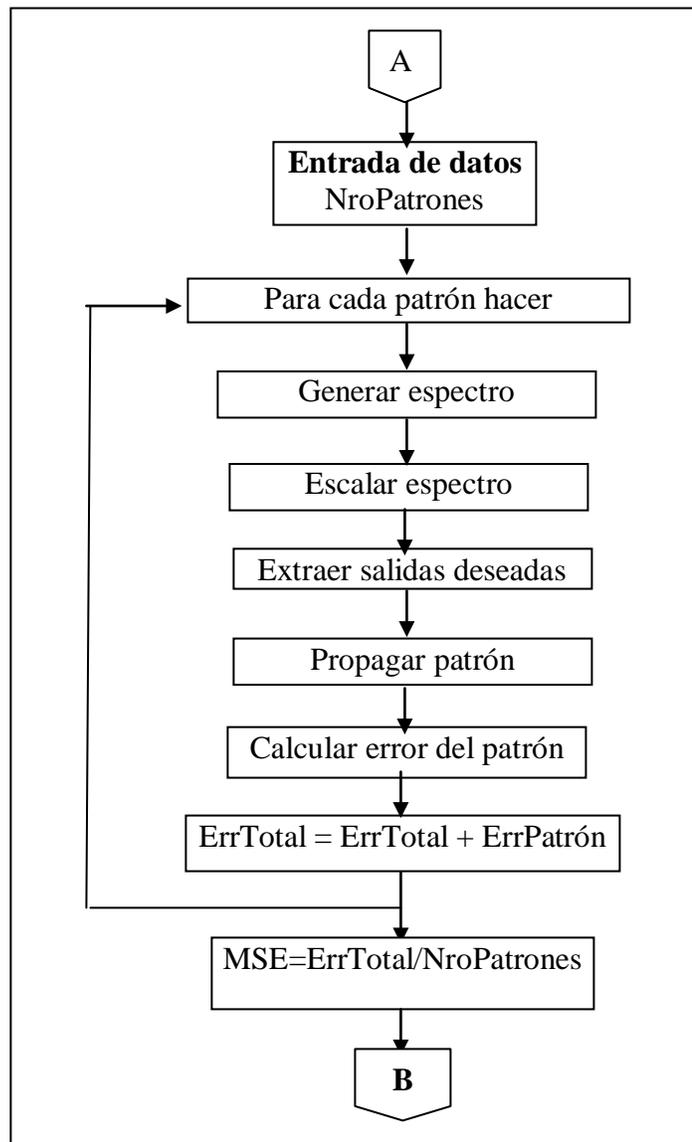


Figura 19. Algoritmo de Prueba. Primer grupo de experimentos.
Córdoba(2005)

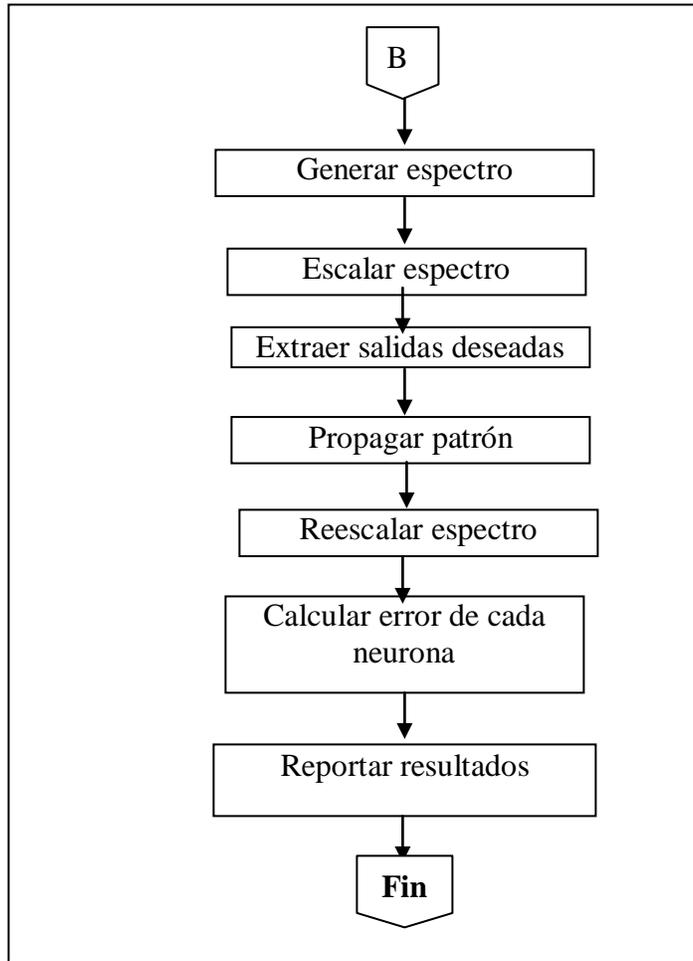


Figura 20. Procesamiento de un espectro. Primer grupo de experimentos Córdoba(2005)

Procesamiento del Segundo Grupo de Experimentos

El segundo grupo de experimentos se realizó entrenando la red con un mismo grupo de espectros previamente generados, escalados y almacenados en un archivo de texto. Esto permitió separar estas tareas en dos programas diferentes, lo cual reduce tanto el tiempo de entrenamiento, como los requerimientos de memoria, permitiendo experimentar con un mayor número de neuronas ocultas y por lo tanto, con varios modelos de red. Los algoritmos de entrenamiento, prueba y procesamiento de los espectros son similares, solo que los procesos de *Generar espectro* y *Escalar*

espectro, se reemplazan por un proceso de *Leer Patrón*, el cual lee el patrón de entrada desde un archivo de texto en el momento que va a ser procesado, tal como se muestra en la figura 21.

