

Reconocimiento de sismicidad volcánica utilizando la técnica “Clustering” a través de la aplicación de Redes Neuronales Artificiales.

Alejandro Salazar Jurado

**Universidad de Nariño
Facultad de Ciencias Exactas y Naturales
Departamento de Física
San Juan de Pasto - Nariño
2019**

Reconocimiento de sismicidad volcánica utilizando la técnica “Clustering” a través de la aplicación de Redes Neuronales Artificiales.

Alejandro Salazar Jurado

Trabajo de grado para optar el título de Físico

Director:

**Oscar Ernesto Cadena Ibarra
MSc. en Ciencias - Geofísica**

**Universidad de Nariño
Facultad de Ciencias Exactas y Naturales
Departamento de Física
San Juan de Pasto - Nariño
2019**

Nota de responsabilidad

Las ideas y conclusiones aportadas en la tesis de grado son responsabilidad exclusiva de los autores.

Artículo 1. del acuerdo No.324 del 11 de Octubre de 1996, emanado por el Honorable Consejo Directivo de la Universidad de Nariño.

Todos los derechos reservados.

Nota de Aceptación

MSc. Oscar Hernesto Cadena

Director

Dr. Alvaro Rugeles Pérez

Jurado

MSc. Sandra Sánchez Sierra

Jurado

San Juan de Pasto, 24 de octubre de 2019

Agradecimientos

A a mis padres, gracias por su confianza, su constante apoyo e incansables esfuerzos que día tras día realizaron con el fin de hacer todo esto posible, además de enseñarme a ser cada día mejor desde su ejemplo.

Al mi asesor de tesis Msc.Oscar Ernesto Cadena por su constante orientación, tiempo y dedicación en la elaboración de este trabajo de grado.

Al Observatorio Vulcanológico y Simológico de Pasto por facilitarme el acceso a información de su base de datos y abrirme sus puertas.

A mis profesores, por la dedicación y horas de empeño en orientar mi aprendizaje.

En general a todos los que formaron parte en mi formación personal y académica.

A mis padres.

Resumen

En este trabajo se aplica una metodología que combina el análisis LPC (coeficientes de predicción lineal) y redes neuronales artificiales (RNA). Se utiliza la técnica “*clustering*” con el objetivo de clasificar de forma no supervisada dos tipos de sismos volcánicos: volcanotectónicos (VT) y de largo periodo (LP). Se usa una herramienta computacional (MATLAB) que está orientada a la clasificación automática de estas señales sísmicas, los archivos que contienen toda la información para el desarrollo de este trabajo fueron suministrados por el Observatorio Vulcanológico y Sismológico de Pasto (OVSP). Tanto señales completas como segmentos cortos de la traza continua se representan utilizando la técnica LPC que también cumple la función de reducir la dimensión del vector que contiene los datos de la forma de onda. Para la fase de clasificación se usa un tipo de red auto-organizada con el fin de optimizar el proceso de clasificación de este tipo de sismos generados en el Volcán Galeras (VG). La efectividad en la clasificación de sismicidad volcánica de tipo LP y VT por parte de determinadas RNA supera el 95 % de aciertos, en el caso de distinción de unidades de ruido y sismo se alcanza hasta el 100 % de aciertos en la discriminación para determinadas muestras.

Palabras claves— clasificación automática, redes neuronales artificiales, predicción lineal, sismos volcánicos.

Abstract

In this paper, a methodology that combines the analysis of the LPC (Linear Predictor Coefficients) and artificial neural networks (ANN) is applied. The clustering technique is used aiming unsupervised classification of two types of volcanic earthquakes: volcano-tectonic (VT) and long-period (LP). A computational tool (MATLAB) is used that is oriented to the automatic classification of these seismic signals, the files that contain all the information for the development of this work were supplied by the Observatorio Vulcanológico y Sismológico de Pasto (OVSP). Both complete signals and short segments of the continuous trace are represented using the LPC technique that also fulfills the function of reducing the dimension of the vector that contains the waveform data. For the classification phase a type of self-organized network is used in order to optimize the classification process of this type of earthquakes generated in the Galeras Volcano (GV). The effectiveness in the classification of volcanic seismicity of LP and VT type by certain ANN exceeds 95 % of correct answers, in the case of distinction of noise and earthquake units up to 100 % of correct answers in the discrimination for certain samples is reached.

keywords— automatic classification, artificial neural networks, linear prediction coding, volcanic seismicity

Tabla de contenido

Introducción	16
1 Sismicidad en el VG	18
1.1 Red de monitoreo del VG	18
1.2 Sismicidad en el volcán Galeras	20
1.3 Clasificación de sismos en el VG	21
1.3.1 Sismos volcánicos de fuente transitoria	22
1.3.2 Señales sísmicas persistentes	26
1.4 Antecedentes	28
2 Definición del Problema	31
2.1 Formulación del problema	31
2.2 Objetivos	32
2.2.1 Objetivo general	32
2.2.2 Objetivos específicos	32
2.3 Justificación	32
3 Marco teórico	33
3.1 Coeficientes de Predicción Lineal	33
3.1.1 Modelo de análisis por LPC	33
3.1.2 Estimación de los coeficientes de predicción lineal	35
3.1.3 Método de autocorrelación	36
3.2 Redes Neuronales Artificiales	37
3.2.1 La Neurona Artificial	38
3.2.2 Arquitectura de las Redes Neuronales	40
3.2.3 Operación de las redes neuronales artificiales	41
3.2.4 Capacidad de generalización	48
3.2.5 Validación cruzada	48
3.3 Técnica Clustering	49

3.3.1	Redes autoorganizables	50
3.3.2	Red de Kohonen	50
3.3.3	El algoritmo SOM	52
4	Metodología	55
4.1	Selección de registros sísmicos	56
4.1.1	Selección de las muestras para de segmentos de ruido y sismo	57
4.1.2	Selección de las muestras para sismos tipo VT y LP	58
4.2	Tratamiento digital de señales	58
4.3	Extracción de propiedades características	60
4.4	Aplicación de los LPC	61
4.4.1	LPC en la representación óptima de señales	62
4.5	Generación de las matrices de entrenamiento	62
4.6	Aplicación de la técnica clustering	62
5	Resultados.	69
5.1	Representación de señales mediante sus LPC	69
5.2	Selección de las muestras	73
5.2.1	Muestras para segmentos de ruido y sismo	73
5.2.2	Muestras de Sismos tipo VT y LP	74
5.3	Generación de las matrices de entrenamiento	75
5.3.1	Clasificación de unidades de ruido vs unidades de Sismo	75
5.3.2	Clasificación de sismos VT vs. sismos LP	76
5.4	Validación de resultados	78
5.4.1	Unidades de sismo vs unidades de ruido	78
5.4.2	Clasificación de sismos tipo VT vs sismos tipo LP	79
5.5	Almacenamiento y operación de las RNA	85
5.5.1	RNA en la clasificación de VT y LP	85
5.5.2	RNA en la clasificación de sismo y ruido	87
6	Discusión	89
7	Conclusiones	92
	Archivos adjuntos	93
	Archivos adjuntos	93
	Bibliografía	94

Lista de tablas

Tabla 1.1.1	Principales características de las estaciones que conforman la red de vigilancia sísmica del VG	20
Tabla 4.1.1	Número de ejemplos extraídos	57
Tabla 5.1.1	Resultados del análisis de trazas de sismos del tipo LP para diferentes órdenes de predicción lineal.	71
Tabla 5.1.2	Resultados del análisis de trazas de sismos del tipo VT para diferentes órdenes de predicción lineal.	72
Tabla 5.2.1	Selección de las muestras para unidades de ruido y sismo	73
Tabla 5.3.1	Datos de las matrices de entrenamiento para la discriminación de segmentos de ruido(R) y sismo(S)	76
Tabla 5.3.2	Configuración de las matrices de entrenamiento para la clasificación entre LP y VT	77
Tabla 5.3.3	Configuración de las matrices de entrenamiento para la clasificación entre LP y VT	78
Tabla 5.4.1	Resultados de la ejecución de las RNA para unidades de sismo y de ruido	79
Tabla 5.4.2	Resultados de la ejecución de las RNA entrenadas con 100 trazas de sismos LP y 100 trazas de sismos VT	80
Tabla 5.4.3	Resultados de la ejecución de las RNA entrenadas con 200 trazas de sismos LP y 200 trazas de sismos VT	81
Tabla 5.4.4	Resultados de la ejecución de las RNA entrenadas con 300 trazas de sismos LP y 300 trazas de sismos VT	82
Tabla 5.4.5	Resultados de la ejecución de las RNA entrenadas con 400 trazas de sismos LP y 400 trazas de sismos VT	83
Tabla 5.4.6	Resultados de la ejecución de las RNA entrenadas con 500 trazas de sismos LP y 500 trazas de sismos VT	84

Lista de figuras

Figura 1.1.1	Red de vigilancia sísmica del Volcán Galeras	19
Figura 1.3.1	Sismograma y espectro de un evento tipo VT-A	23
Figura 1.3.2	Sismograma y espectro de un evento tipo VT-B	23
Figura 1.3.3	Sismograma y espectro de un evento tipo LP.	24
Figura 1.3.4	Sismograma de un evento tipo Tornillo	25
Figura 1.3.5	Componente Norte-Sur para un Sismo tipo HYB	26
Figura 1.3.6	Sismograma y espectro de un TRE armónico	27
Figura 1.3.7	Sismogrma y espectro de un TRE espasmódico	27
Figura 3.1.1	Flujo de la señal desde su exitación hasta su salida	33
Figura 3.1.2	Elementos involucrados en la extracción de los LPC	34
Figura 3.2.1	La neurona artificial	39
Figura 3.2.2	Estructura de un modelo SOM	41
Figura 3.2.3	Descenso del gradiente para una función error	43
Figura 3.2.4	Errores global y de generalización en un problema ideal	49
Figura 3.3.1	Representación gráfica de un mapa SOM mediante la proyec- ción de <i>Sammon</i>	54
Figura 4.0.1	Diagrama de flujo acerca del tratamiento y procesado de la información	55
Figura 4.1.1	Formato de un ejemplo de los archivos en formato ASCII que contiene las trazas de los eventos sísmicos.	56
Figura 4.1.2	Traza de un evento sísmico representado mediante segmentos vectoriales de 300 datos.	58
Figura 4.2.1	Representación de una señal en la que se ha extraído el <i>offset</i>	58
Figura 4.2.2	Representación de una señal normalizada en sus unidades	59
Figura 4.2.3	Representación de una señal filtrada	59
Figura 4.2.4	Segmento de sismo sin ventanear (izquierda), segmento de sis- mo ventaneado (Derecha)	60

Figura 4.3.1	Representación de la envolvente sobre el espectro en una señal	61
Figura 4.6.1	Ventana introductoria del método Clustering	63
Figura 4.6.2	Ventana de selección de los datos	63
Figura 4.6.3	Ventana donde se determina el número total de neuronas	64
Figura 4.6.4	Ventana de entrenamiento de la red neuronal	65
Figura 4.6.5	Ventana posterior al entrenamiento de la red neuronal	66
Figura 4.6.6	Resultados obtenidos en la fase de entrenamiento	66
Figura 4.6.7	Ventana de validación de la red neuronal	67
Figura 4.6.8	Ventana de explotación de resultados en la red neuronal	68
Figura 5.1.1	Varianza vs. orden de LPC para 30 trazas de sismos tipo LP	69
Figura 5.1.2	Varianza vs. orden de LPC para 30 trazas de sismos tipo VT	70
Figura 5.2.1	Representación espectral y envolvente de un sismo de tipo VT	74
Figura 5.2.2	Representación espectral y envolvente de un sismo de tipo LP	75
Figura 5.4.1	Efectividad vs orden LPC para un conjunto de matrices entrenadas con 100 trazas LP y 100 trazas VT	80
Figura 5.4.2	Efectividad vs orden LPC para un conjunto de matrices entrenadas con 200 trazas LP y 200 trazas VT	81
Figura 5.4.3	Efectividad vs orden LPC para un conjunto de matrices entrenadas con 300 trazas LP y 300 trazas VT	82
Figura 5.4.4	Efectividad vs orden LPC para un conjunto de matrices entrenadas con 400 trazas LP y 400 trazas VT	83
Figura 5.4.5	Efectividad vs orden LPC para un conjunto de matrices entrenadas con 500 trazas LP y 500 trazas VT	84
Figura 5.5.1	Forma en la que se almacenaron cada una de las RNA para la clasificación de VT y LP	85
Figura 5.5.2	Segmento del script de una RNA donde se indican un ejemplo sobre las variables “ruta” y “ruta1”	86
Figura 5.5.3	Ejemplo de los resultados que se obtienen al ejecutar una RNA	87
Figura 5.5.4	Forma en la que se almacenaron cada una de las RNA para la clasificación de ruido y sismo	87

Glosario

Coda: Ondas sísmicas que constituyen la parte final del registro de un sismo.

Código ASCII: (Código Estadounidense Estandar para el intercambio de Información) Código definido y establecido en sistemas e informática para representar los caracteres de forma numérica con el fin de trabajar con ordenadores.

Enjambres sísmicos: Es la ocurrencia de un conjunto de eventos sísmicos en un área específica durante un periodo de tiempo relativamente corto.

Envolvente: Curva que pasa por los picos del espectro de frecuencias en una señal modulada.

Espectro: Gráfica de amplitud de energía vs frecuencia.

Factor de calidad: Parámetro utilizado para medir la atenuación de la energía de las ondas sísmicas.

Frecuencia de muestreo: Es el número de muestras por unidad de tiempo que se toman de una señal continua para producir una serie de tiempo discreta.

Fuente: Lugar donde se generan los sismos.

Función de transferencia: Modelo matemático que a través de un cociente relaciona la respuesta de un sistema (modelada o señal de salida) con una señal de entrada o excitación (también modelada)

Hipocentro: Un punto en el interior de la Tierra donde se inicia la ruptura que causa el sismo. Se le localiza mediante la latitud y longitud geográficas, y la profundidad.

Lahares: Es un flujo de sedimento y agua que se moviliza desde las laderas de volcanes.

Magnitud: Parámetro que clasifica los sismos por la amplitud de las ondas sísmicas registradas en los sismógrafos. Da una idea del tamaño del sismo y la cantidad de energía liberada en el hipocentro. La más conocida es la magnitud de Richter.

MATLAB: Abreviatura de MATrix LABoratory, es una herramienta de software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M).

Onda P: Tipo de onda denominada también onda longitudinal o de compresión, en la cual las partículas del medio oscilan en la misma dirección de propagación de la onda.

Onda S: Tipo de onda conocida como onda transversal o de cizalla. Las partículas del medio oscilan perpendicularmente a la dirección de propagación de la onda. Presentan su arribo después de la onda P.

Sismo / Seísmo: Movimiento de la Tierra producido por el paso de las ondas generadas por el desplazamiento repentino del suelo por debajo de la superficie de la Tierra.

Sismógrafo: Instrumento que registra las ondas sísmicas.

Sismograma: Registro hecho por un sismógrafo.

Tornillo: Sismos de largo periodo, monocromático, con lento decaimiento de sus ondas de coda y cuya fuente está asociada con procesos de resonancia.

Traza: Representación gráfica de un sismo.

Introducción

El Volcán Galeras se localiza en el Departamento de Nariño, aproximadamente a 9 km al occidente de la ciudad de San Juan de Pasto, capital de este departamento, en las coordenadas $1^{\circ} 13' 43,8''$ de latitud norte y $77^{\circ} 21' 33,0''$ de longitud oeste y con una altura de 4276 msnm. Es el volcán con mayor actividad en Colombia, el que cuenta con mayores registros históricos. En su zona de influencia se encuentran establecidos nueve municipios y más de siete corregimientos, que en total albergan más de 500.000 habitantes, por lo anterior posee una gran documentación de su actividad y una permanente vigilancia. El volcán Galeras ha registrado diferentes tipos de señales sísmicas, algunas de éstas están asociados a la fractura de roca (volcano-tectónicos) y otras a movimiento transitorio de fluidos (largo periodo) [12].

Una de las tareas del OVSP es la preparación y clasificación de información sísmica como insumo para el análisis e interpretación de los niveles de actividad volcánica que inicia con la generación de histogramas a partir de la clasificación de la sismicidad. Esto se lleva a cabo manualmente, analizando uno a uno los registros sísmicos, lo cual es una tarea laboriosa que requiere la inversión de una gran cantidad de tiempo por parte del personal de turno, dado que la información se adquiere en tiempo real es conveniente un tratamiento, análisis y clasificación más eficaces que el manual, sobre todo en épocas de crisis volcánicas cuando se registran centenares de sismos diariamente.

Con el objetivo de clasificar los diferentes tipos de sismicidad, en este caso tipo VT y LP, se recurre a técnicas computacionales de análisis tales como: el análisis de predicción lineal y redes neuronales artificiales por medio de la técnica “clustering” la cual consiste en un método de minería de datos (data mining) dentro de la disciplina de Inteligencia Artificial que identifica de forma automática agrupaciones o clústeres de elementos de acuerdo a una medida de similitud entre ellos (reconocimiento de patrones) basandose en el uso de mapas auto-organizados. La representación de las señales por medio de LPC se justifica porque reduce la dimensión de los vectores de entrada al sistema clasificador. Las RNA adecuadamente

entrenadas son utilizadas en la fase de clasificación.

El objetivo fundamental de las técnicas de clustering consiste en identificar grupos o clústeres de elementos tal que: La similitud media entre elementos del mismo clúster sea alta. Similitud intra-clúster alta. La similitud media entre elementos de distintos clústeres sea baja. Similitud inter-clúster baja. La técnica de “clustering” se complementa con el uso de redes neuronales artificiales por sus características de sistema inteligente que logra realizar con éxito tareas complejas. Su funcionamiento no está orientado a ser similar al de un ordenador cuando procesa información, sino que persigue aproximarse a la inteligencia artificial.

