

UNIVERSIDAD DON BOSCO



**SISTEMA DE RECONOCIMIENTO DE PATRONES RITMICOS PARA
BATERIA ACÚSTICA**

TRABAJO DE GRADUACIÓN
PREPARADO PARA LA FACULTAD DE INGENIERÍA
PARA OPTAR AL GRADO DE
INGENIERO EN ELECTRÓNICA

POR:
ALFONSO JOSE ARTIGA LANDAVERDE
GERARDO FRANCISCO ESCOBAR FLORES

JULIO DE 2006

CIUDADELA DON BOSCO — EL SALVADOR — AMÉRICA CENTRAL

UNIVERSIDAD DON BOSCO

RECTOR

ING. FEDERICO MIGUEL HUGUET RIVERA

SECRETARIO GENERAL

LIC. MARIO RAFAEL OLMOS

DECANO DE LA FACULTAD DE INGENIERÍA

ING. ERNESTO GODOFREDO GIRÓN

ASESOR DEL TRABAJO DE GRADUACIÓN

ING. EDUARDO RIVERA

JURADO EVALUADOR

ING. JORGE LÓPEZ

ING. NÉSTOR LOZANO

ING. JUAN CARLOS CRUZ DADA

TUTOR DEL TRABAJO DE GRADUACIÓN

ING. WENCESLAO RIVAS

Universidad Don Bosco

Facultad de Ingeniería

Escuela de Ingeniería Electrónica

Jurado Evaluador del Trabajo de Graduación



Ing. Jorge López

Jurado



Ing. Néstor Lozano

Jurado



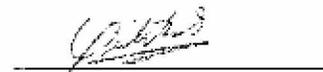
Ing. Juan Carlos Cruz Dada

Jurado



Ing. Eduardo Rivera

Asesor



Ing. Wenceslao Rivas

Tutor

Agradecimientos

Gerardo Francisco Escobar Flores

Deseo agradecer en primer lugar a Dios, quien me brindo la perseverancia para poder terminar la carrera además de la compañía en aquellos momentos difíciles de este proceso y por la protección durante este camino.

A mis padres, Ángela Maria y Francisco, quienes con su esfuerzo me brindaron los medios necesarios para poder culminar esta empresa, por el apoyo moral y económico que me brindaron desinteresadamente a través de estos años, por sus consejos sabios que me ayudaron a tomar mejores decisiones y por las veces que se desvelaron a mi lado.

A mi hermano mayor Diego por los ánimos y sus consejos oportunos y su la ayuda que siempre me presto.

Quisiera agradecer a nuestro asesor Eduardo Rivera que siempre estuvo presto a dirigir este proyecto y proporcionar su tiempo.

Por último deseo compartir este triunfo con Dios y con mi familia ya que por mi cuenta no hubiera podido culminar este proyecto.

Alfonso José Artiga Landaverde

Le agradezco a Dios y a la Virgen por haberme iluminado en el camino y por darme siempre las fuerzas necesarias para sobreponerme a las diferentes pruebas y obstáculos que encontré a lo largo de mi carrera.

A mi papá Alfonso Artiga por sus oportunos y valiosos consejos, dirección y ayuda. A mi ejemplo a seguir, mi fortaleza, consejera y maestra, Mi Madre, Gloria E. Landaverde de Artiga, por todos los sacrificios, desvelos, comprensión, paciencia y sobre todo por el apoyo incondicional que desde siempre me ha regalado.

De igual forma agradezco a mi familia, mis hermanas, Ana Gloria y Flor de María, mis tías y primos por su apoyo y aliento, por todas esas palabras de ánimos que llegaron en los momentos más indicados, por todos aquellos favores con los que de una forma u otra me ayudaron a concretar este triunfo académico.

Asimismo, a la familia Rivera, en especial a Eduardo, por la dedicación, orientación y exigencias que nos brindó para la exitosa conclusión del proyecto.

A mis amigos quienes siguieron el desarrollo del proyecto y encontraron la forma de animarme en los momentos difíciles. Gracias Meli por el diseño y a todos aquellos los que directa o indirectamente colaboraron para que este logro sea una realidad.

Finalmente quiero dedicar este triunfo a Mi Familia y a Dios sin Él nada de esto hubiera sido posible.

INDICE

| | |
|----------------------------------------------------------|----|
| Definición del Tema..... | 1 |
| Justificación del Tema | 1 |
| Objetivos..... | 2 |
| Objetivo General | 2 |
| Objetivo Específicos..... | 2 |
| Alcances y Limitaciones..... | 3 |
| Alcances..... | 3 |
| Limitaciones..... | 4 |
| 1 Conceptos Musicales..... | 6 |
| 1.1 El Pentagrama..... | 6 |
| 1.1.1 Las Líneas Adicionales..... | 6 |
| 1.2 Las Claves..... | 6 |
| 1.2.1 La Clave De Sol..... | 6 |
| 1.2.2 La Clave De Fa..... | 7 |
| 1.3 Notas Musicales..... | 8 |
| 1.4 Figuras Musicales..... | 8 |
| 1.4.1 Silencios..... | 9 |
| 1.5 Alteraciones..... | 10 |
| 1.6 Compases..... | 10 |
| 1.7 Ritmos Secundarios..... | 11 |
| 1.7.1 La Ligadura De Prolongación..... | 11 |
| 1.7.2 El Puntillo y el Doble Puntillo..... | 12 |
| 1.7.2.1 El Puntillo..... | 12 |
| 1.7.2.2 El Doble Puntillo..... | 12 |
| 1.8 Introducción a la Lectura Musical De La Batería..... | 13 |
| 1.8.1 Notación en el Pentagrama..... | 13 |
| 2. Redes Neuronales..... | 14 |
| 2.1 Redes Neuronales Biológicas..... | 14 |
| 2.2 Historia de Las Redes Neuronales..... | 16 |
| 2.3 Redes Neuronales Artificiales..... | 19 |

| | |
|---------------------------------------------------------------------|----|
| 2.3.1 Analogía con las Redes Neuronales Biológicas..... | 19 |
| 2.3.2 Función de Propagación (de red o de base)..... | 20 |
| 2.3.2.1 Función Lineal de Base (LBF) | 21 |
| 2.3.2.2 Función Radial de Base (RBF)..... | 21 |
| 2.3.3 Función de Activación..... | 21 |
| 2.3.3.1 Función Lineal..... | 22 |
| 2.3.3.2 Función Sigmoidal Unipolar..... | 23 |
| 2.3.3.3 Función Tangente Hiperbólica..... | 24 |
| 2.3.4 Estructura y Formas de Interconexión..... | 25 |
| 2.4 Clasificación de las Redes Neuronales..... | 26 |
| 2.4.1 Modelos Supervisados..... | 28 |
| 2.4.1.1 El Perceptrón..... | 28 |
| 2.4.1.2 El Perceptrón Multicapa..... | 28 |
| 2.4.1.3 Adaline y Madaline..... | 30 |
| 2.4.1.4 Retropropagación..... | 30 |
| 2.4.2 Modelos No Supervisados..... | 31 |
| 2.4.2.1 Redes de Hopfield..... | 31 |
| 2.4.2.2 Los Mapas de Kohonen..... | 32 |
| 2.4.2.3 Redes Basadas en la Teoría de la Resonancia Adaptativa..... | 33 |
| 2.5 Resumen de Tipos y Características Importantes..... | 34 |
| 3. Obtención de la Señal y Preprocesado..... | 37 |
| 3.1 Archivos Wave..... | 37 |
| 3.2 Filtros..... | 38 |
| 3.2.1 Sistemas Discretos..... | 38 |
| 3.2.2 Muestreo de Señales..... | 40 |
| 3.2.3 Filtros Digitales..... | 41 |
| 3.3 Compresión de Datos..... | 44 |
| 3.3.1 Transformada de Fourier..... | 45 |
| 3.3.2 Transformada Corta de Fourier..... | 45 |
| 3.3.3 Transformada Wavelet..... | 46 |
| 3.3.4 Familias Wavelet..... | 48 |

| | |
|-------------------------------------------------------------------------------|-----|
| 3.3.5 Descomposición de la Señal usando la Transformada Discreta Wavelet..... | 49 |
| 3.4 Procesado de la Señal | 56 |
| 3.4.1 Coeficientes de Predicción Lineal..... | 56 |
| 3.4.2 Momentos Estadísticos..... | 59 |
| 3.4.3 Máximos y Rangos de Frecuencia..... | 65 |
| 4. Selección de la Red Neuronal a Utilizar | 69 |
| 4.1 Redes seleccionadas para realizar las pruebas..... | 70 |
| 4.1.1 Kohonen..... | 70 |
| 4.1.1.1 Algoritmo de Aprendizaje..... | 72 |
| 4.1.1.2 Pruebas con Red Kohonen..... | 73 |
| 4.1.2 Retropropagación..... | 75 |
| 4.1.2.1 Algoritmo de Aprendizaje..... | 77 |
| 4.1.2.2 Pruebas con Red de Retropropagación..... | 79 |
| 4.1.2.2.1 Validación de Pruebas, Batería en conjunto..... | 80 |
| 4.1.2.2.2 Validación de Pruebas, Instrumentos Individuales..... | 82 |
| 4.1.3 Redes de Teoría de Resonancia Adaptativa 2 (ART2)..... | 88 |
| 4.1.3.1 Algoritmo de Aprendizaje..... | 90 |
| 4.1.3.2 Pruebas con Red de ART2..... | 93 |
| 4.1.3.2.1 Validacion de Pruebas, Bateria en conjunto..... | 93 |
| 4.1.3.2.2 Validacion de Pruebas, Instrumentos Individuales..... | 95 |
| 4.2 Selección de Red Neuronal..... | 101 |
| 4.3 Uso de Redes Neuronales..... | 102 |
| 5. Manual del Usuario de CajaCustica | |
| 5.1 Presentación General..... | 103 |
| 5.2 Pantallas Secundarias..... | 105 |
| 5.3 Calibrador..... | 107 |
| 5.4 Limitantes | 107 |
| Conclusiones..... | 109 |
| BIBLIOGRAFIA..... | 112 |

DEFINICIÓN DEL TEMA

El proyecto consiste en desarrollar un software que sea capaz de reconocer las señales producidas por una persona cuando ejecuta la batería y como producto del procesamiento de dichas señales se tendrá la partitura de lo que se ha ejecutado. Para esto se utilizarán redes neuronales como motor de identificación y algunas técnicas de procesamiento digital de señales para obtener los parámetros que serán las entradas a las redes neuronales.

JUSTIFICACIÓN DEL TEMA

Este proyecto está enfocado hacia aquellas personas interesadas en identificar los errores que cometen al ejecutar una pieza musical en una batería acústica, enfocándonos fundamentalmente en aquellas personas que se encuentran en la etapa de aprendizaje del instrumento. Por otro lado aquellas personas que ya puedan tocar este instrumento podrán ayudar a los principiantes en su proceso de aprendizaje generando diferentes partituras a las predeterminadas en el software aunque éstas no sean agregadas en el tutorial.

Este proyecto tendrá la opción de brindar una comparación entre la pieza ejercicio (la cual estará predefinida en el programa) con la pieza ejecutada.

OBJETIVO GENERAL

- ∅ Adquirir, procesar y reconocer los patrones rítmicos provenientes de la batería para finalmente representarlos en una partitura.

OBJETIVOS ESPECÍFICOS

- ∅ Proveer de una herramienta facilitadora para las personas interesadas en el aprendizaje de la ejecución de la batería.
- ∅ Composición de piezas musicales basadas en los tres instrumentos básicos que se toman en cuenta en una batería acústica (hi-hat, redoblante y bombo)
- ∅ Utilizar como herramienta de reconocimiento Redes Neuronales tales como ART, Kohonen, Retropropagación por mencionar algunos algoritmos de entrenamiento.

ALCANCES

1. Desarrollar un software por medio del cual, aquellas personas que estén interesadas en aprender a tocar la Batería Acústica, logren el aprendizaje básico de una manera más fácil.
2. Obtener la partitura de una pieza musical que contenga únicamente los siguientes elementos: Hi Hat, redoblante y bombo.
3. Se realizarán pruebas con al menos tres algoritmos de entrenamiento de redes neuronales para poder elegir cuál es el que mejor reconoce las señales musicales emitidas por la batería acústica.
4. Utilizaremos la tarjeta de sonido de una PC para poder adquirir y reproducir los sonidos producidos por el instrumento.
5. Será necesario el uso de filtros digitales para el pre-procesamiento de las señales provenientes de la tarjeta de sonido con el fin de poder discriminar aquellas frecuencias que no son de nuestro interés.
6. Se trabajará para obtener un porcentaje de reconocimiento mayor o igual que 90% de la información introducida a la red.
7. Se trabajará en dos lenguajes de programación. Se trabajará en Matlab versión 6.5 para realizar las pruebas pertinentes a cada algoritmo de aprendizaje, una vez encontrado el algoritmo y la topología idónea pasaremos a desarrollarla en Visual Basic 6.0.
8. Trabajaremos con la herramienta matemática de la transformada de Fourier para el tratamiento digital de la señal.

9. En el caso del tutorial se tendrá la opción de poder escuchar la pieza musical ejercicio utilizando el reproductor de música que utilice el usuario.
10. La interfaz gráfica contará con señalización que le marcará el tempo al estudiante.

LIMITACIONES

1. El sistema funcionará bajo el sistema operativo Windows
2. La partitura que se obtendrá al final del proceso se escribirá en compases de cuatro por cuatro (4/4)
3. El sistema sólo será capaz de identificar elementos con dimensiones fijas, para el caso

| | |
|------------|---------------------------------------------|
| Hi Hat | 13 pulgadas |
| Bombo | 22 x 16 pulgadas |
| Redoblante | 14 x 5 ³ / ₄ pulgadas |
4. El software continuará generando la partitura mientras la memoria de la computadora en la cual se está ejecutando se lo permita.
5. Nos limitaremos al uso de un tipo específico de pergaminos (área de impacto del tambor) en el caso del bombo será pinstripe (fabricado por Remo®) el cual es un pergamino de doble capa que contiene en su interior aceite y para el redoblante utilizaremos el pergamino FiberSkin 3 que está hecho de fibra (fabricado por Remo®) ya que cada pergamino posee un sonido característico.
6. La velocidad de ejecución del instrumento estará limitada a la velocidad de procesamiento de la computadora.

7. Como adquisidor de datos utilizaremos la tarjeta de sonido de la computadora.
8. Se probará con distintos micrófonos para poder establecer las diferencias que nos conducirán a seleccionar el micrófono apropiado.
9. Adicionalmente haremos pruebas respecto al requerimiento mínimo del sistema computacional, es decir las características que debe poseer una PC como mínimo para que el sistema funcione correctamente.

Capítulo 1 – Conceptos Musicales

Toda melodía lleva consigo un diagrama, el cual consta de varios símbolos que representan los diferentes sonidos que conforman la melodía, este “mapa” es una guía para la persona que desee ejecutar dicha melodía. Típicamente este “mapa” es mejor conocido como pentagrama y los diferentes símbolos que se escriben sobre él deben ser interpretados por la persona que desee ejecutar dicha melodía. Para entender más sobre cada símbolo y su interpretación, en el presente capítulo desarrollaremos una explicación breve y básica de cada uno de los conceptos fundamentales de la notación musical.

1.1 El Pentagrama

Es un conjunto de cinco líneas que se utiliza para asignar el nombre de las notas a las figuras musicales que se representan sobre ella. Es decir según el espacio o línea en que se encuentre una figura musical dentro del pentagrama, así recibirá un nombre u otro.

Esta asignación de los nombres de las notas, está en función de la clave que se utilice. O sea que, la clave decide qué asignación tiene cada uno de los espacios y las líneas que lo demarcan. [1]

1.1.1 Las Líneas Adicionales

Son líneas que puedes colocar a arriba o abajo del pentagrama para ubicar las notas más agudas o más graves respectivamente. [1]

1.2 Las Claves

Al principio de la plantilla nos vamos a encontrar una clave que nos va a indicar la posición de las notas en el pentagrama, existen dos claves: Clave de Sol y Clave de Fa. [2]

1.2.1 Clave de Sol

Cuando aparece la clave de Sol al principio del pentagrama, nos indica que la nota Sol, se sitúa justo sobre la 2ª línea de la plantilla empezando por abajo, y a partir

de ahí hacia arriba y hacia abajo se colocan el resto de las notas siguiendo la escala. [2]

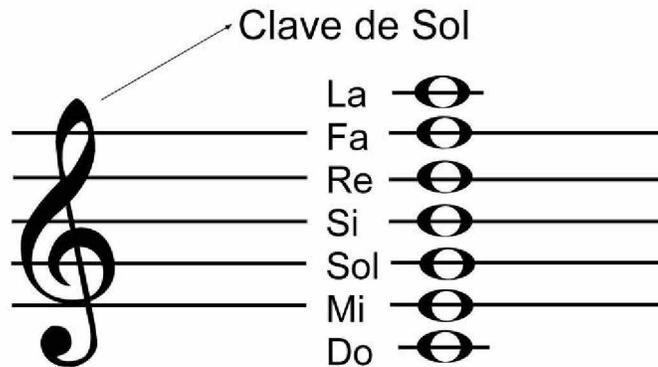


Figura 1.1 Símbolo de que representa la clave de sol y el respectivo orden q llevan las notas cuando se trabaja bajo esta clave.

1.2.2 Clave de Fa

Indica que la nota FA se sitúa sobre la 4ª línea y todas las demás según esa referencia.

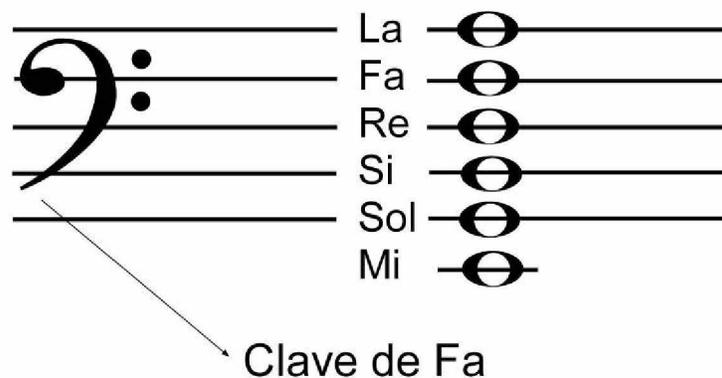


Figura 1.2 Símbolo de que representa la clave de fa y el respectivo orden q llevan las notas cuando se trabaja bajo esta clave.

La clave de sol es utilizada por instrumentos agudos con el violín, la flauta, la trompeta y por otros no tan agudos como la guitarra. La clave de fa es utilizada por instrumentos graves como el contrabajo, el violonchelo y el trombón.

Para guitarra nos vamos a encontrar solamente partituras en clave de Sol, el bajo utiliza la clave de Fa y otros instrumentos como puede ser el piano utiliza ambas claves, la clave de Sol sería para la mano derecha y la clave de Fa para la mano izquierda. [2]

1.3 Notas Musicales

Se representan por medio de 7 nombres: Do Re Mi Fa Sol La Sí. Cada una de estas notas equivale a una altura o frecuencia determinada, en orden de grave a aguda. La nomenclatura de estas notas en inglés es: [1]

Do Re Mi Fa Sol La Si
C D E F G A B

1.4 Figuras Musicales

Son símbolos que nos indican la duración de las notas musicales. A continuación presentamos una imagen con las más usuales, ordenadas según su valor e indicando sus equivalencias [3]

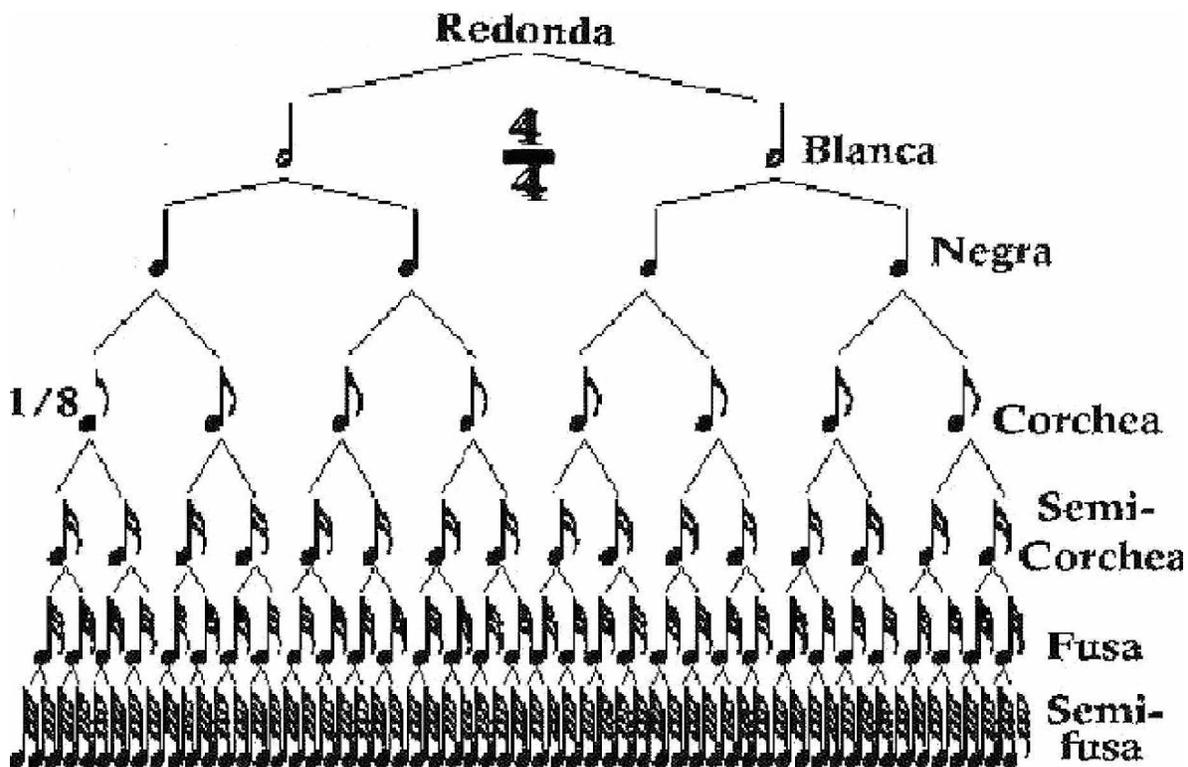


Figura 1.3 Esquema de la subdivisión de tiempos de las nota musicales

La anterior imagen puede ser más entendible guiándonos por medio de la siguiente tabla.

Tabla 1.1 Interpretación de los valores de las figuras musicales [4]

| Figura | Nombre | Valor |
|------------------------------------------------------------------------------------|-------------|------------------------------------------------------|
|  | Cuadrada | El doble de una redonda |
|  | Redonda | El doble de una blanca Mitad de una cuadrada |
|  | Blanca | El doble de una negra Mitad de una redonda |
|  | Negra | El doble de una corchea Mitad de una blanca |
|  | Corchea | El doble de una semicorchea Mitad de una negra |
|  | semicorchea | El doble de una fusa Mitad de una corchea |
|  | Fusa | El doble de una semifusa Mitad de una semicorchea |
|  | Semifusa | Mitad de una fusa |

1.4.1 Silencios

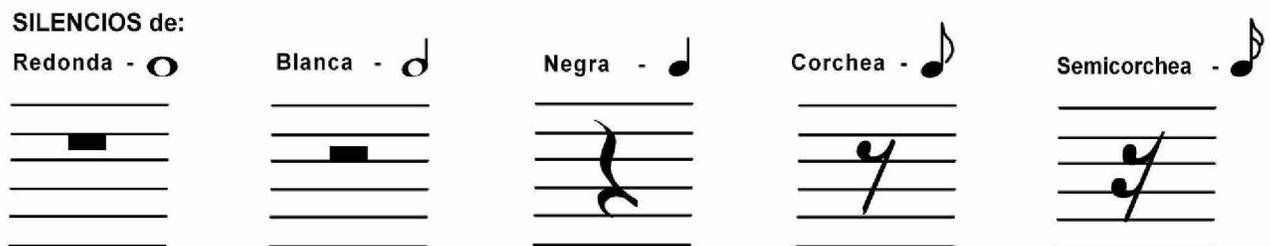


Figura 1.4 Representación de los diferentes silencios existentes

Al igual que las notas tienen su duración, también existen una serie de símbolos que indican un silencio, como su propio nombre indica no existe ningún sonido, pero sí hay que respetar su duración. [2]

1.5 Alteraciones

Las alteraciones tienen el poder de modificar la entonación de los sonidos naturales, subiéndolos o bajándolos.

Estas alteraciones son cinco: sostenido, bemol, doble sostenido, doble bemol y becuadro; se colocan a la izquierda de las figuras, y se leen después de nombrar a las notas. Ejemplo: fa sostenido, si bemol.

Pero en el solfeo (sea cantado, como hablado) se omite el nombre de las alteraciones, pronunciándose solamente las sílabas de las notas.

El sostenido y el bemol son alteraciones simples.

El doble sostenido y el doble bemol son alteraciones dobles.

El becuadro puede conceptuarse como alteración anulatoria. [5][6]

Tabla 1.2 Cuadro explicativo del significado de cada alteración.

| | | |
|----|-----------------|-------------------------------------------------------------------|
| # | Sostenido | Eleva un semitono la altura de la nota. |
| b | bemol | Disminuye un semitono la altura de la nota |
| ## | Doble sostenido | Eleva dos semitonos (un tono) la altura de la nota. |
| bb | Doble bemol | Disminuye dos semitonos (un tono) la altura de la nota |
| □ | Becuadro | Anula las alteraciones de la nota y la deja en su altura natural. |

1.6 Compases

Son fracciones que se colocan al principio del pentagrama y que dividen el tiempo de la pieza musical en partes iguales. El compás nos da las pautas y marcas del tiempo que debemos usar al interpretar rítmicamente la pieza musical.

Podemos apreciar de forma práctica el compás, observando el movimiento de la mano de un director de orquesta. Este utiliza movimientos concretos para cada uno de los compases.

Los compases se clasifican según su división en tres grupos: Binarios, ternarios y cuaternarios, y se representan con una fracción, donde el número de arriba indica

las partes en las que se divide el compás, y el número de abajo el valor de cada parte.

Ej.: 4/4 = Se divide en cuatro partes. Cada parte vale una cuarta de la redonda, es decir una negra.

La división entre un compás y otro se realiza mediante líneas verticales llamadas líneas divisorias. [3]

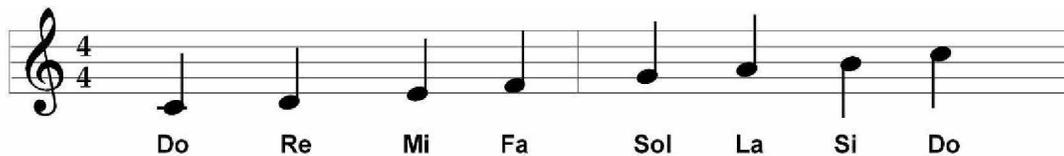


Figura 1.5 Representación de dos compases en un pentagrama, valor del compás 4/4.

1.7 Signos secundarios

Son herramientas auxiliares que determinan la duración de la nota o notas a las cuales afectan. [7]

1.7.1 La ligadura de prolongación

Es una línea arqueada que une dos o más notas de la misma altura e indica que la segunda, y/o siguientes, son prolongación de la primera.

O sea que este sonido NO SE REPITE, Y SE MANTIENE CON LA DURACIÓN QUE INDIQUEN LAS FIGURAS. [7]

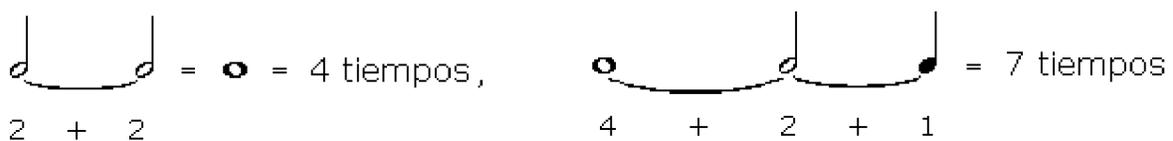


Figura 1.6 Efecto de la ligadura, sumatoria de tiempos



Figura 1.7 Escritura de ligadura en pentagrama, se articula solo la primera nota, sumándole el valor de las restantes.