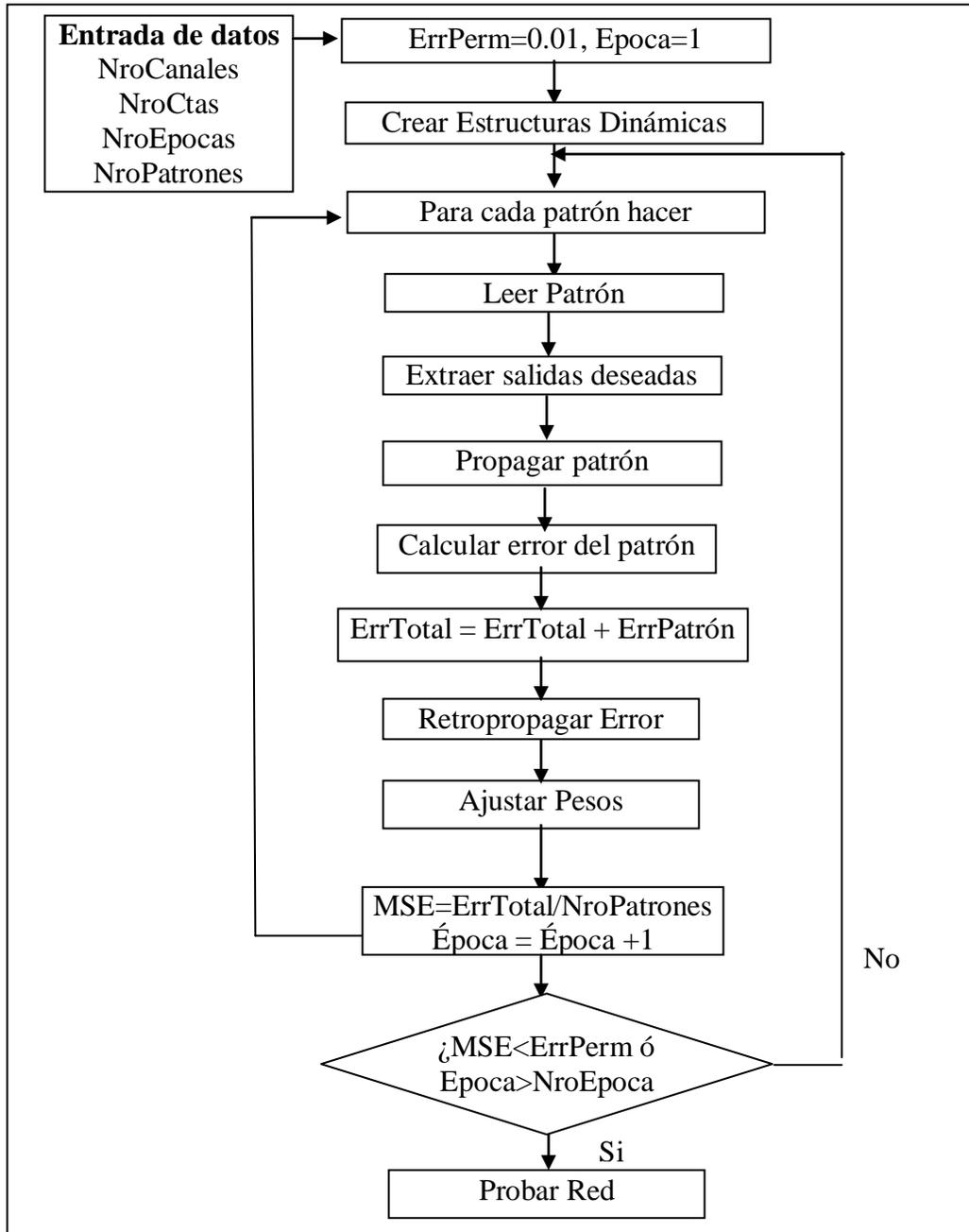


Figura 21. Algoritmo de Entrenamiento. Segundo grupo de experimentos Córdoba(2005)

Experimentos Realizados

Resultados del Primer Grupo de Experimentos

Como se mencionó anteriormente el diseño de la red neuronal establecido es producto de una serie de experimentos realizados variando los parámetros de la red. En todos los casos fueron procesadas 10^6 cuentas y en el proceso de prueba se usaron 50 patrones para validar el rendimiento de la red. En las tablas 3, 4 y 5 se muestran los datos de los experimentos realizados con los resultados obtenidos de los mismos.

Inicialmente, se probó implementar la red con 40 neuronas ocultas, tal como fue experimentado en investigaciones anteriores, sin embargo, los requerimientos de memoria para una red de tal magnitud, sobrepasan los recursos del equipo utilizado, ya que las dimensiones de las estructuras de datos utilizadas para los pesos dependen del número de canales y el número de neuronas ocultas. Es por ello que se comenzó a experimentar para establecer los límites de tales valores.

En el caso de 1500 canales, el valor límite para el número de neuronas ocultas es de una sola unidad, por lo que los experimentos para este caso se diseñaron en base a tales valores. En la tabla 3 se reflejan los resultados de diez de los experimentos más relevantes del total ejecutado. Allí se puede observar que aunque no se logra la convergencia al límite de error permitido, se obtienen errores cercanos a 0.1.

El menor error en el proceso de entrenamiento se logra en el experimento número ocho, con una tasa de aprendizaje de 0.9, el menor error en el proceso de prueba se logra en el experimento número 5 con una tasa de 0.5 y el menor error promedio en el procesamiento de un espectro también se obtiene en ese experimento, notándose sólo una diferencia de 0.004949 en el error de entrenamiento de dichos experimentos; es por ello que se toma como mejor resultado.

Tabla 3

Experimentos Realizados con 1500 canales y una neurona oculta en el primer grupo de experimentos.

Nro.	Nro de Patrones	Épocas	Tasa de Aprendizaje	Época de Convergencia	Error de Entrenamiento	Error de Prueba	Error absoluto						Promedio Error
							T ₁	T ₂	T ₃	I ₁	I ₂	I ₃	
1	575	100	0.25	-	0.137449	0.138045	0.00	0.03	0.91	9.13	9.39	18.13	6.26
2	575	50	0.25	-	0.138546	0.146088	0.00	0.00	0.47	5.40	12.4	4.67	3.80
3	800	50	0.25	-	0.138623	0.140764	0.02	0.07	0.93	11.05	5.41	1.46	3.15
4	700	100	0.25	-	0.140008	0.133821	0.02	0.09	0.43	17.01	11.76	21.46	8.46
5	575	50	0.5	-	0.141984	0.131527	0.02	0.03	0.37	3.79	6.70	2.68	2.26
6	800	50	0.5	-	0.141277	0.131605	0.00	0.02	0.81	2.55	7.20	6.58	2.87
7	1000	100	0.5	-	0.139399	0.141985	0.02	0.08	0.28	14.76	1.17	6.01	3.72
8	575	50	0.9	-	0.137035	0.134284	0.03	0.02	1.06	1.86	10.81	8.62	3.73
9	575	100	0.9	-	0.138190	0.145908	0.02	0.01	1.01	3.59	12.52	5.84	3.83
10	700	100	0.9	-	0.139812	0.147577	0.02	0.01	1.05	3.44	13.45	8.01	4.33

En el caso de 800 canales, el valor límite para las neuronas ocultas es de dos unidades. Igual como en el caso anterior, en la tabla 4 se reflejan los resultados de diez de los experimentos más relevantes.