Capítulo 1

Sismicidad en el VG

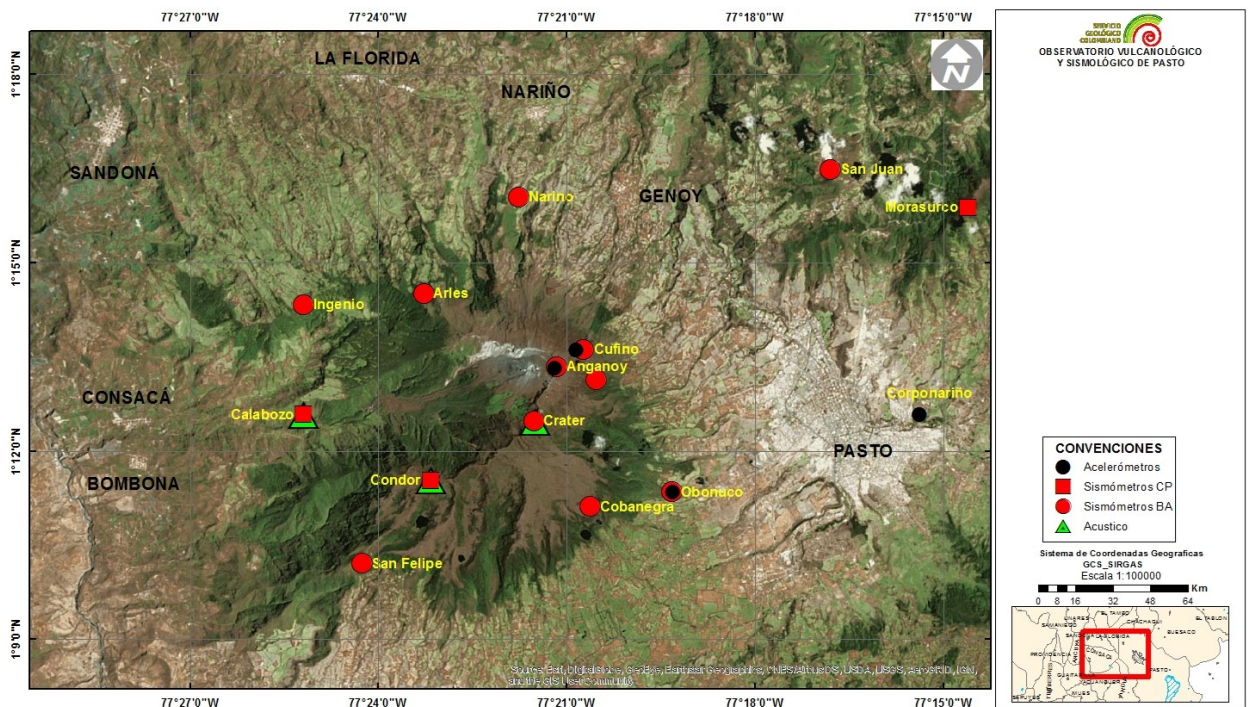
La sismicidad es un reflejo de procesos asociados con liberación de energía en el interior de la Tierra, en regiones volcánicas estos procesos obedecen principalmente a la dinámica del magma. Su movimiento, cambios de fase, cambios de temperatura, desprendimiento de volátiles, son algunos de los fenómenos que producen cambios de presión que son transmitidos por la roca y registrados por los sismómetros. En ocasiones estos cambios de presión son transitorios y en otros se sostienen produciendo sismos de mayor duración. Suele suceder que dichas presiones son suficientes para rebasar el nivel de fragilidad de la roca, fracturándola, originando sismos bastante similares a los sismos tectónicos.

Los sismómetros captan el movimiento del terreno cuya señal dibuja sobre los sismogramas diferentes formas denominadas firmas o registros sísmicos. Cada firma es diferente de otra, sin embargo, existen características comunes que se pueden encontrar en sismos originados por condiciones similares de fuente, de manera que un análisis de los registros y una clasificación que atienda a la fuente que los produjo son claves para lograr una adecuada interpretación del fenómeno volcánico [1].

1.1. Red de monitoreo del VG

La red de vigilancia sísmica (RVSG) para el monitoreo del VG se estableció en 1989, el VG cuenta diferentes estaciones que encuentran ubicadas entre 1 y 14 km de distancia al cráter activo (figura 1.1.1).

Figura 1.1.1: Red de vigilancia sísmica del Volcán Galeras.



Fuente: (SGC-OVSP, 2019)[5]

La red sísmica comprende sismómetros de corto periodo, sismómetros de banda ancha y sensores de movimiento fuerte o acelerómetros. La red de sensores de presión acústica está compuesta por sensores de infrasonido. Entre ellos se destacan dos tipos de sismómetros, de corto periodo si posee una respuesta en velocidad normalmente plana en el rango de frecuencias mayor a 1 Hz, o, de banda ancha si su respuesta en frecuencia es mucho más amplia que la anterior, incluyendo frecuencias desde los 0.08 Hz .

La forma como se transmiten los datos puede ser analógica o digital, actualmente la RVSG cuenta en todas sus estaciones de monitoreo un tipo de transmisión digital (tabla 1.1.1), en donde, un sensor convierte la señal de velocidades en una señal de voltajes, la señal es digitalizada a la misma frecuencia anterior, posteriormente esta señal es enviada por medio de un radio transmisor hacia la sede y pasa por un programa de adquisición.

En la tabla 1.1.1 se presenta información relacionada con su nombre, tipo de transmisión de datos, tipo de sensor que posee, componente, la distancia al cráter (en Km) y altitud.

Tabla 1.1.1: Principales características de las estaciones que conforman la red de vigilancia sísmica del VG.

Estación	Tansmisión	Tipo Sensor	Componente	Distancia cráter	Altitud(m)
Anganoy	Digital	Banda Ancha	Triaxial	0.7	4227
Anganoy	Digital	Acelerómetro	Triaxial	0.7	4227
Cufiño	Digital	Banda Ancha	Triaxial	1.4	3828
Cufiño	Digital	Acelerómetro	Triaxial	1.4	3828
Cráter	Digital	Banda Ancha	Triaxial	1.5	4032
Urcunina	Digital	Banda Ancha	Triaxial	2.4	3494
Arlés	Digital	Banda Ancha	Triaxial	4	3450
Cobanegra	Digital	Banda Ancha	Triaxial	4.3	3625
Cóndor	Digital	Corto Periodo	Triaxial	4.4	3977
Cóndor	Digital	Presión Acústica	4 Entradas	4.4	3977
Nariño-2	Digital	Banda Ancha	Triaxial	5.2	2856
Obonuco	Digital	Banda Ancha	Triaxial	5.4	3021
Obonuco	Digital	Acelerómetro	Triaxial	5.4	3021
Calabozo	Digital	Corto Periodo	Triaxial	6.9	2313
Calabozo	Digital	Infrasonido	4 Entradas	6.9	2313
Ingenio	Digital	Banda Ancha	Triaxial	7	2890
San Felipe	Digital	Banda Ancha	Triaxial	7.6	3491
San Juan	Digital	Banda Ancha	Triaxial	10.6	2898
Corponariño	Digital	Acelerómetro	Triaxial	11.5	2600
Morasurco	Digital	Corto Periodo	Triaxial	13.7	3504
Cráter	Digital	Infrasonido	4 Entradas	1.5	4032
Cóndor	Digital	Infrasonido	4 Entradas	4.4	3977
Calabozo	Digital	Infrasonido	4 Entradas	6.9	2313

Fuente: (SGC-OVSP, 2019) [5]

1.2. Sismicidad en el volcán Galeras

La sismicidad es un reflejo de procesos asociados con liberación de energía en el interior de la Tierra, en regiones volcánicas estos procesos obedecen principalmente a la dinámica del magma. Su movimiento, cambios de fase, cambios de temperatura, desprendimiento de volátiles, son algunos de los fenómenos que producen cambios de presión que son transmiti-

dos por la roca y registrados por los sismómetros. En ocasiones estos cambios de presión son transitorios y en otros se sostienen produciendo sismos de mayor duración. Suele suceder que dichas presiones son suficientes para rebasar el nivel de fragilidad de la roca, fracturándola, originando sismos bastante similares a los sismos tectónicos.

Los sismómetros captan el movimiento del terreno cuya señal dibuja sobre los sismogramas diferentes formas denominadas firmas o registros sísmicos. Cada firma es diferente de otra, sin embargo, existen características comunes que se pueden encontrar en sismos originados por condiciones similares de fuente, de manera que un análisis de los registros y una clasificación que atienda a la fuente que los produjo son claves para lograr una adecuada interpretación del fenómeno volcánico [1].

1.3. Clasificación de sismos en el VG

En los volcanes del mundo se presentan sismos que tienen firmas semejantes y que se han asociado con el tipo de fuente que los origina. Por ejemplo, la fractura de material cortical en el edificio volcánico, sin que importe la causa que la produjo, genera eventos sísmicos que tienen arribos impulsivos y sus fases P y S son fácilmente reconocidas, estos sismos son llamados volcano-tectónicos, se distinguen por las siglas VT. Puede suceder que los sismos VT sean producto de actividad volcánica o de acumulación de esfuerzos tectónicos que se transmiten hacia la región volcánica; el hecho de que su foco esté localizado en inmediaciones de la estructura volcánica es suficiente para clasificarlo como VT.

Se observan también eventos sísmicos volcánicos relacionados con la dinámica del sistema magma-gas e hidrotermal. Este tipo de sismos se divide en dos grupos diferenciados por el tiempo de excitación de la fuente generatriz.

Se llaman sismos de largo periodo o LP aquellos en que su fuente es transitoria, entre otros se encuentra los denominados de tipo tremor o TRE que se producen cuando la fuente es persistente en el tiempo y finalmente sismos que involucra tanto fractura de material sólido como movimiento de fluidos, estos son llamados Híbridos o HYB [1].

1.3.1. Sismos volcánicos de fuente transitoria

Hace referencia a señales cuya fuente actúa en un tiempo relativamente corto, generalmente unos pocos segundos, luego de esto se generan las ondas de coda del evento.

Existen señales transitorias asociadas con fractura de material cortical y también con movimiento de fluidos, dentro de estas señales se encuentran los tipos: Volcano-Tectónico (VT), Largo Periodo (LP), Híbridos (HYB) y Tornillo (TO).

Sismos volcano-tectónicos (VT)

Tienen que ver con el fracturamiento de extensión o de cizalla en la parte sólida del volcán debido a presiones inducidas por magma o cambios bruscos de temperaturas, dentro de éstos tipos de sismos podemos distinguir dos clases, VT-A y VT-B:

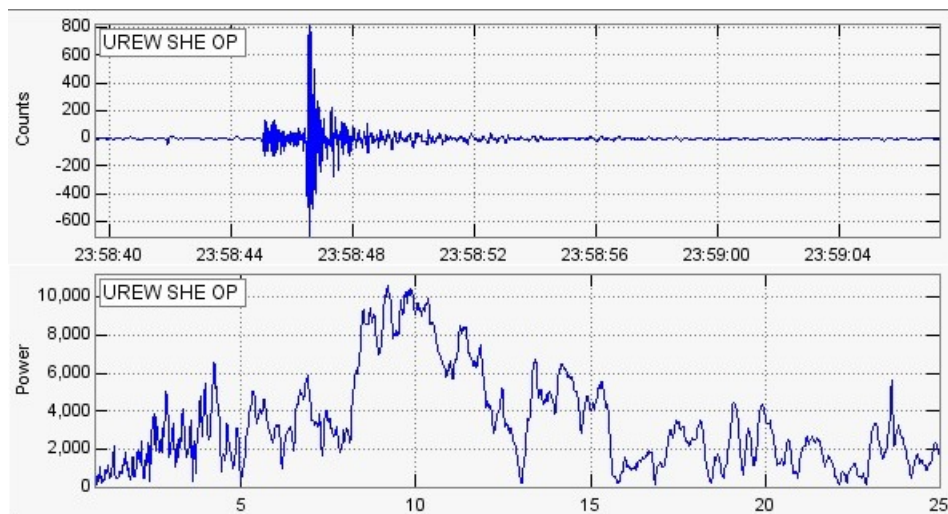
Sismos volcano-tectónicos tipo A o VT-A

En general ocurren a profundidades entre 2 y 10 km desde la cima del volcán. En su registro en el dominio del tiempo es posible distinguir con facilidad las fases P y S (figura 1.3.1). El decaimiento de la envolvente de sus ondas de coda muestra una forma exponencial disminuyendo su amplitud rápidamente con el tiempo. El espectro de este tipo de eventos contiene altas frecuencias, mayores que 5 Hz; los picos que aportan con mayor energía a la señal están generalmente en una banda entre los 6 y 8 Hz y sus ondas de coda pueden contener frecuencias significativamente energéticas, incluso hasta los 15 o 20 Hz.

La distribución de eventos VT en el tiempo es, por lo general, a manera de enjambre, con magnitudes aleatorias, en contraste con los eventos no volcánicos que muestran una secuencia de réplicas que siguen al evento principal y van disminuyendo su magnitud en el tiempo.

Figura 1.3.1: Sismograma y espectro (de arriba hacia abajo) de un evento tipo VT-A registrado por la componente Este-Oeste de la estación Urcunina, ubicada a 2.2 Km al sur del foco del sismo. El eje de las ordenadas, está en cuentas (sin respuesta instrumental) para el sismograma; en el espectrograma está en Hz; el eje de las ordenadas en el espectro es de potencia.

Figura 1.3.1: Sismograma y espectro de un evento tipo VT-A.

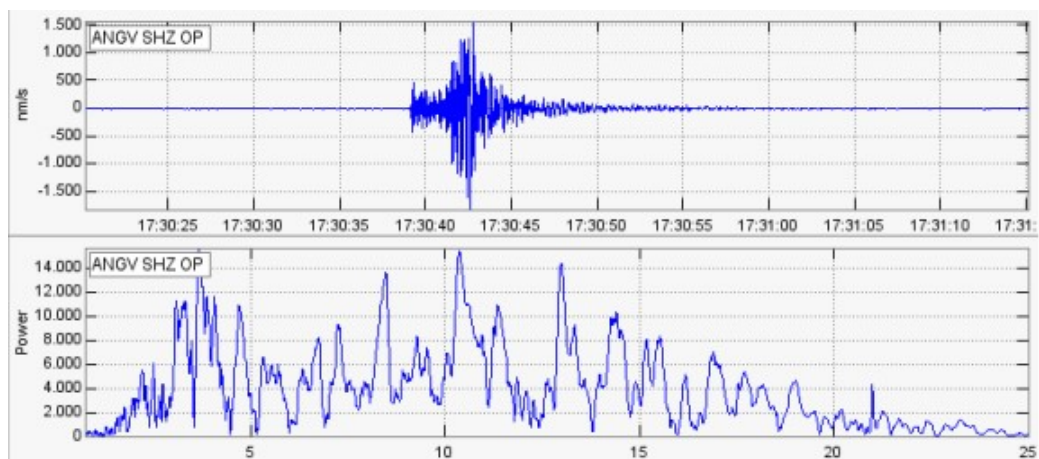


Fuente: (Cadena, 2011) [1]

Sismos volcano-tectónicos tipo B o VT-B

Con ocurrencia a profundidad focal menor a 2 km, muestran larga duración de coda definiendo una dispersión de tren de ondas superficiales, los arribos de las ondas P son emergentes y resulta imposible detectar un arribo claro de fases de ondas S (figura 1.3.2).

Figura 1.3.2: Sismograma y espectro de un evento tipo VT-B.



Fuente: (Cadena, 2011) [1]

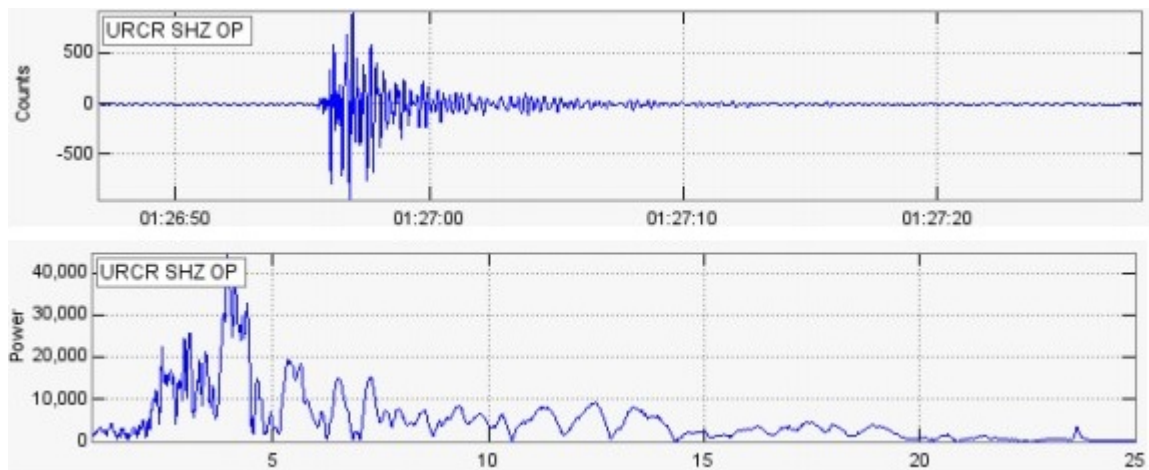
Figura 1.3.2 : Sismogramas de un evento tipo VT-B en dos de las estaciones de la red del Volcán Galeras, una de ellas, Cufiño. Estas estaciones se encuentran ubicadas próximas a la cima del volcán, cerca del cráter principal. En su espectro se observa que la energía del sismo está concentrada en una banda de bajas frecuencias.

Sismos tipo Largo Periodo (LP)

Este tipo de sismo volcánico está asociado al movimiento de fluidos al interior de un edificio volcánico. En general, se encuentran asociados a procesos con origen en la física de fluidos multifase (gases y líquidos) a través de grietas y conductos. Algunas de estas señales se asocian con mecanismos de desgasificación en canales abiertos. El líquido y el gas pueden ser de origen magmático o hidrotermal dependiendo del escenario volcánico, siendo el contenido gaseoso variable de acuerdo a su volatilidad y a la profundidad de la fuente.