1.7.2 El puntillo y el doble puntillo

1.7.2.1 El puntillo

El puntillo es, como su nombre indica, un punto pequeño que, colocado a la derecha de una figura de nota o silencio, prolonga la duración de los mismos la mitad de su valor en tiempo. [8]



Figura 1.8 Notación de puntillo con notas musicales y su respectiva analogía con la ligadura.



Figura 1.9 Notación de puntillo con silencios y su respectiva analogía con la ligadura.

1.7.2.2 El doble puntillo

El doble puntillo prolonga la nota o silencio la mitad de su duración más la mitad de la mitad de su duración. [8]



Figura 1.10 Notación del doble puntillo y su respectiva analogía con la ligadura.

1.8 Introducción a la lectura musical de la Batería

El hecho de "saber" leer partituras parece ser ignorado en estos últimos tiempos por muchos bateristas, quizás porque tocan de oído perfectamente y no necesitan leer música, o bien nunca tuvieron la suerte de que alguien les enseñara. El principal material de estudio de un baterista está basado en partituras, y necesita aprender ese lenguaje musical universal para poder interpretar lo que otros escriben. Y como no todo es leer, también es importante escribir lo que uno toca, y lo ideal es hacerlo mediante un pentagrama como lo hacen los verdaderos músicos. [9]

1.8.1 Notación en el pentagrama

Antes que nada, hay que saber que cada elemento de la batería tiene su lugar determinado en el pentagrama, en general la ubicación es estándar, aunque podemos encontrarnos partituras en donde hay algunas variaciones dependiendo del autor o baterista, pero casi siempre aclaran el tipo de notación usada. Si uno tiene elementos del set que no está dentro de lo estándar, se pueden agregar líneas adicionales al pentagrama. [9]

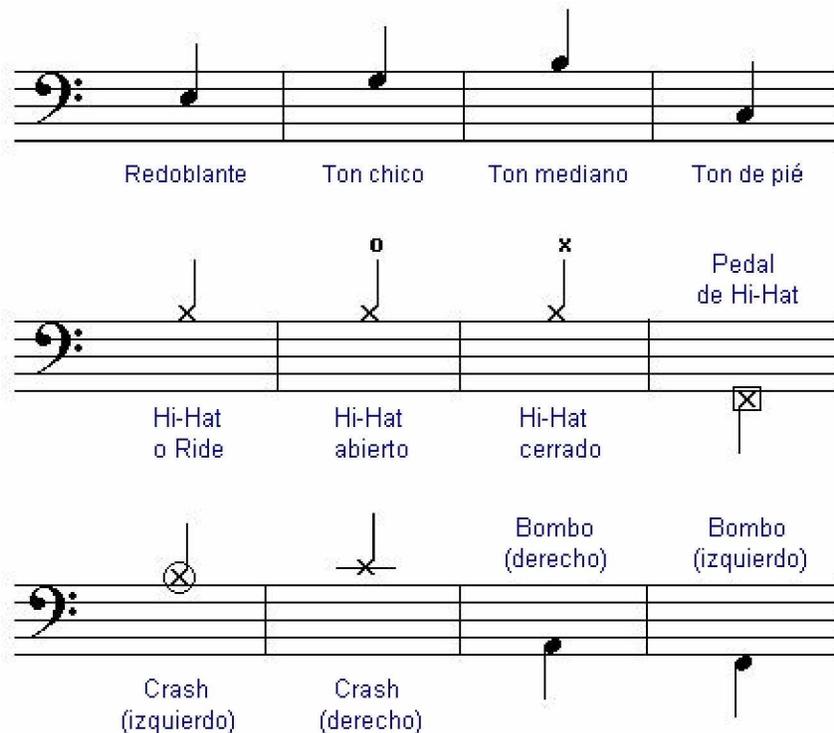


Figura 1.11 Ubicación de cada elemento de la batería en el pentagrama.

Capítulo 2 - Redes Neuronales

A cada instante y para realizar cualquier acción el cerebro humano analiza, procesa y ordena cual es o cual debería ser la respuesta más adecuada según sean las exigencias de la circunstancia, para realizar este proceso el cerebro se vale de una red de neuronas biológicas las cuales se adaptan al evento actual que se vive.

Este complejo comportamiento ha sido estudiado por muchos años y de alguna manera se ha tratado de imitar no en su totalidad pero si en tareas “fáciles” de ejecutar, bajo diferentes algoritmos matemáticos se han desarrollado las Redes Neuronales Artificiales, las cuales tienen diferentes clasificaciones según sea la aplicación que se desee implementar.

En este capítulo se hablará del comportamiento de las redes neuronales biológicas, se desarrollará un poco de historia y finalmente se presentarán la mayoría de las redes neuronales artificiales conocidas.

2.1 Redes Neuronales Biológicas

A grandes rasgos, recordemos que el cerebro humano se compone de decenas de billones de neuronas interconectadas entre sí formando circuitos o redes que desarrollan funciones específicas.

Una neurona típica recoge señales procedentes de otras neuronas a través de una pléyade^{*1} de delicadas estructuras llamadas dendritas. La neurona emite impulsos de actividad eléctrica a lo largo de una fibra larga y delgada denominada axón, que se divide en millares de ramificaciones.

¹ **Pléyade, sinónimo de generación, hornada, legión, conjunto, celebridades*

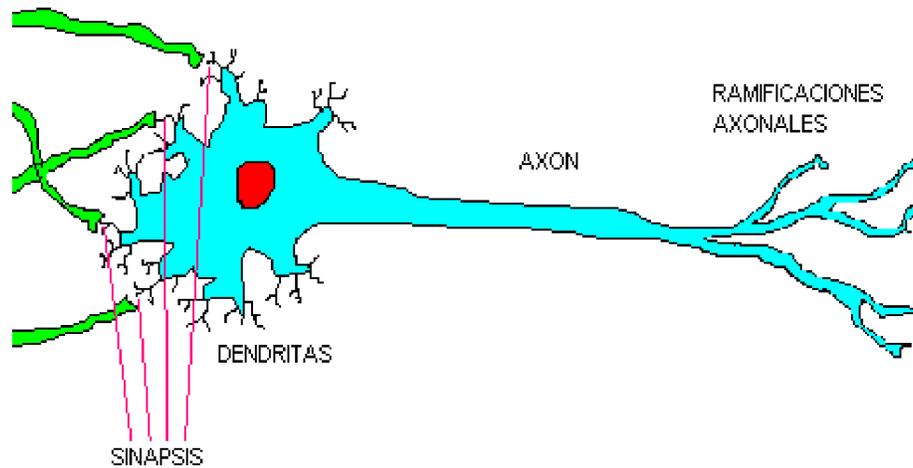


Figura 2.1 Neurona y conexiones sinápticas

Las extremidades de estas ramificaciones llegan hasta las dendritas de otras neuronas y establecen unas conexiones llamadas sinápsis, en las cuales se produce una transformación del impulso eléctrico en un mensaje neuroquímico, mediante la liberación de unas sustancias llamadas neurotransmisor.

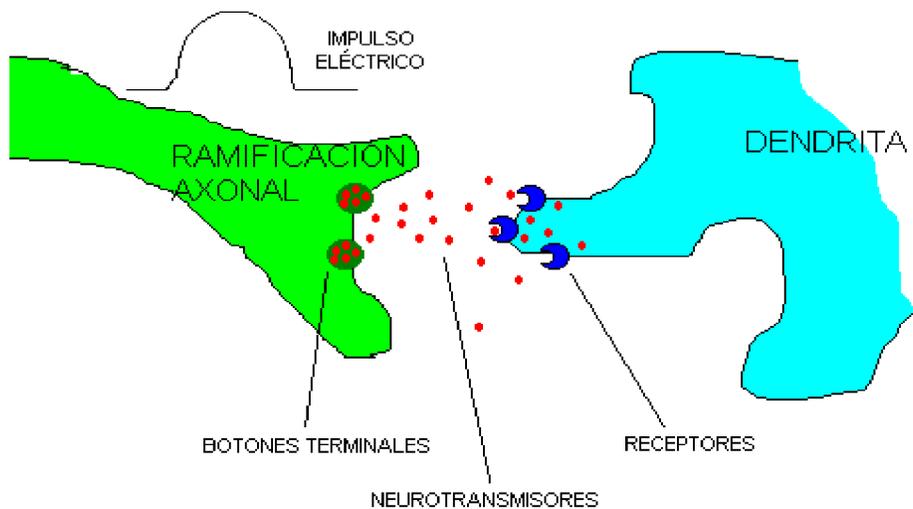


Figura 2.2 Detalle de una sinápsis

El efecto de los neurotransmisores sobre la neurona receptora puede ser excitatorio o inhibitorio, y es variable (la intensidad del efecto depende de numerosos factores que no sería oportuno describir aquí), de manera que podemos hablar de la fuerza o efectividad de una sinápsis. Las señales

excitatorias e inhibitorias recibidas por una neurona se combinan, y en función de la estimulación total recibida, la neurona toma un cierto nivel de activación, que se traduce en la generación de breves impulsos nerviosos con una determinada frecuencia o tasa de disparo, y su propagación a lo largo del axón hacia las neuronas con las cuales sinapta.

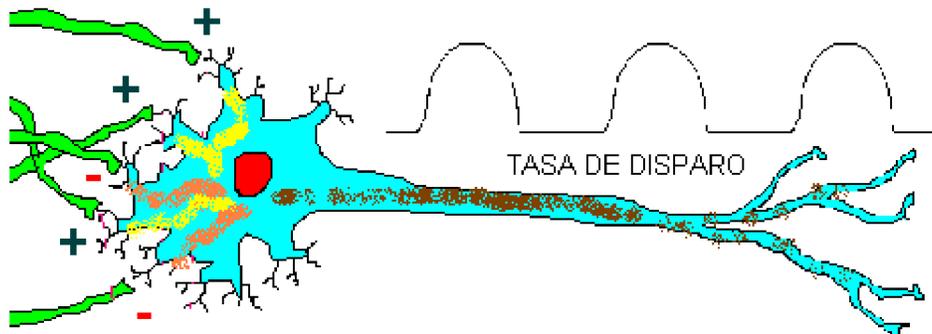


Figura 2.3 Activación y disparo de una neurona

De esta manera la información se transmite de unas neuronas a otras y va siendo procesada a través de las conexiones sinápticas y las propias neuronas. El aprendizaje de las redes neuronales se produce mediante la variación de la efectividad de las sinápsis, de esta manera cambia la influencia que unas neuronas ejercen sobre otras, de aquí se deduce que la arquitectura, el tipo y la efectividad de las conexiones en un momento dado, representan en cierto modo la memoria o estado de conocimiento de la red. [10]

2.2 Historia de las Redes Neuronales Artificiales (RNA)

- 1936 - Alan Turing. Fue el primero en estudiar el cerebro como una forma de ver el mundo de la computación. Sin embargo, los primeros teóricos que concibieron los fundamentos de la computación neuronal fueron Warren McCulloch, un neurofisiólogo, y Walter Pitts, un matemático, quienes, en 1943, lanzaron una teoría acerca de la forma de trabajar de las neuronas (Un Cálculo Lógico de la Inminente Idea de la Actividad Nerviosa - Boletín de Matemática Biofísica 5: 115-133). Ellos modelaron una red neuronal simple mediante circuitos eléctricos. [11]

- 1949 - Donald Hebb. Fue el primero en explicar los procesos del aprendizaje (que es el elemento básico de la inteligencia humana) desde un punto de vista psicológico, desarrollando una regla de como el aprendizaje ocurría. Aún hoy, este es el fundamento de la mayoría de las funciones de aprendizaje que pueden hallarse en una red neuronal. Su idea fue que el aprendizaje ocurría cuando ciertos cambios en una neurona eran activados. También intentó encontrar semejanzas entre el aprendizaje y la actividad nerviosa. Los trabajos de Hebb formaron las bases de la Teoría de las Redes Neuronales. [11]
- 1950 - Karl Lashley. En sus series de ensayos, encontró que la información no era almacenada en forma centralizada en el cerebro sino que era distribuida encima de él. [11]
- 1956 - Congreso de Dartmouth. Este Congreso frecuentemente se menciona para indicar el nacimiento de la inteligencia artificial. [11]
- 1957 - Frank Rosenblatt. Comenzó el desarrollo del Perceptron. Esta es la red neuronal más antigua; utilizándose hoy en día para aplicación como identificador de patrones. Este modelo era capaz de generalizar, es decir, después de haber aprendido una serie de patrones podía reconocer otros similares, aunque no se le hubiesen presentado en el entrenamiento. Sin embargo, tenía una serie de limitaciones, por ejemplo, su incapacidad para resolver el problema de la función OR-exclusiva y, en general, era incapaz de clasificar clases no separables linealmente. [11]
- 1959 - Frank Rosenblatt: Principios de Neurodinámica. En este libro confirmó que, bajo ciertas condiciones, el aprendizaje del Perceptron convergía hacia un estado finito (Teorema de Convergencia del Perceptrón). [11]
- 1960 - Bernard Widroff/Marcian Hoff. Desarrollaron el modelo Adaline (ADaptative LINear Elements). Esta fue la primera red neuronal aplicada a un problema real (filtros adaptativos para eliminar ecos en las líneas

telefónicas) que se ha utilizado comercialmente durante varias décadas. [11]

- 1961 - Karl Steinbeck: Die Lernmatrix. Red neuronal para simples realizaciones técnicas (memoria asociativa). [11]
- 1969 - Marvin Minsky/Seymour Papert. En este año casi se produjo la “muerte abrupta” de las Redes Neuronales; ya que Minsky y Papert probaron (matemáticamente) que el Perceptron no era capaz de resolver problemas relativamente fáciles, tales como el aprendizaje de una función no-lineal. Esto demostró que el Perceptron era muy débil, dado que las funciones no-lineales son extensamente empleadas en computación y en los problemas del mundo real. [11]
- 1974 - Paul Werbos. Desarrolló la idea básica del algoritmo de aprendizaje de *propagación hacia atrás* (Retropropagación); cuyo significado quedó definitivamente aclarado en 1985. [11]
- 1977 - Stephen Grossberg: Teoría de Resonancia Adaptada (TRA). La Teoría de Resonancia Adaptada es una arquitectura de red que se diferencia de todas las demás previamente inventadas. La misma simula otras habilidades del cerebro: memoria a largo y corto plazo. [11]
- 1985 - John Hopfield. Provocó el renacimiento de las redes neuronales con su libro: “Computación neuronal de decisiones en problemas de optimización.” [11]
- 1986 - David Rumelhart/G. Hinton. Redescubrieron el algoritmo de aprendizaje de propagación hacia atrás (Retropropagación). [11]

A partir de 1986, el panorama fue alentador con respecto a las investigaciones y el desarrollo de las redes neuronales. En la actualidad, son numerosos los trabajos que se realizan y publican cada año, las aplicaciones nuevas que surgen (sobre todo en el área de control) y las empresas que lanzan al mercado productos nuevos, tanto hardware como software (sobre todo para simulación). [11]

2.3 Redes Neuronales Artificiales

2.3.1 Analogía con las Redes Neuronales Biológicas

Las neuronas se modelan mediante unidades de proceso. Cada unidad de proceso se compone de una red de conexiones de entrada, una función de red (de propagación), encargada de computar la entrada total combinada de todas las conexiones, un núcleo central de proceso, encargado de aplicar la función de activación, y la salida, por dónde se transmite el valor de activación a otras unidades.

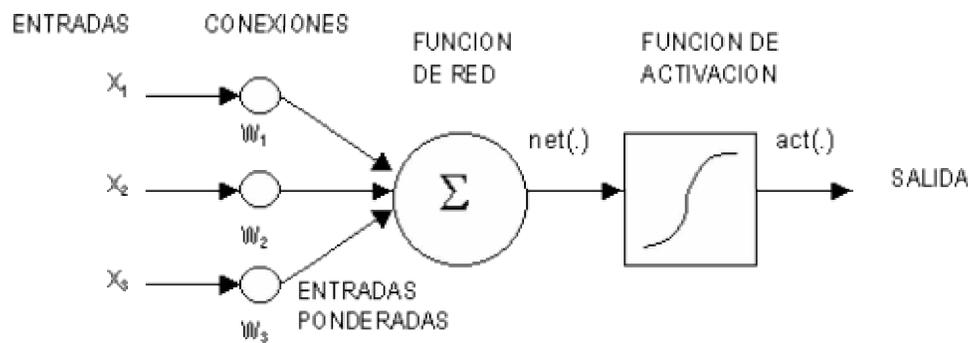


Figura 3.4 Unidad de proceso típica

La función de red es típicamente el sumatorio ponderado, mientras que la función de activación suele ser alguna función de umbral o una función sigmoide.

- **Función de propagación** o de red: Calcula el valor de base o entrada total a la unidad, generalmente como simple suma ponderada de todas las entradas recibidas, es decir, de las entradas multiplicadas por el peso o valor de las conexiones. Equivale a la combinación de las señales excitatorias e inhibitorias de las neuronas biológicas.
- **Función de activación:** Es quizás la característica principal o definitoria de las neuronas, la que mejor define el comportamiento de la misma. Se usan diferentes tipos de funciones, desde simples funciones de umbral a funciones no lineales. Se encarga de calcular el nivel o estado de activación de la neurona en función de la entrada total.

- **Conexiones ponderadas:** hacen el papel de las conexiones sinápticas, el peso de la conexión equivale a la fuerza o efectividad de la sinápsis. La existencia de conexiones determina si es posible que una unidad influya sobre otra, el valor de los pesos y el signo de los mismos definen el tipo (excitatorio/inhibitorio) y la intensidad de la influencia.
- **Salida:** calcula la salida de la neurona en función de la activación de la misma, aunque normalmente no se aplica más que la función identidad, y se toma como salida el valor de activación. El valor de salida cumpliría la función de la tasa de disparo en las neuronas biológicas. [10]

Tabla 2.1 Comparación entre las neuronas biológicas reales y las unidades de proceso artificiales

| Redes Neuronales Biológicas | Redes Neuronales Artificiales |
|--------------------------------------------------|--------------------------------------|
| Neuronas | Unidades de proceso |
| Conexiones sinápticas | Conexiones ponderadas |
| Efectividad de las sinopsis | Peso de las conexiones |
| Efecto excitatorio o inhibitorio de una conexión | Signo del peso de una conexión |
| Efecto combinado de las sinápsis | Función de propagación o de red |
| Activación à tasa de disparo | Función de activación à Salida |

2.3.2 La función de propagación (de red o de base)

Como ya hemos comentado, se encarga de calcular la entrada total de la neurona como combinación de todas las entradas. La función más utilizada con diferencia es la función lineal de base (LBF), que consiste en el sumatorio ponderado de todas las entradas. [10]

2.3.2.1 Función lineal de base (LBF)

Consiste en el sumatorio de las entradas ponderadas. Se trata de una función de tipo hiperplano, esto es, de primer orden.

Dada una unidad j , y n unidades conectadas a ésta, si llamamos X al vector de entradas (que coincide con las salidas de las unidades de la capa anterior) y W_j al vector de pesos de las conexiones correspondientes, esta función quedaría así,

$$net_j(X, W_j) = \sum_{i=1}^n x_i w_{ij}$$

Ecuación 2.1

Al representar los pesos utilizamos dos subíndices para indicar que conectan dos unidades, i y j , dónde j se refiere la unidad actual. [10]

2.3.2.2 Función radial de base (RBF)

Función de tipo hiper esférico, de segundo orden, no lineal. El valor de red representa la distancia a un determinado patrón de referencia. [10]

$$net_j(X, W_j) = \sqrt{\sum_{i=1}^n (x_i - w_{ij})^2}$$

Ecuación 2.2

2.3.3 La función de activación

Una neurona biológica puede estar activa (excitada) o inactiva (no excitada); es decir, que tiene un “estado de activación”. Las neuronas artificiales también tienen diferentes estados de activación; algunas de ellas solamente dos, al igual que las biológicas, pero otras pueden tomar cualquier valor dentro de un conjunto determinado.

La función activación calcula el estado de actividad de una neurona; transformando la entrada global, menos el umbral, en un valor (estado) de activación, cuyo rango normalmente va de 0 a 1 o de -1 a 1. Esto es así, porque una neurona puede estar totalmente inactiva 0 ó -1 o activa 1.

La función activación, es una función de la entrada global gin_i menos el umbral θ_i . Las funciones de activación más comúnmente utilizadas se detallan a continuación: [11]

2.3.3.1 Función Lineal

$$f(x) = \begin{cases} -1 \rightarrow x \leq \frac{-1}{a} \\ a * x \rightarrow \frac{-1}{a} < x < \frac{1}{a} \\ 1 \rightarrow x \geq \frac{1}{a} \end{cases}$$

Ecuación 2.3

Con $x = gin_i - \theta_i$ y $a > 0$ ²

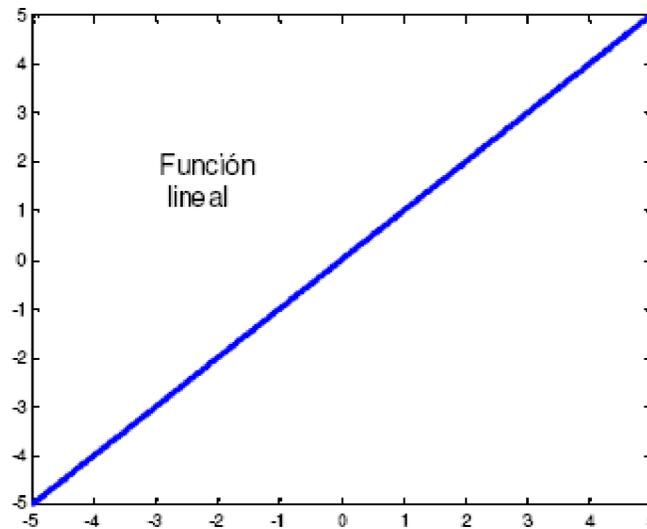


Figura 2.4 Grafica de función lineal

Los valores de salida obtenidos por medio de esta función de activación serán: $(gini - \theta_i)$, cuando el argumento de $(gini - \theta_i)$ esté comprendido dentro del rango $[-1/a, 1/a]$.

² Para entender la nomenclatura de la ecuación favor remitirse a la Pág. 17

Por encima o por debajo de esta zona se fija la salida en 1 ó -1 respectivamente. Cuando $a = 1$ (siendo que la misma afecta la pendiente de la gráfica) la salida es igual a la entrada. [11]

2.3.3.2 Función Sigmoial Unipolar

$$f(x) = \frac{1}{1 + e^{-g*x}}$$

Ecuación 2.4

Con $x = gin_i - \theta_i$ ^{*3}

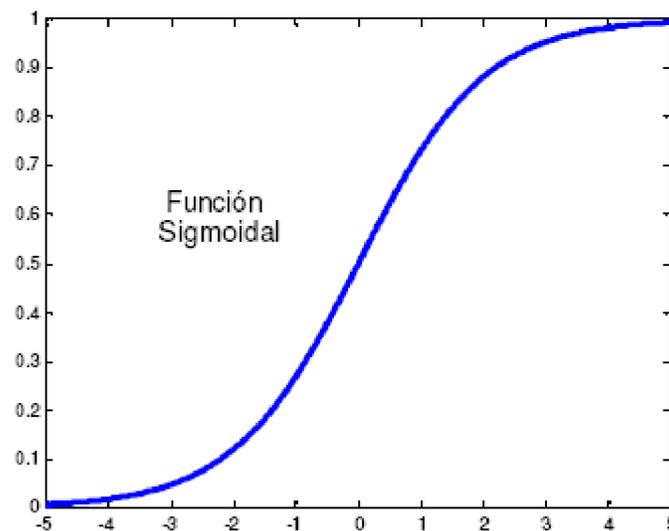


Figura 2.5 Grafica de función Sigmoial

Los valores de salida que proporciona esta función están comprendidos dentro de un rango que va de 0 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación. [11]

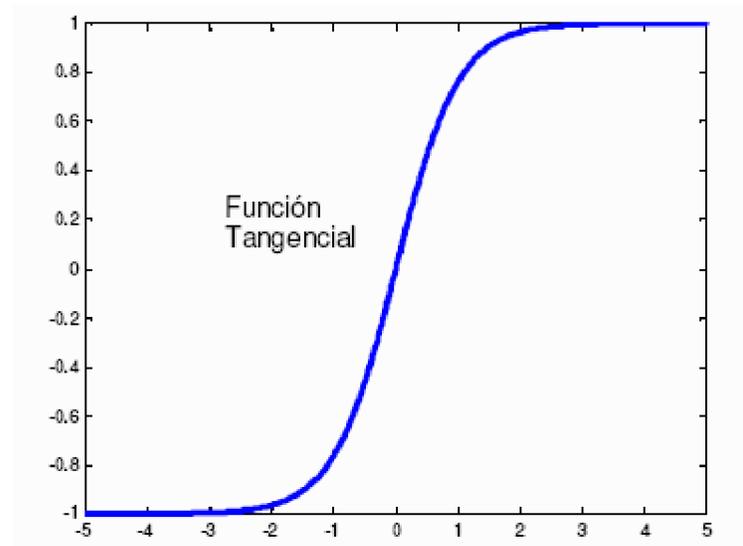
³ Para entender la nomenclatura de la ecuación favor remitirse a la Pág. 17

2.3.3.3 Función Tangente Hiperbólica

$$f(x) = \frac{e^{g*x} - e^{-g*x}}{e^{g*x} + e^{-g*x}}$$

Ecuación 2.5

Con $x = gin_i - \theta_i$ ^{*4}



2.6 Grafica de Función Tangente Hiperbólica

Los valores de salida de la función tangente hiperbólica están comprendidos dentro de un rango que va de -1 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación. [11]

⁴ Para entender la nomenclatura de la ecuación favor remitirse a la Pág. 17

2.3.4 Estructura y formas de interconexión

Para diseñar una red debemos establecer cómo estarán conectadas unas unidades con otras y determinar adecuadamente los pesos de las conexiones. Lo más usual es disponer las unidades en forma de capas, pudiéndose hablar de redes de una, de dos o de más de dos capas, las llamadas *redes multicapa*.

Aunque inicialmente se desarrollaron redes de una sola capa, lo más usual es disponer tres o más capas: la primera capa actúa como buffer de entrada, almacenando la información bruta suministrada a la red o realizando un sencillo pre-proceso de la misma, ésta es llamada capa de entrada; otra capa, la cual actúa como interfaz o buffer de salida, almacenando la respuesta de la red para que pueda ser leída, es llamada capa de salida; y las capas intermedias, principales, encargadas de extraer, procesar y memorizar la información, se les denomina como capas ocultas.

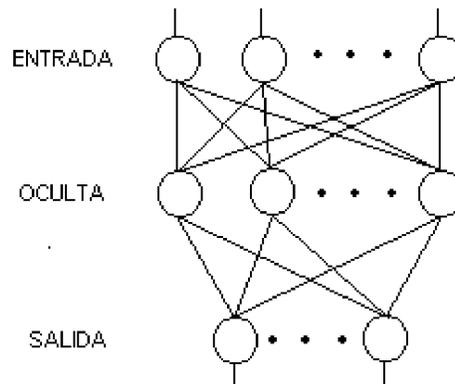


Figura 2.7 Modelo de red en cascada de 3 capas

Además del número de capas de una red, en función de cómo se interconectan unas capas con otras, podemos hablar de *redes recurrentes* (feed-back) y redes no recurrentes o *redes en cascada* (feed-forward). En las redes en cascada la información fluye unidireccionalmente de una capa a otra (desde la capa de entrada a las capas ocultas y de éstas a la capa de salida), y además, no se admiten conexiones intracapa. En las redes recurrentes la información puede volver a lugares por los que ya había pasado, formando bucles, y se admiten las conexiones intracapa (laterales), incluso de una unidad consigo misma.