Allí se puede observar que tampoco se logra la convergencia al límite de error de 0.01, el menor error en el proceso de entrenamiento se obtiene en el experimento número cuatro, con una tasa de aprendizaje de 0.25, el menor error en el proceso de prueba se logra en el experimento número nueve con una tasa de 0.9 y el menor error promedio en el procesamiento de un espectro se observa en el experimento número uno con una tasa de aprendizaje de 0.25, tal como en el número cuatro. Además se nota sólo una diferencia de 0.001059 en el error de entrenamiento de ambos experimentos y en el error de prueba la diferencia es de 0.028431, por lo que se toma el resultado del primer experimento como el mejor.

En el caso de 500 canales, el valor límite para las neuronas ocultas es de tres unidades y los resultados de los ocho experimentos más relevantes se reflejan en la tabla 5.

Aquí se observa que al igual que los experimentos anteriores no se logra la convergencia. El menor error en el proceso de entrenamiento se tiene en el experimento número seis, con una tasa de aprendizaje de 0.5, el menor error en el proceso de prueba se logra también es este experimento, sin embargo, el menor error promedio en el procesamiento de un espectro se obtiene en el experimento número uno con una tasa de aprendizaje de 0.25, notándose sólo una diferencia de 0.008172 en el error de entrenamiento de ambos experimentos y en el error de prueba la diferencia es de 0.022182, por lo que se toma el resultado del primer experimento como el mejor.

Tabla 4

Experimentos Realizados con 800 canales y dos neuronas ocultas en el primer grupo de experimentos

Nro.	Nro de Patrones	Épocas	Tasa de Aprendizaje	Época de Convergencia	Error de Entrenamiento	Error de Prueba	Error absoluto						Promedio Error
							T ₁	T ₂	T ₃	I ₁	I ₂	I ₃	
1	575	100	0.25	-	0.108133	0.135921	0.01	0.04	0.94	4.81	1.98	2.05	1.63
2	575	50	0.25	-	0.111572	0.108543	0.01	0.06	0.92	5.62	9.06	2.49	3.02
3	800	50	0.25	-	0.108392	0.129016	0.02	0.04	0.28	16.76	0.95	10.50	4.75
4	800	100	0.25	-	0.107074	0.129481	0.01	0.02	0.94	19.85	14.82	14.31	8.31
5	575	100	0.5	-	0.140002	0.149879	0.03	0.08	0.84	10.01	9.25	5.14	4.22
6	800	100	0.5	-	0.138807	0.150173	0.02	0.09	0.49	7.00	15.59	3.65	4.47
7	1000	100	0.5	-	0.110779	0.123041	0.02	0.04	1.24	28.5	2.07	0.53	5.40
8	575	50	0.9	-	0.110547	0.112454	0.01	0.08	0.39	3.21	7.01	5.24	2.65
9	575	100	0.9	-	0.111696	0.107490	0.00	0.08	0.34	1.92	7.61	3.87	2.30
10	800	50	0.9	-	0.111150	0.114205	0.01	0.01	0.37	10.03	9.90	0.81	3.52

Tabla 5

Experimentos Realizados con 500 canales y tres neuronas ocultas en el primer grupo de experimentos

Nro	Nro de Patrones	Épocas	Tasa de Aprendizaje	Época de Convergencia	Error de Entrenamiento	Error de Prueba	Error absoluto						Promedio Error
							T ₁	T ₂	T ₃	I ₁	I ₂	I ₃	
1	575	50	0.25	-	0.095972	0.098705	0.01	0.02	0.07	2.23	6.67	0.83	1.63
2	575	100	0.25	-	0.094930	0.083220	0.00	0.02	0.57	5.32	8.90	15.13	4.99
3	575	100	0.5	-	0.096646	0.102065	0.01	0.05	0.08	18.50	1.61	16.69	6.15
4	800	50	0.5	-	0.098101	0.103986	0.01	0.01	1.08	0.55	8.30	1.08	1.83
5	800	100	0.5	-	0.093135	0.112845	0.01	0.08	0.37	8.99	10.58	19.36	6.56
6	1000	100	0.5	-	0.087800	0.076523	0.02	0.03	0.28	7.45	11.51	18.85	6.35
7	800	50	0.9	-	0.101227	0.102772	0.01	0.01	0.41	0.30	10.45	0.98	3.08
8	1000	100	0.9	-	0.100021	0.112154	0.01	0.01	0.42	1.29	9.54	3.24	2.41

Resultados del Segundo Grupo de Experimentos

Como se mencionó en apartados anteriores, el segundo grupo de experimentos se realizó con un mismo conjunto de patrones generados previamente. En las tablas 6, 7 y 8 se muestran los datos de los experimentos realizados con los resultados obtenidos de los mismos.

Inicialmente, se probó implementar la red con la misma cantidad de neuronas ocultas que se usó en el primer grupo de experimentos, posteriormente, se fue incrementando para buscar un mejor rendimiento. Esto fue posible debido a la separación de las funciones de generación de espectros y entrenamiento de la red, con lo que se disminuyeron las necesidades de memoria de la aplicación.

En este grupo de experimentos, se logró la convergencia con errores menores al límite permitido de 0.01 en la mayoría de las corridas, por lo que el trabajo se centró en la búsqueda de la arquitectura que arrojara menor error en los procesos de prueba y procesamiento de espectros, variándose el número de neuronas ocultas, la cantidad de patrones de entrenamiento y la tasa de aprendizaje.

En el caso de 1500 canales, se experimentó con una, dos y tres neuronas ocultas. En la tabla 6 se pueden observar los resultados de los experimentos más relevantes. El menor error en el entrenamiento se obtuvo en el experimento 14, con 3 neuronas ocultas y tasa de aprendizaje de 0.5. El menor error en el proceso de prueba se obtuvo en el experimento 5, con 2 neuronas ocultas y tasa de aprendizaje de 0.25. el menor error promedio en el procesamiento de un espectro se obtuvo en el experimento número 16 con 3 neuronas ocultas y tasa de aprendizaje de 0.9. Si se toma la arquitectura de 3 neuronas ocultas y tasa de aprendizaje intermedia de 0.5, se observa que el error en la prueba difiere del menor obtenido en solo 0.076036, por lo que se seleccionó esta arquitectura para este caso.

Tabla 6

Experimentos Realizados con 1500 canales en el segundo grupo de experimentos.