Los sismos LP son señales de corta duración, no duran más de un minuto y se presentan como enjambres a una profundidad típica menor a 2 km, en un sismograma la onda P se hace emergente y la onda S no es notable (figura 1.3.3).

Figura 1.3.3: Sismograma y espectro de un evento tipo LP.



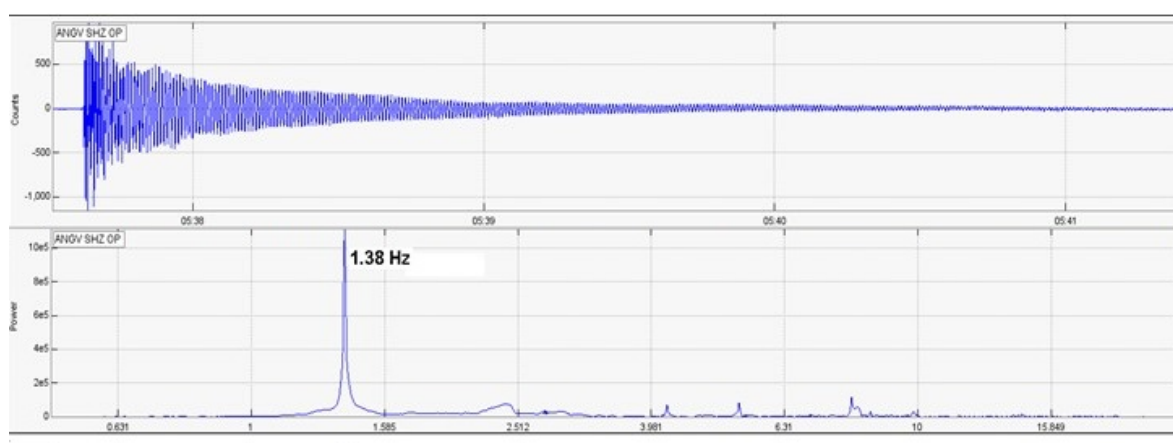
Fuente: (Cadena, 2011) [1]

Figura 1.3.3: Sismograma y espectro de un evento tipo LP en la componente vertical de la estación Urcunina.

Sismos tipo Tornillo (TO)

Este nombre fue acuñado en Colombia por el personal del OVSP para distinguir un tipo especial de sismos de largo periodo cuya fuente está asociada con procesos de resonancia. Su firma muestra un inicio impulsivo tras el cual se observa que uno o varios armónicos alargan su coda por tiempos relativamente largos. La manera en que decae la amplitud del sismograma a lo largo del tiempo es lenta, lo que implica que la tasa de pérdida de energía en la cavidad resonante es lenta también (Figura 1.3.4).

Figura 1.3.4: Sismograma de un evento tipo Tornillo.



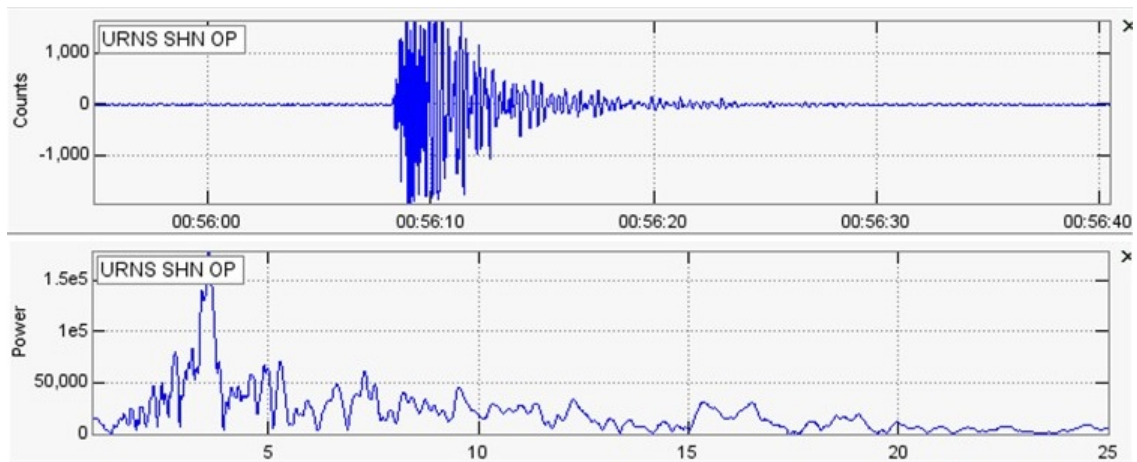
Fuente: (Cadena, 2011) [1]

Figura 1.3.4: Sismograma de un evento tipo Tornillo en la estación Anganoy. Obsérvese el lento decaimiento de sus ondas de coda y el pico dominante en los 1.38 Hz .

Sismos Híbridos (HYB)

Muestran características originadas por la combinación de procesos que involucran movimiento de fluidos y fracturamiento de material cortical por lo cual se ha sugerido que se trata de la combinación de eventos tipo VT-A o VT-B con sismos tipo LP. Generalmente su arribo es impulsivo, dominado altas frecuencias y en muchas ocasiones es posible distinguir la fase S. Sus ondas de coda muestran patrones armónicos no dispersivos, característicos de los eventos tipo LP. Como los sismos VT, los sismos HYB muestran diferentes polaridades en sus arribos en las diferentes estaciones, mientras que los LP tienen igual polaridad en las diferentes estaciones de registro.

Figura 1.3.5: Componente Norte-Sur para un Sismo tipo HYB.



Fuente: (Cadena, 2011) [1]

Figura 1.3.5: Detalles en la componente Norte-Sur de la estación Urcunina para el Sismo tipo HYB en tres estaciones de la red del VG. Sismograma (superior) y espectro (inferior) .

1.3.2. Señales sísmicas persistentes

El sismograma puede estar asociado con una fuente persistente en el tiempo relacionada con dinámica de fluidos o con procesos de deslizamiento de roca o lahares, dentro de estas señales se destaca el Tremor Volcánico (TRE).

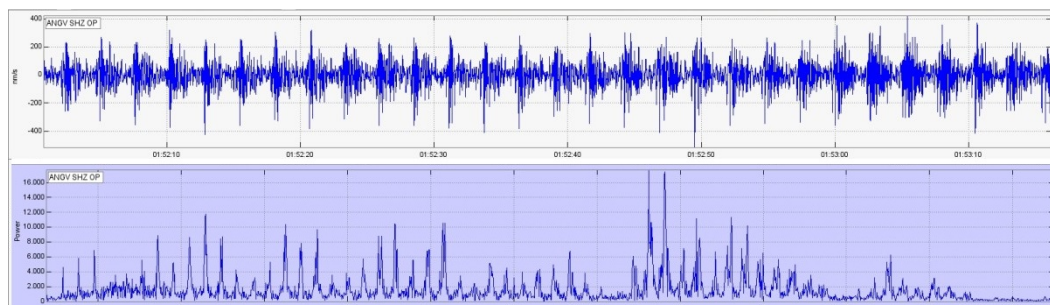
Tremor Volcánico (TRE)

El tremor volcánico es una señal sísmica caracterizada por mantener la amplitud constante durante un largo período de tiempo que puede oscilar entre varios minutos y horas, con contenido espectral centrado en las bandas de frecuencia relativamente estrechas.

En general no se pueden identificar con claridad las fases P debido a su emergencia y no se distinguen fases S, se reconocen dos tipos de TRE:

Armónico: Se caracteriza por sucesiones de onda monocromáticas (una misma frecuencia) o cuasi-monocromáticas en una banda muy angosta, muestra una modulación regular de su envolvente (figura 1.3.6).

Figura 1.3.6: Sismograma y espectro de un TRE armónico.

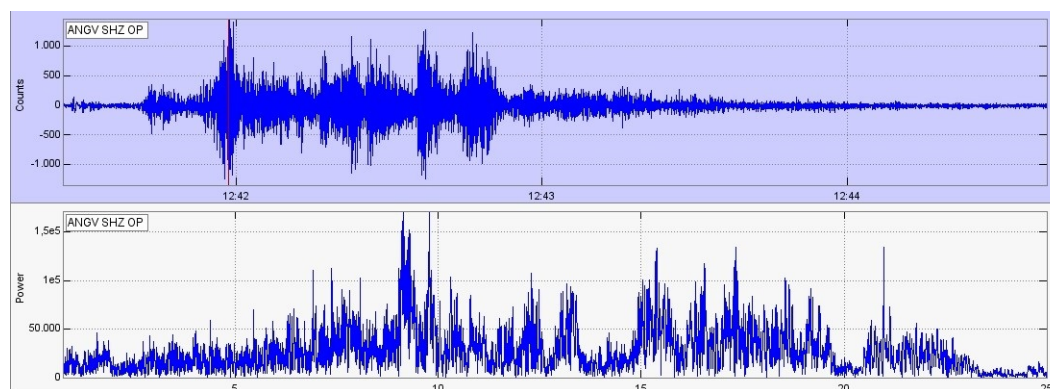


Fuente: (Cadena, 2011) [1]

Figura 1.3.6: Sismograma y espectro de un tremor armónico registrado por la estación Anganoy.

Espasmódico: Su longitud y frecuencia varían significativamente, su banda de contenido espectral es más amplia que en el tremor armónico. (figura 1.3.7)

Figura 1.3.7: Sismograma y espectro de un TRE espasmódico.



Fuente: (Cadena, 2011) [1]

Figura 1.3.7: Sismograma y espectro de un episodio de tremor espasmódico registrado por la estación Anganoy, estación base de clasificación de la red sísmica del VG .

No clasificables

Se considera no clasificable el tipo de sismicidad que debido a su bajo contenido energético imposibilita la ubicación de la fuente de origen, por lo tanto, dificulta la clasificación en

algunos de los tipos de sismos anteriormente mencionados. El principal problema radica en la dificultad para distinguirse del ruido base .

Ruido

Se considera ruido a todo lo que no es una señal útil, por lo que puede deberse a interferencias, clima, temperatura, impurezas, actividad del hombre cerca de las estaciones etc.

1.4. Antecedentes

A continuación se mencionan algunos artículos y trabajos en los que se han utilizado diferentes tipos de métodos para la clasificación de sismicidad volcánica.

Detección y clasificación automática de registros sísmicos en el Observatorio Vulcanológico y Sismológico de Pasto utilizando redes neuronales artificiales

En este trabajo se aplica una metodología que combina Análisis de predicción lineal (LPC) y Redes neuronales artificiales (RNA) con el objetivo de detectar y clasificar tres tipos de sismos volcánicos: volcano-tectónicos (VT) tipo A, largo periodo (LP) y tremor (TRE). Se construye una herramienta computacional orientada a la detección y clasificación automática de estas señales sísmicas, generando información en tiempo real. Segmentos cortos de la traza continua se representan utilizando la técnica LPC que también cumple la función de reducir la dimensión del vector que contiene los datos de la forma de onda. Para las fases de detección y clasificación se utilizan RNA tipo perceptrón multicapa con una capa oculta. Entre las ventajas de la metodología está la posibilidad de detectar sismos con arribos emergentes, de corta duración o de bajo nivel energético. La aplicación del método muestra cerca de 100 de acierto para la distinción entre ruido y señal sísmica (fase de detección), 92 % de acierto para la clasificación entre sismos LP y VT y cerca del 100 % en la detección de sismos tipo TRE.

[1]

Evaluation of Llaima volcano activities for localization and classification of LP, VT and TR events

El objetivo de este documento es la evaluación (eliminación de ruidos, localización y cla-

sificación) y el análisis de las actividades del volcán Llaima, uno de los volcanes más activos en América del Sur. Diferentes métodos ya propuestos, como Butterworth, la sustracción espectral (SS) y el filtro de Wiener (WF) se comparan con la sustracción espectral modificada (MSS) propuesta y el filtro de Wiener modificado (MWF) para encontrar el mejor método para eliminar las señales no deseadas del volcán. Los métodos de eliminación de ruido propuestos hacen que el punto de inicio del evento sea más evidente para aumentar la precisión de la localización de los eventos en los que es difícil encontrar el punto de inicio y se propone un método basado en la novedosa técnica de DNN para clasificar los tres eventos principales que ocurren en el volcán Llaima : TR (Tremor), LP (Período Largo) y VT (Volcáno-Tectónico).[4]

Application of Artificial Neural Networks for the classification of the seismic transients at Soufrie‘re Hills volcano, Montserrat

En este trabajo se propone un método de clasificación supervisada de los eventos transitorios tales como terremotos tectónicos, eventos de largo período, eventos híbridos y desprendimientos de rocas presentados por el volcán Soufrie‘re Hills, basados en redes neuronales artificiales (RNA) sobre la actividad sísmica, ya que este método es útil para procesar los grandes conjuntos de datos acumulados en el pasado, En particular, los conjuntos de datos registrados antes de las erupciones climáticas de 1995 a 2002 pueden permitirnos reconstruir la distribución de las diferentes clases de eventos transitorios sísmicos en el tiempo. Creemos que este análisis puede proporcionar información útil sobre los escenarios eruptivos inminentes. El buen desempeño de la RNA con el 70 % de los eventos transitorios clasificados correctamente en un conjunto de pruebas de 156 datos, hacen que una RNA sea una herramienta poderosa para el procesamiento de datos sismológicos de volcanes.[7]

Machine Learning for Volcano-Seismic Signals: Challenges and Perspectives

Los enfoques actuales para el monitoreo de volcanes se basan principalmente en el análisis manual de varios parámetros, que incluyen mediciones de deformaciones y análisis de señales sísmicas. Sin embargo, debido a la gran cantidad de datos adquiridos por los sensores para el monitoreo a largo plazo, la inspección manual ya no es una opción viable. Como en muchas situaciones de Big Data, ahora se considera los métodos clásicos de Aprendizaje supervisado automatizan el análisis de años de señales grabadas, lo que permite el monitoreo a mayor escala. Este documento se centra en herramientas integradas y operativas dedicadas al análisis automático de señales volcánicas y sísmicas. Es decir, revisamos (i) las herramientas

para la representación óptima de las señales volcán-sísmicas (espacio de características) y los métodos disponibles para la detección de eventos volcánicos-sísmicos (ii) y (iii) la clasificación. Luego proponemos una arquitectura para la clasificación automática de eventos volcán-sísmicos.

Nuestro sistema de predicción se prueba en 6 años de grabaciones que contienen 109434 eventos volcánicos-sísmicos adquiridos del volcán Ubinas (el volcán más activo en Perú).

Nuestro nuevo modelo propuesto se construye utilizando algoritmos supervisados de aprendizaje automático (Máquina de vectores de soporte) y alcanza el 92.2 % de la clasificación correcta en seis clases. Este modelo de predicción se utiliza para analizar completamente los 6 años de señales grabadas. [10]

Machine Learning for Automatic Classification of Volcano-Seismic Signatures

En este artículo se presenta una arquitectura novedosa para la clasificación automática de eventos volcánicos y sísmicos basados en una representación de señal completa con un gran conjunto de características. Según nuestro conocimiento, este es uno de los primeros intentos de automatizar la tarea de clasificación de estas señales. El enfoque propuesto se basa en técnicas de aprendizaje automático supervisadas para construir un modelo de predicción.[9]

Capítulo 2

Definición del Problema

El (VG) es catalogado como uno de los más activos de Colombia, por ésta razón posee una gran documentación y constante vigilancia. Una de las rutinas de operación del (OVSP) se enfoca en la preparación de la información sismológica como herramienta en el análisis de la actividad volcánica. Para ello se generan histogramas a partir de la clasificación de la sismicidad, ésto es una tarea que se realiza manualmente analizando de manera separada cada registro sísmico, en consecuencia se emplea una gran cantidad de tiempo del personal de turno.

En épocas de riesgo cuando se registran una gran cantidad de sismos diariamente se requiere un método más eficiente que el manual ya que se requiere una toma de decisiones rápidas y efectivas con el objetivo de alertar de manera oportuna a las personas que habitan la zona de influencia del volcán.

Este trabajo está dedicado exclusivamente a usar un tipo de técnica que sea capaz de optimizar el tiempo en la clasificación de la sismicidad mediante un proceso automático no supervisado.

2.1. Formulación del problema

¿Es posible aplicar redes neuronales artificiales con aprendizaje no supervisado en una clasificación efectiva de sismicidad volcánica (tipo VT y LP) en el Volcán Galeras?

2.2. Objetivos

2.2.1. Objetivo general

Aplicar la técnica “clustering” para el reconocimiento de sismicidad volcánica.

2.2.2. Objetivos específicos

1. Aplicar RNA como herramienta computacional para el uso de la técnica “clustering”.
2. Evaluar la efectividad de la red neuronal en la clasificación de eventos sísmicos a partir de los datos suministrados por el OVSP.

2.3. Justificación

La actividad volcánica genera un gran impacto económico y social. Actualmente la monitorización de volcanes se fundamenta principalmente en el análisis de la actividad sísmica de los eventos considerados precursores de erupciones. Un sistema automático que sea capaz de clasificar eventos sismo-volcánicos, permitiría una gestión más eficaz al evaluar del riesgo volcánico sobre todo cuando previo a una erupción el incremento de la actividad es tal que compromete la fiabilidad de la clasificación supervisada llevada a cabo por los técnicos de los observatorios. Un análisis detallado de la situación es crucial a la hora de tomar decisiones que pueden ser críticas como la necesidad de evacuación de la población [2].

Con el fin de optimizar el proceso de clasificación sin comprometer la precisión de los resultados, se propone una técnica de clasificación automática que involucra el uso de redes neuronales artificiales cuyo aprendizaje es no supervisado, todo esto, a través de la implementación del software MATLAB.