Las conexiones entre una capa y otra pueden ser totales, es decir, que cada unidad se conecta con todas las unidades de la capa siguiente, o parciales, en las cuales una unidad se conecta con sólo algunas de las unidades de la capa siguiente, generalmente siguiendo algún patrón aleatorio o pseudo-aleatorio (por ejemplo, mediante algoritmos genéticos).[10]

2.4 Clasificación de las Redes Neuronales

Según el criterio que escojamos para clasificar las RNA tendremos dos formas de clasificación, lo más común es usar la arquitectura y el tipo de aprendizaje como criterios de clasificación.

Si nos fijamos en la arquitectura podemos tener dos posibilidades distintas. Si la arquitectura de la red no presenta ciclos, es decir, no se puede trazar un camino de una neurona a sí misma, la red se llama *unidireccional* (feedforward).

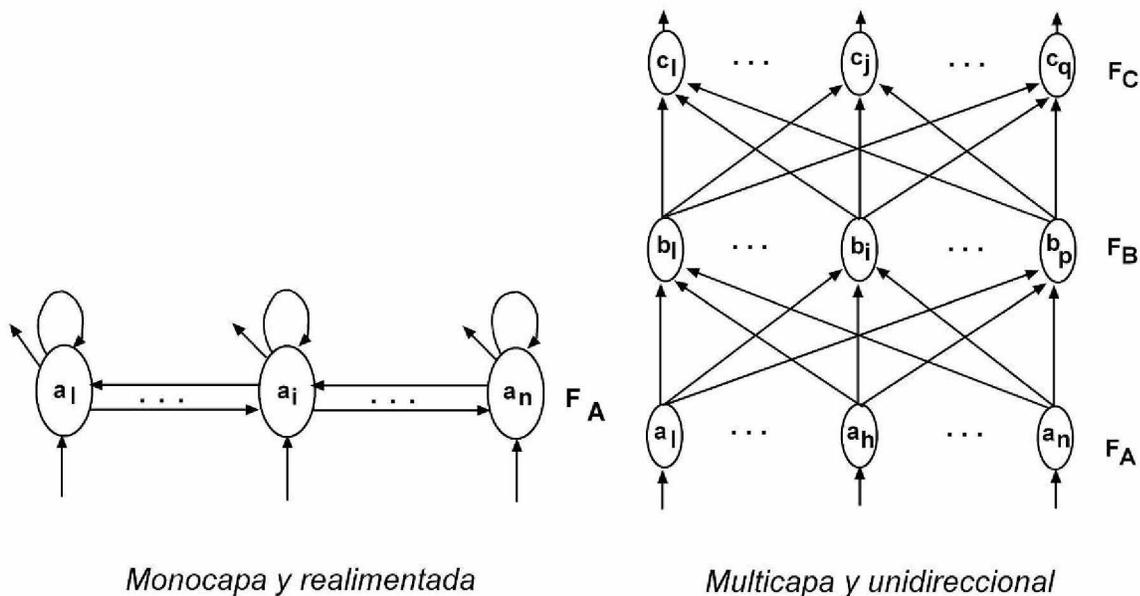


Figura 2.8 Tipos de Arquitectura de Redes Neuronales, Representación de redes unidireccionales y realimentadas

Por el contrario, si podemos trazar un camino de una neurona a sí misma, la arquitectura presenta ciclos. Este tipo de redes se denominan *recurrentes* o *realimentados* (recurrent).

El otro criterio más habitual para clasificar las redes neuronales es el tipo de aprendizaje que se utilice. Hay cuatro clases de aprendizaje distintos:

- a. *Aprendizaje supervisado*: En este tipo de aprendizaje se le proporciona a la RNA una serie de ejemplos consistentes en unos patrones de entrada, junto con la salida que debería dar la red. El proceso de entrenamiento consiste en el ajuste de los pesos para que la salida de la red sea lo más parecida posible a la salida deseada. Es por ello que en cada iteración se use alguna función que nos de cuenta del error o el grado de acierto que está cometiendo la red.
- b. *Aprendizaje no supervisado o autoorganizado*: En este tipo de aprendizaje se presenta a la red una serie de ejemplos pero no se presenta la respuesta deseada. Lo que hace la RNA es reconocer regularidades en el conjunto de entradas, es decir, estimar una función densidad de probabilidad $p(\mathbf{x})$ que describe la distribución de patrones \mathbf{x} en el espacio de entrada R^n .
- c. *Aprendizaje Híbrido*: Es una mezcla de los anteriores. Unas capas de la red tienen un aprendizaje supervisado y otras capas de la red tienen un aprendizaje de tipo no supervisado. Este tipo de entrenamiento es el que tienen redes como las RBF.
- d. *Aprendizaje reforzado (reinforcement learning)*: Es un aprendizaje con características del supervisado y con características del autoorganizado. No se proporciona una salida deseada, pero si que se le indica a la red en cierta medida el error que comete, aunque es un error global. [12]

2.4.1 Modelos supervisados

2.4.1.1 Perceptron

Un Perceptrón unicapa no es más que un conjunto de neuronas no unidas entre sí, de manera que cada una de las entradas del sistema se conectan a cada neurona, produciendo cada una de ellas su salida individual.

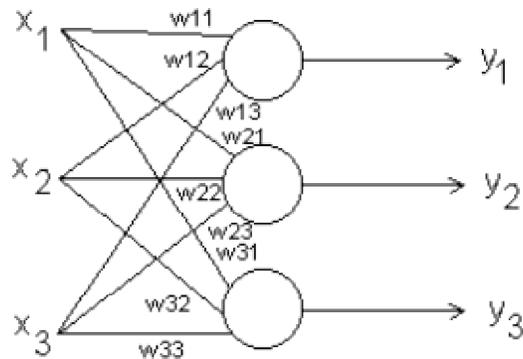


Figura 2.9 Modelo de Perceptrón

Como ya hemos dicho, un conjunto de neuronas no sirve para nada si previamente no le enseñamos qué debe hacer. En el Perceptrón podemos utilizar los métodos que conocemos: Supervisado, Por Refuerzo y No Supervisado.

2.4.1.2 El Perceptrón multicapa

Esta estructura nació con la intención de dar solución a las limitaciones del Perceptrón clásico o unicapa, y supuso el resurgimiento del movimiento conexionista. Como su nombre indica, se trata de un cuantos (dos o tres) perceptrones unicapa conectados en cascada.

El problema de este tipo de Perceptrón está en su entrenamiento, ya que es difícil modificar correctamente los pesos de la capa oculta. Para poder hacer aprender cosas a un Perceptrón de este tipo, se implementó el algoritmo de Retropropagación, que tal como su nombre indica tiene la función de ir propagando los errores producidos en la capa de salida hacia atrás. Esto se puede observar en la figura 2.10, la capa de entrada es la encargada de recibir los datos a reconocer, por lo que dependen del número de patrones que consta cada vector de entrenamiento, éstos se multiplican con los pesos de la capa de entrada (w_{ij})

para obtener las salidas que a su vez pasan a ser las entradas de la capa oculta, finalmente éstos se multiplican por los pesos de la capa oculta (ω_{jy}) para obtener la salida de la red.

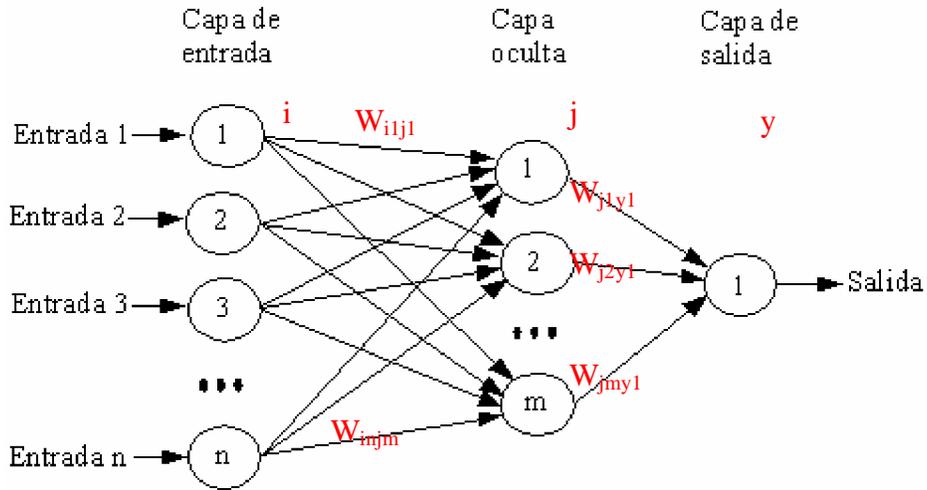


Figura 2.10 Modelo de Perceptrón Multicapa

Un ejemplo práctico de un Perceptrón multicapa podría ser su uso en visión artificial. Dada su capacidad para generalizar, las redes neuronales ya han demostrado su importancia en este campo. El único problema es la lentitud que tiene para aprender. [13]

2.4.1.3 Adaline y Madaline

Es una de las estructuras neuronales más simples junto con el Perceptrón. Está formada por un único elemento de procesado (Adaline: Adaptive Linear Element; "Widrow-Hoff"). Si se combinan varios adalines se obtiene la configuración denominada Madaline. Para su entrenamiento se utiliza un aprendizaje supervisado, concretamente, por corrección de error. El aprendizaje es Off Line.

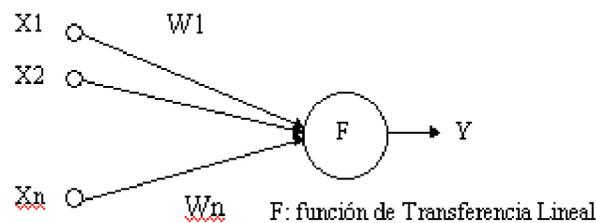


Figura 2.11 Modelo de Adaline

Se trata de modificar los pesos para tratar de reducir la diferencia entre la salida deseada y la actual (para cada patrón). La regla de aprendizaje a utilizar se denomina LMS: minimiza el Error Cuadrático Medio sobre todos los patrones de entrenamiento. [14]

2.4.1.4 Retropropagación

El aprendizaje con redes de neuronas se realiza actualmente siguiendo el enfoque de la Retropropagación de Gradiente, Retropropagación y el de los algoritmos constructivos. El primero necesita introducir a priori el número y conexiones de las unidades ocultas, y determinar los pesos por minimización de un costo. La red así obtenida es eventualmente simplificada eliminando unidades y/o conexiones que parecen inútiles.

El principal defecto de este enfoque consiste en la búsqueda de la mejor arquitectura a través de prueba y error. Por otra parte, con un enfoque constructivo se aprende al mismo tiempo el número de unidades y pesos, en el marco de una arquitectura que comienza generalmente con un perceptron. La característica de estos algoritmos es que construyen una RN adaptada a cada problema particular,

usando la información contenida en el conjunto de aprendizaje y evitando rediseñar la arquitectura.

Una vez construida la red, ésta debe ser capaz de predecir la clase de datos nuevos que no estén presentes en el conjunto de aprendizaje. La calidad del algoritmo de aprendizaje se traduce en la capacidad de predicción de la RN. Esta calidad se mide a través del error de generalización, que es la proporción de clasificaciones correctas realizadas por la red sobre nuevos datos. Esta cantidad se mide empíricamente sobre una serie de problemas estándar (benchmarks) que sirven de prueba. [13]

2.4.2 Modelos No Supervisados

2.4.2.1 Redes de Hopfield

El modelo Hopfield consiste de una red monocapa con N neuronas cuyos valores de salida son binarios 0/1 ó -1/+1. En la versión original del modelo Discrete Hopfield, las funciones de activación de las neuronas eran del tipo escalón. Se trataba, por lo tanto, de una red discreta con entradas y salidas binarias; sin embargo, posteriormente Hopfield desarrolló una versión continua con entradas y salidas analógicas utilizando neuronas de activación tipo sigmoideal.

Una de las características del modelo de Hopfield, es que se trata de una red auto-asociativa. Así como, varios patrones diferentes pueden ser almacenados en la red, como si se tratase de una memoria, durante la etapa de aprendizaje. Posteriormente, si se presenta a la entrada alguna de las informaciones almacenadas, la red evoluciona hasta estabilizarse, ofreciendo entonces en la salida la información almacenada, que coincide con la presentada en la entrada. Si, por el contrario, la información de entrada no coincide con ninguna de las almacenadas, por estar distorsionada o incompleta, la red evoluciona generando como salida la más parecida.

La red Hopfield continua ofrece más posibilidades que la discreta, ya que permite almacenar patrones formados por valores reales (por ejemplo, imágenes en color

o en blanco y negro con diferentes tonalidades de gris) y además facilita la resolución de determinados problemas generales de optimización.

En el caso de un reconocimiento de imágenes en una red Hopfield discreta sería más difícil trabajar con imágenes con tonos de grises o a colores, siendo que para una red Hopfield continua es más sencillo. [13]

2.4.2.2 Los mapas de Kohonen

Una propiedad del cerebro es que las neuronas están ordenadas significativamente. Esto implica que estructuras idénticas en el cerebro funcionan de forma diferente debido a su diferente topología cerebral. Otra propiedad básica es la inhibición lateral. Cuando una neurona se activa, ésta produce un estado excitatorio en las células más cercanas y un efecto inhibitorio en las más lejanas.

En 1982 Teuvo Kohonen propuso el "Kohonen's Feature Map", modelo que pretendía integrar las características de los modelos biológicos. Este nuevo modelo consiste en una Red Neuronal de dos capas, la primera de las cuales es una capa de entrada y la segunda es una "capa de competición". Las células de ambas capas están conectadas entre sí, de tal forma que una célula de la capa de entrada conecta con las de la capa de competición. Estas conexiones son ponderadas, es decir, el estímulo excita la capa de entrada, teniendo ambos la misma dimensión.

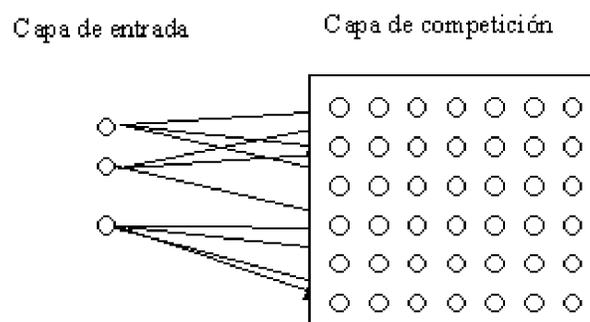


Figura 2.12 Modelo de Kohonen

Se trata de presentar un estímulo, el cual se propaga por la Red dependiendo de la función de distancia utilizada, y finalmente consideraremos como mejor célula (célula ganadora) a aquella que produzca menor señal en la capa de competición.

Podemos definir un espacio alrededor de la célula ganadora, en el cual se incrementarán las eficacias sinápticas de las células. Dicho incremento sería igual para todas las células del vecindario y no existirá dicho incremento en las células externas a dicho vecindario.

El objetivo de este tipo de redes es clasificar los patrones de entrada en grupos de características similares, de manera que cada grupo activará siempre la(s) misma(s) salida(s). Cada grupo de entradas queda representado en los pesos de las conexiones de la unidad de salida triunfante.

Una vez entrenada, podemos usar a la red para clasificar patrones de entrada similares en el espacio n-dimensional. Una clase o grupo de patrones similares tiende a controlar una neurona específica, que representará el centro de una esfera n-dimensional (de radio unitario, pues normalizamos los datos sobre la unidad). Esa neurona resultará la más activada frente a los patrones más parecidos a su vector de pesos. [13] [15]

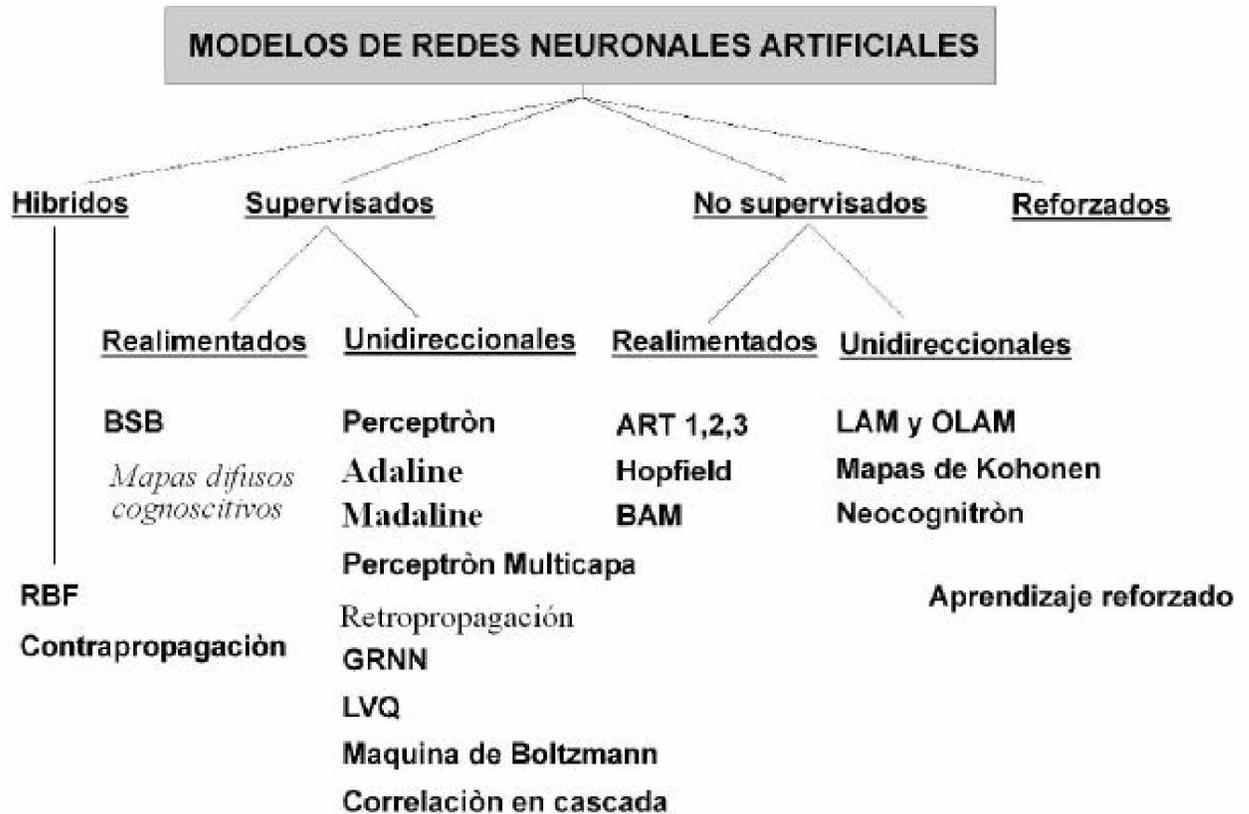
2.4.2.3 Redes basadas en la teoría de la resonancia adaptativa

En los esquemas de aprendizaje competitivo no existe garantía explícita sobre la estabilidad de los grupos de patrones de entrada durante el proceso de decrecimiento del factor de aprendizaje en las sucesivas iteraciones. Las redes basadas en la teoría de la resonancia adaptativa vienen a solucionar este problema.

En este tipo de redes, los vectores de pesos del elemento de procesado seleccionado como ganador sólo se actualizan con los patrones de entrada que "resuenan" con estos, es decir, si son "suficientemente similares". Si no son "suficientemente similares" se genera un nuevo elemento de procesado cuyos pesos son precisamente los del patrón de entrada utilizado. Evidentemente se

hace necesario definir el concepto de "suficiente similitud" que usualmente involucra un parámetro de vigilancia que lo controla. [13]

2.5 Resumen de Tipos y Características Importantes



2.13 Clasificación de las Redes según el tipo de aprendizaje y la arquitectura. [16]

Tabla 2.2.a Datos mas relevantes de los principales Modelos de Redes Neuronales [17]

| Modelo | Año | Aplicaciones | Comentarios | Limitaciones | Desarrollada por |
|-----------------------------------------------|------------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------|----------------------------------------------|
| Avalancha | 1967 | Reconocimiento de habla continua. Control de brazos robots | Ninguna red sencilla puede hacer todo esto | No es fácil alterar la velocidad o interpolar el movimiento | Stephen Grossberg |
| Teoría resonancia adaptativa (ART) | 1986 | Reconocimiento de patrones (radar, sonar, etc.) | Sofisticada. Poco utilizada | Sensible a la traslación, distorsión y escala | Gail Carpenter, Stephen Grossberg |
| Adaline / Madaline | 1960 | Filtrado de señales. Ecuador adaptativo. Modems. | Rápida, fácil de implementar con circuitos analógicos VLSI. | Solo es posible clasificar espacios linealmente separados. | Bernard Widrow |
| Retropropagación | 1974-85 | Síntesis de voz desde texto. Control de robots. Predicción. Reconocimiento de patrones. | Red más popular. Numerosas aplicaciones con éxito. Facilidad de aprendizaje. Potente. | Necesita mucho tiempo para el aprendizaje y muchos ejemplos. | Paul Werbos, David Parker, David Rumelhart |
| Memoria asociativa bidireccional (BAM) | 1985 | Memoria heteroasociativa de acceso por contenido | Aprendizaje y arquitectura simples. | Baja capacidad de almacenamiento. Los datos deben ser codificados | Bart Kosko |
| Máquinas de Boltzman y Cauchy | 1985-86 | Reconocimiento de patrones (imágenes, sonar y radar). Optimización | Redes simples. Capacidad de representación óptima de patrones. | La máquina de Boltzman necesita un tiempo muy largo de aprendizaje | Jeffrey Hinton, Terry Sejnowski, Harold Szu. |
| Brain-State-in-a-Box | 1977 | Extracción de conocimiento de bases de datos | Posiblemente mejor realización que las redes de Hopfield | Realización y potenciales aplicaciones no estudiadas totalmente. | James Anderson |
| Cerebellatron | 1969 | Control del movimiento de los brazos de un robot | Semejante a Avalancha | Requiere complicadas entradas de control | David Marr, James Albus, Andrés Pellionez |

Tabla 2.2.b Datos mas relevantes de los principales Modelos de Redes Neuronales [17]

| Modelo | Año | Aplicaciones | Comentarios | Limitaciones | Desarrollada por |
|-----------------------------------------------------------------|------------|-----------------------------------------------------------------|-------------------------------------------------------------------|------------------------------------------------------------|-------------------------|
| Counter-propagation | 1986 | Comprensión de imágenes | Combinación de perceptrón y TPM. | Numerosas neuronas y conexiones | Robert Hecht-Nielsen |
| Hopfield | 1982 | Reconstrucción de patrones y optimización | Puede implementarse en VLSI. Fácil de conceptualizar. | Capacidad y estabilidad. | John Hopfield. |
| Neocognitron | 1978-84 | Reconocimiento de caracteres manuscritos | Insensible a la traslación, rotación y escala. | Requiere muchos elementos de proceso, niveles y conexiones | K. Fukushima. |
| Perceptron | 1957 | Reconocimiento de caracteres impresos | La red más antigua. Construida en H.W. | No puede reconocer caracteres complejos | Frank Rosenblat |
| Self-organizing-map (SOM). Topology-preserving-map (TPM) | 1980-84 | Reconocimiento de patrones, codificación de datos, optimización | Realizan mapas de características comunes de los datos aprendidos | Requiere mucho entrenamiento | Teuvo Kohonen |

Capítulo 3 – Obtención de la Señal y Preprocesado

3.1 Archivos Wave

Para que el programa pueda reconocer la señal es necesario que éste pueda leer” esta información la cual es representada por bytes en una computadora. Para tal fin se hará uso de la tarjeta de sonido, que hoy en día es bastante común en cualquier equipo de cómputo. La tarjeta de sonido será utilizada como un CAD, es decir, un Codificador Analógico Digital que tendrá la tarea de digitalizar las muestras de sonido. Para poder almacenar esta información haremos uso de archivos WAVE (más comúnmente conocidos como WAV), que es un formato estándar de almacenamiento de sonido digital.

La forma de los archivos WAV de representar las señales analógicas es por medio de una serie de secuencias numéricas conocida como Modulación por Codificación de Pulso (PCM). El formato en el cual los archivos WAV son usualmente guardados es conocido como RIFF (Resource Interchange File Format) el cual, dentro de su estructura de datos, contiene toda la información relevante a la señal codificada: el número de bits por muestra, la frecuencia de muestreo, el número de canales, por mencionar los más importantes. [18] [19]

Los archivos WAV están estructurados como un conjunto, “trozos”, de información. Estos “trozos” pueden ser de distintos tipos pero los más importantes son: el trozo de formato y el trozo de datos. El trozo de formato contiene los parámetros que definen a la señal muestreada tales como la frecuencia de muestreo, el número de bits por muestra, el número de canales y el tamaño del archivo. El trozo de datos, como su nombre lo dice, contiene los datos de las muestras codificadas de la señal analógica. Los datos están arreglados en formato “little endian”, es decir, para el caso de 16 bits por muestra, que el primer byte será el menos significativo y el que le sigue será el byte más significativo. [19]

3.2 Filtros

El proyecto consiste en el reconocimiento de 3 instrumentos diferentes, esto implica que cuando los tres sean ejecutados simultáneamente las ondas sonoras emitidas se mezclarán siendo imposible diferenciarlas y por lo tanto reconocerlas. Para poder hacer una separación de las señales individuales se hará uso de filtros digitales.

Un filtro digital es un sistema discreto que trata de extraer alguna característica de una señal muestreada. Antes de entrar de lleno a los filtros será necesario definir qué es un sistema discreto y qué es una señal muestreada. [20]

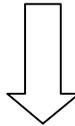
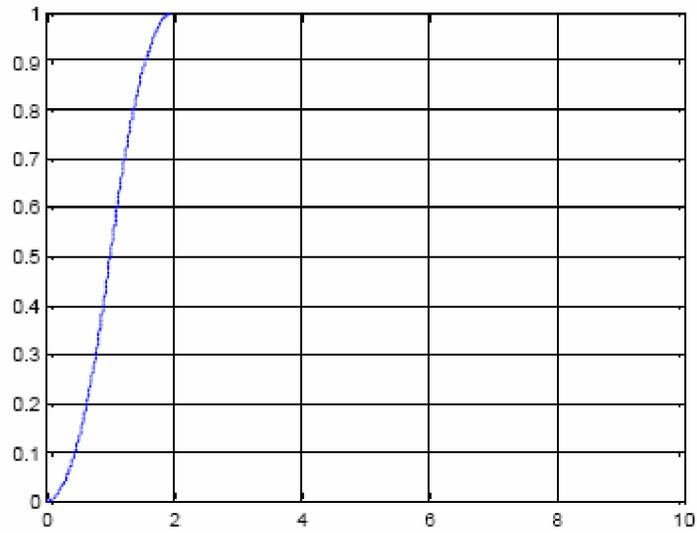
3.2.1 Sistemas Discretos

Los sistemas discretos se caracterizan porque sus entradas y salidas son secuencias de valores numerados, es decir, que cada elemento de entrada/salida es identificado por un número correlativo. Una secuencia entonces puede ser representada así: $\{X_k\}$ donde k es un entero asociado al elemento, de manera que:

$$\{X_k\} = \{X_0, X_1, X_2, \dots, X_N\} \quad \text{Ecuación 3.1}$$

La figura 3.1 muestra un sistema discreto el cual posee secuencias de valores tanto en las entradas como en las salidas. En otras palabras un sistema discreto es un algoritmo que permite transformar una secuencia (entrada) en otra (salida).

$$\{Y_k\} = f\{X_k\} \quad \text{Ecuación 3.2}$$



Sistema Discreto

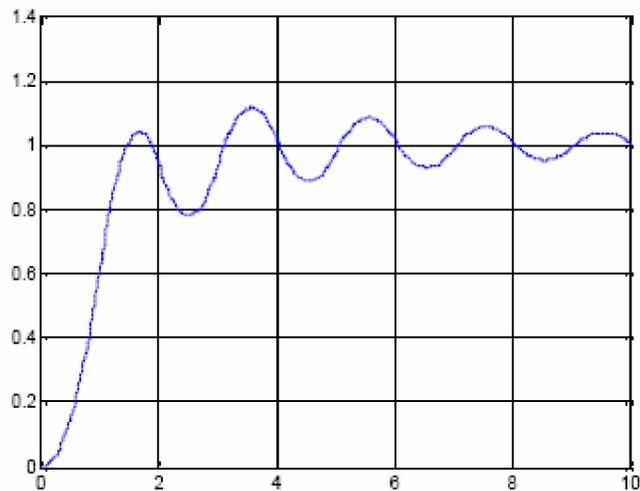
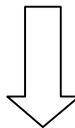


Figura 3.1 Un sistema discreto transforma una secuencia de entrada por una secuencia de salida diferente de la original

Los sistemas discretos pueden clasificarse de dos formas:

- *Los sistemas discretos estáticos* son aquellos en los cuales el elemento de salida, cuyo índice es “i”, depende únicamente del elemento de entrada “i”.
- *Los sistemas discretos dinámicos* son aquellos en los cuales los elementos de salida de cierto índice dependen de los elementos de entrada y salida diferentes al suyo.