Nro.	Neuronas Ocultas	Nro de Patrones	Épocas	Tasa Aprendizaje	Época de Convergencia	Error Entrenamiento	Error de Prueba	Error absoluto						Promedio Error
								T ₁	T ₂	T ₃	I ₁	I ₂	I ₃	
1	1	800	50	0.25	3	0.009289	0.559193	0.01	0.17	0.67	6.69	13.23	6.69	4.57
2	1	1000	50	0.25	2	0.008692	0.557250	0.01	0.17	0.67	6.63	13.17	6.64	4.54
3	1	800	50	0.5	2	0.009254	0.566594	0.01	0.17	0.68	6.79	13.33	6.77	4.62
4	1	800	100	0.9	2	0.008595	0.577684	0.01	0.17	0.69	6.93	13.47	6.89	4.69
5	2	575	50	0.25	22	0.007133	0.180408	0.02	0.08	0.44	7.24	0.71	4.91	2.23
6	2	800	50	0.25	3	0.089880	0.566961	0.00	0.15	0.43	4.02	10.68	4.34	3.27
7	2	575	50	0.5	14	0.009985	0.513185	0.02	0.09	0.41	6.59	0.12	4.32	1.92
8	2	800	50	0.5	2	0.008379	0.526300	0.02	0.08	0.40	6.78	0.21	4.53	2.00
9	2	575	50	0.9	8	0.009989	0.517109	0.02	0.08	0.47	7.56	1.01	5.20	2.39
10	2	800	50	0.9	2	0.008817	0.618445	0.01	0.17	0.63	6.35	12.80	6.21	3.30
11	3	575	50	0.25	6	0.009915	0.476200	0.02	0.09	0.28	5.06	1.15	3.13	1.62
12	3	800	50	0.25	2	0.009014	0.533396	0.00	0.16	0.53	4.64	11.43	5.41	3.69
13	3	575	50	0.5	4	0.009698	0.477533	0.02	0.09	0.40	6.74	0.22	4.50	1.99
14	3	800	50	0.5	4	0.004952	0.256524	0.02	0.08	0.50	7.82	1.47	5.39	2.54
15	3	575	50	0.9	3	0.009520	0.474375	0.02	0.08	0.44	7.20	0.66	4.89	2.21
16	3	800	50	0.9	1	0.009429	0.525872	0.01	0.12	0.04	0.91	5.36	0.31	1.12

En el caso de 800 canales, se experimentó con dos, tres y cuatro neuronas ocultas. En la tabla 7 se pueden observar los resultados de los experimentos realizados. El menor error en el entrenamiento se obtuvo en el experimento 4, con 2 neuronas ocultas y tasa de aprendizaje de 0.5. El menor error en el proceso de prueba y en el procesamiento del espectro se obtuvo en el experimento 14, con 4 neuronas ocultas y tasa de aprendizaje de 0.25, por lo que se toma esta arquitectura como la mejor para 800 canales.

En el caso de 500 canales los resultados de los experimentos más relevantes se reflejan en la tabla 8. Se experimentó con 3, 4, 5, 6, 7, 8 y 9 neuronas. Aunque no se logró la convergencia al límite de error permitido de 0.01, se obtuvieron valores muy cercanos al mismo, tanto para el error del proceso de entrenamiento, como en el error en los procesos de prueba.

El menor valor para el error de entrenamiento se observó en el experimento 8 con 8 neuronas ocultas y una tasa de aprendizaje de 0.25. Por otra parte, el menor error del proceso de prueba se obtuvo en el experimento número 6 con 7 neuronas ocultas y tasa de 0.25. el menor valor para el error en el procesamiento del espectro se observa en el experimento 3 con 5 neuronas ocultas y tasa de aprendizaje de 0.9, aunque el error de prueba es mayor que en los anteriores..

Entre el experimento 6 y el 8, la diferencia entre los errores es mínima, por lo que se toma la arquitectura del experimento 8 como la mejor, de acuerdo al comportamiento en el procesamiento de un espectro.

Tabla 7

Experimentos Realizados con 800 canales en el segundo grupo de experimentos

Nro.	Neuronas Ocultas	Nro de Patrones	Épocas	Tasa de Aprendizaje	Época de Convergencia	Error de Entrenamiento	Error de Prueba	Error absoluto						Promedio Error
								T ₁	T ₂	T ₃	I ₁	I ₂	I ₃	
1	2	575	50	0.25	2	0.008123	0.629011	0.03	0.11	1.33	4.24	10.11	11.46	4.54
2	2	800	50	0.25	2	0.005739	0.604044	0.03	0.11	1.31	3.88	9.93	11.16	4.40
3	2	575	50	0.5	2	0.006072	0.589302	0.03	0.10	1.28	3.75	9.61	11.18	4.32
4	2	800	50	0.5	2	0.004314	0.597764	0.03	0.10	1.31	3.94	4.78	11.04	3.53
5	2	575	50	0.9	3	0.006143	0.639317	0.03	0.11	1.34	4.34	10.29	11.56	4.61
6	2	800	50	0.9	1	0.093390	0.647956	0.03	0.11	1.35	4.43	10.38	11.64	4.65
7	3	575	50	0.25	2	0.007136	0.547392	0.03	0.10	1.25	3.18	9.13	10.94	3.50
8	3	800	50	0.25	2	0.005142	0.568934	0.03	0.10	1.24	3.08	9.53	10.49	4.07
9	3	575	50	0.5	2	0.005616	0.548748	0.03	0.13	1.22	3.26	9.25	10.79	4.11
10	3	800	50	0.5	2	0.004512	0.657853	0.03	0.11	1.36	4.57	10.52	11.75	4.49
11	3	575	50	0.9	2	0.005252	0.581096	0.03	0.10	1.28	3.51	9.24	10.86	4.17
12	3	800	50	0.9	1	0.008781	0.632643	0.03	0.11	1.34	4.33	10.09	11.45	4.55
13	4	575	50	0.25	4	0.072740	0.547311	0.03	0.10	1.25	3.00	8.45	10.43	3.87
14	4	800	50	0.25	4	0.004712	0.390138	0.02	0.08	1.03	0.12	5.67	8.39	2.55
15	4	575	50	0.5	2	0.005302	0.538610	0.03	0.10	1.24	2.90	8.59	10.27	3.85
16	4	800	50	0.5	2	0.003991	0.555658	0.03	0.10	1.23	3.43	8.99	10.77	4.09
17	4	575	50	0.9	1	0.009696	0.627188	0.03	0.11	1.33	4.14	10.21	11.34	4.52
18	4	800	50	0.9	1	0.007474	0.607843	0.03	0.11	1.31	3.94	9.88	11.24	4.20