Capítulo 3

Marco teórico

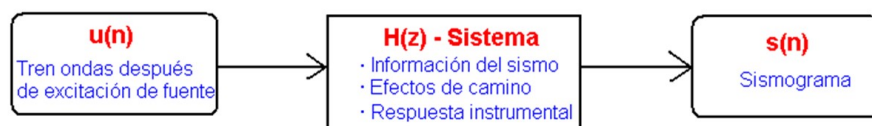
Algunas de las siguientes definiciones han sido tomadas del trabajo de titulación denominado “*Detección y clasificación automática de registros sísmicos en el Observatorio Vulcanológico y Sismológico de Pasto utilizando redes neuronales artificiales*” [1]

3.1. Coeficientes de Predicción Lineal

3.1.1. Modelo de análisis por LPC

Supongamos una señal de excitación $u(n)$; esta señal está afectada por un sistema que modelaremos a través de su función de transferencia $H(z)$; la señal de salida del sistema es $s(n)$. El sistema, representado por $H(z)$, incluye características de fuente y todos los factores que afectan la señal hasta obtener el registro sismográfico (figura 3.1.1).

Figura 3.1.1: Flujo de la señal desde su excitación hasta su salida.

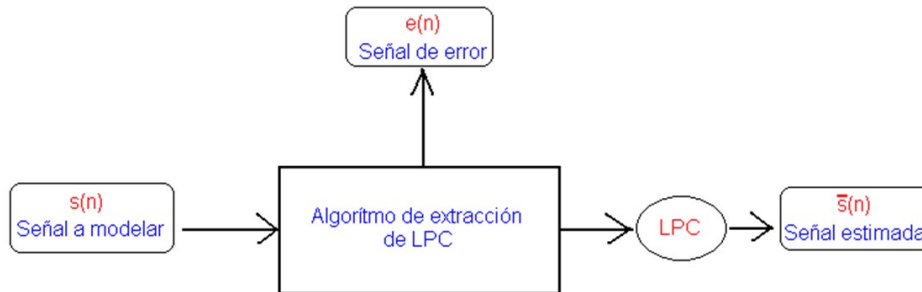


Fuente: (Cadena, 2011) [1]

Los LPC se calculan utilizando un algoritmo que actúa sobre la señal $s(n)$ y arroja los vectores que contienen los coeficientes de predicción y una señal de error $e(n)$. Es posible recuperar la señal estimada $\tilde{s}(n)$ a partir de la señal de error aplicando sobre $e(n)$ un filtro inverso

que contenga los LPC como coeficientes del filtro (figura 3.1.2).

Figura 3.1.2: Elementos involucrados en la extracción de los LPC.



Fuente: (Cadena, 2011) [1]

La muestra $\tilde{s}(n)$ se obtiene mediante la combinación lineal de las salidas pasadas y las entradas presente y pasadas. Esta representación corresponde a la siguiente ecuación de diferencias:

$$\tilde{s}(n) = \sum_{k=1}^p a_k s(n-k) + G \sum_{l=0}^q b_l u(n-l), \quad b_0 = 1 \quad (3.1)$$

donde G , a_k y b_l , y son los parámetros del sistema. El valor de p es el número de muestras de salidas pasadas involucradas en el cálculo e indica el orden de la predicción lineal. La función de transferencia del sistema se obtiene aplicando la transformada \mathbf{Z} a la ecuación (3.1) y calculando la relación entre la salida y la entrada en este dominio:

$$H(z) = \frac{S(z)}{U(z)} = G \frac{1 + \sum_{l=1}^q b_l z^{-l}}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (3.2)$$

La transformada \mathbf{Z} de una señal de tiempo discreto está definida como la serie de potencias [13]:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} \quad (3.3)$$

donde \mathbf{z} es una variable compleja.

La ecuación (3.2) representa un modelo *polo-cero*. Sin embargo, la información más importante se encuentra en las bandas espectrales en que se concentra la mayor cantidad de energía y estas están representadas por los polos de $H(z)$. En teoría de filtros [13] existen las siguientes posibilidades:

Si $b_l = 0$, para $1 \leq l \leq q$, $H(z)$ se reduce a un modelo *todo-polo* que se conoce como modelo autoregresivo.

Si $a_k = 0$, para $1 \leq k \leq p$, $H(z)$ se reduce a un modelo *todo-cero* o de promedio móvil.

Para resolver un modelo *polo-cero* es necesario dar solución a un conjunto de ecuaciones no lineales, mientras que la solución de un modelo *autoregresivo* requiere resolver un conjunto de ecuaciones lineales, facilitando su cálculo.

La función de transferencia para un modelo *todo-polo* es [13]:

$$H(z) = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (3.4)$$

y su transformada \mathbf{Z} inversa es:

$$\tilde{s}(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n) \quad (3.5)$$

Si imponemos que el factor G de ganancia sea igual a 1, tenemos:

$$H(z) = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} = \frac{1}{A(z)} \quad (3.6)$$

donde el polinomio $(1 - \sum_{k=1}^p a_k z^{-k})$ está representado por $A(z)$. Los coeficientes a_k son llamados *coeficientes de predicción lineal* ó LPC

El error $e(n)$ asociado a la estimación de los coeficientes está dado por la diferencia entre la señal adquirida y la señal estimada mediante el filtro

$$e(n) = \tilde{s}(n) - \sum_{k=1}^p a_k s(n-k) \quad (3.7)$$

que en el dominio \mathbf{z} es equivalente a:

$$E(z) = S(z)A(z) \quad (3.8)$$

3.1.2. Estimación de los coeficientes de predicción lineal

Para estimar los LPC en este estudio se utiliza el método de *autocorrelación*. En este método se buscan los coeficientes a_k tales que la energía residual $e(n)$ sea mínima, esto se hace mediante la técnica de *mínimos cuadrados*.

3.1.3. Método de autocorrelación

Se pretende minimizar la energía de la señal residual que está definida como:

$$E = \sum_{n=-\infty}^{\infty} e^2(n) = \sum_{n=-\infty}^{\infty} \left(s_w(n) - \sum_{k=1}^p a_k s_w(n-k) \right)^2 \quad (3.9)$$

Donde s_w es la señal que se ha sometido a un proceso de *ventaneo*. Para encontrar los valores de que minimizan la función E es necesario encontrar las derivadas parciales $\frac{\partial E}{\partial a_k}$ para $k = 1, \dots, p$ e igualarlas a cero. Lo anterior da lugar a p ecuaciones con p variables por determinar, así:

$$\sum_{k=1}^p a_k \sum_{n=-\infty}^{\infty} s_w(n-i)s_w(n-k) = \sum_{n=-\infty}^{\infty} s_w(n-i)s_w(n), \quad 1 \leq i \leq p \quad (3.10)$$

Estas ecuaciones lineales pueden expresarse en términos de la función de *autocorrelación*. La *función de autocorrelación* del segmento ventaneado $s_w(n)$ está definida como:

$$R(i) = \sum_{n=i}^{N_w-1} s_w(n)s_w(n-i), \quad 0 \leq i \leq p. \quad (3.11)$$

Se sabe que la función de *autocorrelación* es una función par $R(i) = R(-i)$. Aplicando esta condición en la ecuación (3.11) tenemos:

$$\sum_{k=1}^p R(|i-k|) a_k = R(i), \quad 1 \leq i \leq p \quad (3.12)$$

Que en forma matricial se representa como:

$$\begin{bmatrix} R(0) & R(1) & \dots & R(p-1) \\ R(1) & R(0) & \dots & R(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(p-1) & R(p-2) & \dots & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix} \quad (3.13)$$

o que en notación matricial sería igual a:

$$\mathbf{R}\mathbf{a} = \mathbf{r} \quad (3.14)$$

La matriz \mathbf{R} es una matriz de *Toeplitz* en la que los elementos de la diagonal principal son iguales. Esta característica permite que las ecuaciones puedan resolverse a través del algoritmo de *Levinson-Durbin*.

El algoritmo de *Levinson-Durbin* resuelve sistemas de ecuaciones de la forma $\mathbf{Ax}=\mathbf{b}$, donde es una matriz de *Toeplitz*, simétrica y positiva y \mathbf{b} es un vector cualquiera. *Durbin* encontró un algoritmo más eficiente para resolver el sistema en el caso en que \mathbf{b} tenga algunos elementos de \mathbf{A} y las ecuaciones de *autocorrelación* también satisfagan esta condición.

Sea $a_k(m)$ el k -ésimo coeficiente para un segmento en particular en la m -ésima iteración. El algoritmo de *Levinson-Durbin* resuelve el siguiente conjunto de ecuaciones ordenadas recursivamente por $m = 1, 2, \dots, p$:

$$k(m) = R(m) - \sum_{k=1}^{m-1} a_k(m-1)R(m-k) \quad (3.15)$$

$$a_m(m) = k(m) \quad (3.16)$$

$$a_k(m) = a_k(m-1) - k(m)a_{m-k}(m-1), \quad 1 \leq k \leq m \quad (3.17)$$

$$E(m) = (1 - k(m)^2)E(m-1) \quad (3.18)$$

donde inicialmente $E(0) = R(0)$ y $a(0) = 0$. En cada iteración el m -ésimo coeficiente $a_k(m)$ para $k = 1, 2, \dots, m$ describe el m -ésimo predictor lineal óptimo y el error $E(m)$ se reduce en un factor de $(1 - k(m)^2)$ [6].

3.2. Redes Neuronales Artificiales

Las *redes neuronales artificiales* (en adelante RNA) nacen como el producto de un intento por emular ciertas características del cerebro humano desde el punto de vista sensorial. Las neuronas, procesadores elementales y relativamente lentos se agrupan formando capas y estas a su vez forman redes que funcionan como un todo que nosotros conocemos como cerebro. El estilo de procesamiento que el cerebro utiliza es en paralelo, esto es, varias neuronas pueden activarse al mismo tiempo sin que el sistema colapse por ello. El conocimiento está ampliamente distribuido en la red lo que evita grandes pérdidas de información en caso de daños parciales. El cerebro humano tiene la capacidad de adaptarse a las condiciones cambiantes del medio exterior, aprendiendo nuevas cosas y aplicándolas sobre ese nuevo entorno. En el cerebro existe una gran tolerancia a los fallos ya que si una neurona deja de funcionar no es motivo para que el sistema completo interrumpa su funcionamiento. El tipo de control de los procesos es *auto-organizado*, lo que implica que existen grupos de neuronas especializadas

en tipos específicos de procesos.

En la descripción anterior, de las características del cerebro que se pretende emular mediante RNA se encuentran las siguientes: procesamiento en paralelo, almacenamiento distribuido, adaptabilidad, tolerancia a fallos y control de proceso *auto-organizado*.

Las RNA se han convertido en una alternativa para solucionar problemas con las siguientes características

- Falta de un conjunto de reglas sistemáticas que describan completamente el problema.
- Disponibilidad de muchos ejemplos o casos históricos.
- Datos imprecisos o ruidosos.
- Elevada *dimensionalidad*.
- Condiciones de trabajo cambiantes.

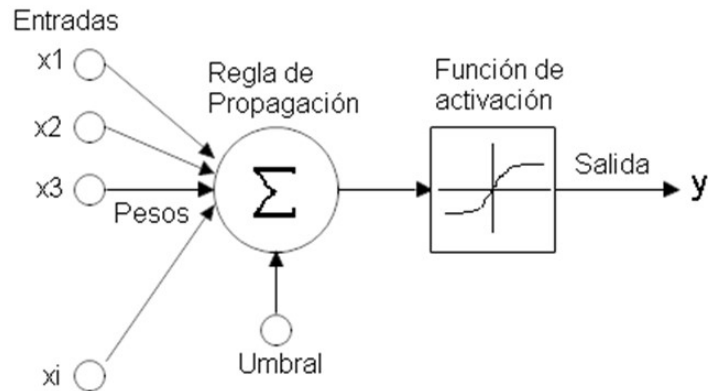
3.2.1. La Neurona Artificial

Una neurona artificial es una unidad elemental de proceso que realiza cálculos a partir de un vector de entrada procedente del exterior o de otras neuronas. La neurona artificial es la base estructural de las redes neuronales artificiales.

En una neurona pueden distinguirse los siguientes componentes (Figura. 3-1):

- Conjunto de **entradas** x_j
- **Pesos sinápticos** de la neurona i , w_{ij} que representan la fuerza de la interacción entre cada neurona *presináptica* j (anterior a los pesos sinápticos) y la neurona *postsináptica* i (posterior a los pesos sinápticos).
- **Regla de propagación** que proporciona el valor del potencial *postsináptico* de la neurona i en función de sus pesos sinápticos y entradas.
- **Función de activación** que proporciona el estado de activación actual de la neurona i en función de su estado anterior y de su potencial *postsináptico*.
- **Función de salida** que proporciona la salida actual de la neurona i en función de su estado de activación.

Figura 3.2.1: La neurona artificial.



Fuente: (Cadena, 2011) [1]

Frecuentemente se añade al conjunto de pesos sinápticos de la neurona un parámetro adicional que se denomina umbral Θ_i , que reduce el valor de entrada a la *función de activación*.

Entradas y salidas

La forma de las entradas y salidas en las redes neuronales puede variar dependiendo del tipo de aplicación. Cuando se quiere resolver un problema de clasificación es conveniente utilizar como entradas un vector que contiene información acerca del patrón que se desea discriminar y como salidas vectores de baja *dimensionalidad* cuyos elementos suelen estar en los rangos $[0, +1]$ o $[-1, +1]$. Los vectores de entrada pueden también normalizarse teniendo en cuenta los conjuntos completos de los diferentes tipos de patrones a clasificar.

Pesos sinápticos

Los pesos sinápticos condicionan el poder de la interacción entre la neurona presináptica j y la neurona postsináptica i . Si la entrada es positiva el peso sináptico que interactúa con dicha entrada producirá efectos que realzarán la excitación (peso positivo) o inhibición (peso negativo) de la neurona dependiendo de si el peso es positivo o negativo respectivamente.

Regla de propagación

La regla de propagación determina el potencial postsináptico a partir del vector de entrada y los pesos sinápticos, generalmente se utiliza el producto escalar entre el vector de entrada y

el de pesos sinápticos:

$$h_i(t) = \sum_j w_{ij}x_j \quad (3.19)$$

Función de activación

De esta función depende el estado de activación de la neurona, su argumento es el potencial postsináptico. Las funciones utilizadas suelen ser continuas y monótonas crecientes; en problemas de clasificación los algoritmos utilizados para el aprendizaje de la red exigen también que la función de activación sea derivable. El estado de activación actual de la red sería entonces:

$$a_i(t) = f_i(h_i(t)) \quad (3.20)$$

Función de salida

Esta función es la que presenta el estado global de la neurona, depende del estado de activación; lo más común es utilizar este estado como la propia salida de la neurona, por tanto:

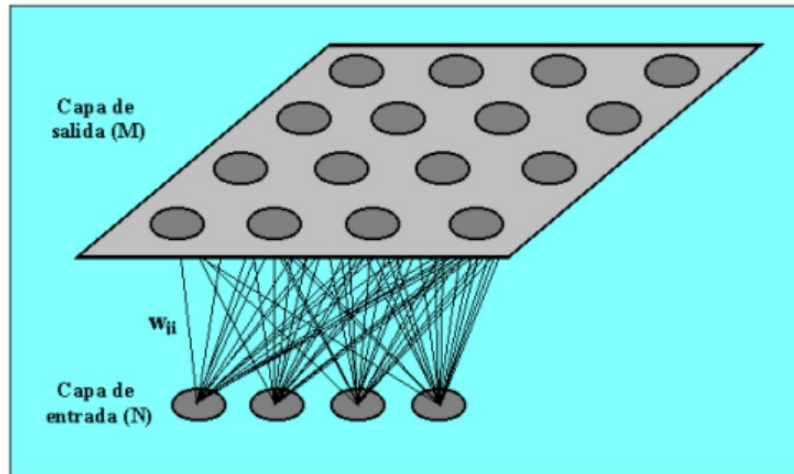
$$y_i(t) = F_i(a_i(t)) = a_i(t) \quad (3.21)$$

3.2.2. Arquitectura de las Redes Neuronales

Se denomina *arquitectura* de una red neuronal a la estructura que conecta las neuronas entre sí para formar una red neuronal. Las arquitecturas de red son de diversos tipos, unas son más apropiadas que otras para la solución de determinados problemas. La arquitectura de red que se escoja depende entre, otras cosas, de las entradas que se usen para representar los patrones y de la manera como se desee obtener las salidas. En nuestro caso, en que se necesita clasificar entre conjuntos con diferentes características, se suele utilizar la arquitectura del SOM (Self-Organizing Maps)[14]

Un modelo SOM está compuesto por dos capas de neuronas. La capa de entrada (formada por N neuronas, una por cada variable de entrada) se encarga de recibir y transmitir a la capa de salida la información procedente del exterior. La capa de salida (formada por M neuronas) es la encargada de procesar la información y formar el mapa de rasgos. Normalmente, las neuronas de la capa de salida se organizan en forma de mapa bidimensional (figura3.2.2).

Figura 3.2.2: Estructura de un modelo SOM.



Fuente: (Universidad Carlos III de Madrid, 2019) [8]

Las conexiones entre las dos capas son siempre hacia adelante, es decir, la información siempre se propaga desde la capa de entrada hacia la capa de salida. Cada neurona de entrada i está conectada con una de las neuronas j mediante el peso w_{ij} . De esta forma, las neuronas de salida tienen asociado un vector de pesos W_j llamado vector de referencia, debido a que constituye el vector prototipo de la categoría representada por la neurona de salida j . Así, el SOM define una proyección desde un espacio de datos en alta dimensión a un mapa bidimensional de neuronas. Entre las neuronas de la capa de salida, puede decirse que existen conexiones laterales de excitación o inhibición implícitas, pues aunque no estén conectadas, cada una de estas neuronas va a tener influencia sobre sus vecinas. Esto se consigue a través de un proceso de competición entre las neuronas y de la aplicación de una función denominada de vecindad, que produce la topología o estructura del mapa. Las topologías más frecuentes son la rectangular y la hexagonal.

3.2.3. Operación de las redes neuronales artificiales

Existen dos modos en los que operan las RNA, aprendizaje o entrenamiento y recuerdo o ejecución. En la fase de entrenamiento la RNA aprenderá una tarea deseada hasta que el entrenador (quien opera la red) considere que la tarea puede ser ejecutada sin la necesidad de ser supervisada, es entonces cuando se pone en marcha la fase de ejecución en que la RNA entrenada aplica los conocimientos adquiridos para realizar la labor para la que fue creada.