Un sistema discreto será causal si el elemento de salida depende de los elementos de la secuencia de entrada y salida con índice menor al suyo como puede verse en la ecuación 3.3.

$$Y_K = f(Y_{K-1}, Y_{K-2}, \dots, X_K, X_{K-2}, \dots) \quad \text{Ecuación 3.3}$$

Si la función que relaciona cada elemento de salida es lineal, entonces el sistema se denomina así mismo lineal tal y como lo muestra la ecuación 3.4. [20]

$$Y_K = a_1 * Y_{K-1} + a_2 * Y_{K-2} + \dots + a_n * Y_{K-n}, b_0 * X_K + b_1 * X_{K-1} + \dots + b_m * X_{K-m} \quad \text{Ecuación 3.4}$$

3.2.2 Muestreo de Señales

El muestreo de señales consiste en la construcción de secuencias a partir de una señal continua. Estas secuencias son construidas a partir de la toma de valores de una señal en instantes de tiempo sucesivos.

Si estos instantes de toma de muestra son espaciados regularmente en el tiempo, entonces se dice que el muestreo es periódico y al instante de tiempo se le denomina período de muestreo T y a su inversa, frecuencia de muestreo f .

El período de muestreo jugará un papel muy importante en esta etapa. Si bien es cierto conviene muestrear datos con la menor frecuencia posible ésta deberá ser lo suficientemente alta como para poder evitar la pérdida de información para que sea posible la reconstrucción de la señal mediante la secuencia obtenida. [20]

3.2.3 Filtros Digitales

Un filtro digital es un sistema discreto utilizado para extraer características desde el dominio de la frecuencia sobre una señal muestreada. Algunas de las ventajas que ofrecen los filtros digitales sobre los analógicos son:

- *Respuesta Dinámica*, el ancho de banda de los filtros digitales está limitado por la frecuencia de muestreo, mientras que en los filtros analógicos suelen estar restringidos por un amplificador operacional.
- *Conmutabilidad*, si los parámetros de un filtro se conservan en un registro, estos parámetros se pueden modificar a voluntad. Además estos filtros se pueden conmutar pudiendo multiplexar en el tiempo para procesar varias entradas a la vez
- *Adaptabilidad*, un filtro digital se puede implementar físicamente (hardware) o mediante un programa de computadora (software).
- *Ausencia de problemas de componentes*, el filtro digital está compuesto por una secuencia de números invariantes en el tiempo. Al no haber componentes se elimina el problema de las tolerancias de los elementos, de impedancias de entrada y de salida y por ende del acople de impedancias.

Una de las diferencias entre los sistemas discretos dinámicos lineales e invariantes es sobre la duración de su respuesta ante el impulso. Es por esto que se habla de sistemas de respuesta de impulso finito o no recursivo (FIR, *Finite*

Impulse Response) y de sistemas de respuesta infinita o recursiva (IIR, *Infinite Impulse Response*). Tomando la ecuación de los sistemas discretos dinámicos:

$$Y_K = a_1 * Y_{K-1} + a_2 * Y_{K-2} + \dots + a_n * Y_{K-n}, b_0 * X_K + b_1 * X_{K-1} + \dots + b_m * X_{K-m}$$

Ecuación 3.4

Si todos los coeficientes a_i se igualan a 0 se tendrá un filtro FIR, siendo m el orden del filtro:

$$Y_K = b_0 * X_K + b_1 * X_{K-1} + \dots + b_m * X_{K-m}$$

Ecuación 3.5

Propiedades:

- La secuencia de ponderación es siempre infinita para un filtro IIR mientras que para un filtro no recursivo la respuesta es siempre finita e igual al orden del filtro.
- Los filtros FIR son siempre estables, es decir, la secuencia tiene todos sus lados acotados, mientras que la estabilidad de los filtros recursivos dependerá de su función de transferencia y habrá que utilizar algún método algebraico para analizar su estabilidad.
- Cualquier filtro recursivo podrá ser reemplazado por un filtro no recursivo con infinitos coeficientes, sus valores vendrán dados por la misma secuencia de ponderación del IIR. La conclusión inversa no se cumple.

Se seleccionó el tipo de filtros FIR ya que poseen las siguientes sobre los IIR que los hacen adaptarse mejor a las necesidades del sistema:

- Poseen la característica de tener una respuesta en frecuencia de fase lineal, esto es, que la señal a filtrar no será distorsionada por el filtro.
- Los FIR son por naturaleza estables, no existen polos, facilitando su diseño e implementación.

Para filtrar las señales de interés para el proyecto se utilizan tres filtros pasa banda. Un filtro pasa banda es aquel que sólo permite “pasar” un determinado ancho de banda de una señal, rechazando las demás frecuencias que no están contempladas en el ancho de banda antes mencionado. Se seleccionaron tres filtros debido a que se quiere distinguir tres diferentes instrumentos: un hi hat, un redoblante y un bombo. Los anchos de banda con los que opera cada uno de los filtros son los siguientes [21]:

- Para el bombo: de 10 a 200 Hz
- Para el redoblante: de 1,620 a 3,000 Hz
- Para el hi hat: de 5,500 a 10,000 Hz

El tipo de filtro utilizado en el sistema, es un filtro tipo FIR de fase lineal utilizando una ventana de truncamiento *Hamming*.

La ecuación 3.5 y 3.7 describen la función de transferencia del filtro utilizado. La ecuación 3.8 muestra cómo es utilizado el proceso de truncamiento mediante ventanas [21] [22].

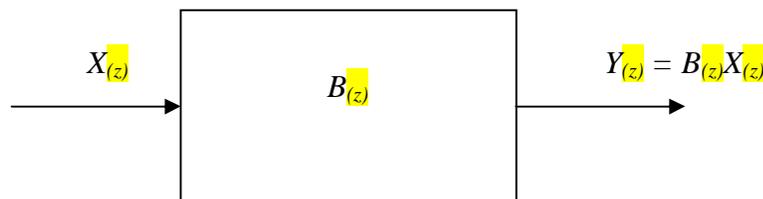


Figura 3.2 Diagrama a bloques del filtro implementado

$$Y_{(z)} = B_{(z)}X_{(z)}$$

Ecuación 3.6

Donde

ü $Y_{(z)}$ es la señal de salida del filtro

ü $X_{(z)}$ es la señal de entrada al filtro

ü $B_{(z)}$ son los coeficientes del filtro digital

$$B_{(z)} = b_{(1)} + b_{(2)}z^{-1} + \dots + b_{(n+1)}z^{-n}$$

Ecuación 3.7

$$b_{(n)} = w_{(n)}h_{(n)}$$

Ecuación 3.8

Donde

ü $b_{(n)}$ es el elemento n de la función característica del filtro

ü $w_{(n)}$ denota la ventana utilizada en el filtro

ü $h_{(n)}$ es la respuesta al impulso del filtro ideal

3.3 Compresión de Datos

En este punto surge un problema: la gran cantidad de información que se obtiene como resultado del muestreo y filtrado de la señal. Esto genera una gran cantidad de información a la entrada de la red tanto en la etapa de entrenamiento como en la de funcionamiento, lo cual no es deseado y no es eficiente y hace que la red sea un poco lenta siendo esto perjudicial al momento de trabajar el sistema en tiempo real.

Para superar esta situación se probó utilizando una herramienta llamada Transformada *Wavelet*.

El análisis de Wavelet es una técnica empleada para analizar señales en el tiempo, pero para poder entender mejor su naturaleza empezaremos por explicar a sus predecesores: la Transformada de Fourier y la Transformada Corta de Fourier.

3.3.1 Transformada de Fourier

El análisis de Fourier consiste en describir una señal mediante la suma de una serie de sinusoides cada una con diferente frecuencia. También es posible transformar la señal en dominio del tiempo al dominio de la frecuencia (figura 3.3).



Figura 3.3 A la izquierda, la señal en el dominio del tiempo. A la derecha la misma señal representada en dominio de la frecuencia.

El análisis de Fourier a pesar de ser una herramienta bastante útil posee una deficiencia: al transformar del dominio del tiempo al dominio de la frecuencia se pierde toda la información referente al tiempo. Al observar la gráfica de la señal (figura 3.3) en el dominio de la frecuencia es imposible saber cuando un evento particular ocurrió. Si la señal analizada no cambiara mucho en el tiempo, es decir una señal estacionaria, esta deficiencia no sería muy notoria. Sin embargo la información más interesante de las señales es que poseen características no estacionarias: tendencias, cambios abruptos, inicio y final de eventos. Estas características no pueden ser detectadas por el análisis de Fourier. [21]

3.3.2 Transformada Corta de Fourier

Para corregir esta deficiencia Dennis Gabor (1946) adaptó el análisis de Fourier para que sólo analizara una pequeña sección en el tiempo de la señal, creando una "ventana". A esta adaptación se le denominó TRANSFORMADA CORTA DE FOURIER, el cual mapea la función en 2 dimensiones: frecuencia contra tiempo (figura 3.4).



Figura 3.4 La transformada corta de Fourier puede relacionar en que momento estuvo presente una frecuencia en particular

Este análisis provee cierta información acerca de cuándo y qué frecuencias aparecieron durante un período de tiempo dado. Sin embargo la precisión de esta información es limitada, esta precisión está dada por el tamaño de la ventana, limitando así el análisis a señales cuya frecuencia esté dentro de los límites establecidos por el tamaño de la ventana. [21]

3.3.3 Transformada Wavelet

Este análisis fue desarrollado para superar la deficiencia anterior conservando siempre la característica de dar información tanto del tiempo como de la frecuencia. Esta técnica involucra “ventanas” de tiempo de tamaños variables de forma que para analizar las componentes de baja frecuencia utilizamos ventanas con intervalos de tiempo grandes y para las componentes de alta frecuencia utilizamos ventanas con intervalos de tiempo más cortos (figura 3.4).



Figura 3.5 La transformada Wavelet no tiene la limitante de las ventanas “estáticas”.

La transformada Wavelet continua se define como la suma de la señal multiplicada por una función Wavelet que ha sido escalada y desplazada a lo largo de toda la señal.

Cuando se habla de *escalar* un Wavelet se refiere a estrechar o comprimir la señal (figura 3.6).

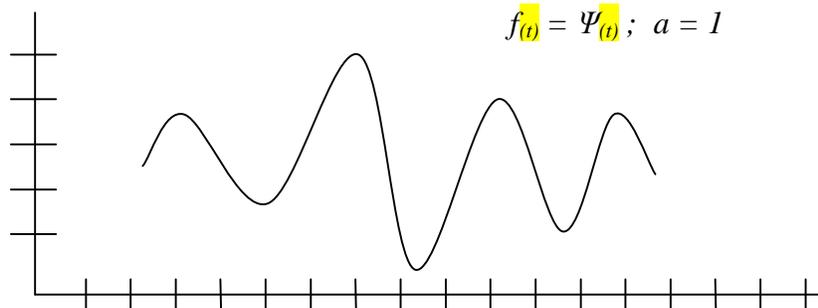


Figura 3.6a

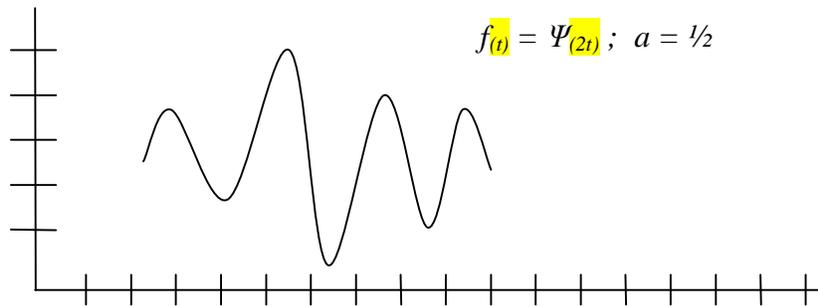


Figura3.6b

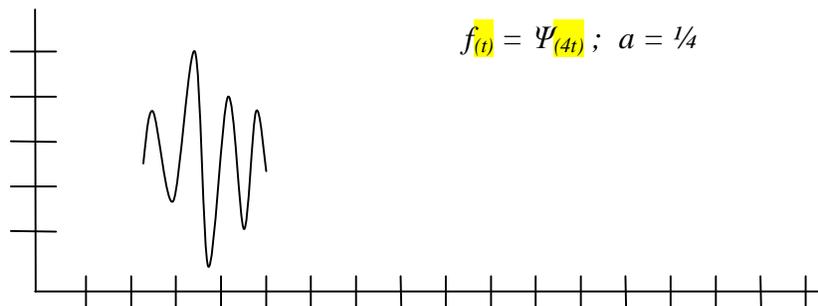


Figura 3.6c

Figura 3.6a, 3.6b y 3.6c Proceso de escalamiento: el periodo de la señal aumenta o disminuye dependiendo del factor a

En el cual a es el factor de escalamiento.

Otro concepto que hay que definir es el de *desplazamiento* que no es más que desplazar el Wavelet a lo largo del tiempo (figura 3.7).

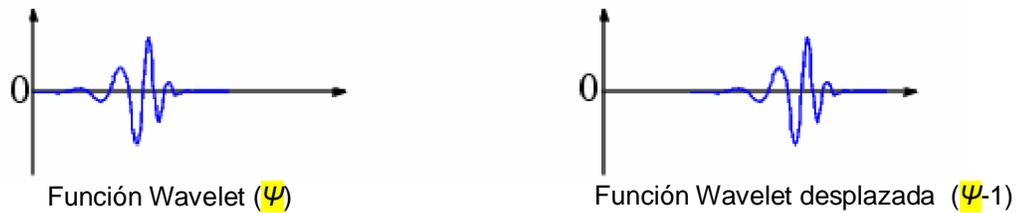


Figura 3.7 Desplazamiento del Wavelet a lo largo del tiempo

Los resultados de la transformada Wavelet continua son varios coeficientes C , los cuales son función de la escala y la posición.

La transformada Wavelet consiste básicamente en comparar la señal de análisis con otra señal: la Wavelet. El desplazamiento en el análisis de la Transformada Wavelet sirve para recorrer la señal que deseamos analizar y la escala sirve para poder obtener comparaciones con un rango amplio de frecuencias ya que al tener una escala grande podemos analizar señales de baja frecuencia y al tener escalas pequeñas podemos analizar señales de alta frecuencia.

[21]

3.3.4. Familias Wavelet

Existen diferentes tipos de Wavelet los cuales se agrupan por familias. Entre las distintas familias Wavelet podemos mencionar [21]:

- Wavelet Haar
- Wavelet Daubechies
- Wavelet Symlets
- Wavelet Morlet
- Wavelet Meyer

A continuación se muestra una Wavelet de la familia Daubechies.



Figura 3.8 Wavelet Daubechies

3.3.5 Descomposición de una Señal utilizando la transformada discreta Wavelet

La compresión utilizando Wavelets se basa en la descomposición. Esto se realiza haciendo pasar la señal por 2 filtros: uno pasa bajos y uno pasa altos.

En muchos de los casos la información más relevante de la señal se encuentra en las componentes de baja frecuencia mientras que las componentes de alta frecuencia son por lo general ruido.

En el análisis Wavelet se habla de aproximaciones A y detalles D . Las aproximaciones son las componentes de gran escala y de baja frecuencia. Por otro lado los detalles son las componentes de pequeña escala y alta frecuencia.

Como se mencionó anteriormente el proceso de descomposición se lleva a cabo mediante 2 filtros uno pasa bajos y uno pasa altos (figura 3.9).

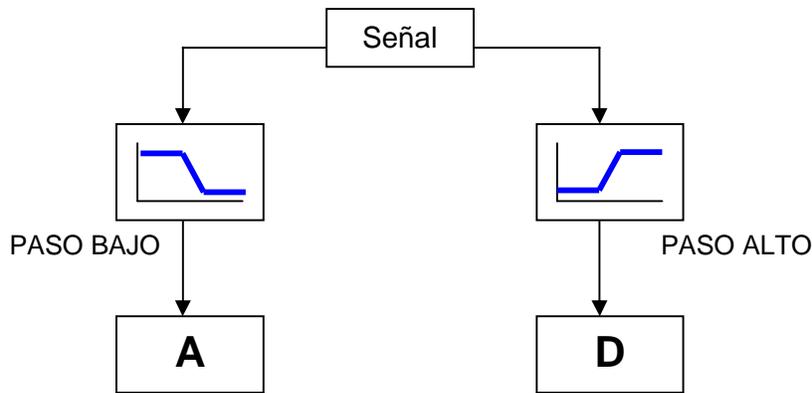


Figura 3.9 Esquema del proceso de descomposición

La señal pasa por los dos filtros y como resultado tendremos 2 señales diferentes. Al realizar esta operación el número de datos con los que se trabaja se ve incrementado al doble, ya que si la señal original poseía 1000 datos, el resultado de este procedimiento serían 2000 datos; 1000 de la señal resultado del filtro pasa bajos y 1000 de la señal del filtro pasa altos tal y como se ilustra en la figura 3.10.

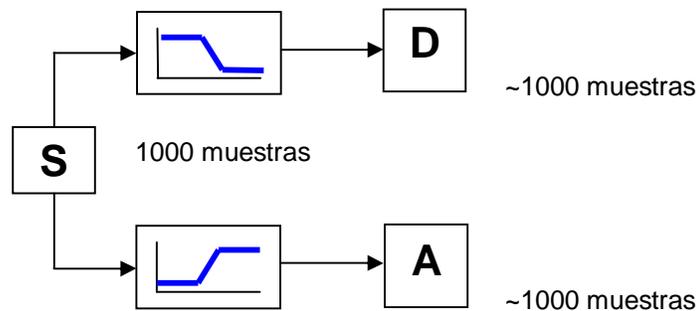


Figura 3.10 Con una señal de entrada de 1000 muestras se tienen al final 2000 muestras aproximadamente

En este punto es necesario introducir un nuevo concepto que será el que nos ayudará a disminuir el número de datos, este concepto es el sub-muestreo, el cual consiste en escoger únicamente las muestras con índice impar obteniendo así dos secuencias llamadas cA y cD como se muestra en la figura 3.11.

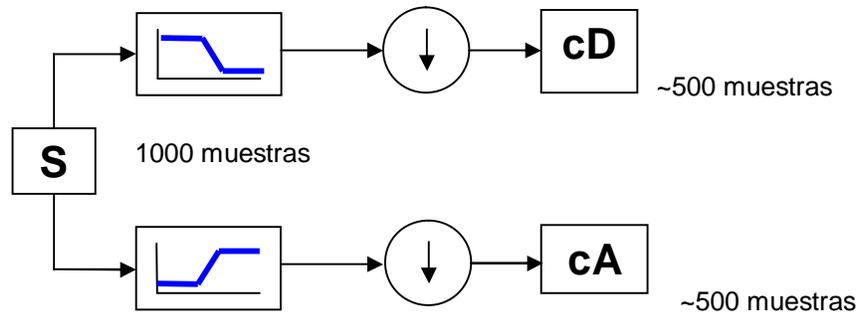


Figura 3.11 El sub-muestreo permite reducir el número de salidas a la mitad aproximadamente

Este es el proceso que se realiza en una descomposición simple, es decir de un solo nivel. Si nosotros incrementamos el nivel, o sea las veces que se repite el proceso de descomposición, estaríamos comprimiendo aún más los datos gracias al sub-muestreo. Esto es mostrado en lo que se llama “el árbol de descomposición Wavelet” (figura 3.12).

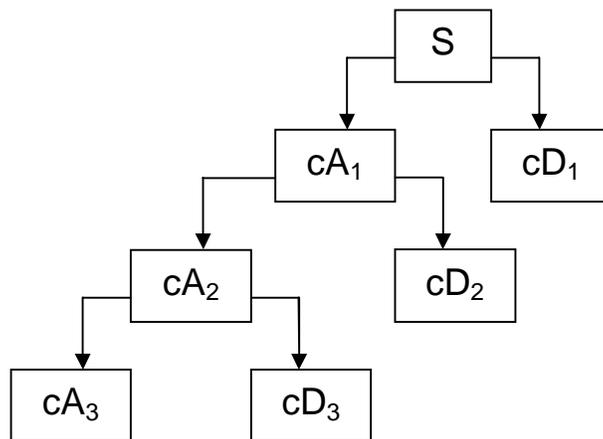


Figura 3.12 Árbol Wavelet de descomposición de nivel: al incrementar el nivel los datos se comprimen aun más

Como se mencionó anteriormente, para el análisis de la Transformada Wavelet se cuenta con una serie de familias con las cuales poder realizar el análisis. Para la transformada que es utilizada por el pre-procesado de la señal se optó por utilizar la Wavelet *symlet 4* con 3 niveles de descomposición.

[21]

A continuación se ilustrará de forma gráfica el procesamiento de cada una de las señales de los tres instrumentos con los cuales se trabajará.

- **BOMBO**

- Señal de entrada al sistema

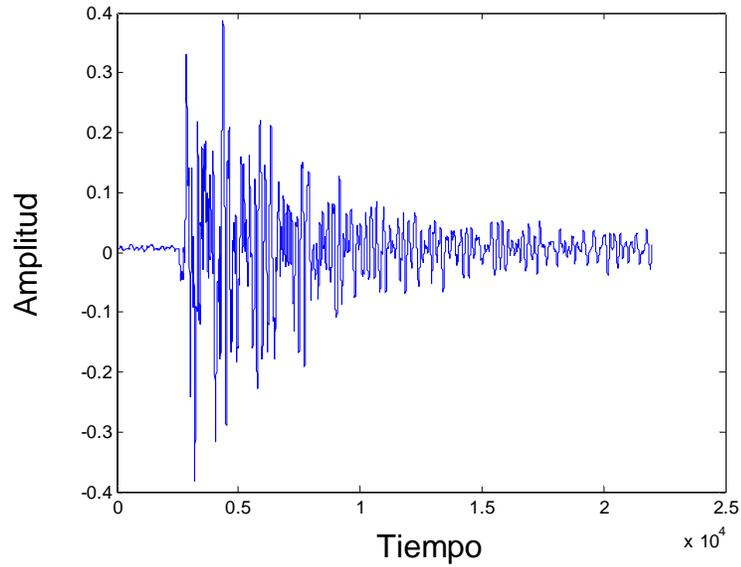


Figura 3.13 Señal adquirida de bombo.

- Señal filtrada

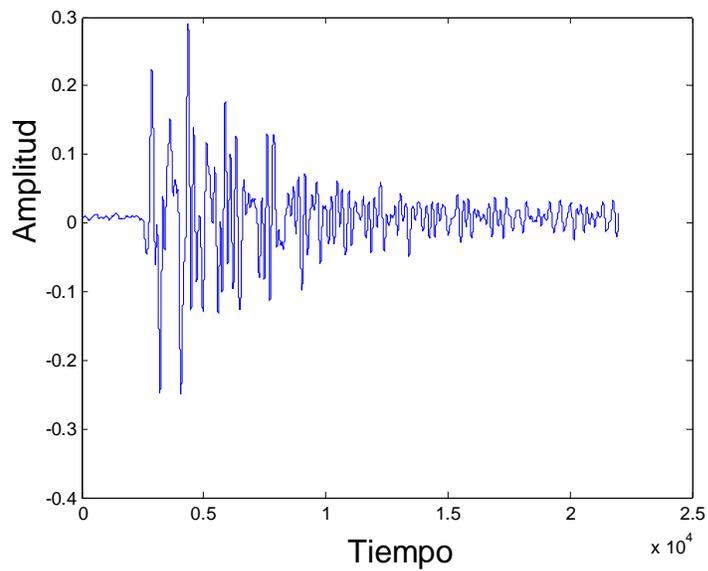


Figura 3.14 Señal obtenida después de aplicar un filtro pasa banda.