Tabla 8

Experimentos Realizados con 500 canales en el segundo grupo de experimentos

Nro.	Neuronas Ocultas	Nro de Patrones	Épocas	Tasa de Aprendizaje	Época de Convergencia	Error de Entrenamiento	Error de Prueba	Error absoluto						Promedio Error
								T ₁	T ₂	T ₃	I ₁	I ₂	I ₃	
1	3	800	500	0.9	-	0.075773	0.078639	0.01	0.03	0.19	1.93	2.94	0.66	0.96
2	4	800	500	0.9	-	0.055972	0.050341	0.02	0.03	0.03	1.34	2.07	0.54	0.67
3	5	800	500	0.9	-	0.047818	0.043605	0.02	0.04	0.05	0.57	1.08	0.25	0.33
4	6	800	500	0.9	-	0.043580	0.040997	0.02	0.04	0.10	1.07	1.07	0.40	0.45
5	7	800	500	0.9	-	0.042916	0.038105	0.02	0.04	0.57	1.09	2.22	0.56	0.75
6	7	800	500	0.25	-	0.034706	0.030279	0.01	0.03	0.28	1.67	2.61	0.48	0.84
7	8	800	500	0.9	-	0.043719	0.040496	0.02	0.03	0.00	1.43	1.05	0.35	0.48
8	8	800	500	0.25	-	0.033661	0.030622	0.01	0.04	0.20	2.09	1.56	0.48	0.73
9	9	800	500	0.25	-	0.035102	0.031401	0.01	0.04	0.10	1.70	1.64	0.00	0.58
10	9	800	500	0.9	-	0.035838	0.031382	0.02	0.03	0.05	1.36	1.55	0.16	0.52
11	9	1000	500	0.9	-	0.359910	0.040758	0.00	0.03	0.20	1.84	2.06	0.69	0.80

Prueba de Generalización

Luego de establecer la configuración de la red más adecuada para cada caso de acuerdo a los resultados de los experimentos, se generaron algunos espectros con ruido, estableciendo características diferentes a los utilizados en el proceso de entrenamiento, esto con el propósito de analizarlos con la red y verificar el poder de generalización de la misma. Los parámetros variados fueron los rangos de las intensidades y tiempos de vida, el canal de tiempo cero y el ancho de canal. En la figura 22 se observa uno de los espectros generados en los cuales el canal de tiempo cero fue el canal 25 y los rangos de los parámetros se muestran en la tabla 9. En la tabla 10 se muestran los resultados de estas pruebas, donde las primeras seis se realizaron con espectros generados con los primeros rangos de la tabla 9 y los restantes con los otros rangos de la misma tabla.

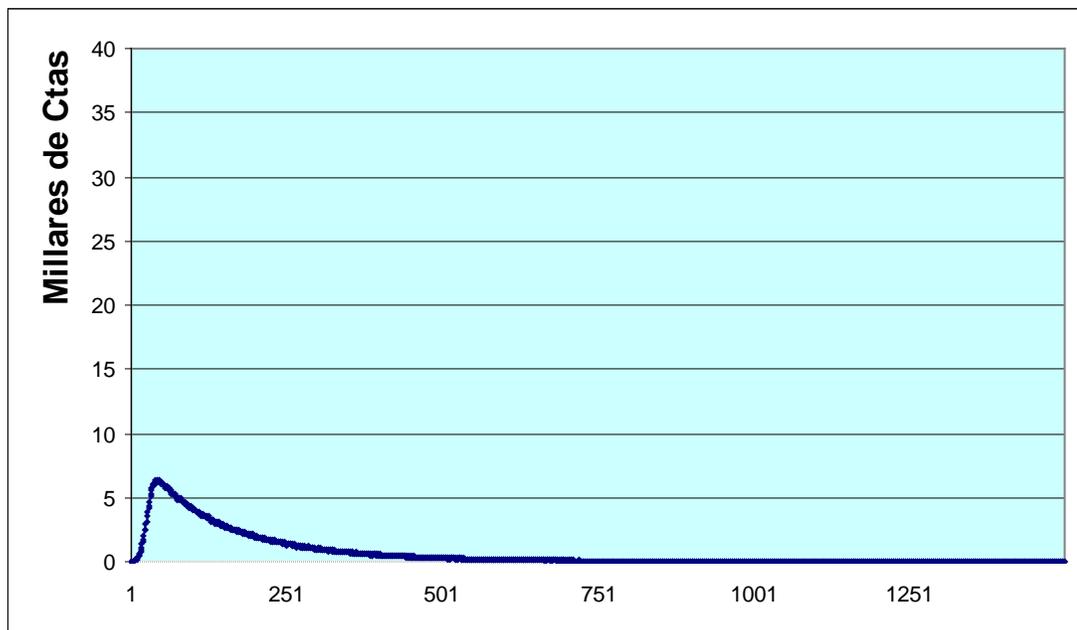


Figura 22 Espectro con ruido.
Córdoba(2005)

Tabla 9

Rango de Variación de los parámetros de los espectros con ruido

		Rango 1		Rango 2	
Primer	Intensidad	05	20	10	50
Componente	Tiempo de Vida	0.8	2.1	4.2	8.0
Segundo	Intensidad	15	30	40	45
Componente	Tiempo de Vida	2.25	4.44	0.4	1.5
Tercer	Intensidad	30	60	45	80
Componente	Tiempo de Vida	3.6	5.0	0.1	3.6

Tabla 10

Resultados del procesamiento de espectros con ruido.

Nro.	Número de Canales	Canal de tiempo cero	Ancho de Canal	Error absoluto						Promedio Error
				T ₁	T ₂	T ₃	I ₁	I ₂	I ₃	
1	1500	25	25.0	0.79	0.44	1.24	1.01	6.22	17.47	4.58
2	800	25	25.0	0.67	0.02	1.09	0.62	5.06	19.84	4.55
3	500	25	25.0	0.74	1.02	1.17	8.67	5.75	9.80	4.52
4	1500	10	25.0	0.72	0.03	0.33	0.49	6.58	17.19	4.22
5	800	10	25.0	0.25	1.75	1.07	1.01	6.28	17.27	4.60
6	500	10	25.0	0.30	1.18	0.44	9.43	6.92	12.62	5.14
7	1500	30	30.0	1.44	0.87	2.15	7.12	2.05	13.17	4.46
8	800	30	30.0	2.59	0.84	2.94	8.81	2.0	11.68	4.81
9	500	30	30.0	1.12	0.41	0.72	17.43	1.08	3.79	4.09
10	1500	40	30.0	2.56	0.68	1.41	4.55	2.03	15.93	4.52
11	800	40	30.0	2.51	0.83	1.18	5.52	2.00	15.05	4.51
12	500	40	30.0	0.25	0.37	0.13	20.89	1.08	6.91	4.93

Análisis y Comparación de Resultados

En los resultados de los experimentos realizados se puede observar en general una considerable mejoría con respecto a los resultados de investigaciones previas, tanto en reducción de la complejidad, como en la exactitud de los experimentos.