Fase de aprendizaje

A nivel operativo la fase de aprendizaje o entrenamiento consiste en ajustar los pesos sinápticos w_{ij} de la red a través de la aplicación de un algoritmo que minimiza una función de error hasta satisfacer un criterio establecido anticipadamente. Este algoritmo llamado también regla de aprendizaje es el encargado de tomar las señales del exterior o ejemplos y en función de estos ajustar en un valor Δw_{ij} el estado actual de los pesos sinápticos:

$$w_{ij}(t + 1) = w_{ij}(t) + \Delta w_{ij}(t) \quad (3.22)$$

donde $w_{ij}(t)$ representa el valor actual del peso sináptico que une la neurona presináptica j con la postsináptica i y Δw_{ij} representa el ajuste. Este proceso se realiza de manera iterativa hasta que se considere que la red ha alcanzado cierto nivel de aprendizaje. En este punto del proceso la RNA tiene una percepción del entorno que utilizará para realizar la tarea que le fue asignada. De los tipos de aprendizaje que existen los más utilizados son el supervisado y el no supervisado.

Aprendizaje supervisado

En este tipo de aprendizaje el entrenador le presenta a la red patrones de un tipo junto con una etiqueta del conjunto al que pertenece el patrón. Por ejemplo, se presenta un conjunto de vectores como entrada y un objetivo común como salida; este último suele ser un vector de menor dimensión con el cual se desea asociar las entradas. El proceso de presentación de los patrones es iterativo y la regla de aprendizaje ajusta los pesos sinápticos de la red en cada iteración hasta encontrar una relación entre las entradas y las salidas.

Aprendizaje no supervisado

Este tipo de aprendizaje se basa en el cálculo de la densidad de probabilidad $P(x)$ dada por la distribución de los patrones de entrada x . La regla de aprendizaje busca características similares entre los ejemplos que se le presentan y los agrupa en conjuntos dependiendo de sus características. La escogencia de una arquitectura de red y una regla de aprendizaje adecuados evita problemas de convergencia en el algoritmo al momento de minimizar la función de error.

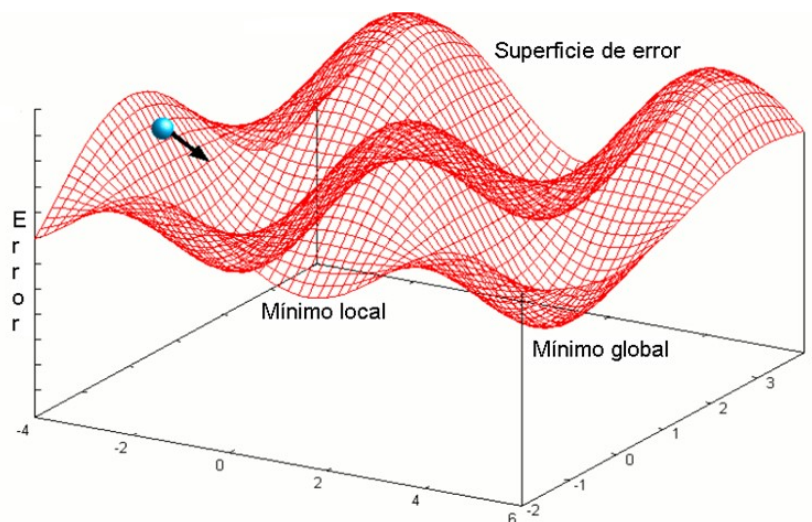
Retropropagación del error

Atendiendo al tipo de problema a resolver existe una variedad de algoritmos de aprendizaje, entre los más utilizados están: el aprendizaje Jeddiano, la regla LMS (Least Mean Squares), la regla de aprendizaje de Boltzman, el de corrección de errores que incluye la Adalina y el de retropropagación del error.

El algoritmo por retropropagación del error o *BP* por sus siglas en inglés (Backpropagation) está basado en la regla *Widrow-Hoff* o LMS que conduce a actualizaciones de los pesos sinápticos proporcionales al error de la red neuronal.

Sea E una función de error $E = E[W], E = R^n \rightarrow R$, función de los pesos sinápticos que mide la eficacia de la red. Se requiere encontrar un proceso que optimice la función de error buscando un mínimo. Para ello se utiliza el método de descenso por el gradiente que se ilustra en la Figura (3.2.3). Si la función utilizada dependiera solamente de dos variables (en problemas reales su dimensión es mucho mayor) su gráfica sería semejante a una sábana plegada con crestas y valles; estar sobre uno de sus valles corresponde a estar ubicado en un mínimo local óptimo. El objetivo, que no siempre es posible de lograr, es encontrar un mínimo global; sin embargo un buen mínimo local muy frecuentemente es suficiente para abordar el problema. Este proceso se lleva a cabo iterativamente con ayuda de los patrones de aprendizaje y los objetivos de cada conjunto de entrenamiento.

Figura 3.2.3: Descenso del gradiente para una función error.



Fuente: (Cadena, 2011) [1]

Figura 3.2.3: Esquema de descenso por el gradiente para una función de error. La esfera azul indica el punto donde se evalúa el error y la flecha la dirección hacia donde se debe mover la esfera para alcanzar un mínimo de la función.

En el método de descenso por el gradiente los pesos sinápticos para $t = 0$, $W(0)$ (inicializados con valores aleatorios pequeños) se ajustan una vez se ha calculado el gradiente de la función de error $E(W)$. El gradiente proporciona la dirección de la máxima variación de la función, por tanto el ajuste se hace en sentido contrario a esta dirección y en un valor adecuado, no tan grande de tal manera que no se salten características importantes de esa superficie tales como mínimos locales cercanos, ni tan pequeño que el tiempo de computo sea excesivamente largo. Cada una de estas actualizaciones tendría la forma:

$$W(t + 1) = W(t) - \epsilon \cdot \nabla E(w) \quad (3.23)$$

donde ϵ indica el tamaño del paso tomado en cada iteración. Para la iteración la variación en la función de error $E(W)$ es:

$$\delta(E(w_{ij})) = \sum_{ij} \frac{\delta E(w_{ij})}{\delta w_{ij}} \delta w_{ij} \quad (3.24)$$

donde los w_{ij} son los elementos de la matriz de pesos sinápticos. Al multiplicar por el gradiente obtenemos:

$$\delta(f(w_{ij})) = \sum_{ij} \frac{\delta E(w_{ij})}{\delta w_{ij}} \left(-\epsilon \frac{\delta E(w_{ij})}{\delta w_{ij}} \right) = -\epsilon \sum_{ij} \left(\frac{\delta E(w_{ij})}{\delta w_{ij}} \right)^2 \leq 0 \quad (3.25)$$

Así se demuestra que la actualización es siempre menor que cero, asegurando que a través de este método se encontrarán mínimos ya sean locales o globales.

En este método la forma como se alcanza un mínimo es asintótica, ya que con el transcurrir de las iteraciones las variaciones son cada vez más pequeñas dado que son proporcionales al error asociado al paso por la red.

La desventaja de la regla LMS reside en que la función de error debe estar en función únicamente de los pesos sinápticos de la capa de salida. En el caso de una arquitectura neuronal con capas ocultas es necesario generalizar la regla, esto se logra con el algoritmo *BP*. Debido a que en el proceso de ajuste de los pesos sinápticos es necesario derivar las funciones según la regla de la cadena se exige que las funciones de transferencia de las neuronas sean derivables.

Fase de recuerdo

Una vez ha concluido la fase de aprendizaje, la RNA está lista para responder a estímulos con los que nunca ha trabajado, pero que tienen características comunes a los conjuntos con los que se entrenó. Inicia entonces la fase de recuerdo, que depende de la manera en que están memorizadas las características de los patrones con que se entrenó a la red.

Memoria asociativa

La simulación de una red neuronal biológica conlleva al desarrollo de una memoria asociativa. La principal propiedad de este tipo de memoria es su habilidad para relacionar (asociar) un patrón de salida, como consecuencia de cierta actividad neuronal, con un patrón de entrada.

Durante la fase de aprendizaje se presenta a la red un estímulo de entrada o patrón de entrada y la memoria lo transforma en uno memorizado, mientras que en la fase de recuerdo se presenta a la red una versión ruidosa del patrón presentado anteriormente y la memoria se encarga de asociarlo con el memorizado.

La memoria asociativa puede ser lineal o no lineal, esto depende del tipo de neuronas que se utilicen en la arquitectura de red. Sean \mathbf{a} y \mathbf{b} los vectores de entrada y salida respectivamente. La memoria asociativa es de tipo lineal cuando la relación entre estos vectores es de la forma:

$$\mathbf{b} = \mathbf{M}\mathbf{a} \quad (3.26)$$

donde \mathbf{M} es la matriz de memoria. Si la relación entre \mathbf{b} y \mathbf{a} tiene la forma:

$$\mathbf{b} = \varphi(\mathbf{M}, \mathbf{a})\mathbf{a} \quad (3.27)$$

donde φ es una función que depende de la matriz de memoria \mathbf{M} y del vector de entradas \mathbf{a} , entonces se dice que la memoria es de tipo no lineal. Supongamos que como respuesta a un vector de estímulo de entrada \mathbf{a}_k se obtiene simultáneamente un vector \mathbf{b}_k donde el subíndice k denota el número de patrón o vector que se asociará. Para nuestro problema contaremos con q patrones:

$$\mathbf{a}_k = [a_{k1}, a_{k2}, \dots, a_{kp}]^T \quad (3.28)$$

$$\mathbf{b}_k = [b_{k1}, b_{k2}, \dots, b_{kp}]^T \quad (3.29)$$

Por simplicidad convenimos que la dimensión de los vectores de entrada y salida es igual, digamos p . Si asumimos un tipo de memoria lineal podemos escribir:

$$\mathbf{b}_k = \mathbf{W}(k)\mathbf{a}_k, \quad k = 1, 2, \dots, q \quad (3.30)$$

donde $W(k)$ es la matriz de pesos sinápticos. La salida b_{ki} de la neurona i debida a la acción combinada de los elementos del conjunto de patrones \mathbf{a}_k aplicados como estímulo en la capa de entrada de la red es:

$$b_{ki} = \sum_{j=1}^p w_{ij}(k)a_{ij}, \quad i = 1, 2, \dots, q \quad ; \quad j = 1, 2, \dots, p \quad (3.31)$$

Los w_{ij} son los pesos sinápticos de la neurona i correspondientes al k –esimo par de patrones asociados. La ecuación (3.31) en notación matricial es:

$$b_{ki} = \begin{bmatrix} w_{i1}(k) & w_{i2}(k) & \dots & w_{ip}(k) \end{bmatrix} \begin{bmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kp} \end{bmatrix} \quad i = 1, 2, \dots, q \quad (3.32)$$

que se puede escribir como:

$$\begin{bmatrix} b_{k1} \\ b_{k2} \\ \cdot \\ \cdot \\ \cdot \\ b_{kp} \end{bmatrix} = \begin{bmatrix} w_{11}(k) & w_{12}(k) & \dots & w_{1p}(k) \\ w_{21}(k) & w_{22}(k) & \dots & w_{2p}(k) \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ w_{p1}(k) & w_{p2}(k) & \dots & w_{pp}(k) \end{bmatrix} \begin{bmatrix} a_{k1} \\ a_{k2} \\ \cdot \\ \cdot \\ \cdot \\ a_{kp} \end{bmatrix} \quad (3.33)$$

expandiendo las matrices de pesos sinápticos:

$$\mathbf{W}(k) = \begin{bmatrix} w_{11}(k) & w_{12}(k) & \dots & w_{1p}(k) \\ w_{21}(k) & w_{22}(k) & \dots & w_{2p}(k) \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ w_{p1}(k) & w_{p2}(k) & \dots & w_{pp}(k) \end{bmatrix} \quad (3.34)$$

La asociación de los pares individuales produce matrices del tipo $\mathbf{W}(k)$. De tal manera que se puede definir una matriz que contenga la suma de las matrices de pesos sinápticos que asocian cada par. Esta matriz de memoria tendrá la forma:

$$\mathbf{M} = \sum_{k=1}^P \mathbf{W}(k) \quad (3.35)$$

que define la conectividad total entre las capas de entrada y de salida de la memoria asociativa.

La ecuación (3.35) puede escribirse en forma recursiva, para hacer más evidente la forma como se construye la matriz de memoria asociativa \mathbf{M} acudiendo a las actualizaciones de la matriz a medida que los pares de entrada y salida entran en juego. Así tenemos:

$$\mathbf{M}_k = \mathbf{M}_{k-1} + \mathbf{W}(k), \quad k = 1, 2, \dots, q \quad (3.36)$$

donde \mathbf{M}_k es la matriz $\mathbf{W}(k)$ ajustada por la matriz que contiene la información la nueva (k - esima) asociación.

Consideremos ahora un estimado de la matriz \mathbf{M} que denotaremos por $\hat{\mathbf{M}}$ como sigue:

$$\hat{\mathbf{M}} = \sum_{k=1}^q \mathbf{b}_k \mathbf{a}_k^T \quad (3.37)$$

donde el producto $\mathbf{b}_k \mathbf{a}_k^T$ es un estimado de la matriz. En forma matricial $\hat{\mathbf{M}}$ puede escribirse como:

$$\hat{\mathbf{M}} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_q \end{bmatrix} \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_q^T \end{bmatrix} \quad (3.38)$$

finalmente podemos expresar la ecuación (3.38) en forma recursiva:

$$\hat{\mathbf{M}}_k = \hat{\mathbf{M}}_{k-1} + \mathbf{b}_k \mathbf{a}_k^T \quad k = 1, 2, \dots, q \quad (3.39)$$

que es una ecuación análoga a (3.36), pero con la diferencia de que en este caso tenemos un estimado de la matriz \mathbf{M} .

Entonces, el problema en la fase de recuerdo es solucionar:

$$\mathbf{b} = \hat{\mathbf{M}} \mathbf{a}_j \quad (3.40)$$

para dar respuesta a un patrón nuevo, que no fue utilizado por el conjunto de ejemplos de entrenamiento a partir del que se estimó dicha matriz.

3.2.4. Capacidad de generalización

Si el problema a solucionar fuese el de verificar que el patrón desconocido es exactamente uno de los presentados en la fase de entrenamiento, sería mejor utilizar una base de datos que contuviera todos los vectores con los que se habría aplicado el entrenamiento y un algoritmo que comparara elemento a elemento los vectores memorizados con el que se desea dar respuesta. Una de las ventajas de las RNA es que no necesita memorizar uno a uno los patrones de entrenamiento sino el de reconocer características claves en ellos y extender ese conocimiento a estímulos desconocidos, esto es lo que se conoce como generalización. Para detener el entrenamiento es necesario medir el nivel de generalización de la red en función de su nivel de aprendizaje.

3.2.5. Validación cruzada

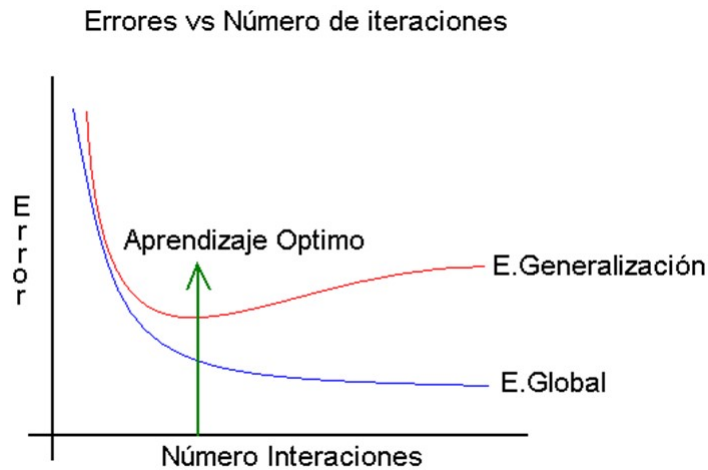
Durante el proceso de entrenamiento aparecen dos tipos de error:

1. El error de ajuste de cada patrón a su respectivo objetivo, llamado error global, que se calcula como el error cuadrático medio del conjunto de patrones y objetivos (para el caso del aprendizaje supervisado).
2. El error de respuesta a patrones desconocidos, llamado error de generalización.

Se considera que una RNA ha tenido una fase de aprendizaje exitosa cuando extiende la aplicación de lo aprendido a patrones desconocidos, esto se puede evaluar realizando pruebas de la respuesta de la red a ejemplos de entrada no conocidos. Es preferible seleccionar una RNA que muestre un error de generalización aceptable a una que arroje errores globales pequeños, ya que estos últimos pueden reducirse tanto como se quiera con solo incrementar el número de iteraciones. Un hecho observado en experimentos con simuladores de RNA es que cuando se alcanza un error global bastante pequeño, digamos menor que 19 %, la generalización de la red se degrada.

En la Figura (3.2.4) se observa como el error global disminuye asintóticamente a medida que el número de iteraciones crece, mientras que el error de generalización alcanza un mínimo. Detener el proceso de entrenamiento o guardar la configuración de pesos sinápticos y umbrales en ese instante en que el aprendizaje es óptimo, nos asegura que la configuración de la matriz de memoria es la más adecuada para utilizarla en la fase de recuerdo, esta técnica se conoce como validación cruzada [1].

Figura 3.2.4: Errores global y de generalización en un problema ideal, la flecha verde indica el momento en que el aprendizaje es óptimo.



Fuente: (Cadena, 2011) [1]

3.3. Técnica Clustering

La técnica “clustering” generalmente se utiliza en el análisis de datos estadísticos, éste método se utiliza en muchos campos, incluido el aprendizaje automático, la minería de datos, el reconocimiento de patrones, el análisis de imágenes entre otros. La agrupación en clúster es el proceso de agrupar objetos similares en diferentes grupos, o más precisamente, la partición de un conjunto de datos en subconjuntos, de modo que los datos en cada subconjunto de acuerdo con una medida de distancia definida, El general, ésta técnica resuelve el siguiente problema: A partir de un conjunto de datos de entrada (de N elementos) caracterizados por la información de n variables X_j , donde ($j = 1, 2, \dots, n$), la red neuronal debe ser capaz de clasificarlos de manera que los datos pertenecientes a un grupo (cluster) (y siempre con respecto a la información disponible) sean tan similares entre sí como sea posible, y los distintos grupos entre ellos difieran en gran medida.