- Señal después de descomposición Wavelet

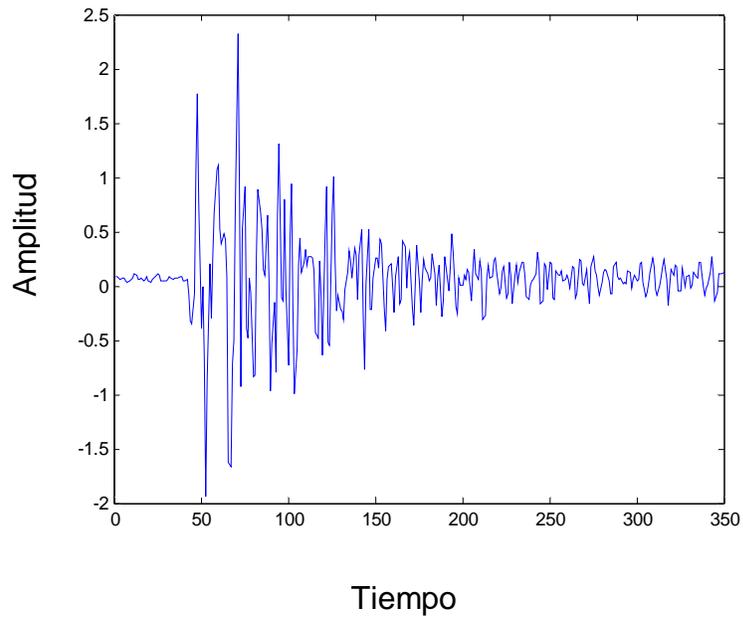


Figura 3.15 Señal obtenida al aplicar Wavelet

- **REDOBLANTE**

- Señal de entrada al sistema

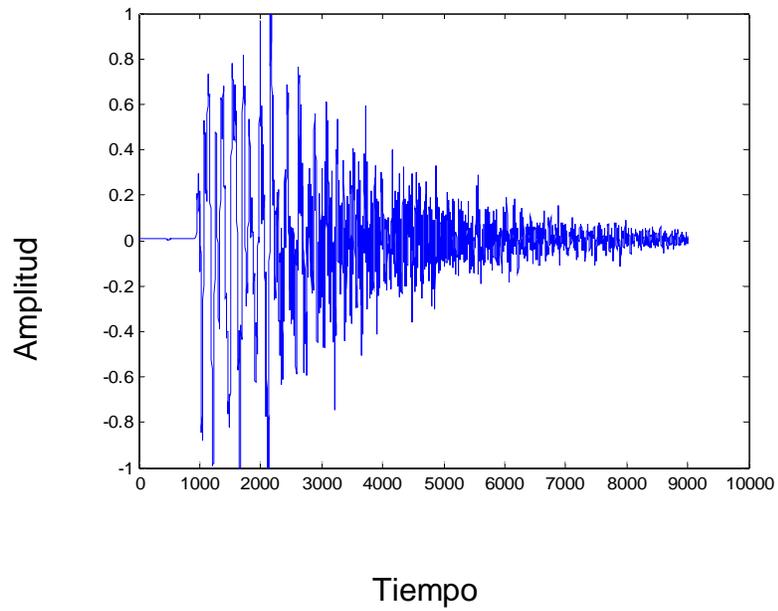


Figura 3.16 Señal adquirida de redoblante

- Señal filtrada

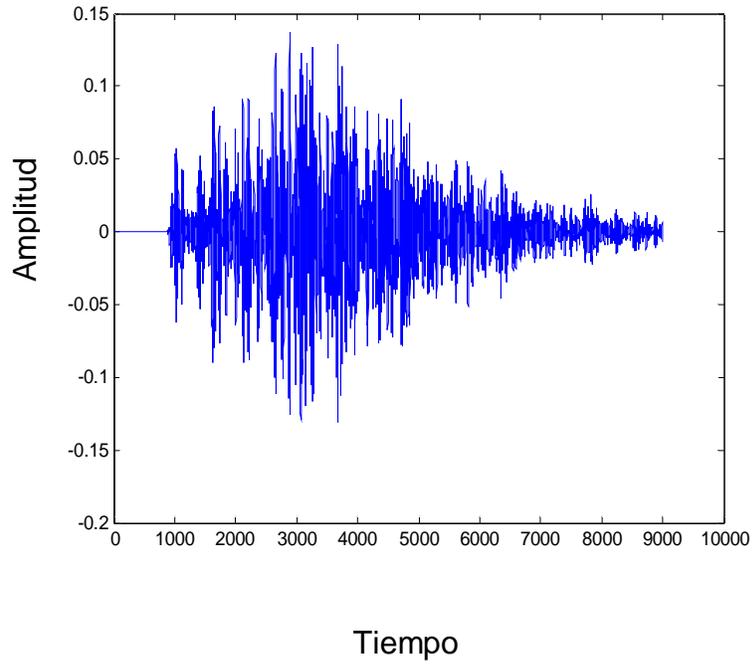


Figura 3.17 Señal de redoblante filtrada

- Señal después de la descomposición Wavelet

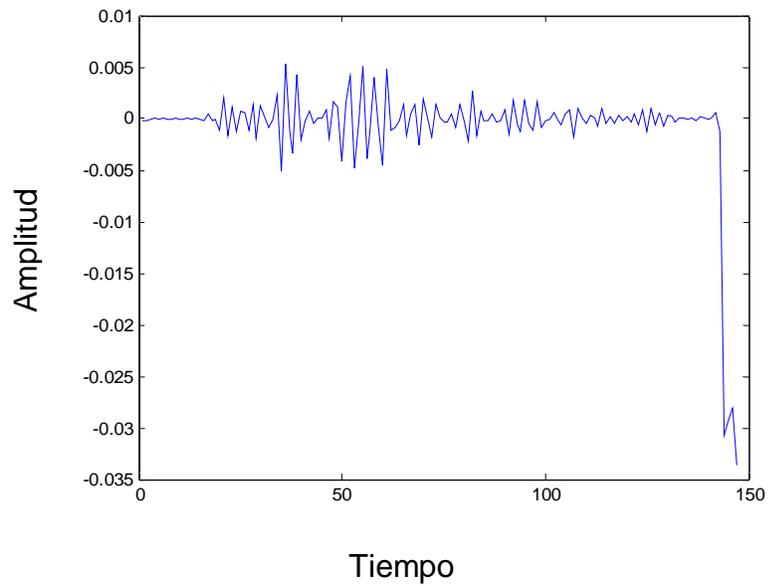


Figura 3.18 Señal resultado de aplicar Wavelet a la señal filtrada de redoblante

- **HI HAT**

- Señal de entrada al sistema

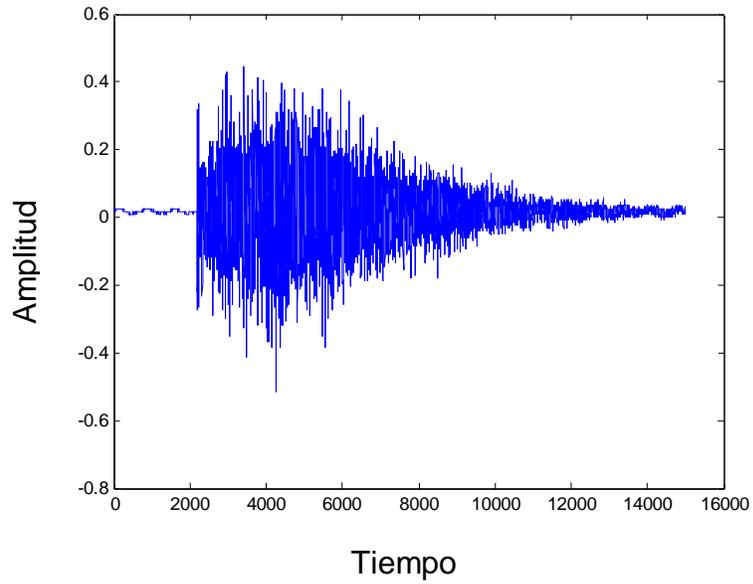


Figura 3.19 Señal adquirida de Hi Hat

- Señal filtrada

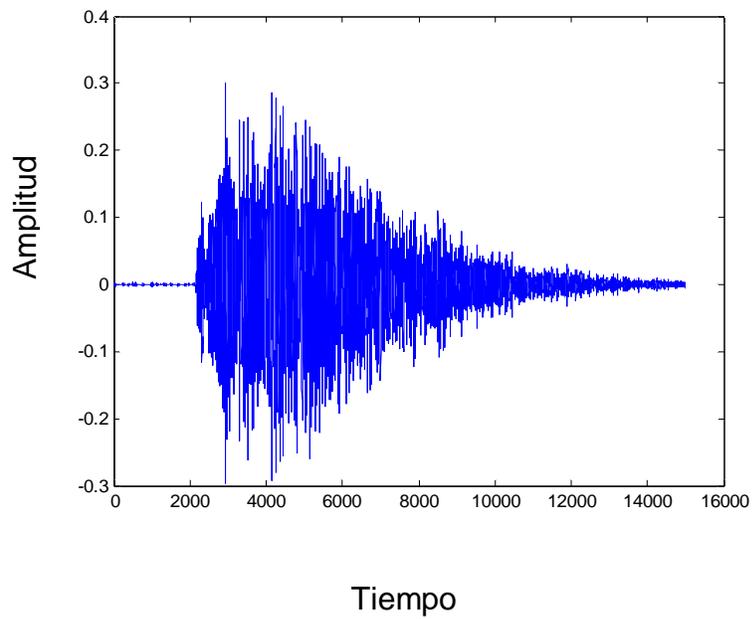


Figura 3.20 Señal de Hi Hat filtrada

- Señal después de descomposición Wavelet

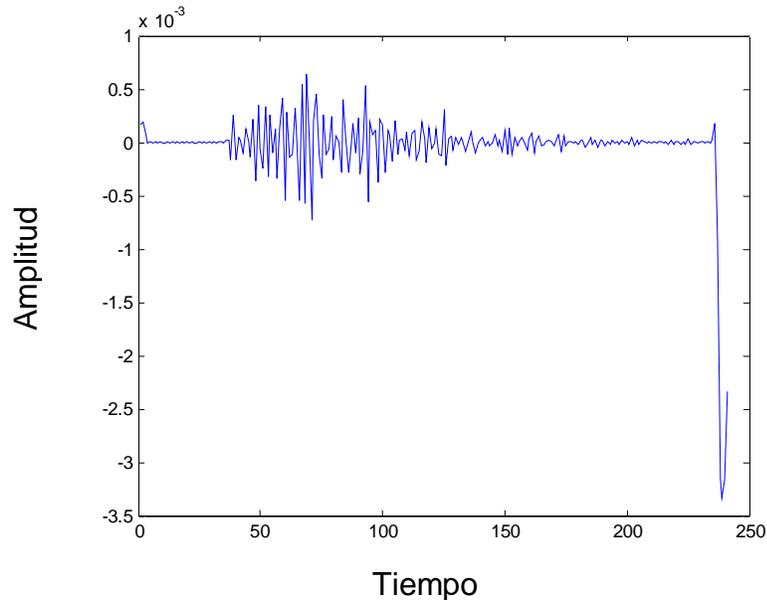


Figura 3.21 Resultado de aplicar Wavelet a la señal de Hi Hat filtrada

3.4 Procesado de la señal

Un sistema de reconocimiento está compuesto por una etapa de extracción de características y otra de identificación de dichas características. El procesado de la señal es la etapa en la cual se extraen las características propias de cada señal.

3.4.1 Coeficientes de Predicción Lineal.

Esta es una técnica usada frecuentemente en el reconocimiento de voz y que se trató de implementar en este proyecto para extraer de la señal de entrada un número reducido de características que permitieran a la red reconocer una señal de una mejor manera.

El modelo matemático establece que el tracto bucal puede modelarse mediante un filtro digital, siendo los parámetros los que determinan la función de transferencia. Los parámetros obtenidos son los coeficientes del filtro, los cuales hacen posible que al hacer pasar por el sistema una señal, la cual puede ser un tren de pulsos o una señal de ruido aleatorio, el sistema sea capaz de volver a generar la señal original.

La función de transferencia del filtro puede ser expresada de la siguiente manera:

$$s_{(n)} = -\sum_{k=1}^p a_k * s_{(n-k)} * G * u_{(n)} \quad \text{Ecuación 3.9}$$

Donde $s_{(n)}$ es la señal a analizar, G es la ganancia y $u_{(n)}$ es la entrada del filtro que bien puede ser un tren de pulsos o una fuente de ruido aleatorio. Ahora el trabajo consistirá en determinar los coeficientes de predicción y la ganancia, ésta última dependerá de la naturaleza de la señal. Serán los coeficientes del filtro los parámetros a ser reconocidos por la red neuronal. La determinación de estos coeficientes se hará minimizando el error que se comete cuando se intenta realizar la aproximación de la señal original.

Sea s_p la señal predicha y s la señal original, entonces:

$$s_{p(n)} = -\sum_{k=1}^p a_k * s_{(n-k)} \quad \text{Ecuación 3.10}$$

Y el error entre la señal predicha y la señal original será:

$$e_{(n)} = s_{(n)} - s_{p(n)} = s_{(n)} + \sum_{k=1}^p a_k * s_{(n-k)} \quad \text{Ecuación 3.11}$$

Utilizando el método de los mínimos cuadrados se calculan los coeficientes de predicción minimizando el error cuadrático medio con respecto a cada uno de los coeficientes.

Sea el error cuadrático total:

$$E = \sum_n e^2_{(n)} = \sum_{k=1}^p a_k * s^2_{p(n-k)} \quad \text{Ecuación 3.12}$$

Se realiza la minimización con respecto a a_k :

$$\sum_{k=1}^p a_k * \sum_n s_{p(n-k)} * s_{p(n-1)} = -\sum_n s_{p(n)} * s_{p(n-i)} \quad 1 < i < p \quad \text{Ecuación 3.13}$$

A continuación se procede a realizar un análisis de auto correlación. La función de auto correlación proporciona una medida de la correlación de la señal con una copia desfasada en el tiempo de si misma. De esta manera se busca identificar los coeficientes que minimizan los errores en la estimación de la señal predicha, es decir, permiten encontrar aquellos coeficientes que generan una señal mas parecida a la original.

Los resultados de aplicar esta técnica a los sonidos adquiridos de cada uno de los instrumentos en cuestión no fueron muy satisfactorios. Los tres instrumentos fueron tratados mediante esta técnica pero al analizar el resultado de cada una de las salidas de los filtros no se pudo hacer distinción entre ellos. Esto imposibilitaría la identificación de cada instrumento al ser la separabilidad casi imposible. A continuación se muestran los resultados graficados de aplicar esta técnica a la salida del filtro para cada uno de los tres instrumentos.

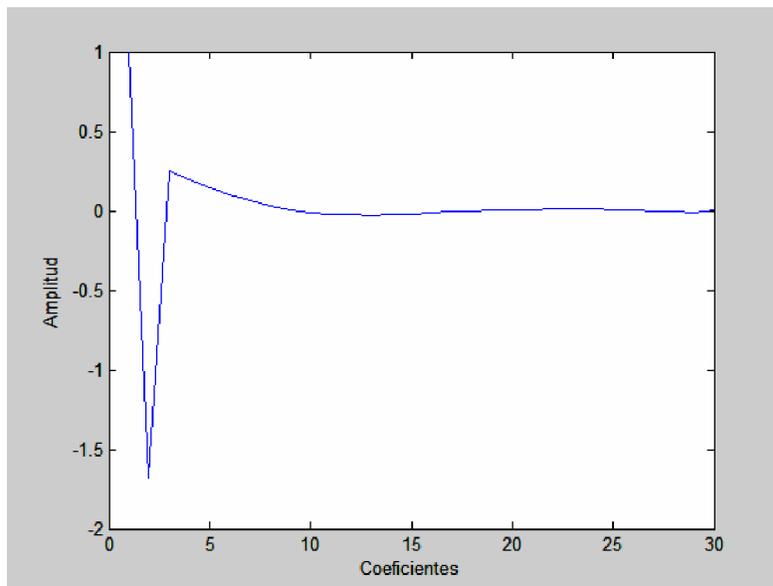


Figura 3.21 Análisis LPC a la salida del filtro de bombo al pasar una señal de bombo.

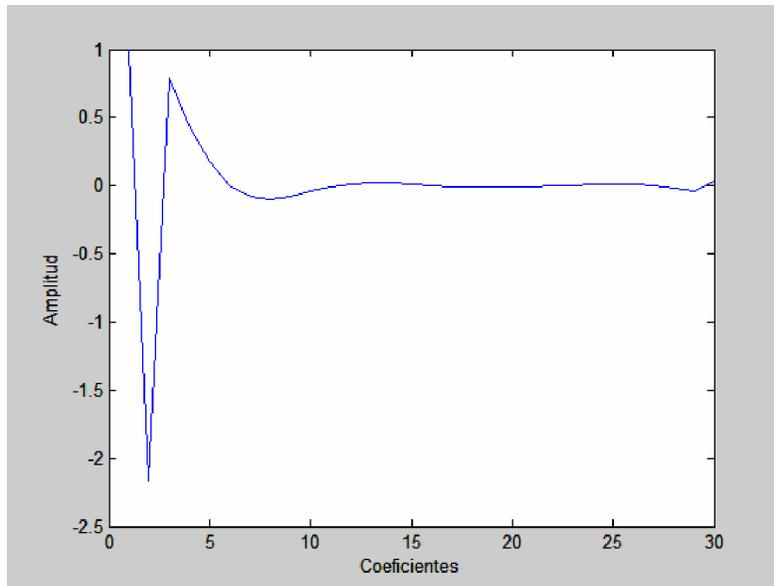


Figura 3.22 Análisis LPC a la salida del filtro de bombo al analizar una señal de redoblante.

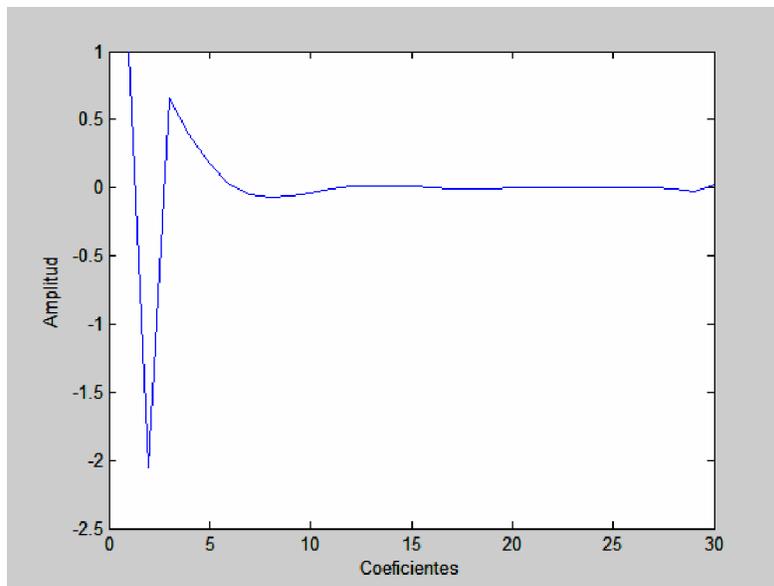


Figura 3.22 Análisis LPC a la salida del filtro de bombo al analizar una señal de hi hat.

Esta condición se mantiene para los demás instrumentos. [22]

3.4.2 Momentos Estadísticos

Al aplicar transformada de Fourier a la salida de cada uno de los filtros utilizados por el sistema, podemos observar que existe una clara distinción entre los

diferentes espectros de frecuencia en cuanto a su forma. Esto puede apreciarse en las siguientes figuras al aplicar el banco de filtros a un sonido de bombo.

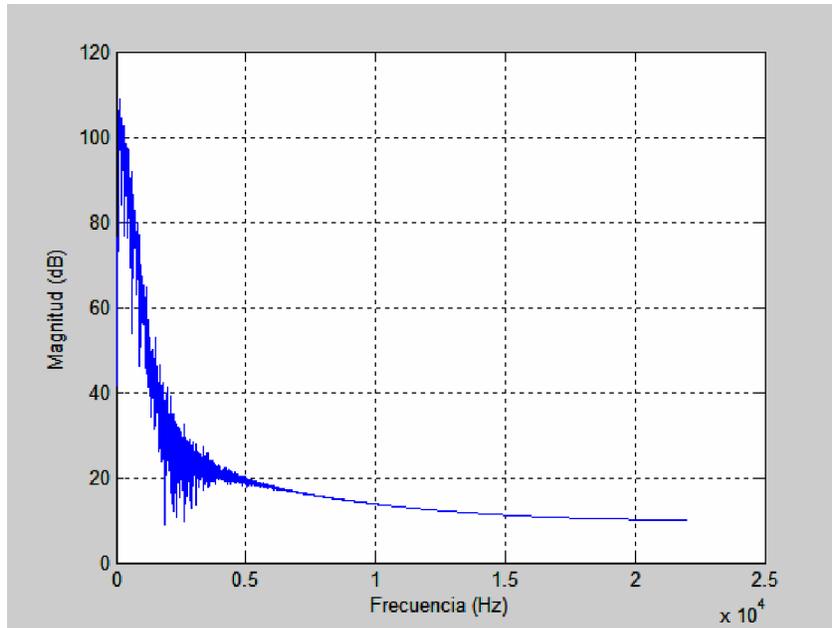


Figura 3.23 Salida filtro para bombo

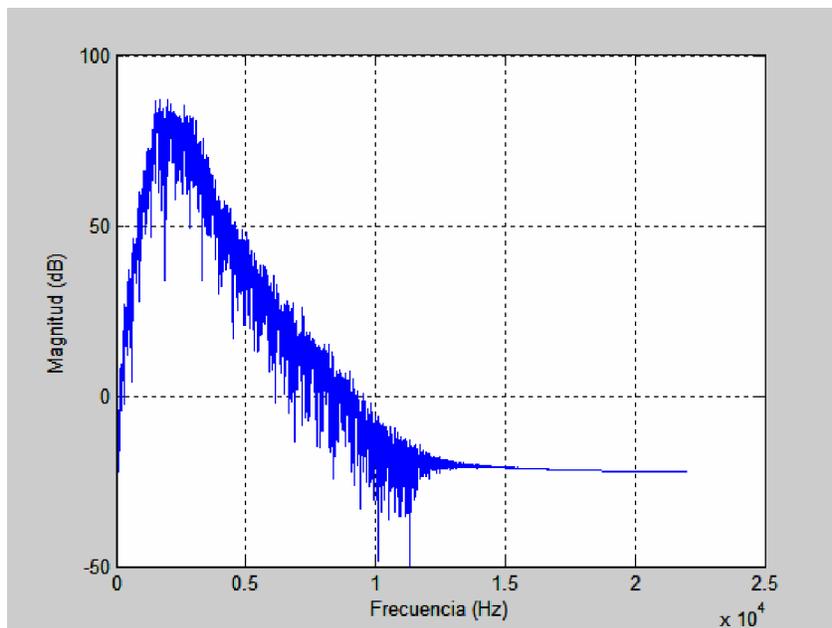


Figura 3.24 Salida filtro para redoblante

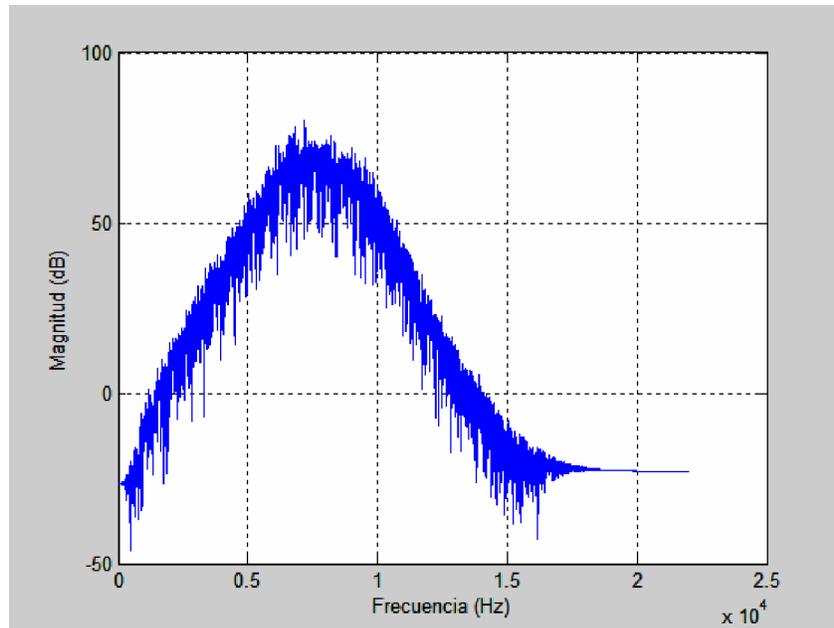


Figura 3.25 Salida filtro para hi hat

Los momentos estadísticos son herramientas que permiten describir una serie de datos por medio de su media (primer momento), su varianza (segundo momento) y demás características (momentos de orden mayor) que dependen de la forma que describan los datos en cuestión.

Dado que cada una de las formas anteriores son diferentes, se utilizó esta técnica para poder identificar el espectro de frecuencia que corresponde a cada uno de los instrumentos.

Para el cálculo de los momentos considerar el siguiente ejemplo: se desea describir mediante momentos la señal de la figura 3.26.

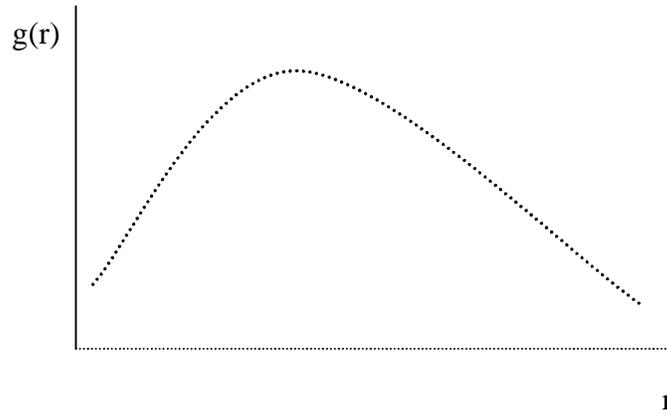


Figura 3.26 Señal a ser descrita por momentos estadísticos

Tomamos la amplitud g como una variable aleatoria de v y se forma un histograma de amplitud $p_{(v_i)}$, $i=0,1,2,\dots, A-1$, donde A es el número de incrementos discretos en los cuales se ha dividido la escala de amplitud. Luego tomando en cuenta que $p_{(v)}$ es un estimado de la probabilidad de ocurrencia de v se tiene que el n ésimo momento de v alrededor de su media es:

$$\mu_{n(v)} = \sum_{i=0}^{A-1} (v_i - m)^n p_{(v_i)} \quad \text{Ecuación 3.14}$$

Donde

$$m = \sum_{i=0}^{A-1} v_i p_{(v_i)} \quad \text{Ecuación 3.15}$$

Otra manera de poder calcular el n ésimo momento es normalizando $g_{(r)}$ de manera que el área interna descrita por la grafica sea unitaria. En otras palabras $g_{(r)}$ es tratado como la probabilidad de ocurrencia del valor r_i . En este caso, r es tratada como la variable aleatorio y los momentos son determinados por la ecuación 3.16:

$$\mu_{n(r)} = \sum_{i=0}^{K-1} (r_i - m)^n g_{(r_i)} \quad \text{Ecuación 3.16}$$

Donde

$$m = \sum_{i=0}^{K-1} r_i g_{(t_i)}$$

Ecuación 3.17

Para poder aplicar esta técnica de procesamiento estadístico es necesario suavizar las señales de salida del banco de filtros buscando obtener únicamente una línea que describe el contorno del espectro.

Para este procesamiento se trabajaron con 10 momentos para poder tener un número de descriptores que pudieran brindar una mejor información a la red.

Con esta técnica surgió un problema: los valores devueltos tenían una tendencia lineal para los tres instrumentos dificultando así la separabilidad de éstos. Además surgió el inconveniente de que los valores eran extremadamente grandes, especialmente los momentos de orden mayor, lo cual no es conveniente para el entrenamiento de la red. A manera de ejemplificar se aplicará la técnica a un sonido de redoblante.

La señal de salida del filtro de redoblante es la siguiente:

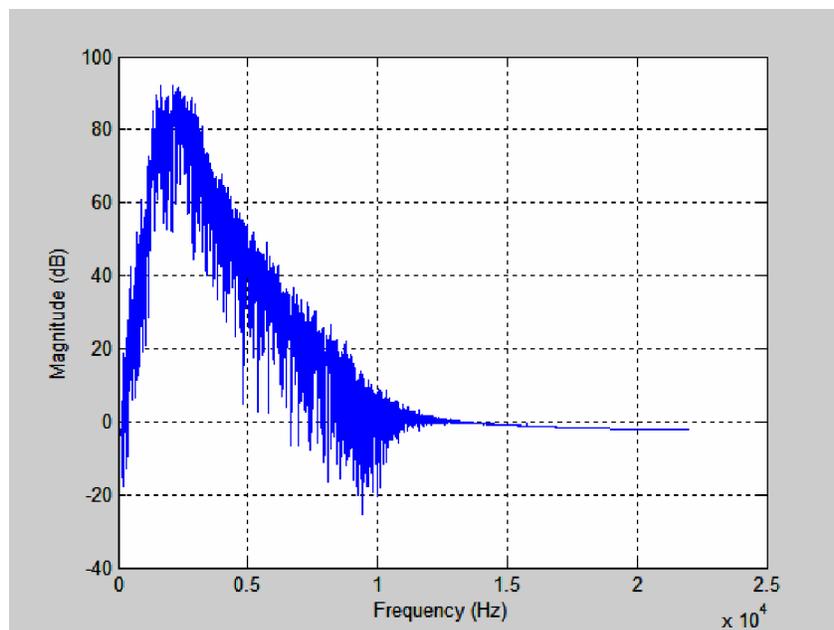


Figura 3.27 Espectro de frecuencia resultado de pasar un sonido de redoblante a través del filtro pasa banda.