Específicamente, se redujo la complejidad de la red a no más de nueve neuronas ocultas, a diferencia de los trabajos de Pázsit y otros (1999) y Avdic y otros (2003), donde se utilizaron redes de retropropagación con 40 neuronas en la capa oculta. Esta reducción de la complejidad conlleva también a la reducción del tiempo de entrenamiento e igualmente a la menor necesidad de requerimientos de recursos de memoria.

En cuanto a la exactitud de la red, se logra la convergencia en la mayoría de los experimentos al límite permitido de 1%. Sólo en el caso de 500 canales el error mínimo obtenido fue de 3%, sin embargo, aún así el rendimiento es mayor que en el trabajo de Pázsit y otros (1999), donde el límite de error fue establecido en 5%, y en el trabajo de Avdic y otros (2003) donde el error es de 10%.

La convergencia se logra rápidamente en las primeras épocas, esto puede ser originado debido al preprocesamiento al que son sometidos los espectros antes de introducirlos en la red, para llevarlos a una escala común entre los valores cero y uno. El rendimiento obtenido también es producto de la variante utilizada añadiendo un término de momento para acelerar el proceso de aprendizaje. Éste término hace que la actualización de los pesos en una iteración dependa parcialmente de la actualización de la iteración anterior para mantener la misma dirección del descenso de gradiente en la superficie del error. Esto puede tener el efecto de incrementar el tamaño del paso de la búsqueda en regiones donde el gradiente cambia poco.

Como pudo ser observado en apartados anteriores, la red generaliza la relación aprendida para procesar espectros con otras distribuciones de datos con respecto a los rangos de intensidades y tiempos de vida, así como también con las variaciones del canal de tiempo cero y ancho de canal, demostrando además que la red es tolerante a datos con ruido.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

Este trabajo logró cumplir con todos los objetivos propuestos de la investigación realizada, dando así solución al problema estudiado mediante el diseño y desarrollo de una red neuronal artificial tipo retropropagación para el análisis de espectros simulados, tal como fue planteado en el objetivo general del proyecto.

La red diseñada permite estimar las intensidades y tiempos de vida de los tres componentes de espectros analizados, con un rendimiento que mejora los resultados de investigaciones anteriores, tanto en complejidad como en exactitud, contribuyendo así a nutrir la novedosa línea de investigación de aplicación de redes neuronales artificiales para el análisis de espectros de tiempo de vida de aniquilación de positrones. Para lograr el objetivo general anteriormente descrito, se cumplieron con todos los objetivos específicos planteados.

Se desarrolló un programa de simulación que permite generar espectros de forma automática, tomando en cuenta todas las características asociadas a los mismos, de tal forma que dichos espectros representan con bastante similitud la distribución de datos de espectros reales. Esto fue posible por la utilización del método de Montecarlo, verificándose que el mismo es adecuado para la simulación de procesos de naturaleza aleatoria y caracterizados por gran cantidad de eventos, tal como es el caso de los espectros estudiados. De igual forma, se concluye acerca de la gran utilidad de los procesos de simulación cuando no se tienen datos de la fuente primaria.

Se logró diseñar una red neuronal artificial sencilla con pocas neuronas ocultas, con lo que se pudo constatar la posibilidad de aplicación de las mismas en el análisis de espectros simulados de tiempos de vida de aniquilación de positrones, demostrándose que esta red neuronal perceptrónica basada en retropropagación del error aproxima la función no lineal asociada a la generación de los espectros.

Se diseñaron y llevaron a cabo una cantidad suficiente de experimentos en cada uno de los casos planteados, es decir, para 1500, 800 y 500 canales, variando diversos parámetros tales como el número de neuronas ocultas, la tasa de aprendizaje, cantidad de patrones de entrenamiento, entre otros. Del total de experimentos realizados se tomaron en cuenta para los procesos de entrenamiento y prueba, alrededor de 73 experimentos que se consideraron de gran relevancia para la selección de la arquitectura de red adecuada. También se diseñaron experimentos para verificar la capacidad de generalización de la red y la tolerancia al ruido, generando y procesando espectros con características diferentes a los aprendidos por la red. Esto fue posible, variando los rangos de valores para las intensidades y tiempos de vida de aniquilación, cambiando también el canal de tiempo cero y el ancho de canal.

Se logró realizar el entrenamiento de la red con los espectros generados, en tiempos aceptables, mejorando su desempeño respecto a trabajos anteriores. Igualmente, se pudo determinar que el rendimiento de la red permite establecer las intensidades y tiempos de vida de los espectros con un margen de error igualmente aceptable, disminuyendo considerablemente con respecto a otras investigaciones de la misma índole.

En cada uno de los experimento se llevó a cabo la validación de la red, sometiénola al proceso de espectros no vistos en el entrenamiento, logrando obtener errores que permiten comparar las diferentes arquitecturas, para realizar la selección de la más adecuada en cada caso.

La realización de los diferentes experimentos y la comparación de sus resultados llevó a establecer la configuración óptima de la red neuronal, en base al número de neuronas ocultas y la tasa de aprendizaje. De igual forma, se verificó la utilidad del preprocesamiento de los datos y de agregar un factor de momento para acelerar el proceso de convergencia de la red.

Con respecto a las características de microcomputador usado, se pudo probar la posibilidad de su utilización sin menoscabar el rendimiento obtenido, lo cual permite suponer que en máquinas con mayor disponibilidad de recursos, esta aplicación puede arrojar mejores resultados.

Recomendaciones

Se recomienda el entrenamiento y prueba de la red con datos experimentales para verificar su eficacia y eficiencia en el procesamiento de datos reales, y así poder aplicarla en casos prácticos.

De igual forma, la herramienta debe ser probada en otras plataformas de sistema operativo, como por ejemplo Unix, para verificar su portabilidad y funcionamiento. También puede ser probada en máquinas más potentes donde pueda mejorar su rendimiento aún más.

La herramienta puede ser mejorada, en cuanto a su aspecto para presentar una interfaz gráfica mas amena.

Se puede mejorar la flexibilidad de la aplicación, permitiendo que el usuario pueda analizar espectros con diferentes números de canales y con diferentes rangos de valores para las intensidades y tiempos de vida, entre otros, que puedan ser establecidos de acuerdo al material que se someta a estudio.