Básicamente, el análisis constará de un algoritmo de clasificación que nos permitirá la obtención de una o varias particiones, de acuerdo con los criterios establecidos. En esta metodología se involucran redes neuronales de aprendizaje no supervisado, es decir, no requieren influencia externa para ajustar los pesos de las conexiones entre sus neuronas, dado que no se recibe ninguna información por parte del entorno, esta redes son capaces de autoorganizarse.

3.3.1. Redes autoorganizables

Las redes neuronales auto-organizables ó de *Kohonen*, son diseñadas basadas en el principio organizacional de las rutas sensoriales del cerebro. Según este principio, el posicionamiento de las neuronas es ordenado y refleja algunas características físicas de los estímulos externos que son sensados. En éstas redes neuronales con aprendizaje no supervisado se destacan dos tipos:

Aprendizaje asociativo: Este tipo de aprendizaje mide la familiaridad o extrae características de los datos de entrada. De acuerdo con el estudio psicológico del aprendizaje asociativo, *Donald Hebb* propuso la Regla de *Hebb* [3]: “*Cuando un axón de una celda A está lo suficientemente cerca de otra celda B como para excitarla y repetidamente ocasiona su activación, un cambio metabólico se presenta en una o ambas celdas, tal que la eficiencia de A como celda excitadora de B, se incrementa*”. Para *Hebb*, una celda es una red neuronal con conexiones de estructura compleja y la eficiencia es el peso de la conexión, mientras que la actividad coincidente es crucial para fortalecer las conexiones. Este postulado marcó el inicio del aprendizaje no supervisado.

Aprendizaje competitivo: En este tipo de aprendizaje las neuronas compiten para activarse quedando finalmente una, o una por grupo, como neurona vencedora; el resto quedarán anuladas y siendo forzadas a sus valores de respuesta mínimos. La competencia se realiza en todas las capas de la red, existiendo conexiones de autoexcitación y conexiones de inhibición. El objetivo de la competencia es categorizar o agrupar los datos de entrada, de tal manera que las entradas de la misma categoría activen la misma neurona de salida

3.3.2. Red de Kohonen

Existen evidencias que demuestran que en el cerebro hay neuronas organizadas en zonas, de tal forma que la información captada del entorno se representa internamente en forma de mapas bidimensionales. Estas teorías se sustentan en que algunas áreas del cerebro pueden crear y ordenar neuronas especializadas y construir mapas especiales para atributos y características.

El modelo propuesto por Tuevo Kohonen en los años ochenta tiene dos variantes:

- Learning Vector Quantization (LVQ)

- Topology Preserving Map (TPM) o Self Organizing Map (SOM)

Las dos variantes se basan en la formación de mapas topológicos para establecer características comunes entre la información de la entrada (vectores). En el caso del LVQ los mapas son de una dimensión, y en el SOM son bidimensionales o tridimensionales. El aprendizaje en las redes de *Kohonen* es de tipo *offline*, es decir existe una etapa de aprendizaje donde se fijan los pesos de las conexiones y otra de funcionamiento.

Un concepto muy importante es la zona de vecindad o vecindario alrededor de la neurona vencedora i^* , ya que los pesos de las neuronas que se encuentran en esa zona denominada $X(q)$ serán actualizados en conjunto con el de la vencedora (aprendizaje cooperativo). El algoritmo de aprendizaje es el siguiente:

1. Se inicializan los pesos con valores aleatorios pequeños y se fija la vecindad entre las neuronas de salida.
2. Se presenta la entrada p en forma de vector con valores continuos.
3. Se determina la neurona vencedora de la capa de salida, cuyos pesos de conexión con las entradas son los más parecidos a la entrada p . Para esto se calculan las distancias euclidianas entre los dos vectores, entradas y pesos:

$$d_i = \sum_{j=1}^N (p_j - w_{ij})^2, \quad 1 \leq i \leq M \quad (3.41)$$

4. Con la neurona vencedora se actualizan los pesos entre las entradas y ésta, y entre las vecinas y las entradas, según la siguiente expresión:

$$w(q) = w(q-1) + \alpha(q) (p(q) - w(q-1)), \quad \text{para } i \in X(q) \quad (3.42)$$

El tamaño de $X(q)$ puede reducirse en cada iteración y el término $\alpha(q)$ es el coeficiente de aprendizaje, con un valor entre 0 y 1 el cual decrece con el número de iteraciones q del proceso de entrenamiento. Para encontrar α se utilizan:

$$\alpha(q) = \frac{1}{q} \quad ; \quad \alpha(q) = \alpha_1 \left(1 - \frac{q}{\alpha_2} \right) \quad (3.43)$$

5. El proceso se repite con todos los patrones de aprendizaje hasta obtener la salida deseada. El aprendizaje ocurre cuando la neurona i sea miembro del conjunto $X(q)$.

Este tipo de redes es especialmente útil para establecer relaciones desconocidas previamente entre conjuntos de datos [11].

Mapas autoorganizados

Mapas de Autoorganización (SOM) Estas redes comprenden un patrón de conexión entre neuronas donde cada una se refuerza a sí misma (*on-center*) mientras inhibe a todas las neuronas a su alrededor (*off-surround*). Se determina primero la neurona ganadora i^* y luego los vectores de pesos de las neuronas que se encuentren en una región cercana denominada vecindario N_{i^*} , dentro de un radio d , serán actualizados mediante la Regla de *Kohonen*:

$$w_i(q) = w_i(q - 1) + \alpha(q) (p(q) - w_i(q - 1)), \text{ para } i \in N_{i^*}(d), N_{i^*}(d) = \{j, d_{ij} \leq d\} \quad (3.44)$$

3.3.3. El algoritmo SOM

El proceso de aprendizaje del SOM es el siguiente:

1. Un vector x es seleccionado al azar del conjunto de datos y se calcula su distancia (similitud) a los demás vectores, usando, por ejemplo, la distancia euclídea:

$$\|x - m_c\| = \min(j) \{\|x - m_j\|\} \quad (3.45)$$

2. Una vez que se ha encontrado el vector más próximo o BMU (*best matching unit*) el resto de vectores es actualizado. El BMU y sus vecinos (en sentido topológico) se mueven cerca del vector x en el espacio de datos. La magnitud de dicha atracción está regida por la *tasa de aprendizaje*. Mientras se va produciendo el proceso de actualización y nuevos vectores se asignan al mapa, la tasa de aprendizaje decrece gradualmente hacia cero. Junto con ella también decrece el radio de vecindad. La regla de actualización para el vector de referencia dado i es la siguiente:

$$m_j(t + 1) = \begin{cases} m_j(t) + \alpha(t)(x(t) - m_j(t)) & j \in N_c(t) \\ m_j(t) & j \notin N_c(t) \end{cases} \quad (3.46)$$

En cada iteración las ecuaciones (3.45) y (3.46) se van repitiendo hasta que el entrenamiento termina. El número de pasos de entrenamiento se debe fijar con anticipación para calcular la tasa de convergencia de la función de vecindad y de la tasa de aprendizaje.

Una vez terminado el entrenamiento, el mapa ha de ordenarse en sentido topológico: n vectores topológicamente próximos se aplican en n neuronas adyacentes o incluso en la misma neurona.

Precisión del mapa

Una vez que se ha entrenado el mapa, es importante saber si se ha adaptado adecuadamente a los datos de entrenamiento. Como medidas de calidad de los mapas se considera la precisión de la proyección y la preservación de la topología. La medida de precisión de la proyección describe cómo se adaptan o responden las neuronas a los datos. Habitualmente, el número de datos es mayor que el número de neuronas y el error de precisión es siempre diferente de 0. Para calcular la precisión de la proyección se usa el error medio de cuantificación sobre el conjunto completo de datos:

$$\varepsilon_q = \frac{1}{N} \sum_{i=1}^N \|x_i - m_c\| \quad (3.47)$$

La medida de preservación de la topología describe la manera en la que el SOM preserva la topología del conjunto de datos. Esta medida considera la estructura del mapa.

En un mapa que esté *retorcido* de manera extraña, el error topográfico es grande incluso si el error de precisión es pequeño.

Una manera simple de calcular el error topográfico es:

$$\varepsilon_t = \frac{1}{N} \sum_{k=1}^N u(x_k) \quad (3.48)$$

donde $u(x_k)$ es igual a 1 si el primer y segundo BMUs de x_k no están próximos el uno al otro. De otro modo, $u(x_k)$ es igual a 0.

Visualización del SOM

El SOM es fácil de visualizar y, además, en los últimos años se han desarrollado diferentes técnicas de visualización, tanto para los vectores de referencia como para los histogramas de datos. La proyección de *Sammon* representa el SOM de la manera gráfica que se muestra en la figura 3.3.1. Esta trata de encontrar una proyección no lineal óptima para los datos en alta

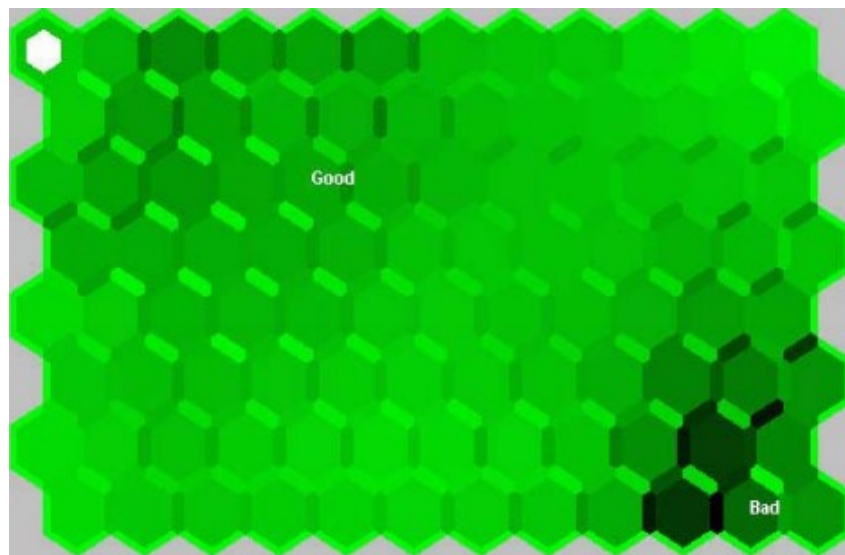
dimensión, de manera que los vectores que se proyectan en la superficie bidimensional, conservan la misma distancia euclídea relativa entre ellos que la que tenían en alta dimensión. La matriz unificada de distancias, o matriz U , es el método más popular para mostrar el SOM. Representa el mapa como una rejilla regular de neuronas, el tamaño y topología del mapa se puede observar en el gráfico donde cada elemento representa una neurona.

Cuando se genera la matriz U se calcula, a su vez, una matriz de distancias entre los vectores de referencia de neuronas adyacentes en el mapa bidimensional.

Después se selecciona algún tipo de representación gráfica, por ejemplo una escala de grises.

Los colores en la figura se seleccionan de modo que cuanto más oscuro es el color entre dos neuronas, menor es la distancia entre ellas [8].

Figura 3.3.1: Representación gráfica de un mapa SOM mediante la proyección de *Sammon*.



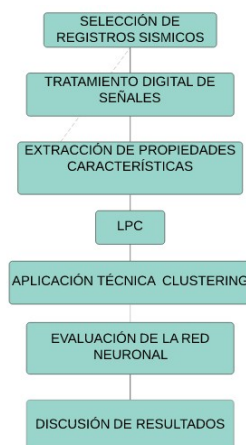
Fuente: (Universidad Carlos III de Madrid, 2019) [8]

Capítulo 4

Metodología

La estrategia implementada en el desarrollo de éste trabajo obedece a la siguiente descripción. Inicialmente se selecciona las trazas de los sismos a evaluar, luego se realiza rutinas computacionales para el tratamiento y pre-proceso de la información, se usa una representación alternativa de las señales con el fin de reducir la *dimensionalidad* de los vectores de entrada mediante LPC, posteriormente dicha información ya procesada se ingresada al sistema clasificador, el siguiente paso se enfoca principalmente en el estudio de los mapas auto organizados (éstos mapas se generan automáticamente por Matlab en la fase de entrenamiento), luego se etiqueta las regiones del mapa (o clúster) que se activan dependiendo de la información de entrada, y finalmente se evalúa la efectividad del sistema en la clasificación de nueva información.

Figura 4.0.1: Diagrama de flujo acerca del tratamiento y procesado de la información.



Fuente: (Ésta investigación)

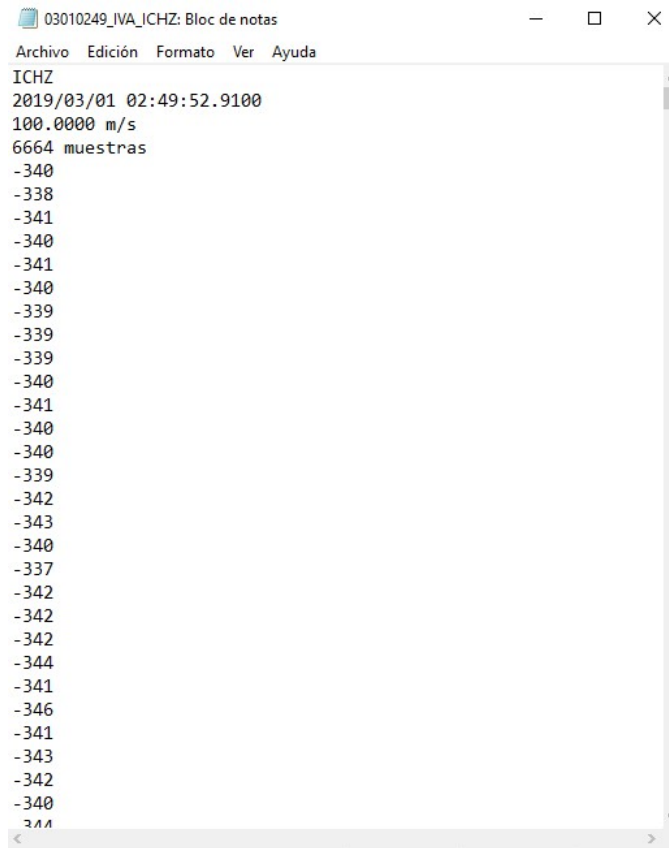
4.1. Selección de registros sísmicos

Los registros sísmicos fueron proporcionados por el SGC-OVSP de su base de datos, entre ellos se encuentran las formas de ondas de sismos del tipo VT, LP además de trazas que contienen segmentos de ruido y sismo, éstos archivos están comprendidos en periodos entre el 1 de enero de 2006 y el 31 de diciembre de 2009, los de tipo VT tienen magnitud local mayor o igual a 1 en la escala de Richter.

Estos registros en formato ASCII (*American Standard Code for Information Interchange*) provienen directamente de las estaciones de Anganoy y Cufiño (estaciones que forman parte de la red de monitoreo del VG)

En la figura (4.1.1) se muestra un ejemplo de un archivo en formato ASCII

Figura 4.1.1: Formato de un ejemplo de los archivos en formato ASCII que contiene las trazas de los eventos sísmicos.



```
03010249_IVA_ICHZ: Bloc de notas
Archivo Edición Formato Ver Ayuda
ICHZ
2019/03/01 02:49:52.9100
100.0000 m/s
6664 muestras
-340
-338
-341
-340
-341
-340
-339
-339
-339
-340
-341
-340
-340
-339
-342
-343
-340
-337
-342
-342
-342
-344
-341
-346
-341
-343
-342
-340
-344
```

Fuente: (Ésta investigación)

Los archivos tienen 4 renglones de cabecera, en la primera línea se indica el código de la estación, la segunda muestra la fecha en formato año/mes/día y hora del evento sísmico, la tercera es la tasa de muestreo, en la cuarta el número total de datos que contiene el evento y enseguida se encuentran los datos de las amplitudes que conforman el registro.

En la siguiente tabla se muestra la cantidad de registros que fueron utilizados en el desarrollo de éste trabajo, se menciona el tipo de señal y el número de ejemplos extraídos de cada una de ellas.

Tabla 4.1.1: Número de ejemplos extraídos.

Tipo de Señal	Número de ejemplos
Segmentos de sismo	169
Segmentos de ruido	1733
Sismos tipo LP	1046
Sismos tipo VT	713

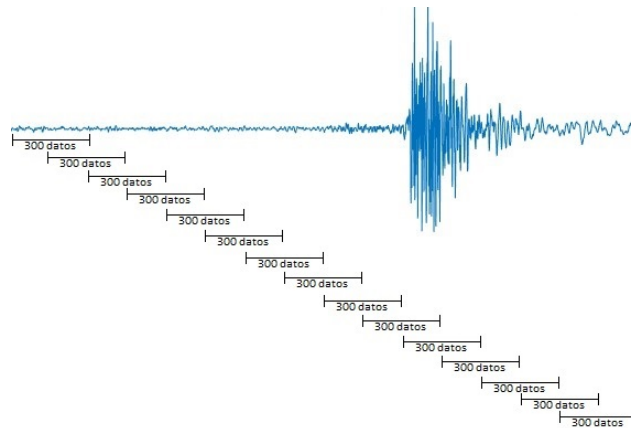
Fuente: (Ésta investigación)

4.1.1. Selección de las muestras para de segmentos de ruido y sismo

La selección de estas muestras se realiza a partir de trazas completas de diferentes eventos sísmicos, la manera para obtener los vectores que contienen información tanto de ruido y sismo es definiendo regiones específicas sobre cada traza que contengan las suficientes unidades para mencionadas muestras, una vez identificadas dichas regiones se procede a segmentarlas a manera vectorial, cada vector tendrá un número determinado de unidades que será igual para todos, además de un porcentaje de traslape entre cada vector y el siguiente, todo con el objetivo de obtener vectores con las mismas dimensiones.