Como se dijo anteriormente esta señal debe ser suavizada y al mismo tiempo su área debe ser unitaria. La señal resultante es como sigue:

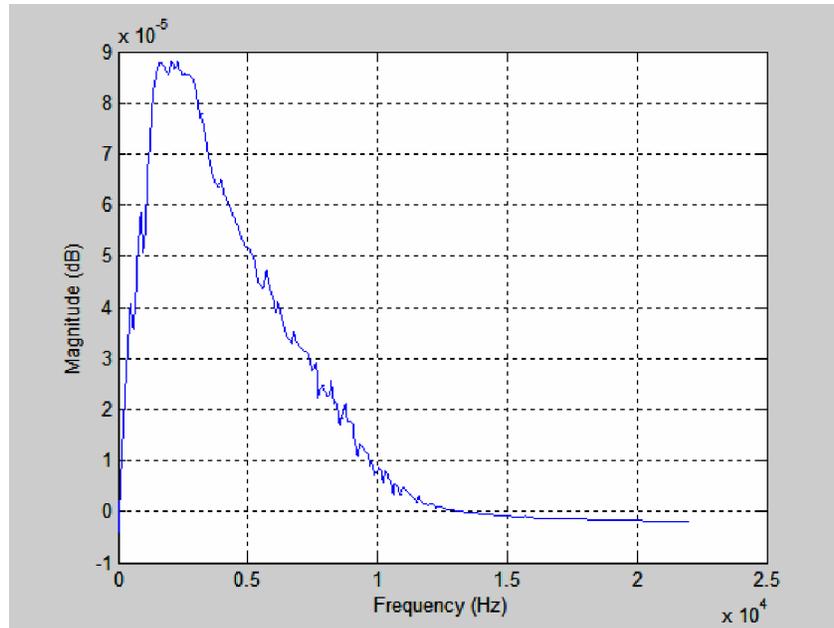


Figura 3.28 Espectro de frecuencias suavizado

Una vez se tiene la señal normalizada es posible aplicar momentos estadísticos para poder extraer características de la señal. Para el caso, los valores que se obtienen para esta señal en particular fueron:

$$M1 = 9.68232716225259e-011$$

$$M2 = 3587003.78345418$$

$$M3 = -693599380664.279$$

$$M4 = -2.68670833441572e+016$$

$$M5 = -9.91099921441895e+020$$

$$M6 = -3.52193874109478e+025$$

$$M7 = -1.25169666825544e+030$$

$$M8 = -4.46869992152639e+034$$

$$M9 = -1.60571711613833e+039$$

$$M10 = -5.80661318963124e+043$$

La falta de repetitividad de estos momentos respecto a otros sonidos del mismo instrumento fue otro inconveniente. Es por esto que se optó por descartar el procesado basado en momentos estadísticos. [23]

3.4.3 Máximos y Rangos de Frecuencia

Tal y como se expresó en el literal anterior, cada instrumento genera un espectro de frecuencias y magnitudes propio. Esto quiere decir que cada instrumento tiene mayor presencia en determinados rangos del espectro de frecuencia. Esta presencia depende también de la intensidad con que se ejecute cada instrumento ya que la batería es un instrumento de percusión.

Bajo este contexto se realizó un estudio tratando de determinar la mayor magnitud de las componentes de frecuencia así como la frecuencia a la cual ocurre. En otras palabras se buscó la frecuencia fundamental para cada instrumento haciendo uso de la transformada de Fourier.

Para la realización de este estudio se tomaron 4 muestras de cada instrumento y se hicieron pasar por el banco de filtros. A estas señales filtradas se les buscó cual era la componente con mayor amplitud y a que frecuencia se encontraba tal componente. Los resultados se muestran a continuación:

- Muestra Bombo
 - Valores máximos

Tabla 3.1a Valores de potencia para Bombo

| Filtro / muestra | 1 | 2 | 3 | 4 |
|------------------|---------------|---------------|---------------|---------------|
| Bombo | 110.83 | 107.86 | 110.88 | 109.08 |
| Redoblante | 80.46 | 82.03 | 82.03 | 82.63 |
| Hi hat | 61.71 | 63.76 | 65.5 | 67.47 |

- Valores de frecuencia

Tabla 3.1b Valores de frecuencia para Bombo

| Filtro / muestra | 1 | 2 | 3 | 4 |
|------------------|--------|--------|--------|--------|
| Bombo | 169.86 | 137.32 | 140.68 | 137.43 |
| Redoblante | 1595.7 | 2229.2 | 2350.8 | 2278.1 |
| Hi hat | 6749.3 | 7405.2 | 7273.4 | 7363.1 |

- Muestra Redoblante

- Valores máximos

Tabla 3.2a Valores de potencia para Redoblante

| Filtro / muestra | 1 | 2 | 3 | 4 |
|------------------|--------------|--------------|--------------|--------------|
| Bombo | 109.85 | 111.97 | 115.07 | 113.84 |
| Redoblante | 91.98 | 93.90 | 98.46 | 95.74 |
| Hi hat | 87.13 | 88.43 | 92.82 | 91.05 |

- Valores de frecuencia

Tabla 3.2b Valores de frecuencia para Redoblante

| Filtro / muestra | 1 | 2 | 3 | 4 |
|------------------|--------|--------|--------|--------|
| Bombo | 208.88 | 201.43 | 201.88 | 202.19 |
| Redoblante | 1639.7 | 2591.1 | 1979.8 | 1617.2 |
| Hi hat | 7651.9 | 6763.9 | 7076.1 | 7317.5 |

- Muestra Hi Hat

- Valores máximos

Tabla 3.3a Valores de potencia para Hi Hat

| Filtro / muestra | 1 | 2 | 3 | 4 |
|------------------|--------------|--------------|--------------|--------------|
| Bombo | 80.88 | 82.74 | 82.57 | 87.35 |
| Redoblante | 81.74 | 80.45 | 82.76 | 83.04 |
| Hi hat | 92.29 | 90.16 | 92.76 | 98.76 |

- Valores de frecuencia

Tabla 3.3b Valores de frecuencia para Hi Hat

| Filtro / muestra | 1 | 2 | 3 | 4 |
|------------------|--------|--------|--------|--------|
| Bombo | 342.32 | 379.01 | 269.09 | 454.5 |
| Redoblante | 2223.8 | 2193.8 | 1173.6 | 2183.9 |
| Hi hat | 7012.5 | 9311.5 | 7394.6 | 8163 |

Se puede ver claramente una diferencia en cuanto a los máximos que proporciona cada salida de filtro y el sonido que se está analizando.

De esta forma se puede determinar si la salida de cada filtro posee una señal válida, mediante la verificación de los valores máximos de los componentes y la componente frecuencial a la cual pertenecen dichos máximos.

Por ejemplo, si se tiene un sonido que posea una componente mayor de 90 dB en el rango de frecuencias de 6,900 a 9,400 Hz quiere decir que el sonido pertenece a un Hi Hat.

3.5 Modulo de calibración

Se ha desarrollado una aplicación basada en el sistema de reconocimiento descrito en la sección 3.4.3 el cual consiste en calibrar el equipo en busca de un mejor reconocimiento. Esta calibración busca registrar 10 muestras por cada instrumento y luego obtener los máximos y sus respectivas frecuencias. El objetivo que se persigue es tener un promedio de las muestras adquiridas y establecer rangos dentro de los cuales podamos tener la certeza de que si una lectura esta dentro de dichos rangos poder afirmar que corresponde al instrumento ejecutado.

Para establecer dichos rangos se toman los siguientes criterios:

- Bombo
 - Umbral: debe ser mayor que el promedio de los máximos obtenidos
 - Rango de frecuencias: se establecerá como $\pm 30\%$ del promedio de las frecuencias

- Redoblante
 - Umbral: debe ser mayor que el promedio de los máximos obtenidos
 - Rango de frecuencias: se establecerá como $\pm 15\%$ del promedio de las frecuencias

- Hi Hat
 - Umbral: debe ser mayor que el promedio de los máximos obtenidos
 - Rango de frecuencias: se establecerá como $\pm 18\%$ del promedio de las frecuencias

Capítulo 4 – Selección de la Red Neuronal a Utilizar

En base a ciertas características descritas a continuación se seleccionaron tres redes neuronales para realizar pruebas, Retropropagación, Kohonen y ART2, de estas se elegirá una tomando en cuenta su porcentaje de reconocimiento y algunas características propias de la señal a reconocer.

A continuación se detallan los diferentes sonidos para el entrenamiento de las redes y sus diferencias.

Tabla 4.1 Diferencias de los sonidos empleados para el entrenamiento de las redes

| SONIDO | DESCRIPCIÓN |
|--------------|----------------------------------------------------|
| Bombo 1 | El instrumento se tocó fuerte |
| Bombo 2 | El instrumento se tocó suave |
| Bombo 3 | Este sonido se grabó mientras una persona hablaba |
| Bombo 4 | Este sonido se grabó mientras una persona silbaba |
| Bombo 5 | Este sonido se grabó con la tv de fondo |
| Bombo 6 | Este sonido se grabó mientras una persona aplaudía |
| Redoblante 1 | El instrumento se tocó fuerte |
| Redoblante 2 | El instrumento se tocó suave |
| Redoblante 3 | Este sonido se grabó mientras una persona hablaba |
| Redoblante 4 | Este sonido se grabó mientras una persona silbaba |
| Redoblante 5 | Este sonido se grabó con la tv de fondo |
| Redoblante 6 | Este sonido se grabó mientras una persona aplaudía |
| Hi hat 1 | El instrumento se tocó fuerte |
| Hi hat 2 | El instrumento se tocó suave |
| Hi hat 3 | Este sonido se grabó mientras una persona hablaba |
| Hi hat 4 | Este sonido se grabó mientras una persona silbaba |
| Hi hat 5 | Este sonido se grabó con la tv de fondo |
| Hi hat 6 | Este sonido se grabó mientras una persona aplaudía |

4.1 Redes Seleccionadas para realizar las pruebas

La selección de las redes neuronales que se utilizaron estuvo basada en el análisis de las tablas 2.2 a y 2.2 b, en el propósito de reconocer patrones rítmicos y en características particulares de la señal a reconocer.

Los sonidos que se desean reconocer son descritos en función del tiempo por medio de valores analógicos, por lo cual son necesarias redes que tengan la capacidad de reconocer entradas analógicas, en base a esto se escogió a la red ART2 entre las redes de Teoría de Resonancia Adaptativa (ART), ya que es la única de esta familia capaz de trabajar con este tipo valores. Esta capacidad también la poseen las redes Kohonen y Retropropagación.

4.1.1 Kohonen

El aprendizaje en el modelo de Kohonen es de tipo Off-line, por lo que se distingue una etapa de aprendizaje y otra de funcionamiento. En la etapa de aprendizaje se fijan los valores de las conexiones (feedforward) entre la capa de entrada y la salida. Esta red utiliza un aprendizaje no supervisado de tipo competitivo, las neuronas de la capa de salida compiten por activarse y sólo una de ellas permanece activa ante una determinada información de entrada a la red, los pesos de las conexiones se ajustan en función de la neurona que haya resultado vencedora.

Durante la etapa de entrenamiento, se presenta a la red un conjunto de informaciones de entrada (vectores de entrenamiento) para que ésta establezca en función de la semejanza entre los datos las diferentes categorías (una por neurona de salida), que servirían durante la fase de funcionamiento para realizar clasificaciones de nuevos datos que se presenten a la red. Los valores finales de los pesos de las conexiones entre cada neurona de la capa de salida con las de entrada se corresponderán con los valores de los componentes del vector de aprendizaje que consigue activar la neurona correspondiente. En el caso de existir más patrones de entrenamiento que neuronas de salida, más de uno deberá asociarse con la misma neurona, es decir pertenecerán a la misma clase.

En este modelo el aprendizaje no concluye después de presentarle una vez todos los patrones de entrada, sino que habrá que repetir el proceso varias veces para refinar el mapa topológico de salida, de tal forma que cuantas más veces se presenten los datos, tanto más se reducirán las zonas de neuronas que se deben activar ante entradas parecidas, consiguiendo que la red pueda realizar una clasificación mas selectiva.

Un concepto muy importante en la red de Kohonen es la *zona de vecindad*, o vecindario alrededor de la neurona vencedora i^* . Los pesos de las neuronas que se encuentren en esta zona, a la que se le dará el nombre de $X_{(q)}$, serán actualizados junto con el peso de la neurona ganadora en un ejemplo de aprendizaje cooperativo. [24]

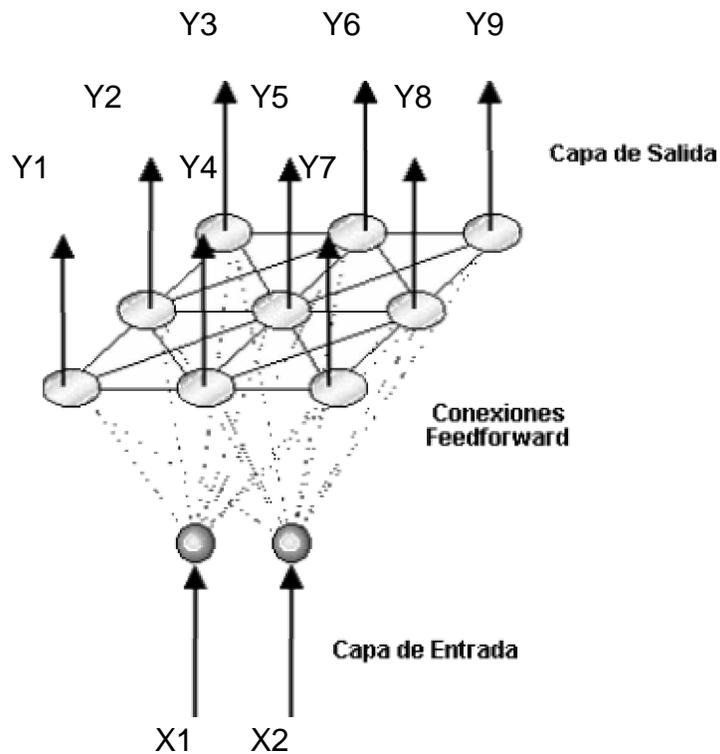


Figura 4.1 Arquitectura de Red Kohonen

4.1.1.1 Algoritmo de aprendizaje:

1. Inicializar los pesos con valores relativamente pequeños.
2. Se define la zona de vecindad entre las neuronas (normalmente esto se llama radio).
3. Poner un vector de entrada.
4. Se determina la neurona vencedora de la etapa de salida.

Esto se realiza aplicando los siguientes pasos:

- a) Calcular la salida de todas las neuronas según la fórmula

$$d_j = \sum_{i=1}^N (e_i^k - w_{ji})^2 \quad \text{Ecuación 4.1}$$

- b) Se determina cual de estos valores es el menor y ese correspondería a la neurona ganadora.

5. Se modifican los pesos de la neurona ganadora y de las neuronas con vecindad

$$w_{ji}(t+1) = w_{ji}(t) + \alpha(t) * [e_i^k - w_{ji}(t)] \quad \text{Ecuación 4.2}$$

α es un valor de aprendizaje, algunos autores recomiendan que el valor inicial de α sea $1/t$; donde t es el número total de iteraciones que se realizarán.

6. Se repite el proceso para todos los patrones de entrada.
7. se repite el proceso como mínimo unas 500 veces, se recomienda disminuir α en pequeños valores.
8. El proceso termina después del número de iteraciones esperado.

Algoritmo de funcionamiento.

1. Poner un vector de entrada.
2. Encontrar la neurona ganadora.
 - a) Calcular la salida de todas las neuronas según la fórmula

$$d_j = \sqrt{\sum_{i=1}^N (e_i^k - w_{ji})^2} \quad \text{Ecuación 4.3}$$

b) Se determina cual de estos valores es el menor y ese correspondería a la neurona ganadora.

3. En caso de querer probar un nuevo patrón regresar al paso 1. [27]

4.1.1.2 Pruebas con Red Kohonen

Las pruebas realizadas consisten en introducir la información obtenida del preprocesado para que sean reconocidas por la red. La cantidad de datos que se introduce a la red son 2205, esto es así debido a que para poder describir frecuencias de hasta 10 Hz con una frecuencia de 22050 Hz se necesitan 2205 muestras. Se escogió 10 Hz ya que en el caso del bombo, éste posee las frecuencias mas bajas de entre los tres instrumentos a reconocer. Para el entrenamiento se utilizaron una muestra de cada uno de los tres instrumentos que serán reconocidos. El número de salidas por lo tanto será de tres lo cual permitirá saber que instrumento fue reconocido.

Para el proceso de entrenamiento se utilizaron los primeros dos vectores de cada instrumento (6 en total) y para las pruebas se utilizaron los doce vectores restantes.

Tabla 4.2a Resultados de entrenamiento Red Kohonen, Arreglo inicial de vectores en el que no se obtuvo un buen porcentaje de aprendizaje.

| | 700 épocas y ventana = 0 | | | | 700 épocas y ventana = 1 | | | |
|------------------------------|--------------------------|-----------|------------|---------------|--------------------------|-----------|------------|---------------|
| | alpha 0.3 | alpha 0.1 | alpha 0.05 | alpha 0.00018 | alpha 0.3 | alpha 0.1 | alpha 0.05 | alpha 0.00018 |
| Bombo 3 | X | X | X | X | X | X | X | X |
| Redoblante 3 | - | - | - | X | - | - | - | X |
| Hi Hat 3 | X | - | X | - | X | - | X | - |
| Bombo 4 | X | X | - | - | X | X | - | - |
| Redoblante 4 | - | - | X | X | - | - | X | X |
| Hi Hat4 | - | X | - | - | - | X | - | - |
| Bombo 5 | X | X | X | X | X | X | X | X |
| Redoblante 5 | X | - | - | - | X | - | - | - |
| Hi Hat 5 | - | - | - | X | - | - | - | X |
| Bombo 6 | X | X | X | X | X | X | X | X |
| Redoblante 6 | - | - | X | - | - | - | X | - |
| Hi Hat 6 | - | X | - | X | - | X | - | X |
| Porcentaje De Reconocimiento | 50% | 50% | 50% | 58.3% | 50% | 50% | 50% | 58.3% |

Tabla 4.2b Resultados de entrenamiento Red Kohonen, Arreglo Final de Vectores.

| | Para un vecindario nulo, ventana = 0 (solo se actualiza la neurona ganadora) | | | | | | | |
|------------------------------|------------------------------------------------------------------------------|-----------|------------|---------------|------------|-----------|------------|---------------|
| | 700 épocas | | | | 500 épocas | | | |
| | alpha 0.3 | alpha 0.1 | alpha 0.05 | alpha 0.00018 | alpha 0.3 | alpha 0.1 | alpha 0.05 | alpha 0.00018 |
| Bombo 3 | x | X | x | x | x | x | X | x |
| Bombo 4 | x | X | x | x | x | x | X | x |
| Bombo 5 | x | X | x | x | x | x | X | x |
| Bombo 6 | x | X | x | x | x | x | X | x |
| Redoblante 3 | x | X | x | x | x | x | X | x |
| Redoblante 4 | x | X | x | x | x | x | X | x |
| Redoblante 5 | x | X | x | x | x | x | X | x |
| Redoblante 6 | x | X | x | x | x | x | X | x |
| Hi Hat 3 | - | - | - | - | - | - | - | - |
| Hi Hat 4 | - | - | - | - | - | - | - | - |
| Hi Hat 5 | - | - | - | - | - | - | - | - |
| Hi Hat 6 | - | - | - | - | - | - | - | - |
| Porcentaje De Reconocimiento | 66.67% | 66.67% | 66.67% | 66.67% | 66.67% | 66.67% | 66.67% | 66.67% |

Tabla 4.2c Cuadro Resumen de Pruebas con Red Kohonen

| | Para una ventana=1 | | | |
|------------------------------|--------------------|-----------|------------|---------------|
| | 700 épocas | | | |
| | alpha 0.3 | alpha 0.1 | alpha 0.05 | alpha 0.00018 |
| Bombo 3 | x | x | x | X |
| Bombo 4 | x | x | x | X |
| Bombo 5 | x | x | x | X |
| Bombo 6 | x | x | x | X |
| Redoblante 3 | x | x | x | X |
| Redoblante 4 | x | x | x | X |
| Redoblante 5 | x | x | x | X |
| Redoblante 6 | x | x | x | X |
| Hi Hat 3 | - | - | - | - |
| Hi Hat 4 | - | - | - | - |
| Hi Hat 5 | - | - | - | - |
| Hi Hat 6 | - | - | - | - |
| Porcentaje De Reconocimiento | 66.67% | 66.67% | 66.67% | 66.67% |

x: patrón reconocido
-: patrón no reconocido

4.1.2 Retropropagación (Backpropagation)

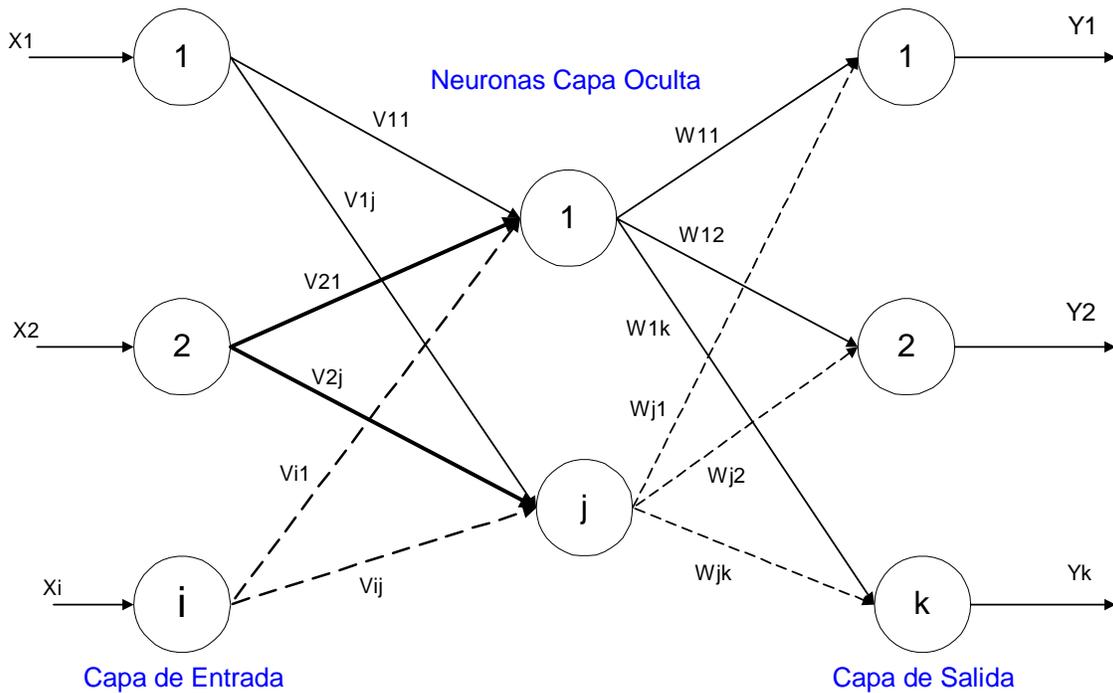


Figura 4.2 Arquitectura Red Retropropagación [27]

Es un tipo de red de aprendizaje supervisado, que emplea un ciclo propagación-adaptación de dos fases. Una vez que se ha aplicado un patrón a la entrada de la red como estímulo, éste se propaga desde la primera capa a través de las capas superiores de la red, hasta generar una salida. La señal de salida se compara con la salida deseada y se calcula una señal de error para cada una de las salidas.

Las salidas de error se propagan hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa oculta que contribuyen directamente a la salida. Sin embargo las neuronas de la capa oculta solo reciben una fracción de la señal total del error, basándose aproximadamente en la contribución relativa que haya aportado cada neurona a la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido una señal de error que describa su contribución relativa al error total. Basándose en la señal de error percibida, se actualizan los pesos de conexión de cada neurona, para hacer que la red converja hacia un estado que permita clasificar correctamente todos los patrones de entrenamiento.

La importancia de este proceso consiste en que, a medida que se entrena la red, las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada. Después del entrenamiento, cuando se les presente un patrón arbitrario de entrada que contenga ruido o que esté incompleto, las neuronas de la capa oculta de la red responderán con una salida activa si la nueva entrada contiene un patrón que se asemeje a aquella característica que las neuronas individuales hayan aprendido a reconocer durante su entrenamiento. Y a la inversa, las unidades de las capas ocultas tienen una tendencia a inhibir su salida si el patrón de entrada no contiene la característica para reconocer, para la cual han sido entrenadas.

Varias investigaciones han demostrado que, durante el proceso de entrenamiento, la red de Retropropagación tiende a desarrollar relaciones internas entre neuronas con el fin de organizar los datos de entrenamiento en clases. Esta tendencia se puede extrapolar, para llegar a la hipótesis consistente en que todas las unidades de la capa oculta de una red de Retropropagación están asociadas de alguna manera a características específicas del patrón de entrada como consecuencia del entrenamiento. Lo que sea o no exactamente la asociación puede no resultar evidente para el observador humano, lo importante es que la red ha encontrado una representación interna que le permite generar las salidas deseadas cuando se le dan las entradas, en el proceso de entrenamiento. Esta misma representación interna se puede aplicar a entradas que la red no haya visto antes, y la red clasificará estas entradas según las características que compartan con los ejemplos de entrenamiento. [25]

Por lo general se usan las funciones:

- Lineal cuya derivada es 1
- Sigmoidal unipolar

$$f_{(x)} = \frac{1}{1+e^{-x}} \text{ cuya derivada es } f'_{(x)} = f_{(x)} [1 - f_{(x)}]$$

- Sigmoidal Bipolar

$$f_{(x)} = \frac{2}{1 + e^{-x}} - 1 \text{ cuya derivada es } f'_{(x)} = \frac{1}{2} [1 + f_{(x)}] * [1 - f_{(x)}]$$

Función a utilizar en el entrenamiento del sistema a implementar

Observaciones:

La cantidad de neuronas ocultas o de capas ocultas depende de la aplicación en la que se esté trabajando, por lo general la mayoría de autores coinciden en que basta con una capa de neuronas ocultas para realizar el aprendizaje de esta red y recomiendan que se pruebe con varias cantidades de neuronas en la capa oculta. [27]

4.1.2.1 Algoritmo de aprendizaje.

1. Inicializar los pesos con valores aleatorios pequeños (**sugerencia: entre 0.05 y 0.25**). También asignar valores a los ajustes de cada neurona oculta y de salida.
2. Poner un patrón a reconocer en la entrada.
3. Determinar la salida de la(s) neurona(s) oculta(s)

$$z_{in_j} = v_{0_j} + \sum_{i=1}^n x_i v_{ij} \quad \text{Ecuación 4.4}$$

donde

Z_{in_j} son las neuronas de la capa oculta

v_{0_j} es el ajuste para cada una de las neuronas de la capa oculta.

x_i son las entradas a la red

v_{ij} son los pesos que influyen en las neuronas de la capa oculta.

n es el número de neuronas de la capa oculta.

4. Aplicar a la salida de las neuronas ocultas la función de salida.

$$z_j = f_{(z_{in_j})} \quad \text{Ecuación 4.5}$$

5. Propagar z_j a las siguientes neuronas (por lo general estas neuronas son las de salida, pero en algunos casos son otra capa de neuronas ocultas, en cuyo caso se repiten los pasos 3 y 4 para dicha(s) capa(s)).

6. Determinar la salida de las neuronas de salida.

$$y_{in_k} = w_{0k} + \sum_{j=1}^p z_j w_{jk} \quad \text{Ecuación 4.6}$$

Donde

y_{in_k} son las neuronas de la capa de salida.

w_{0k} es el ajuste para cada una de las neuronas de la capa de salida.

z_k son las entradas a la red

w_{jk} son los pesos que influyen en las neuronas de la capa de salida.

p es el número de neuronas de la capa de salida.