INDICE DE TABLAS

Tabla	Pag
1. Estimación de gastos.....	45
2. Rango de Variación de los parámetros de los espectros.....	49
3. Experimentos Realizados con 1500 canales y una neurona oculta. Primer grupo de experimentos.....	65
4. Experimentos Realizados con 800 canales y dos neuronas ocultas Primer grupo de experimentos.....	67
5. Experimentos Realizados con 500 canales y tres neuronas ocultas Primer grupo de experimentos.....	68
6. Experimentos Realizados con 1500 canales. Segundo grupo de experimentos.....	70
7. Experimentos Realizados con 800 canales. Segundo grupo de experimentos.....	72
8. Experimentos Realizados con 500 canales. Segundo grupo de experimentos.....	73
9. Rango de Variación de los parámetros de los espectros con ruido.....	75
10. Resultados del Procesamiento de espectros con ruido.....	75

INDICE DE FIGURAS

Figura	Pag
1. Proceso de creación de Pares.....	22
2. Aniquilación electrón-positrón.....	22
3. Espectro de tiempo de vida de positrones.....	25
4.- Distribución de probabilidad de Poisson para $\mu=10 / 6$	26
5.- Distribución de probabilidad de Gauss.	27
6. Red Neuronal Artificial Multicapa de conexión hacia delante.	29
7. Funciones de activación más comunes.....	30
8. Red Neuronal Artificial con una sola capa oculta.	31
9. Red con conexiones hacia atrás.....	32
10. Ejemplos de Espectros simulados.	51
11. Algoritmo de Simulación de los Espectros.....	52
12.. Red Neuronal Artificial para 1500 canales. Primer grupo de experimentos.....	54
13.. Red Neuronal Artificial para 800 canales. Primer grupo de experimentos.....	54

14.. Red Neuronal Artificial para 500 canales. Primer grupo de experimentos.....	55
15.. Red Neuronal Artificial para 1500 canales. Segundo grupo de experimentos.....	55
16.. Red Neuronal Artificial para 800 canales. Segundo grupo de experimentos.....	56
17.. Red Neuronal Artificial para 500 canales. Segundo grupo de experimentos.....	56
18.. Algoritmo de Entrenamiento. Primer grupo de experimentos.....	59
19.. Algoritmo de Prueba. Primer grupo de experimentos.....	61
20.. Procesamiento de un espectro. Primer grupo de experimentos.....	62
21.. Algoritmo de Entrenamiento. Segundo grupo de experimentos.....	63
22.. Espectro con ruido.....	74

INDICE

	Página
DEDICATORIA.....	iv
AGRADECIMIENTOS.....	v
INDICE GENERAL.....	vi
INDICE DE TABLAS.....	ix
INDICE DE FIGURAS	x
RESUMEN	xii
INTRODUCCIÓN.....	01
CAPITULO	
I. EL PROBLEMA	04
Planteamiento del Problema.....	04
Objetivo General.....	13
Objetivos Específicos	13
Justificación e Importancia.....	14
Alcance y Limitaciones	15
II. MARCO TEÓRICO.....	17
Antecedentes.....	17
Bases Teóricas	20
Caracterización de Defectos	20
Aniquilación Positrón/Electrón	21
Espectroscopia por Aniquilación de Positrones	23
Señal de Fondo o Background	25
Redes Neuronales Artificiales	27

Paradigma de Aprendizaje	32
Algoritmo de Retropropagación de Error	33
Método de Montecarlo	36
Glosario de Términos Básicos.....	39
III. MARCO METODOLÓGICO.....	42
Tipo de Investigación	42
Fase de Diagnóstico.....	42
Diseño de la Investigación o Procedimiento	42
Técnicas e Instrumentos de Recolección de Información	43
Técnicas de Análisis de los Datos	43
Estudio de Factibilidad	44
Factibilidad Técnica.....	44
Factibilidad Operativa.....	44
Factibilidad Económica.....	45
IV. PROPUESTA DEL ESTUDIO.....	46
Justificación.....	46
Objetivos	47
Objetivo General.....	47
Objetivos Específicos.....	47
Descripción de la Propuesta.....	47
Aspectos Generales.....	47
Procedimiento de Simulación.....	49
Diseño de la Red.....	53
Procedimientos de Entrenamiento y Prueba	58
Procesamiento del primer grupo de experimentos	58
Procesamiento del segundo grupo de experimentos	62
Experimentos Realizados	64
Resultados del primer grupo de experimentos	64

Resultados del segundo grupo de experimentos	69
Prueba de Generalización.....	74
Análisis y Comparación de Resultados.....	76
V. CONCLUSIONES Y RECOMENDACIONES.....	78
Conclusiones.....	78
Recomendaciones.....	80
 REFERENCIAS BIBLIOGRAFICAS.....	 81
ANEXOS.....	84
Carta de Aceptación del Tutor	
Listado del Programa..	

ANEXOS

UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"

ANÁLISIS DE ESPECTROS SIMULADOS DE TIEMPO DE VIDA DE
POSITRONES MEDIANTE REDES NEURONALES ARTIFICIALES

SONIA MARÍA CÓRDOBA FREITEZ

Barquisimeto, Febrero de 2005

UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCIÓN: INTELIGENCIA ARTIFICIAL

ANÁLISIS DE ESPECTROS SIMULADOS DE TIEMPO DE VIDA DE
POSITRONES MEDIANTE REDES NEURONALES ARTIFICIALES

Trabajo presentado para optar al grado de
Magíster Scientiarum

Por: SONIA MARÍA CÓRDOBA FREITEZ

Barquisimeto, Febrero de 2005

INDICE

DEDICATORIA

A la Esencia Divina, fuente de todo conocimiento

A mi amado hijo Gilbertson, principal motivo de mis luchas

A la familia que me cobijó en su seno con Amor

A mis amigos que me brindan su cariño

RESUMEN

UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCIÓN: INTELIGENCIA ARTIFICIAL

**ANÁLISIS DE ESPECTROS SIMULADOS DE TIEMPO DE VIDA DE
POSITRONES MEDIANTE REDES NEURONALES ARTIFICIALES PARA
EL ESTUDIO DE DEFECTOS EN MATERIALES**

Autor: Ing. Sonia María Córdoba Freitez

Tutores: Ing. Maritza Bracho

Dr. Freddy Torrealba

RESUMEN

La presente investigación tiene como objetivo desarrollar un modelo de Red Neuronal Artificial que permita estimar los tiempos de vida e intensidades de aniquilación a partir de espectros simulados de tiempo de vida de partículas llamadas positrones, técnica que es usada para la caracterización de defectos en materiales. Estos espectros obtenidos como el número de aniquilaciones (cuentas) por canal de tiempo, serán generados usando el método para simulación de variables aleatorias de Montecarlo, con un máximo de tres componentes, que esta asociado a un máximo de dos defectos, registrando aproximadamente 10^6 cuentas en 1500 canales por espectro. A través de una investigación documental y de entrevistas no estructuradas con expertos se pudo constatar que el uso de redes neuronales en este tipo de problema ha sido poco investigado, por lo que constituye un área de investigación relativamente novedosa. Dado que el trabajo se enmarca dentro de la modalidad de proyecto factible, se verifica su viabilidad técnica, económica y operativa; posteriormente se procede a la elaboración de la propuesta para lo cual, se diseña una red neuronal multicapa con un máximo de 1500 entradas, 6 salidas y aprendizaje supervisado basado en el algoritmo de retropropagación del error; que se entrenó y probó usando los datos simulados. La red se entrena para procesar espectros de 1500, 800 y 500 canales. La red converge en cada caso usando periodos de tiempo no mayores a 20 minutos con una exactitud aceptable, mejorando los resultados de investigaciones anteriores.