La Figura 4.1.2 muestra la traza de un evento sísmico en la cual se seleccionaron segmentos vectoriales de 300 datos que contienen ya sea información de ruido ó sismo, cada vector tiene un traslape del 50 % con el adyacente, cabe mencionar que el número total de vectores crece a medida que el porcentaje de traslape entre cada vector aumenta, ésto es válido tanto para segmentos de ruido como de sismo.

Figura 4.1.2: Traza de un evento sísmico representado mediante segmentos vectoriales de 300 datos.



Fuente: (Ésta investigación)

4.1.2. Selección de las muestras para sismos tipo VT y LP

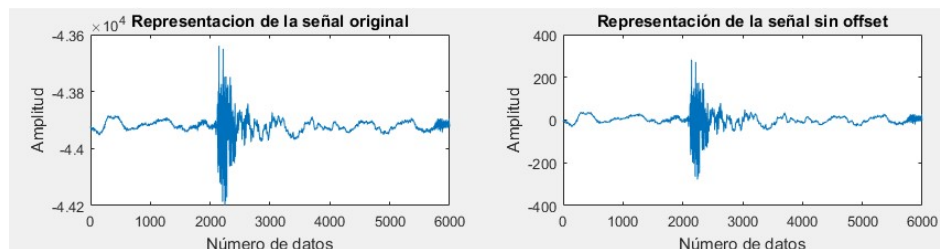
En este caso se seleccionaron trazas completas de sismos de tipo VT y de tipo LP, el número total de ejemplos extraídos se muestran en la Tabla 4.1.1

4.2. Tratamiento digital de señales

Para la fase de pre-proceso de la información se han diseñado e implementado rutinas computacionales convenientes para realizar las siguientes actividades:

- **Remoción del offset:** En ésta etapa la señal se desplaza al origen de un sistema coordenado.

Figura 4.2.1: Representación de una señal en la que se ha extraído el *offset*.

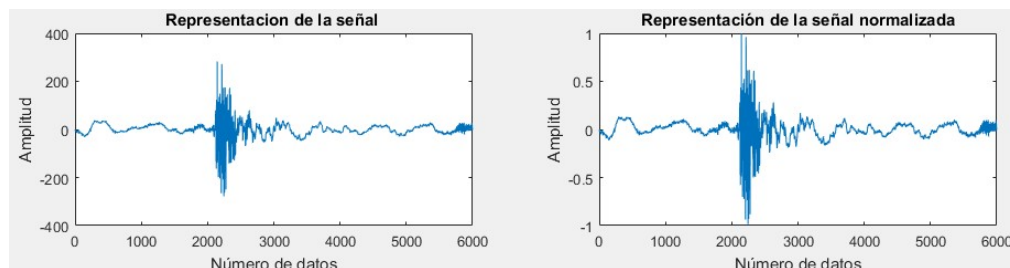


Fuente: (Ésta investigación)

En la figura 4.2.1 Se muestra la representación de la señal original con *offset* (lado izquierdo), y sin *offset* (lado derecho)

- **Normalización de las unidades:** La señal que es sometida a éste proceso, oscila sobre un intervalo comprendido entre $[-1, 1]$ ó $[0, 1]$.

Figura 4.2.2: Representación de una señal normalizada en sus unidades.

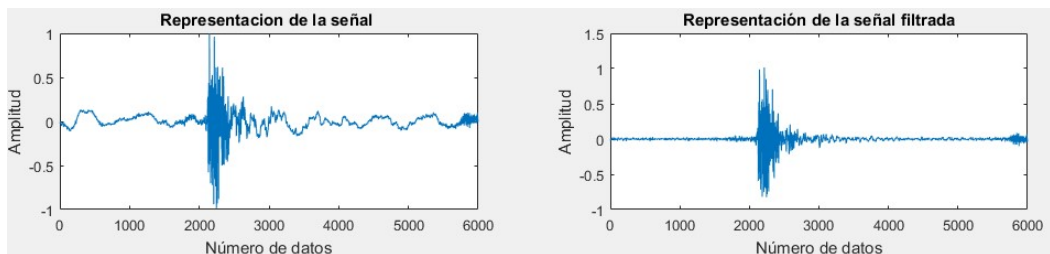


Fuente: (Ésta investigación)

La figura 4.2.2 muestra una señal original en sus unidades (lado izquierdo) y su representación, donde sus elementos se han normalizado (lado derecho).

- **Filtrado de señales:** Se eliminan señales con frecuencias (en éste caso inferiores a 1 hz) mediante un filtro digital, dado que que en éste rango de frecuencia, las señales no poseen información de nuestro interés.

Figura 4.2.3: Representación de una señal filtrada.



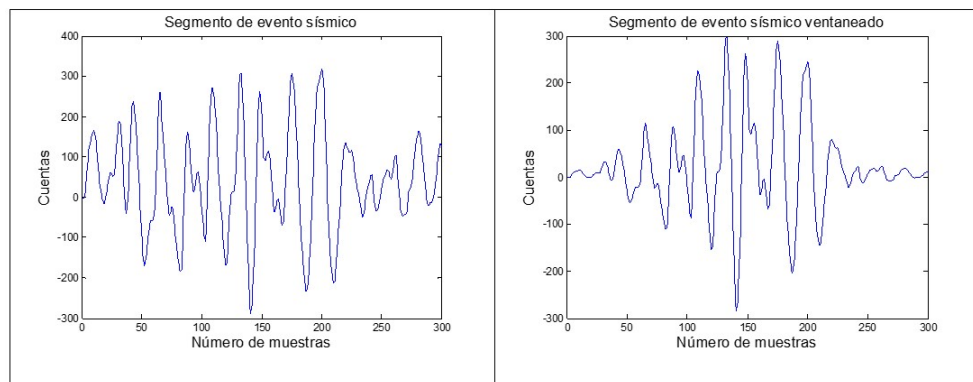
Fuente: (Ésta investigación)

La figura 4.2.3 representa una señal de entrada (lado izquierdo), y la misma señal sometida al filtro digital de “*Butterworth*” tipo “*Pasa Altas*” que permite el paso a frecuencias mayores a una frecuencia de corte fija, en este caso superiores a 1Hz (lado derecho).

- **Proceso de ventaneo:** Si deseamos representar una señal con la ayuda de una herramienta de procesamiento digital, es necesario hacerlo de manera discreta. De la elección de cada una de las ventanas variarían los resultados. En este trabajo se emplea un tipo de ventana suavizada en los extremos como la ventana de *Hamming* para evitar los efectos de borde.

En la figura (4.2.4) se muestra un tramo de evento sísmico sin ventanear y el mismo segmento ventaneado con la función de *Hamming*.

Figura 4.2.4: Segmento de sismo sin ventanear (izquierda), segmento de sismo ventaneado (Derecha).



Fuente: (Cadena, 2011) [1]

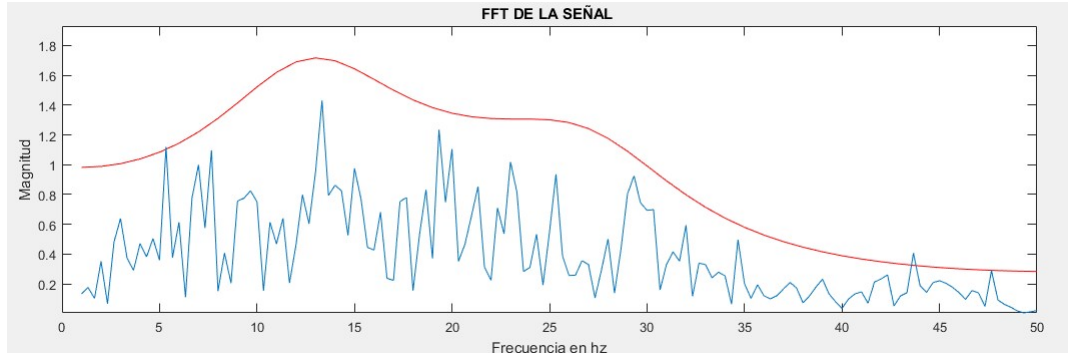
4.3. Extracción de propiedades características

En la búsqueda de encontrar propiedades características en las señales, se realiza el estudio basado en las bandas espectrales donde se concentra la mayor cantidad de energía, esto se debe a que la información más importante en este análisis se encuentra en éstos rangos.

Dado que las señales contienen gran cantidad de información, es conveniente su estudio en una representación más sencilla pero que guarde las características generales, esto conlleva a beneficios en el procesado de la información por parte del ordenador logrando optimizar el tiempo. Para lograr éste objetivo en concreto se procedió a dividir la señal de entrada en segmentos más cortos a manera de vectores, posteriormente se realiza la representación del segmento de señal en el dominio de la frecuencia y finalmente se calcula la envolvente para cada vector.

A continuación, se indica un ejemplo de lo descrito anteriormente.

Figura 4.3.1: Representación de la envolvente sobre el espectro en una señal.



Fuente: (Ésta investigación)

La figura 4.3.1 muestra la representación de la envolvente del fragmento de una señal en el dominio de la frecuencia. Este fragmento corresponde a un vector de 300 datos tomados de la señal original, y la envolvente está asociada al grado de predicción lineal que en este caso es 5.

4.4. Aplicación de los LPC

La representación de una señal por medio de los LPC se justifica ya que reduce la dimensionalidad de los vectores al sistema clasificador.

La representación de una señal en términos de sus LPC requiere determinar el número de orden adecuado de la predicción lineal sujeta a ciertas condiciones de nuestro interés, entre ellas están:

- El número de coeficientes no debe ser tan grande como para incluir en detalles de la señal que entorpecería la clasificación.
- El número de coeficientes no debe ser tan pequeño tal que esto involucre una pérdida de información relevante.

Una manera para determinar el orden de predicción más adecuado, es el análisis del error en la representación. Para ello se estudia el comportamiento de la varianza de cada unidad en función del orden de LPC, el objetivo es determinar el orden en el cual la varianza cambia relativamente poco, es decir entre un 5 % y 10 %

4.4.1. LPC en la representación óptima de señales

En el estudio para obtener la representación más adecuada de una señal por medio de sus LPC, es necesario establecer un orden óptimo en la predicción lineal.

Este análisis se realiza a partir de un número determinado de trazas de señales correspondientes a sismos del tipo LP y de tipo VT, se determinan los coeficientes de predicción lineal asociado a cada señal variando los órdenes en la predicción lineal desde un valor relativamente pequeño a uno más amplio, se calcula la varianza para cada traza, se mide el promedio de las varianzas asociadas a un determinado orden de LPC, se calcula la respectiva desviación estándar y se cuantifica la diferencia entre un orden y el siguiente en términos porcentuales.

4.5. Generación de las matrices de entrenamiento

Las matrices de entrenamiento se crean a partir de vectores que contiene la información espectral de los sismos, los vectores se organizan de forma intercalada entre los ejemplos que pertenecen a un grupo y otro. Adicional a las matrices de entrenamiento es necesario generar matrices de prueba, estas pueden crearse de las matrices de entrenamiento o a partir nueva información en la que hay certeza del grupo al que pertenece, mencionadas matrices se aplican en la fase de validación de las redes seleccionadas.

Cabe destacar que las dimensiones de los vectores que conforman cada matriz deben ser iguales, el tamaño de cada matriz es proporcional al número de ejemplos utilizados y al orden de predicción lineal.

4.6. Aplicación de la técnica clustering

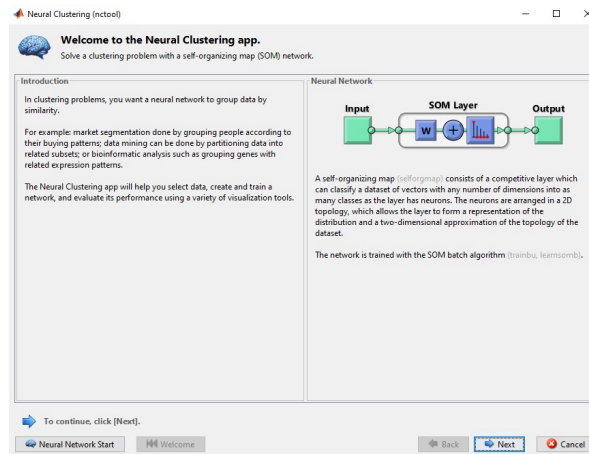
Esta es la técnica encargada de la clasificación de los datos y está basada en el uso en redes neuronales artificiales con aprendizaje no supervisado.

En este procedimiento, la información ya debe estar organizada a manera de arreglos matriciales para el posterior análisis, éstas matrices deben contener ya sea en sus filas o columnas los vectores creados a partir de los coeficientes de predicción asociados a cada ejemplo.

Adicionalmente es conveniente la generación de matrices de prueba como requisito en la fase de validación de resultados, éstas pueden ser creadas a partir de las matrices de entrenamiento. Para iniciar la aplicación que contiene esta metodología se realiza los siguiente pasos:

1. Se escribe en la ventana *Command Window* de MATLAB el comando “*nctool*”.

Figura 4.6.1: Ventana introductoria del método Clustering.

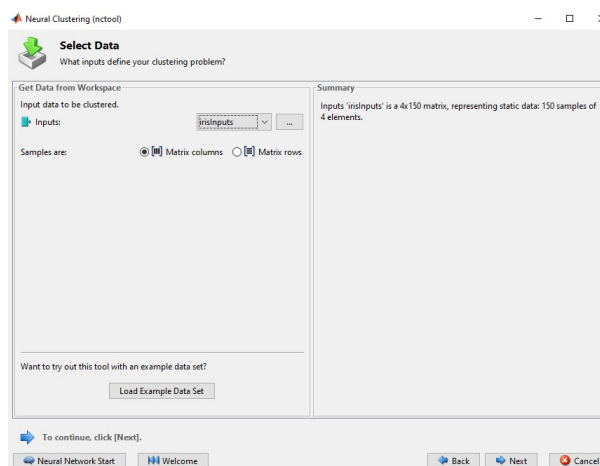


Fuente: Programa de prueba de MATLAB

En la figura 4.6.1 se muestra una guía introductoria acerca del método y la arquitectura que posee la red neuronal.

2. Al pulsar “*Next*” se despliega una nueva ventana donde se selecciona la matriz de entrenamiento, se debe tener en cuenta como se organizaron los ejemplos en la matriz (en filas o columnas).

Figura 4.6.2: Ventana de selección de los datos.



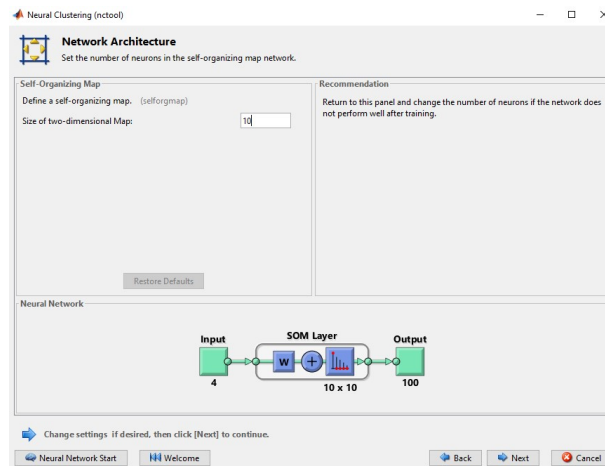
Fuente: Programa de prueba de MATLAB

En la figura 4.6.2 se muestra la ventana de selección de la información a manera matricial, en este caso se escogió un ejemplo de la base de datos de MATLAB, la matriz seleccionada consta de 150 ejemplos de 4 elementos organizados en columnas.

A continuación pulsamos “Next”

3. En ésta ventana se establece el número de neuronas de la red en el mapa autoorganizado (esto depende de la complejidad del problema).

Figura 4.6.3: Ventana donde se determina el número total de neuronas.



Fuente: Programa de prueba de MATLAB

La figura 4.6.3 nos permite determinar el número total de neuronas presente en un mapa bidimensional de tipo SOM.

Para problemas de agrupamiento, el mapa SOM es la red más utilizada, ya que después de que la red ha sido entrenada, hay muchas herramientas de visualización que se pueden usar para analizar los grupos resultantes.

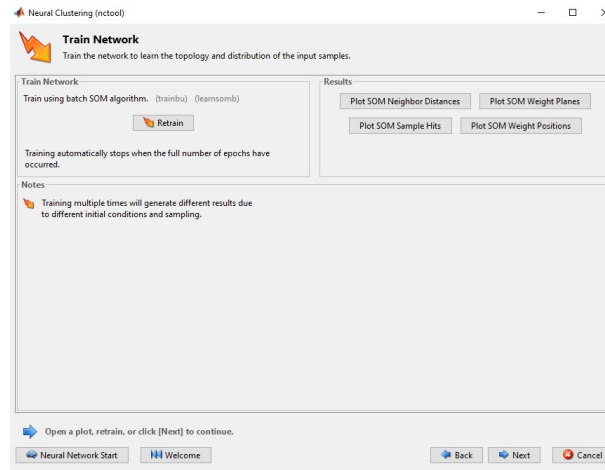
Esta red tiene una capa, con neuronas organizadas en una cuadrícula. Al crear la red, es necesario especificar el número de filas y columnas que estarán presentes en el SOM. Aquí, el número de filas y columnas está predeterminado en un valor de 10.

El número total en éste caso es de 100 neuronas. Se puede cambiar este número en otra ejecución si se desea.

Posteriormente pulsamos *Next*.

4. Ventana para el entrenamiento de la red neuronal usando el algoritmo SOM.

Figura 4.6.4: Ventana de entrenamiento de la red neuronal.