7. Aplicar a la salida de las neuronas la función de salida.

$$y_k = f_{(y_{in_k})} \quad \text{Ecuación 4.7}$$

8. Se determina el error de la entrada basándose en la salida esperada (para la capa de salida).

$$\delta_k = (t_k - y_k) * f'_{(y_{in_k})} \quad \text{Ecuación 4.8}$$

9. Se cambian los pesos de la capa de salida.

$$\Delta w_{jk} = w_{jk} + \alpha \delta_k z_j \quad \text{Ecuación 4.9}$$

$$\Delta w_{0k} = w_{0k} + \alpha \delta_k \quad \text{Ecuación 4.10}$$

NOTA: α es el factor de aprendizaje que puede tomar valores entre 0 y 1

10. Se determina el error para la capa oculta.

$$\delta_{in_j} = \sum_{k=1}^m \delta_k w_{jk} \quad \text{Ecuación 4.11}$$

$$\delta_j = \delta_{in_j} * f'_{(z_{in_j})} \quad \text{Ecuación 4.12}$$

11. Se cambian los pesos de la capa de salida.

$$\Delta v_{ij} = v_{ij} + \alpha \delta_j x_i \quad \text{Ecuación 4.13}$$

$$\Delta v_{0j} = v_{0j} + \alpha \delta_j \quad \text{Ecuación 4.14}$$

12. Con los nuevos pesos se regresa al numeral 2 y solo se avanza al siguiente numeral hasta que se hallan introducido todos los patrones que se quiere reconocer.

13. Se determina el error cuadrático medio de los errores de la capa de salida y se compara con un valor predeterminado como válido para el aprendizaje. Si es menor o igual a dicho valor se termina el aprendizaje en caso contrario se vuelve al numeral 2.

$$\delta^2 = \frac{1}{2L} \sum_{k=1}^L \delta_k^2 \quad \text{Ecuación 4.15}$$

[27]

4.1.2.2 Pruebas con Red de Retropropagación

Para retropropagación se introducen solamente 4 entradas a la red, estas eran el resultado del preprocesado y de un agrupamiento realizado también en el preprocesado, la arquitectura de nuestra red fue la siguiente, 4 neuronas de entrada, 4 en la capa oculta y 4 salidas, los parámetros para la elección de este número de neuronas en la capa oculta se basan en ejemplos encontrados y estudiados en el libro “Fundamentals of Neural Networks”. Se utiliza la función de activación Sigmoidal bipolar para cada salida de las neuronas de la capa de entrada, capa oculta y capa de salida. Los valores de las salidas de estas neuronas serán interpretadas para poder conocer que instrumento fue reconocido. La información es interpretada de la siguiente manera:

| Salida 1 | Salida 2 | Salida 3 | Salida 4 | Reconoce |
|----------|----------|----------|----------|------------|
| 1 | -1 | 1 | 1 | Silencio |
| 1 | -1 | -1 | 1 | Bombo |
| -1 | -1 | 1 | 1 | Redoblante |
| -1 | 1 | 1 | -1 | Hi Hat |

Se utilizaron 11 vectores (según la tabla: bombo 2, 3 y 4 redoblante 1, 2, 4 y 5 y Hi Hat 1, 2, 4 y 5) de 4 muestras cada uno para el proceso de entrenamiento. En el proceso de validación de resultados se utilizaron 15 patrones adquiridos directamente de los instrumentos.

El hecho de reconocer más de un instrumento al mismo tiempo se debe también al preprocesado, ya que este siempre regresará 3 grupos de valores en los que cada grupo hace referencia a cada instrumento posible de reconocer. Luego la red evalúa en tres ocasiones para determinar si en alguno de estos grupos existe un dato válido, de manera que, si en más de un grupo existe información, la red lo detecta y genera un mensaje de combinación de instrumentos.

4.1.2.2.1 Validación de Pruebas, Batería en conjunto

Tabla 4.3a Cuadro Resumen de Pruebas Retropropagación.

| Retropropagacion Baqueta punta de Plástico | | | |
|--------------------------------------------|-------------|---|---------------------|
| No. De Muestra | Instrumento | | Patrón Reconocido |
| 1 | Hi Hat | - | Redoblante y Hi Hat |
| 2 | Hi Hat | x | Hi Hat |
| 3 | Hi Hat | - | Redoblante y Hi Hat |
| 4 | Hi Hat | x | Hi Hat |
| 5 | Hi Hat | x | Hi Hat |
| 6 | Redoblante | x | Redoblante |
| 7 | Redoblante | x | Redoblante |
| 8 | Redoblante | x | Redoblante |
| 9 | Redoblante | - | Redoblante y Hi Hat |
| 10 | Redoblante | x | Redoblante |
| 11 | Bombo | x | Bombo |
| 12 | Bombo | x | Bombo |
| 13 | Bombo | x | Bombo |
| 14 | Bombo | x | Bombo |
| 15 | Bombo | x | Bombo |

Tabla 4.3b Cuadro Resumen de Pruebas Retropropagación

| Retropropagación Baqueta punta de Madera | | | |
|------------------------------------------|-------------|---|---------------------|
| No. De Muestra | Instrumento | | Patrón Reconocido |
| 1 | Hi Hat | x | Hi Hat |
| 2 | Hi Hat | - | Redoblante y Hi Hat |
| 3 | Hi Hat | - | Redoblante y Hi Hat |
| 4 | Hi Hat | x | Hi Hat |
| 5 | Hi Hat | - | Redoblante y Hi Hat |
| 6 | Redoblante | x | Redoblante |
| 7 | Redoblante | x | Redoblante |
| 8 | Redoblante | x | Redoblante |
| 9 | Redoblante | x | Redoblante |
| 10 | Redoblante | x | Redoblante |
| 11 | Bombo | x | Bombo |
| 12 | Bombo | x | Bombo |
| 13 | Bombo | x | Bombo |
| 14 | Bombo | x | Bombo |
| 15 | Bombo | x | Bombo |

x: patrón reconocido
-: patrón no reconocido

Se realizaron pruebas con diferentes tipos de baquetas y diferentes usuarios, generando los mismos resultados de reconocimiento pero en diferente orden. El porcentaje de reconocimiento fue de 80%^{*5} con la ejecución de los instrumentos en vivo.

⁵ Este valor de porcentaje de reconocimiento se obtuvo de la sumatoria de muestras satisfactorias entre el número total de muestras ejecutada, multiplicadas por 100. Es de hacer notar que este es un porcentaje global, corresponden a la batería en conjunto

4.1.2.2.2 Validación de Pruebas, Instrumentos individuales

Tabla 4.4a Validación de Bombo con Ruido Ambiental constante ^{*6}

| Retropropagación | | |
|-------------------------------------------------|---|-------------------|
| Ambiente: Ruido Ambiental Constante | | |
| Obtención de Porcentaje de Reconocimiento Bombo | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Bombo |
| 2 | x | Bombo |
| 3 | x | Bombo |
| 4 | x | Bombo |
| 5 | x | Bombo |
| 6 | x | Bombo |
| 7 | x | Bombo |
| 8 | x | Bombo |
| 9 | x | Bombo |
| 10 | x | Bombo |
| 11 | x | Bombo |
| 12 | x | Bombo |
| 13 | x | Bombo |
| 14 | x | Bombo |
| 15 | x | Bombo |
| 16 | x | Bombo |
| 17 | x | Bombo |
| 18 | x | Bombo |
| 19 | x | Bombo |
| 20 | x | Bombo |
| 21 | x | Bombo |
| 22 | x | Bombo |
| 23 | x | Bombo |
| 24 | x | Bombo |
| 25 | x | Bombo |
| 26 | x | Bombo |
| 27 | x | Bombo |
| 28 | x | Bombo |
| 29 | x | Bombo |
| 30 | x | Bombo |

Porcentaje de Reconocimiento

100.00%

⁶ El escenario fue dentro de un Cuarto de Equipos de Comunicaciones, un cuarto de Servidores, donde se puede encontrar con ruido de ventiladores encendidos todo el tiempo.

Tabla 4.4b Validación de Bombo con Ruido Ambiental Variable

| Retropropagación | | |
|-------------------------------------------------|---|-------------------|
| Ambiente: Cochera, Aire Libre | | |
| Obtención de Porcentaje de Reconocimiento Bombo | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Bombo |
| 2 | x | Bombo |
| 3 | x | Bombo |
| 4 | x | Bombo |
| 5 | x | Bombo |
| 6 | x | Bombo |
| 7 | x | Bombo |
| 8 | x | Bombo |
| 9 | x | Bombo |
| 10 | x | Bombo |
| 11 | x | Bombo |
| 12 | x | Bombo |
| 13 | x | Bombo |
| 14 | x | Bombo |
| 15 | x | Bombo |
| 16 | x | Bombo |
| 17 | x | Bombo |
| 18 | x | Bombo |
| 19 | x | Bombo |
| 20 | x | Bombo |
| 21 | x | Bombo |
| 22 | x | Bombo |
| 23 | x | Bombo |
| 24 | x | Bombo |
| 25 | x | Bombo |
| 26 | x | Bombo |
| 27 | x | Bombo |
| 28 | x | Bombo |
| 29 | x | Bombo |
| 30 | x | Bombo |

Porcentaje de Reconocimiento

100.00%

Tabla 4.5a Validación de Hi Hat con Ruido Ambiental constante ^{*7}

| Retropropagación | | |
|--------------------------------------------------|---|---------------------|
| Ambiente: Ruido Ambiental Constante | | |
| Obtención de Porcentaje de Reconocimiento Hi Hat | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Hi Hat |
| 2 | x | Hi Hat |
| 3 | x | Hi Hat |
| 4 | x | Hi Hat |
| 5 | x | Hi Hat |
| 6 | x | Hi Hat |
| 7 | x | Hi Hat |
| 8 | x | Hi Hat |
| 9 | x | Hi Hat |
| 10 | x | Hi Hat |
| 11 | x | Hi Hat |
| 12 | x | Hi Hat |
| 13 | x | Hi Hat |
| 14 | x | Hi Hat |
| 15 | x | Hi Hat |
| 16 | x | Hi Hat |
| 17 | x | Hi Hat |
| 18 | x | Hi Hat |
| 19 | - | Hi Hat y Redoblante |
| 20 | x | Hi Hat |
| 21 | x | Hi Hat |
| 22 | x | Hi Hat |
| 23 | x | Hi Hat |
| 24 | x | Hi Hat |
| 25 | - | Hi Hat y Redoblante |
| 26 | x | Hi Hat |
| 27 | x | Hi Hat |
| 28 | x | Hi Hat |
| 29 | x | Hi Hat |
| 30 | x | Hi Hat |

Porcentaje de Reconocimiento

93.33%

⁷ El escenario fue dentro de un Cuarto de Equipos de Comunicaciones, un cuarto de Servidores, donde se puede encontrar con ruido de ventiladores encendidos todo el tiempo.

Tabla 4.5b Validación de Hi Hat con Ruido Ambiental Variable

| Retropropagación | | |
|--------------------------------------------------|---|---------------------|
| Ambiente: Cochera, Aire Libre | | |
| Obtención de Porcentaje de Reconocimiento Hi Hat | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Hi Hat |
| 2 | - | Hi Hat y Redoblante |
| 3 | x | Hi Hat |
| 4 | x | Hi Hat |
| 5 | x | Hi Hat |
| 6 | x | Hi Hat |
| 7 | x | Hi Hat |
| 8 | x | Hi Hat |
| 9 | x | Hi Hat |
| 10 | x | Hi Hat |
| 11 | x | Hi Hat |
| 12 | - | Hi Hat y Redoblante |
| 13 | x | Hi Hat |
| 14 | x | Hi Hat |
| 15 | x | Hi Hat |
| 16 | x | Hi Hat |
| 17 | x | Hi Hat |
| 18 | x | Hi Hat |
| 19 | x | Hi Hat |
| 20 | x | Hi Hat |
| 21 | x | Hi Hat |
| 22 | x | Hi Hat |
| 23 | x | Hi Hat |
| 24 | x | Hi Hat |
| 25 | x | Hi Hat |
| 26 | x | Hi Hat |
| 27 | - | Redoblante |
| 28 | x | Hi Hat |
| 29 | x | Hi Hat |
| 30 | x | Hi Hat |

Porcentaje de Reconocimiento

90.00%

Tabla 4.6a Validación de Redoblante con Ruido Ambiental constante ^{*8}

| Retropropagación | | |
|------------------------------------------------------|---|---------------------|
| Ambiente: Ruido Ambiental Constante | | |
| Obtención de Porcentaje de Reconocimiento Redoblante | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Redoblante |
| 2 | - | Bombo |
| 3 | x | Redoblante |
| 4 | x | Redoblante |
| 5 | x | Redoblante |
| 6 | x | Redoblante |
| 7 | x | Redoblante |
| 8 | x | Redoblante |
| 9 | - | Bombo |
| 10 | x | Redoblante |
| 11 | x | Redoblante |
| 12 | x | Redoblante |
| 13 | x | Redoblante |
| 14 | x | Redoblante |
| 15 | x | Redoblante |
| 16 | x | Redoblante |
| 17 | x | Redoblante |
| 18 | x | Redoblante |
| 19 | x | Redoblante |
| 20 | - | Redoblante y Hi Hat |
| 21 | x | Redoblante |
| 22 | x | Redoblante |
| 23 | x | Redoblante |
| 24 | x | Redoblante |
| 25 | x | Redoblante |
| 26 | x | Redoblante |
| 27 | - | Hi Hat |
| 28 | x | Redoblante |
| 29 | x | Redoblante |
| 30 | x | Redoblante |

Porcentaje de Reconocimiento

86.67%

⁸ El escenario fue dentro de un Cuarto de Equipos de Comunicaciones, un cuarto de Servidores, donde se puede encontrar con ruido de ventiladores encendidos todo el tiempo.

Tabla 4.6b Validación de Redoblante con Ruido Ambiental Variable

| Retropropagación | | |
|------------------------------------------------------|---|---------------------|
| Ambiente: Cochera, Aire Libre | | |
| Obtención de Porcentaje de Reconocimiento Redoblante | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | - | Redoblante y Hi Hat |
| 2 | x | Redoblante |
| 3 | x | Redoblante |
| 4 | x | Redoblante |
| 5 | x | Redoblante |
| 6 | - | Redoblante y Hi Hat |
| 7 | x | Redoblante |
| 8 | x | Redoblante |
| 9 | x | Redoblante |
| 10 | x | Redoblante |
| 11 | x | Redoblante |
| 12 | x | Redoblante |
| 13 | x | Redoblante |
| 14 | x | Redoblante |
| 15 | x | Redoblante |
| 16 | - | Hi Hat |
| 17 | x | Redoblante |
| 18 | - | Hi Hat |
| 19 | x | Redoblante |
| 20 | x | Redoblante |
| 21 | - | Redoblante y Hi Hat |
| 22 | x | Redoblante |
| 23 | x | Redoblante |
| 24 | x | Redoblante |
| 25 | x | Redoblante |
| 26 | x | Redoblante |
| 27 | x | Redoblante |
| 28 | x | Redoblante |
| 29 | x | Redoblante |
| 30 | x | Redoblante |

Porcentaje de Reconocimiento

83.33%

x: patrón reconocido
-: patrón no reconocido

4.1.3 Redes de Teoría de Resonancia Adaptativa 2 (ART2)

La red ART 2 es una red heteroasociativa de aprendizaje no supervisado que acepta valores continuos en su entrada. A fin de conseguir clasificar patrones de valor continuo debe realizar una serie de cálculos que vuelven más complejo el algoritmo de entrenamiento. Para comenzar, la capa de entrada es una combinación de funciones de normalización y supresión de ruido.

La arquitectura típica de la red ART 2 se muestra en la figura 4.3. La capa $F1$ está formada por seis tipos de unidades (W , X , U , V , P y Q) de modo que por cada una de las entradas de la red hay seis unidades involucradas en la generación de las señales enviadas a la siguiente capa. Un elemento adicional entre W y X recibe señales de todas las unidades W y calcula la norma o magnitud del vector \mathbf{w} y envía ese resultado a todas la unidades de la subcapa X . Mecanismos similares de normalización se encuentran entre las unidades P a Q y V a U . Las unidades P actúan como interfaces hacia la capa $F2$ de esta red, y cada una de las unidades P_i está conectada a todas las neuronas de salida por medio de dos conexiones de pesos, una ascendente (bottomà up) y una descendente (topà down) y es en estas conexiones dobles que se almacenan los prototipos de cada uno de los grupos formados durante el funcionamiento de la red. La acción de la capa $F2$ es de competencia del tipo *ganador-toma-todo* es decir que solamente una de las neuronas de salida (la ganadora) tendrá su salida activa a la vez.

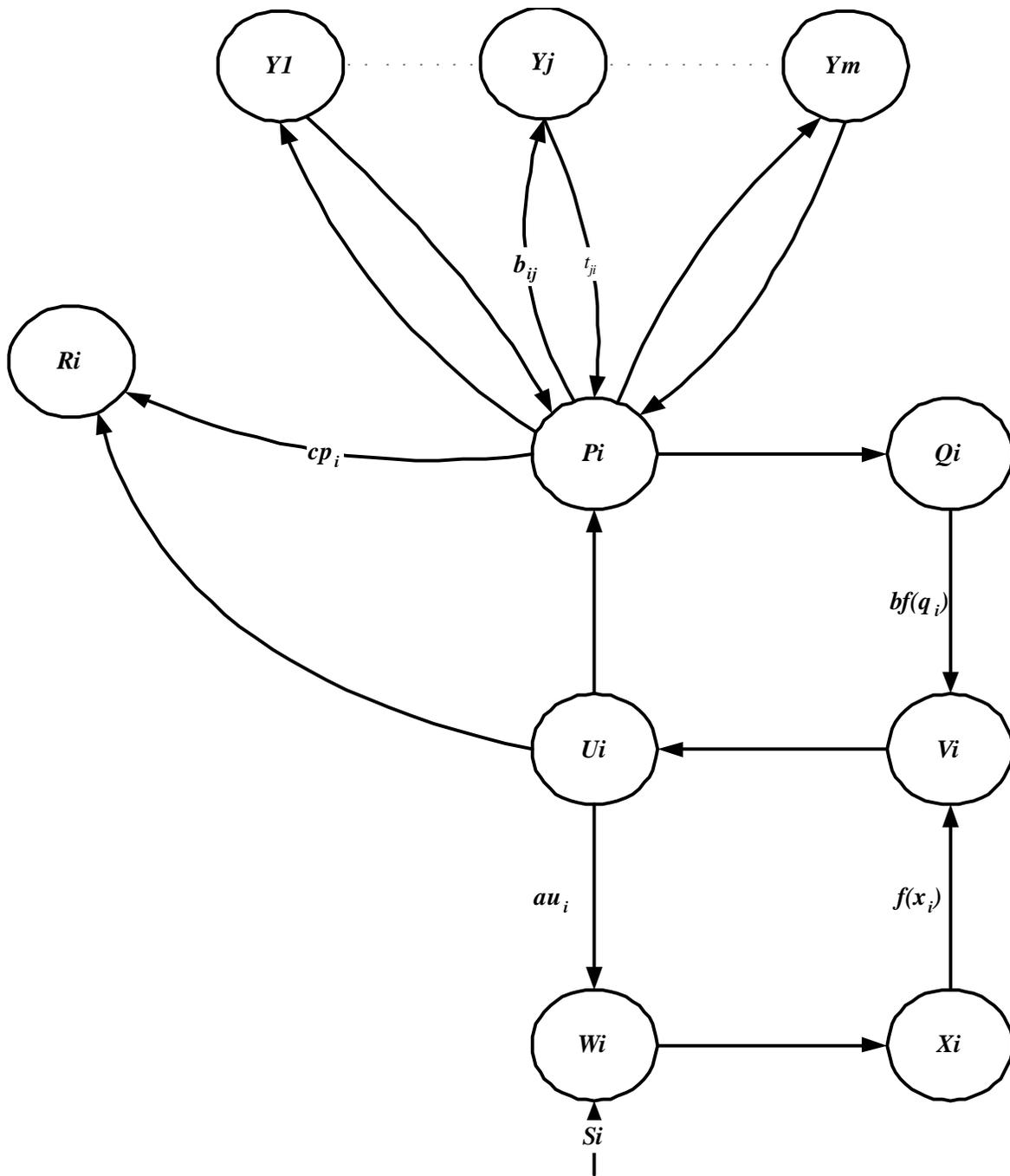


Figura 4.3 Arquitectura Red ART2

4.1.3.1 Algoritmo de Aprendizaje

El ciclo de aprendizaje para un patrón de entrada inicia con el cálculo de la función de activación de la subcapa U (que no es más que la salida normalizada de la subcapa V). Cada elemento U_i envía una señal a sus respectivos elementos W_i y P_i , entonces se calcula el nivel de activación de estas unidades. Los elementos W_i suman la señal recibida de U_i y S_i (la señal de entrada). P_i suma las señales provenientes de U_i y la señal descendente generada, en caso de hallarse activa alguna de las neuronas de la capa $F2$. La activación de X_i y Q_i son, respectivamente, las versiones normalizadas de W_i y P_i . A estas unidades se les aplica una función de activación y el resultado es enviado a las unidades V_i , donde es realizada la adición de las señales provenientes de la subcapa X y Q , con lo que se completa el ciclo de actualización de la capa $F1$.

La función de activación utilizada en todas las pruebas realizadas fue la siguiente:

$$f(x) = \begin{cases} x; & \Leftrightarrow x \geq \theta \\ 0; & \Leftrightarrow x < \theta \end{cases} \quad \text{Ecuación 4.16}$$

La razón que justifica el uso de esta función es que, de acuerdo a los creadores de esta red (Carpenter & Grossbreg 1987), las unidades de las capas U y P alcanzan el equilibrio después de tan solo 2 ciclos de actualización de la capa $F1$, y además es recomendada por los autores que han realizado publicaciones sobre este tema (cfg Fausett 1994, Hilera)

Una vez que la capa $F1$ ha alcanzado el equilibrio, las unidades P envían sus señales a la capa $F2$, donde tiene lugar la competencia para escoger al candidato a aprender el patrón de entrada. También las unidades de las subcapas U y P envían señales hacia las unidades del bloque de control de reset y es en base a la respuesta de ese bloque que la neurona de salida de un grupo o cluster puede ser aceptada o rechazada como no válida por presentar un grado de semejanza pequeño entre los patrones. Si ese es el caso, la unidad rechazada es inhibida, (es decir, su salida es fijada en -1 para excluirla de futuras competencias) y la unidad que tenga la siguiente salida más grande será propuesta como candidata a aprender el patrón de entrada. Ese proceso se repite hasta que el patrón es admitido en un clúster o cuando ya se ha agotado la cantidad máxima de clusters

permitidos. El aprendizaje ocurre únicamente si un patrón es aceptado como parte de un clúster de la red.

Descrito en forma más clara el proceso de aprendizaje de una red ART 2 es como sigue:

Paso 0. Inicializar parámetros de la red:

n: número de entradas de la red (capa $F1$)

m: número máximo de clusters (capa $F2$)

a, b: pesos fijos de la capa $F1$. Inicializarlos a cero produce inestabilidad

c: Peso fijo usado para verificar Reset. Un valor pequeño de c proporciona un rango mayor de operación al parámetro de vigilancia ρ .

d: activación de unidad ganadora de la capa $F2$. Los valores escogidos de c y d deben satisfacer la siguiente inecuación

$$\frac{c d}{1 - d} \leq 1 \quad \text{Ecuación 4.17}$$

e: un valor muy pequeño que se usa para prevenir la división entre cero

θ : parámetro de supresión de ruido utilizado en la función de activación de las neuronas. Un valor típico es $1/\sqrt{n}$

ρ : parámetro de vigilancia. Junto con los valores de los pesos ascendentes, b_{ij} , determina la cantidad de clusters que se forman durante el aprendizaje. Valores por debajo de 0.7 no difieren de los resultados obtenidos al fijarlo a valores tan bajos como $\rho = 0$.

t_{ji} : pesos descendentes ($F2 \rightarrow F1$), deben ser inicializados a 0 para asegurarse que no ocurrirá un Reset para el primer patrón admitido en un clúster.

b_{ij} : pesos ascendentes ($F1 \rightarrow F2$), deben ser escogidos cumpliendo la siguiente desigualdad: $b_{ij} \leq 1 / ((1 - d) \sqrt{n})$ Ecuación 4.18

Paso 1. Realizar pasos del 2 al 12 el número especificado de veces de acuerdo a la cantidad de patrones de ejemplo disponibles.

Paso 2. Actualizar los valores de las unidades de la capa $F1$ con $u_i = q_i = p_i = 0$

$$x_i = s_i / (e + \| \mathbf{s} \|)$$

$$w_i = s_i$$

$$v_i = f_{(x)}$$

Paso 3. Actualizar nuevamente los valores en la capa $F1$ con los resultados del paso anterior

$$\begin{aligned} u_i &= v_i / (e + \|\mathbf{v}\|) \\ w_i &= s_i + a^*u_i & p_i &= u_i + d^*t_{ji} \\ x_i &= w_i / (e + \|\mathbf{w}\|) & q_i &= p_i / (e + \|\mathbf{p}\|) \\ v_i &= f_{(x)} + b^*f_{(q)} \end{aligned}$$

Paso 4. Calcular las señales de la capa $F2$:

$$y_j = \sum_i b_{ij} p_i$$

Ecuación 4.19

Paso 5. Mientras *ResetFlag* sea igual a “verdadero” repetir los pasos 6 à 7

Paso 6. Encontrar la unidad ganadora, Y_J , de la capa $F2$; donde el valor de J es tal que $y_J \geq y_j$ para $j = 1, 2, \dots, m$)

Paso 7. Verificar el Reset

$$\begin{aligned} u_i &= v_i / (e + \|\mathbf{v}\|) \\ p_i &= u_i + d^*t_{ji} \\ r_i &= (u_i + c^*p_i) / (e + \|\mathbf{u}\| + c^*\|\mathbf{p}\|) \\ \text{Si } \|\mathbf{r}\| &< \rho - e, \text{ entonces} \end{aligned}$$

Resetflag = “verdadero”

$Y_J = -1$. Repetir Paso 6.

Si $\|\mathbf{r}\| \geq \rho - e$, entonces

$$\begin{aligned} w_i &= s_i + a^*u_i \\ x_i &= w_i / (e + \|\mathbf{w}\|) \\ q_i &= p_i / (e + \|\mathbf{p}\|) \\ v_i &= f_{(x)} + b^*f_{(q)} \end{aligned}$$

Resetflag = “falso”; continuar con el paso 8

Paso 8. Actualizar los pesos de la unidad ganadora J

$$\begin{aligned} t_{ji} &= u_i / (1 - d) \\ b_{iJ} &= u_i / (1 - d) \end{aligned}$$

Paso 9. Verificar si se ha alcanzado el número de iteraciones previstas [26]

4.1.3.2 Pruebas con Red ART2

Para el caso particular de ART2 se realiza un paso menos en el preprocesado y éste regresa un rango de frecuencias, 30 muestras, que son las que finalmente ingresan a la red. Se creó una red con 4 posibles categorías, detalladas a continuación:

| Categoría | Reconoce |
|-----------|------------|
| 1 | Bombo |
| 2 | Redoblante |
| 3 | Hi Hat |
| 4 | Silencio |

4.1.3.2.1 Validación de Pruebas, Batería en conjunto

Tabla 4.7a Cuadro Resumen de Pruebas ART2

| ART2 Baqueta punta de Madera | | | |
|------------------------------|-------------|---|---------------------------|
| No. De Muestra | Instrumento | | Patrón Reconocido |
| 1 | Hi Hat | x | Hi Hat |
| 2 | Hi Hat | x | Hi Hat |
| 3 | Hi Hat | x | Hi Hat |
| 4 | Hi Hat | x | Hi Hat |
| 5 | Hi Hat | x | Hi Hat |
| 6 | Redoblante | x | Redoblante |
| 7 | Redoblante | - | Bombo Redoblante y Hi Hat |
| 8 | Redoblante | - | Redoblante y Hi Hat |
| 9 | Redoblante | x | Redoblante |
| 10 | Redoblante | x | Redoblante |
| 11 | Bombo | - | Bombo y Redoblante |
| 12 | Bombo | - | Bombo y Redoblante |
| 13 | Bombo | x | Bombo |
| 14 | Bombo | x | Bombo |
| 15 | Bombo | x | Bombo |

Tabla 4.7b Cuadro Resumen de Pruebas ART2

| ART2 Baqueta punta de Plástico | | | |
|--------------------------------|-------------|---|---------------------|
| No. De Muestra | Instrumento | | Patrón Reconocido |
| 1 | Hi Hat | - | Redoblante y Hi Hat |
| 2 | Hi Hat | x | Hi Hat |
| 3 | Hi Hat | x | Hi Hat |
| 4 | Hi Hat | x | Hi Hat |
| 5 | Hi Hat | - | Silencio |
| 6 | Redoblante | x | Redoblante |
| 7 | Redoblante | x | Redoblante |
| 8 | Redoblante | x | Redoblante |
| 9 | Redoblante | x | Redoblante |
| 10 | Redoblante | x | Redoblante |
| 11 | Bombo | - | Bombo y Redoblante |
| 12 | Bombo | x | Bombo |
| 13 | Bombo | - | Silencio |
| 14 | Bombo | - | Bombo y Redoblante |
| 15 | Bombo | x | Bombo |

x: patrón reconocido
-: patrón no reconocido

Se realizaron pruebas con diferentes tipos de baquetas y diferentes usuarios, generando resultados similares de reconocimiento pero en diferente orden. El porcentaje de reconocimiento promedio fue de 70%^{*9} con la ejecución de los instrumentos en vivo.