Palabras clave: espectros de tiempo de vida del positrón, redes neuronales artificiales, espectroscopia de tiempo de vida de aniquilación de positrones, método de Montecarlo.

REFERENCIAS BIBLIOGRAFICAS

Avdic Senada, Chakarova Roumina, Pazsit Imre.(2003). Analysis of the Experimental Positron Lifetime Spectra By Neural Networks. Nuclear Technologies & Radiation Protection. Vol. 18, Issue 1, pp.16-21

Bondarenko Vladimir (2004). Positron Spectroscopy 1.2. Positron laboratory. Positron Annihilation at Martin-Luther-University Halle. Germany. URL:<http://positronAnnihilation.net>. Consulta (Octubre 2004)

Daintith, John (1998). Diccionario de Física. Grupo Editorial Norma S.A. Colombia.

Eisberg – Resnick(2002). Física Cuántica. Átomos, Moléculas, Sólidos, Núcleos y Partículas. Limusa Noriega Editores. México.

Fantoni Paolo, Mazzola Alessandro.(1995). Accuracy Estimate of Artificial Neural Networks Based Models for Industrial Applications. Artificial Intelligence in the Petroleum Industry. Symbolic and Computational Applications 2, cap. 17. B.Braunschweig & B.Bremdal Editions. Halden, Norway.

Garcia Carmen, Villatoro Francisco (2004). Técnicas de simulación mediante el método de Montecarlo. URL: <http://polaris.lcc.uma.es/~villa/mmtc/tema14.pdf>. Consulta (Junio 2004)

Gebauer J., Krause-Rehberg R, Domke C., Ph. Ebert, and K. Urban (1997). Positron annihilation and scanning tunneling microscopy used to characterise defects in highly Si-doped GaAs. Materials Science Forum, Vol. 258-263. Issue 19, pp. 885-892

Geocities (2004). Dirac y la Antimateria. URL:http://www.geocities.com/id_imaginedream/antimateria.htm (Consulta, Junio 2004).

Hilera González, José R. y Martínez Hernando, Víctor J. (1995). Redes Neuronales Artificiales. Fundamentos, modelos y aplicaciones. Addison-Wesley Iberoamericana. España.

Hines Wesley, Gribok Andrei, Attieh Ibrahim, and Uhrig Robert (1999). Regularization Methods for Inferential Sensing in Nuclear Power Plants. Real-Time Systems Vol. 27, Issue 1, pp. 7 - 26

Joran, Peterson, Holmstrom Kenneth. (1998). A Review of the Parameters Estimation Problem of Fitting Positive Exponential Sums to Empirical Data. Applied Mathematics. Technical Report 08. pp. 2-30.

LFN (2003), Espectroscopía con un detector NaI(Tl). Laboratorio de Física Nuclear. Facultad de Física - Universidad de Valencia. URL: http://evalu29.ific.uv.es/lab_fan_03/espectroscopia.pdf

MathWorks, Inc. (1998), URL: <http://www.MathWorks.com>. (Consulta Marzo, 2003)

Mendenhall W, Sincich T.(1997) Probabilidad y Estadística para Ingeniería y Ciencias. Prentice Hall Hispanoamericana, S.A. México.

Mitchell, Tom. (1997), Machine Learning. Artificial Neural Networks. Cap. 4 McGraw-Hill.

MLUH (2004). Martin-Luther-University Halle. Positron Lifetime Spectroscopy. URL:positron.physik.uni-halle.de/panet/techniques/Positron%20Lifetime%20Spectroscopy.pdf. (Consulta, Junio 2004)

Montesinos Gaspar (2001). Un algoritmo simple para la simulación del paso de fotones ($E=1$ MeV) a través de la materia. URL: <http://www.gae.ucm.es/fisatom/docencia/trabajos/gaspar/trabajo-gaspar.pdf>. (Consulta Julio, 2004).

Morales Alvaro. (2004). La Ciencia de Simular Propiedades de Materiales. Rev. Acad. Colomb. Cienc. Vol. 28 Issue 107, pp. 253-263.

Müller, B. and Reinhardt, J.,(1990) Neural Networks: An Introduction, Springer-Berlin.

Nanopos (2004). Positron research website. University of michigan. URL:<http://positrons.physics.lsa.umich.edu/nanopos/PALS-intro/index.htm>. Consulta Octubre 2004)

Nilsson. Nils J (2001). " Inteligencia Artificial. Una nueva síntesis". McGrawHill. España

Pázsit, I.; Chakarova, R.; Lindén, P. Maurer (1999). Unfolding positron lifetime spectra with neural networks- Applied Surface Science, Volume: 149, Issue: 1-4, pp.97-102.

Pázsit, Imre. Garis, Ninos. Sandberg Urban, Andersson Tell. (1998). Determination of PWR Control Rod Position By Core Physics And Neural Network Methods. Nuclear Technology Vol. 123, pp. 278 – 295.

PDG (2000). Particule Data Group. Aniquilación Positron/electron. URL:<http://particleadventure.org/particleadventure/spanish/edds.html>. (Consulta, Junio 2004)

Rich, Elaine. Knight, Kevin. Inteligencia Artificial (1994). McGrawHill. España.

Shiva, M. Kishi. T. (2003) NDT-based Assessment of Damage: An Overview. Encyclopaedia of Materials: Science and Technology. Editors-in-Chief. pp. 1-8.

SOBOL, L.M. Lecciones populares de matemáticas. Método de Montecarlo. Editorial MIR. Moscú, 1976.

U.C.L.A, 2002. Manual para la presentación de Trabajos de Especialización, Maestría y Doctorado.

Wikipedia. (2004). Enciclopedia libre online. URL: <http://en.wikipedia.org/wiki>. Consulta (Julio 2004)