⁹ Este valor de porcentaje de reconocimiento se obtuvo de la sumatoria de muestras satisfactorias entre el número total de muestras ejecutada, multiplicadas por 100. Es de hacer notar que este es un porcentaje global, corresponden a la batería en conjunto

4.1.3.2.2 Validación de Pruebas, Instrumentos Individuales

Tabla 4.8a Validación de Bombo con Ruido Ambiental constante ^{*10}

| ART2 | | |
|-------------------------------------------------|---|--------------------|
| Ambiente: Ruido Ambiental Constante | | |
| Obtención de Porcentaje de Reconocimiento Bombo | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Bombo |
| 2 | x | Bombo |
| 3 | x | Bombo |
| 4 | x | Bombo |
| 5 | x | Bombo |
| 6 | - | Bombo y Redoblante |
| 7 | - | Bombo y Redoblante |
| 8 | x | Bombo |
| 9 | x | Bombo |
| 10 | x | Bombo |
| 11 | x | Bombo |
| 12 | x | Bombo |
| 13 | x | Bombo |
| 14 | x | Bombo |
| 15 | x | Bombo |
| 16 | x | Bombo |
| 17 | x | Bombo y Redoblante |
| 18 | x | Bombo |
| 19 | x | Bombo |
| 20 | x | Bombo |
| 21 | x | Bombo |
| 22 | x | Bombo |
| 23 | x | Bombo |
| 24 | x | Bombo |
| 25 | x | Bombo |
| 26 | x | Bombo |
| 27 | x | Bombo |
| 28 | x | Bombo |
| 29 | x | Bombo |
| 30 | x | Bombo |

Porcentaje de Reconocimiento

93.33%

¹⁰ El escenario fue dentro de un Cuarto de Equipos de Comunicaciones, un cuarto de Servidores, donde se puede encontrar con ruido de ventiladores encendidos todo el tiempo.

Tabla 4.8b Validación de Bombo con Ruido Ambiental Variable

| ART2 | | |
|-------------------------------------------------|---|--------------------|
| Ambiente: Cochera, Aire Libre | | |
| Obtención de Porcentaje de Reconocimiento Bombo | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Bombo |
| 2 | - | Bombo y Redoblante |
| 3 | x | Bombo |
| 4 | x | Bombo |
| 5 | x | Bombo |
| 6 | x | Bombo |
| 7 | x | Bombo |
| 8 | x | Bombo |
| 9 | - | Bombo |
| 10 | x | Bombo y Redoblante |
| 11 | x | Bombo |
| 12 | x | Bombo |
| 13 | x | Bombo |
| 14 | x | Bombo |
| 15 | x | Bombo |
| 16 | x | Bombo |
| 17 | - | Bombo y Redoblante |
| 18 | x | Bombo |
| 19 | x | Bombo |
| 20 | x | Bombo |
| 21 | x | Bombo |
| 22 | x | Bombo |
| 23 | - | Bombo |
| 24 | x | Bombo y Redoblante |
| 25 | x | Bombo |
| 26 | x | Bombo |
| 27 | x | Bombo |
| 28 | x | Bombo |
| 29 | x | Bombo |
| 30 | x | Bombo |

Porcentaje de Reconocimiento

86.67%

Tabla 4.9a Validación de Hi Hat con Ruido Ambiental constante ^{*11}

| ART2 | | |
|--------------------------------------------------|---|---------------------|
| Ambiente: Ruido Ambiental Constante | | |
| Obtención de Porcentaje de Reconocimiento Hi Hat | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Hi Hat |
| 2 | - | Hi Hat y Redoblante |
| 3 | x | Hi Hat |
| 4 | x | Hi Hat |
| 5 | x | Hi Hat |
| 6 | - | Hi Hat y Redoblante |
| 7 | x | Hi Hat |
| 8 | x | Hi Hat |
| 9 | - | Redoblante |
| 10 | x | Hi Hat |
| 11 | - | Hi Hat y Redoblante |
| 12 | x | Hi Hat |
| 13 | x | Hi Hat |
| 14 | x | Hi Hat |
| 15 | x | Hi Hat |
| 16 | x | Hi Hat |
| 17 | x | Hi Hat |
| 18 | x | Hi Hat |
| 19 | x | Hi Hat |
| 20 | x | Hi Hat |
| 21 | x | Hi Hat |
| 22 | - | Redoblante |
| 23 | x | Hi Hat |
| 24 | - | Hi Hat y Redoblante |
| 25 | x | Hi Hat |
| 26 | x | Hi Hat |
| 27 | x | Hi Hat |
| 28 | - | Hi Hat y Redoblante |
| 29 | x | Hi Hat |
| 30 | x | Hi Hat |

Porcentaje de Reconocimiento

76.67%

¹¹ El escenario fue dentro de un Cuarto de Equipos de Comunicaciones, un cuarto de Servidores, donde se puede encontrar con ruido de ventiladores encendidos todo el tiempo.

Tabla 4.9b Validación de Hi Hat con Ruido Ambiental Variable

| ART2 | | |
|--------------------------------------------------|---|---------------------|
| Ambiente: Cochera, Aire Libre | | |
| Obtención de Porcentaje de Reconocimiento Hi Hat | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | - | Redoblante |
| 2 | x | Hi Hat |
| 3 | x | Hi Hat |
| 4 | x | Hi Hat |
| 5 | - | Redoblante |
| 6 | - | Hi Hat y Redoblante |
| 7 | x | Hi Hat |
| 8 | - | Redoblante |
| 9 | - | Hi Hat |
| 10 | x | Hi Hat |
| 11 | x | Hi Hat |
| 12 | x | Hi Hat |
| 13 | x | Hi Hat |
| 14 | x | Hi Hat |
| 15 | x | Hi Hat |
| 16 | x | Hi Hat |
| 17 | - | Hi Hat |
| 18 | x | Hi Hat |
| 19 | x | Hi Hat |
| 20 | x | Hi Hat |
| 21 | x | Hi Hat |
| 22 | x | Hi Hat |
| 23 | - | Hi Hat |
| 24 | x | Hi Hat |
| 25 | x | Hi Hat |
| 26 | x | Hi Hat |
| 27 | x | Hi Hat |
| 28 | x | Hi Hat |
| 29 | x | Hi Hat |
| 30 | x | Hi Hat |

Porcentaje de Reconocimiento

76.67%

Tabla 4.10a Validación de Redoblante con Ruido Ambiental constante ^{*12}

| ART2 | | |
|------------------------------------------------------|---|---------------------|
| Ambiente: Ruido Ambiental Constante | | |
| Obtención de Porcentaje de Reconocimiento Redoblante | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Redoblante |
| 2 | x | Redoblante |
| 3 | x | Redoblante |
| 4 | x | Redoblante |
| 5 | - | Redoblante y Hi Hat |
| 6 | x | Redoblante |
| 7 | - | Redoblante y Hi Hat |
| 8 | x | Redoblante |
| 9 | x | Redoblante |
| 10 | x | Redoblante |
| 11 | x | Redoblante |
| 12 | - | Redoblante y Hi Hat |
| 13 | - | Redoblante y Hi Hat |
| 14 | x | Redoblante |
| 15 | x | Redoblante |
| 16 | x | Redoblante |
| 17 | - | Hi Hat |
| 18 | x | Redoblante |
| 19 | x | Redoblante |
| 20 | x | Redoblante |
| 21 | - | Hi Hat |
| 22 | - | Hi Hat |
| 23 | x | Redoblante |
| 24 | x | Redoblante |
| 25 | x | Redoblante |
| 26 | x | Redoblante |
| 27 | - | Hi Hat |
| 28 | x | Redoblante |
| 29 | x | Redoblante |
| 30 | x | Redoblante |

Porcentaje de Reconocimiento

73.33%

¹² El escenario fue dentro de un Cuarto de Equipos de Comunicaciones, un cuarto de Servidores, donde se puede encontrar con ruido de ventiladores encendidos todo el tiempo.

Tabla 4.10b Validación de Redoblante con Ruido Ambiental Variable

| ART2 | | |
|------------------------------------------------------|---|---------------------|
| Ambiente: Cochera, Aire Libre | | |
| Obtención de Porcentaje de Reconocimiento Redoblante | | |
| No. De Muestra | | Patrón Reconocido |
| 1 | x | Redoblante |
| 2 | x | Redoblante |
| 3 | - | Redoblante y Hi Hat |
| 4 | x | Redoblante |
| 5 | x | Redoblante |
| 6 | x | Redoblante |
| 7 | x | Redoblante |
| 8 | - | Redoblante y Hi Hat |
| 9 | x | Redoblante |
| 10 | - | Redoblante y Hi Hat |
| 11 | x | Redoblante |
| 12 | - | Redoblante y Hi Hat |
| 13 | - | Redoblante y Hi Hat |
| 14 | - | Redoblante y Hi Hat |
| 15 | x | Redoblante |
| 16 | x | Redoblante |
| 17 | x | Redoblante |
| 18 | - | Redoblante y Hi Hat |
| 19 | x | Redoblante |
| 20 | x | Redoblante |
| 21 | x | Redoblante |
| 22 | - | Redoblante y Hi Hat |
| 23 | x | Redoblante |
| 24 | x | Redoblante |
| 25 | x | Redoblante |
| 26 | x | Redoblante |
| 27 | x | Redoblante |
| 28 | - | Redoblante y Hi Hat |
| 29 | x | Redoblante |
| 30 | - | Redoblante y Hi Hat |

Porcentaje de Reconocimiento

66.67%

x: patrón reconocido
-: patrón no reconocido

4.2 Selección de Red

En base a las pruebas realizadas con las tres redes seleccionadas para experimentar, se observa que, la que mejores resultados da es la red de **Retropropagación**. Las tablas 4.3a y 4.3b muestran una validación de pruebas globales que proporciona un porcentaje general de reconocimiento del 80%. De la misma forma se realizaron pruebas de validación con los instrumentos individuales para esta Red en dos ambientes diferentes, uno con ruido ambiental constante y el otro con ruido ambiental variable, similar a la mayoría de ambientes conocidos donde se podría ejecutar el sistema. Los porcentajes de reconocimiento para Bombo se muestran en las tablas 4.4a (100%) y 4.4b (100%), Hi Hat se muestran en las tablas 4.5a (93.33%) y 4.5b (90%) y finalmente los resultados de Redoblante se muestran en las tablas 4.6a (86.67%) y 4.6b (83.33%).

Kohonen presenta el problema de generar nuevas categorías sin que estas puedan ser limitadas por el usuario por lo que para la aplicación que se implementó no puede ser utilizada debido a la variabilidad de las frecuencias entrantes y la confusión que se genera. Los resultados de las pruebas con esta Red se muestran en las tablas 4.2a, 4.2b y 4.2c.

ART2 aun siendo una de las redes más poderosas, ya que da la posibilidad de limitar las categorías que se desean crear, presenta el problema de la complejidad en su algoritmo, aun con la red ya entrenada el algoritmo de reconocimiento genera un retraso importante para la finalidad de la aplicación, la cual produce reconocimientos muchas veces erróneos. Las tablas 4.7a y 4.7b muestran una validación de pruebas globales que proporciona un porcentaje general de reconocimiento del 70%. De la misma forma se realizaron pruebas de validación con los instrumentos individuales para esta red en dos ambientes diferentes: uno con ruido ambiental constante y el otro con ruido ambiental variable, similar a la mayoría de ambientes conocidos donde se podría ejecutar el sistema. Los porcentajes de reconocimiento para Bombo se muestran en las tablas 4.8a (93.33%) y 4.8b (86.67%), para Hi Hat se muestran en las tablas 4.9a (76.67%) y 4.9b (76.67%) y finalmente los resultados de Redoblante se muestran en las tablas 4.10a (73.33%) y 4.10b (66.67%).

Por lo anterior la red a utilizar es la de retropropagación.

4.3 Uso de las Redes Neuronales

Para el caso específico de la Red de Retropropagación surgió un problema de reconocimiento cuando se ingresaron los 3 grupos de 30 datos, que proporcionaba el preprocesado directamente. Para tratar de solventar esto, se realizó una normalización de los datos para que se ubicaran en un rango entre -1 y 1, pero los problemas de reconocimiento continuaron.

Finalmente en la parte del procesado se incluyó un post procesado el cual agrupa los rangos adquiridos y los asocia con un vector específico, a continuación se muestran los rangos junto con el respectivo vector asignado.

| Instrumento | Rangos de Frecuencia (Hz) | Vector Asociado |
|-------------|---------------------------|-----------------|
| Bombo | 105 a 190 | 1 -1 -1 1 |
| Redoblante | 1600 a 2600 | -1 -1 1 1 |
| Hi Hat | 6900 a 9400 | -1 1 1 -1 |
| Silencio | Todos las demás | -1 -1 -1 -1 |

Tabla 4.11 Asociación entre Rangos de Frecuencia y Vectores Salida (Post Procesado)

Los vectores finales de este post procesado son los que se introducen como entradas a la red. Como se puede observar la Red para este caso puede ser omitida ya que prácticamente el reconocimiento se realiza en el preprocesado, sin embargo, se decidió continuar con ella debido a que inicialmente se definió como el motor de reconocimiento del proyecto. Por otro lado, en el caso de la red ART2 no es necesario realizar la adaptación que se realizó en Retropropagación, por lo que aquí esta red neuronal *S/* es el motor de reconocimiento, ya que recibe directamente los 3 grupos de 30 datos producto del algoritmo de preprocesado.

Capítulo 5 – Manual del Usuario de CajaCustica

5.1 Presentación General



Figura 5.1 Menú Principal de CajaCustica

En la figura 5.1 se presenta la pantalla principal de CajaCustica con las siguientes opciones representadas en botones:



Este botón lleva a la ayuda del software. Ofrece una explicación sobre la navegación a través del mismo



Este botón proporciona la opción de seleccionar con qué método se desean realizar las pruebas

Este despliega las siguientes opciones:

BASADOS EN RESULTADOS DE ART 2

Este botón conducirá a un ambiente diferente, en el cual ya se puede hacer el reconocimiento. El sistema realizará la adquisición y el reconocimiento en base a los pesos que se obtuvieron cuando se realizó el entrenamiento de las muestras con la Red Neuronal ART2

BASADOS EN RESULTADOS BACK PROPAGATION

Este botón conducirá a un ambiente diferente, en el cual ya se puede hacer el reconocimiento. El sistema realizará la adquisición y el reconocimiento en base a los pesos que se obtuvieron cuando se realizó el entrenamiento de las muestras con la Red Neuronal Retropropagación

MODO DE EJECUCION

Este botón conducirá a realizar la adquisición y reconocimiento en base a la red que resultó ser la que mejor porcentaje de reconocimiento nos brinda.

Calibración

Este botón conducirá a un módulo de calibración para la personalización del software, es decir, ajustará la forma en que cada usuario toca cada instrumento para lograr un mejor porcentaje de reconocimiento.

Salir

Con este botón se sale del sistema definitivamente

5.2 Pantallas Secundarias

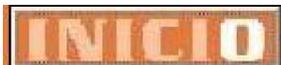
Se cuenta con un solo formato de pantallas secundarias aun cuando estas representan diferentes métodos de reconocimiento.

A continuación se muestra el formato de las pantallas secundarias:



Figura 5.2 Formato de Pantallas de reconocimiento de CajaCustica

En la figura 5.2 se muestra el formato general de las pantallas secundarias, en las cuales se realiza el reconocimiento por los distintos métodos seleccionados. A continuación se detallan las características de esta pantalla:



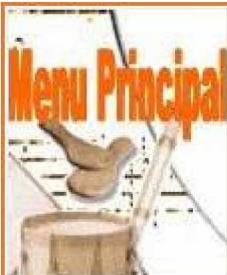
Botón que permite dar inicio al reconocimiento, los resultados se mostrarán en la pequeña pantalla que se encuentra debajo de la leyenda “Instrumento Reconocido”

Detener

Botón que permite detener la adquisición. Esta es la mejor manera para realizar una pausa en la adquisición ya que cierra todo el proceso de adquisición y no cierra la pantalla actual.

Valores Calibrados

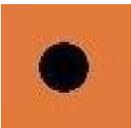
Botón que se activa solo si en el menú principal se accede a la rutina de calibración (botón calibración). Permite cargar valores personalizados para el programa



Botón que permite regresar el Menú Principal para elegir nuevamente alguna opción

SALIR

Botón que permite salir del programa completamente



Metronomo, su función es la de mostrar el estado del sistema, cuando es verde indica que el sistema está adquiriendo y cuando está en negro indica que el sistema está procesando o está apagado.

5.3 Calibrador

Este subsistema permite adaptarse a las condiciones específicas de cada usuario, tales como fuerza del golpe, características del microfono, condiciones del entorno (niveles de eco y tamaño de la habitacion).

Para cumplir dicha funcion se registran 10 golpes de cada instrumento para luego obtener un promedio de la potencia maxima y su correspondiente frecuencia. El orden de adquisicion de los instrumentos es Bombo, Redoblante y Hi Hat.

A continuación se muestra la pantalla de dicho subsistema



Figura 5.3 Formato de Modulo de Calibración de CajaCustica

5.4 Limitantes

Cabe mencionar que el programa tiene ciertas limitantes en cuanto a su funcionamiento en tiempo real. El programa para adquirir las señales hace uso de archivos WAVE los cuales son guardados cíclicamente por el programa como parte de su proceso de adquisición de datos, esto implica un espacio de tiempo durante el cual no hay ni adquisición ni procesamiento. A este hecho se suma otro factor importante: después de la adquisición, el sistema se ocupa del procesamiento, esto se traduce en pérdida de información ya que durante el tiempo de procesamiento cualquier información relevante no podrá ser tomada en cuenta por el sistema, debido a que durante este tiempo el sistema no es capaz de

adquirir datos nuevos. Se ha calculado que este tiempo “muerto” es de aproximadamente un segundo.

Este tiempo “muerto” influye grandemente en el reconocimiento del sistema debido a que si se ejecuta cualquier instrumento durante este lapso de tiempo, el sistema no será capaz de captar la información y mucho menos procesarla.

Conclusiones

- Las redes neuronales son solo una parte de cualquier sistema de reconocimiento. Para ejecutar completamente esta tarea es necesario que previamente se hayan extraído aquellas características propias de lo que se desee identificar.
- La idea de comprimir la señal usando Wavelet no es una forma útil de procesar la señal para que pueda ser reconocida por una red neuronal. Un procesamiento consiste en extraer características propias de la señal, en lugar de trabajar para que la red asimile la señal entera o una fracción de la misma, en el caso de una compresión.
- La frecuencia de muestreo debe ser tal que permita adquirir los sonidos de más alta frecuencia que se deseen reconocer.
- Identificar un instrumento de percusión por medio de su magnitud y frecuencia no es lo más indicado ya que, en el caso del hi hat, que es un instrumento que posee componentes de frecuencia en el rango de trabajo del redoblante, puede dar lugar a reconocimientos erróneos.
- Para el tipo de sistema que se ha construido, influirán grandemente la posición del micrófono respecto de los instrumentos y el espacio en el que se encuentren ambos, ya que el reconocimiento se basa en la intensidad de los sonidos captados por el sistema.
- Los coeficientes de predicción lineal poseen el inconveniente de que no brindan claridad en cuanto a si el sonido que se ha analizado es de un instrumento determinado. En el caso particular del filtro pasa banda con frecuencias de corte para aislar el bombo, la señal resultante es la misma cuando se hace pasar por él ya sea una señal de bombo, de hi hat o de redoblante.

- Los momentos estadísticos son descriptores que son capaces de interpretar la forma de una señal. Estos presentan el inconveniente de no ser manejables fácilmente por la red ya que sus valores son demasiado grandes. Adicional a esto se intentó obtener su logaritmo natural pero no resultó ya que todos los valores para todos los instrumentos representaban una línea recta.
- Las Redes Neuronales, aun siendo una de las mejores técnicas de reconocimiento que existen, deben ser alimentadas por datos coherentes para su entrenamiento, es decir la creencia que pueden reconocer cualquier tipo de patrón con solo proporcionar una vaga referencia es falsa. Para su buen funcionamiento es necesario alimentarlas con datos significativos y depurados de cada señal, he aquí la importancia de un buen pre-procesado.
- En la Red de Retropropagación, la arquitectura ideal se adquiere en base a la aplicación realizada y a la práctica, por lo que es imposible definir una sola arquitectura que sea estándar para un buen reconocimiento.
- Existen dos tipos de parámetros de paro para la Red de Retropropagación, por LMS (error medio cuadrático) y por el número de épocas realizadas. A pesar de su vital importancia para un buen entrenamiento, en la mayoría de bibliografía se contradicen mucho respecto a la forma de calcular el LMS a tal punto que algunos libros lo mencionan pero no dejan clara su forma de cálculo. La mejor forma de detener el entrenamiento es en base a pruebas modificando el número de épocas.
- La Red ART2 proporciona la característica de agrupar patrones en categorías según el número de categorías definidas por el entrenador, esto sumado a la característica de verificación redundante que esta posee la hace la red más fuerte en el área de datos análogos.

- La Red de Kohonen presenta el inconveniente que generar categorías sin que estas puedan ser reguladas por el entrenador, por lo cual fue descartada.
- La selección de la Red de Retropropagación se debe a la simplicidad que el preprocesado genera para este tipo de Red.
- La red de ART2 nos genera un retraso en el reconocimiento debido a la complejidad de su algoritmo de reconocimiento, este es un factor que la hace inefectiva para nuestra aplicación.
- El tipo de baquetas, instrumento para ejecución de algunos instrumentos de percusión, no presenta mayor influencia en el reconocimiento.
- El sistema necesita una tarjeta de sonido con capacidad para trabajar con frecuencias de muestreo de 44,100 Hz, con entrada para micrófono y una resolución de mínima de 16 bits.
- El sistema operativo bajo el cual puede correr se recomienda como mínimo Windows 98.

Bibliografía

- [1] Conceptos Musicales
Disponible:
<http://usuarios.lycos.es/sonymusica/solfeo/solfeo.htm>
Última visita: Febrero 2006
- [2] Solfeo
Disponible:
<http://www.guitarra.net/curso/Guitarra/solfeo.htm>
Última visita: Febrero 2006
- [3] Teoría de la Música
Disponible:
<http://www.emsia.com.ar/teoria.htm>
Última Visita: Febrero 2006
- [4] Las Figuras Musicales
Disponible:
<http://www.teoria.com/referencia/lectura/10.htm>
Última Visita: Febrero 2006
- [5] Alteraciones y Escalas
Disponible:
<http://mx.geocities.com/baldemusic/apuntes/alteraciones.html>
Última Visita: Febrero 2006
- [6] Teoría Musical – Instrumentos Musicales – Música Clásica y Moderna
Disponible:
www.mabeto.com/teoria/teoria.asp?L=397
Última Visita: Febrero 2006

- [7] Teoría Musical gratis – Pentagrama
Disponible:
<http://usuarios.lycos.es/pianoclasses/aulalec2.htm>
Última Visita: Febrero 2006
- [8] Escalas Menores
Disponible:
alufis35.uv.es/~jose/teoria/IV/IV.html
Última Visita: Febrero 2006
- [9] Batería – Lecciones – Conceptos sobre tiempos
Disponible:
http://www.guitarraonline.com.ar/bateria_partituras.htm
Última Visita: Febrero 2006
- [10] Redes Neuronales Artificiales
Disponible:
http://www.iiia.csic.es/~mario/rna/tutorial/RNA_marcos.html
Última Visita: Febrero 2006
- [11] Redes Neuronales y algoritmos de primer orden
Disponible:
<http://html.rincondelvago.com/redes-neuronales-y-algoritmos-de-primer-orden.html>
Última Visita: Febrero 2006
- [12] Clasificación de las RNA
Disponible:
http://es.tldp.org/Presentaciones/200304curso-glisa/redes_neuronales/curso-glisa-redes_neuronales-html/x84.html
Última Visita: Febrero 2006

[13] Redes Neuronales

Disponible:

http://html.rinconelvago.com/redes-neuronales_1.html

Última Visita: Febrero 2006

[14] Adaline y Perceptrón

Disponible:

<http://www.lfcia.org/~cipenedo/cursos/scx/Tema3/nodo3-1.html>

Última Visita: Febrero 2006

[15] Kohonen

Disponible:

<http://electronica.com.mx/neural/informacion/kohonen.html>

Última Visita: Febrero 2006

[16] Conceptos Básicos Redes Neuronales

Disponible:

<http://www.us.es/gtocom/pid/pid10/RedesNeuronales.htm>

Última Visita: Febrero 2006

[17] Procesamiento de Imágenes con Redes Neuronales

Disponible:

<http://www.sav.us.es/formaciononline/asignaturas/asigpid/apartados/textos/recursos/redesneuronales/PIRN.htm>

Última Visita: Febrero 2006

[18] Using WAV data, Ayuda en línea MSDN

Disponible:

<http://msdn1.microsoft.com/en-us/default.aspx>

Última visita: Febrero 2006

- [19] WAVE file Format
Disponible:
<http://www.borg.com/~jglatt/tech/wave.htm>
Última visita: Febrero 2006
- [20] Introducción al Procesamiento Digital de Señales, Carlos Platero,
Universidad Politécnica de Madrid
Disponible:
www.elai.upm.es/spain/Publicaciones/pub01/intro_procsdig.pdf
Última visita: Enero 2006
- [21] Wavelet Toolbox Documentation, ayuda en línea de Mathworks
Disponible:
www.mathworks.com
Última visita: Febrero 2006
- [22] Análisis de Predicción Lineal
Disponible:
<http://alek.pucp.edu.pe/~dflores/lpc.html>
Última visita: Junio 2006
- [23] Digital Image Processing, Second Edition, R. González, R. Woods
Disponible:
Biblioteca Universidad Don Bosco, San Salvador, El Salvador
- [24] Kohonen
Disponible:
<http://ohm.utp.edu.co/neuronales/Capitulo2/Competitivas/Kohonen.htm>
Última visita: Marzo 2006

[25] Retropropagación

Disponible:

<http://ohm.utp.edu.co/neuronales/Capitulo2/Backpropagation/AntecedentesB.htm>

Última visita: Marzo 2006.

[26] Trabajo de Graduación "*Sistema de reconocimiento de notas musicales para flauta dulce soprano*" - Universidad Don Bosco 2004

Disponible:

Biblioteca Central Universidad Don Bosco, San Salvador, El Salvador

Última Visita: Marzo 2006

[27] Material de Clase – Curso "Redes Neuronales", Universidad Don Bosco

[28] Fundamentals or Neural Networks "Architectures, Algorithms and Applications. – Laurene Fausett (Florida Institute of Technology)

Disponible:

Biblioteca Universidad Don Bosco, San Salvador, El Salvador