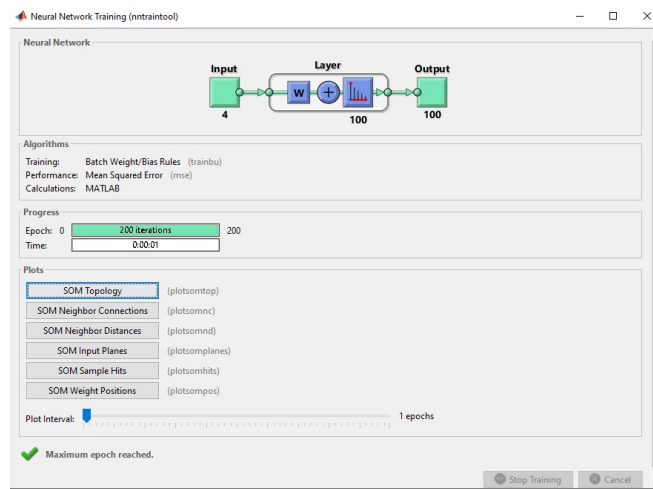
Fuente: Programa de prueba de MATLAB

La figura 4.6.4 contiene las opciones necesarias para el entrenamiento de la red neuronal usando el algoritmo SOM.

Durante la fase de entrenamiento, el vector de peso asociado con cada neurona se mueve para convertirse en el centro de un grupo de vectores de entrada. Además, las neuronas que están adyacentes entre sí en la topología también deben acercarse entre sí en el espacio de entrada, por lo tanto, es posible visualizar un espacio de entradas de alta dimensión en las dos dimensiones de la topología de red.

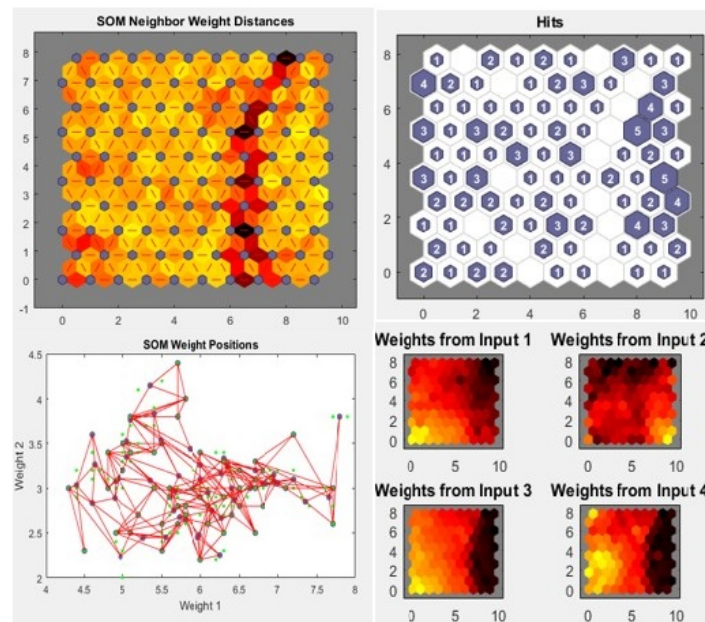
Al pulsar “*Train*” se despliega información adicional sobre el procedimiento llevado a cabo en ésta fase (Figura 4.6.5), al seleccionar estas opciones se obtiene a manera gráfica el tipo de arquitectura utilizada y los resultados obtenidos (Figura 4.6.6).

Figura 4.6.5: Ventana posterior al entrenamiento de la red neuronal.



Fuente: Programa de prueba de MATLAB

Figura 4.6.6: Resultados obtenidos en la fase de entrenamiento.



Fuente: Programa de prueba de MATLAB

La Figura 4.6.6 está compuesta por 4 subfiguras que a continuación se describen:

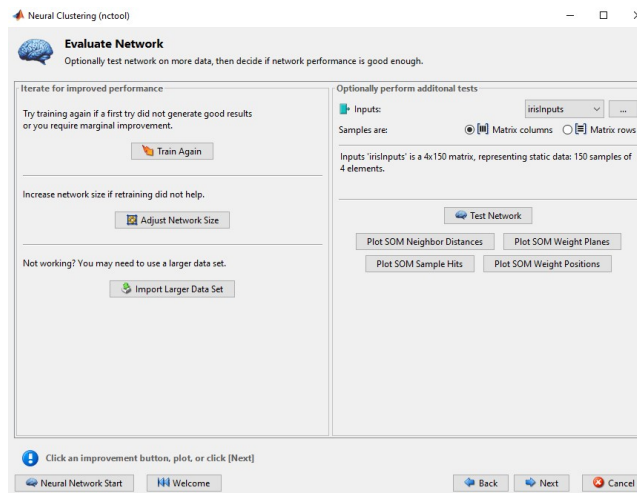
- La subfigura 1 (en la parte superior izquierda), muestra el mapa autoorganizativo en el cual cada hexágono en color grisáceo representa una neurona. Aquí se indica

un plano de peso para cada elemento del vector de entrada (cuatro, en este caso). Son visualizaciones de los pesos que conectan cada entrada a cada una de las neuronas. (Los colores más oscuros representan pesos más grandes). Si los patrones de conexión de dos entradas eran muy similares, puede suponer que las entradas están altamente correlacionadas. En este caso, la entrada 1 tiene conexiones que son muy diferentes a las de la entrada 2, la entrada 3 tiene conexiones claramente similares a la entrada 4

- La subfigura 2 (en la parte superior derecha), representa las neuronas que se activan mediante los datos ingresados a través de la matriz de entrenamiento. La topología predeterminada del SOM es hexagonal. Ésta subfigura muestra las ubicaciones de las neuronas en la topología e indica cuántos datos de entrenamiento están asociados con cada una de las neuronas (centros de agrupación). La topología es una cuadrícula de 10 por 10, por lo que hay 100 neuronas. El número máximo de “Hits” asociados con cualquier neurona es 5 en este caso. Por lo tanto, hay 5 vectores de entrada en ese grupo
- La subfigura 3 (en la parte inferior izquierda), representa la distribución de los registros en clusters y las neuronas asociadas a cada patrón de similitud.
- La subfigura 4 (en la parte inferior derecha), representa los patrones encontrados en cada elemento de todos los registros ingresados.

5. Se evalúa los resultados que arroja el algoritmo .

Figura 4.6.7: Ventana de validación de la red neuronal.



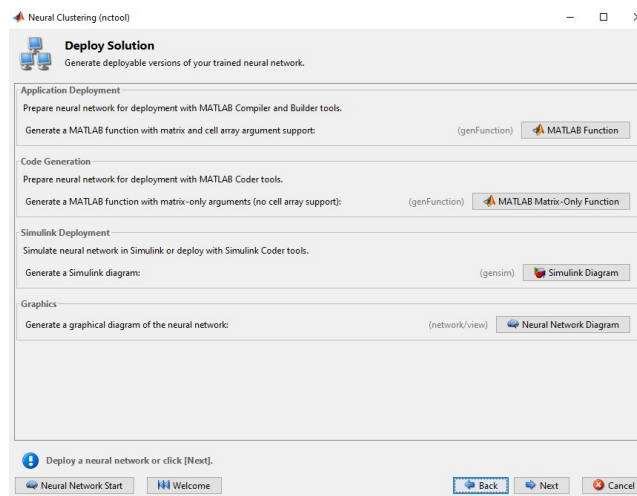
Fuente: Programa de prueba de MATLAB

La figura 4.6.7, permite la evaluación de la red neuronal, para ello hay 2 posibilidades, ya sea, usar ejemplos que no se incluyeron en las matrices de entrenamiento (pero en la que hay certeza del tipo al que pertenecen) ó usar ejemplos que estén incluidos en dicha matriz.

Si los resultados no son los deseados, es posible modificar ciertos parámetros como el tamaño del mapa auto-organizado y posteriormente re-entrenar la red neuronal.

6. Se explota los resultados.

Figura 4.6.8: Ventana de explotación de resultados en la red neuronal.



Fuente: Programa de prueba de MATLAB

La figura 4.6.8 muestra opciones de vital importancia en el desarrollo de este trabajo.

En las últimas versiones MATLAB, incorpora herramientas de generación automática de funciones cuyo objetivo radica en determinar la posición de la neurona vencedora en el mapa auto-organizado.

Al hacer clic en Simple Script o Advanced Script se crea un código MATLAB que se puede usar para reproducir todos los pasos anteriores desde la línea de comandos. La creación de código MATLAB puede ser útil si desea aprender a usar la funcionalidad de línea de comandos de la caja de herramientas para personalizar el proceso de capacitación. Cabe destacar que en los scripts generados ya se encuentran ajustados todos los parámetros de la red que se obtuvieron como resultado en la fase de entrenamiento.

Cuando se hayan generado los scripts y guardado sus resultados, se finaliza el procedimiento.

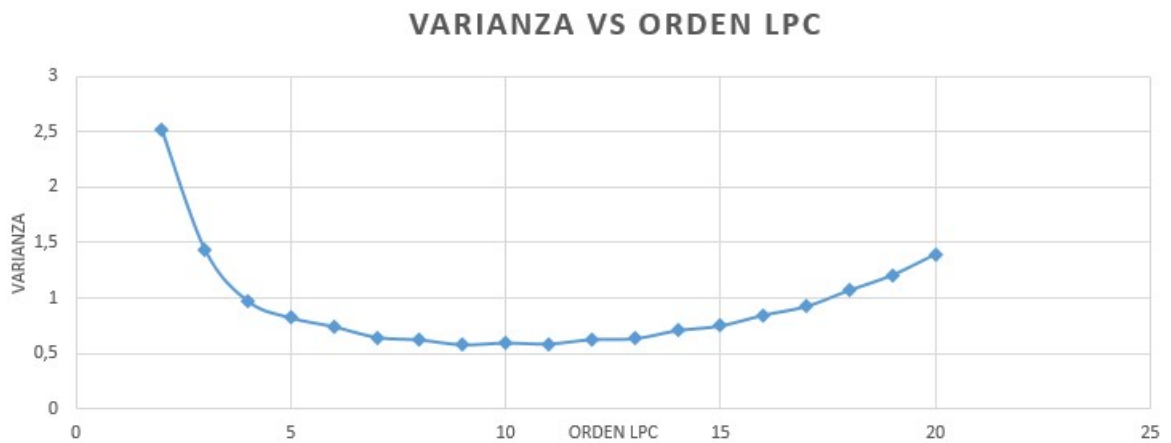
Capítulo 5

Resultados.

5.1. Representación de señales mediante sus LPC

En éste análisis se escogieron 30 trazas de señales correspondientes a sismos del tipo LP y 30 trazas para el tipo VT donde el orden de predicción lineal varía entre 2 y 20. Los resultados obtenidos se describen en las tablas 5.1.1 y 5.1.2 .

Figura 5.1.1: Varianza vs. orden de LPC para 30 trazas de sismos tipo LP.

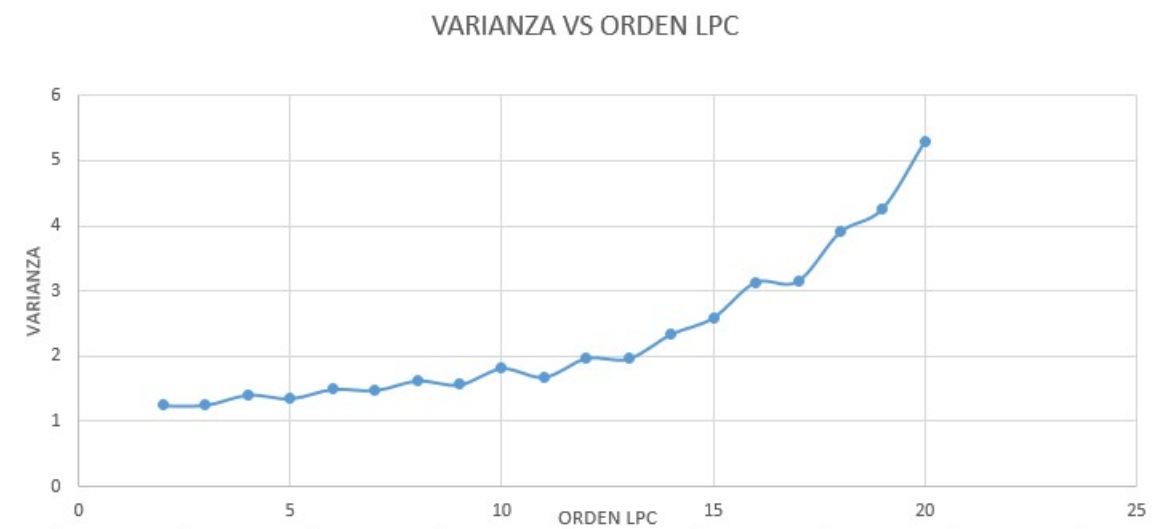


Fuente: (Ésta investigación)

En la Figura 5.1.1 se muestra el comportamiento de las varianzas promedio sucesivas en función al orden de predicción (entre 2 y 20) para 30 muestras de sismos tipo LP .

Obsérvese que en el orden 10 la curva presenta un punto de inflexion, es importante destacar el comportamiento de las diferencias entre las varianzas sucesivas cercanas a éste punto, ya que se tornan relativamente pequeñas en magnitud.

Figura 5.1.2: Varianza vs. orden de LPC para 30 trazas de sismos tipo VT.



Fuente: (Ésta investigación)

En la Figura 5.1.2 se observa el crecimiento de la varianza de forma proporcional al orden de LPC, la curva presenta un comportamiento relativamente estable en el intervalo [2 ,10], donde las diferencias de las varianzas sucesivas son pequeñas y, por lo tanto, de nuestro interés.

Tabla 5.1.1: Resultados del análisis de trazas de sismos del tipo LP para diferentes órdenes de predicción lineal.

TRAZAS DE SISMOS TIPO LP				
Orden LPC	Promedio varianza	Desviación estándar	Diferencia	Diferencia (%)
2	2.5125	1.5851		
3	1.4346	1.1977	1.0778	100.0
4	0.9713	0.9855	0.4633	42.9
5	0.8213	0.9062	0.1500	13.9
6	0.7416	0.8611	0.0796	7.4
7	0.6455	0.8034	0.0960	8.9
8	0.6221	0.7887	0.0234	2.2
9	0.5794	0.7612	0.0426	4.0
10	0.5953	0.7715	0.0158	1.5
11	0.5864	0.7658	0.0088	0.8
12	0.6254	0.7908	0.0390	3.6
13	0.6336	0.7960	0.0082	0.8
14	0.7070	0.8408	0.0733	6.8
15	0.7493	0.8656	0.0423	3.9
16	0.8462	0.9199	0.0969	9.0
17	0.9248	0.9617	0.0786	7.3
18	1.0709	1.0348	0.1460	13.5
19	1.2073	1.0987	0.1363	12.6
20	1.3960	1.1815	0.1887	17.5

Fuente: (Ésta investigación)

En las Tablas 5.1.1 y 5.1.2 se muestra el orden de predicción lineal asociado a cada señal, el promedio de las varianzas para los segmentos entre los órdenes 2 y 20, la desviación estándar correspondiente, la diferencia de las varianzas entre un orden y el siguiente y su representación en términos porcentuales.

Tabla 5.1.2: Resultados del análisis de trazas de sismos del tipo VT para diferentes órdenes de predicción lineal.

TRAZAS DE SISMOS TIPO VT				
Orden LPC	Promedio varianza	Desviación estándar	Diferencia	Diferencia (%)
2	1.2443	1.1155		
3	1.2553	1.1204	0.0109	1.1
4	1.4089	1.1869	0.1535	15.0
5	1.3537	1.1635	0.0551	5.4
6	1.4977	1.2238	0.1440	14.1
7	1.4799	1.2165	0.0178	1.7
8	1.6263	1.2752	0.1463	14.3
9	1.5675	1.2520	0.0587	5.7
10	1.8154	1.3473	0.2478	24.2
11	1.6829	1.2972	0.1324	13.0
12	1.9721	1.4043	0.2892	28.3
13	1.9656	1.4023	0.0064	0.6
14	2.3405	1.5298	0.3748	36.7
15	2.5938	1.6105	0.2532	24.8
16	3.1370	1.7711	0.5432	53.1
17	3.1524	1.7755	0.0153	1.5
18	3.9046	1.9760	0.7521	73.6
19	4.2657	2.0653	0.3610	35.3
20	5.2878	2.2995	1.0221	100.0

Fuente: (Ésta investigación)

Para la creación de matrices de entrenamiento es necesario el uso de ejemplos con las mismas dimensiones, esto implica que el orden de predicción debe ser el mismo tanto en trazas de sismos LP como de VT. Se necesita buscar el rango de orden de LPC donde las varianzas sean pequeñas simultáneamente en ambos casos. En la Figura 5.1.1 el rango más adecuado es el que contiene los puntos más cercanos al punto de inflexión donde el orden de lpc es 10, a razón de, como se mencionó anteriormente la diferencia de varianzas es pequeña, por otro lado en la Figura 5.1.2, el rango de predicción superior a 10 corresponde a varianzas considerables. Teniendo en cuenta esto, el orden de predicción adecuado esta comprendido en el intervalo definido entre 4 y 10, donde se observa que ambas curvas tienen un comportamiento estable y cuyas diferencias de varianzas sucesivas son relativamente pequeñas.

5.2. Selección de las muestras

5.2.1. Muestras para segmentos de ruido y sismo

Tabla 5.2.1: Selección de las muestras para unidades de ruido y sismo.

Nombre Traza	(%) Traslapo	Uds de Ruido	Uds de Sismo	# Total de Uds
sismo0	50	28	3	31
sismo0	60	35	5	40
sismo0	70	48	6	54
sismo0	80	67	8	75
sismo0	90	132	20	152
sismo1	50	35	2	37
sismo1	60	35	2	37
sismo1	70	58	4	62
sismo1	80	86	6	92
sismo1	90	172	11	183
sismo2	50	30	3	33
sismo2	60	36	4	40
sismo2	70	49	5	54
sismo2	80	72	7	79
sismo2	90	143	12	155
sismo3	50	25	4	29
sismo3	60	33	4	37
sismo3	70	42	5	47
sismo3	80	65	10	75
sismo3	90	128	20	148
sismo4	50	36	2	38
sismo4	60	45	3	48
sismo4	70	60	4	64
sismo4	80	91	6	97
sismo4	90	182	13	195

Fuente: (Ésta investigación)

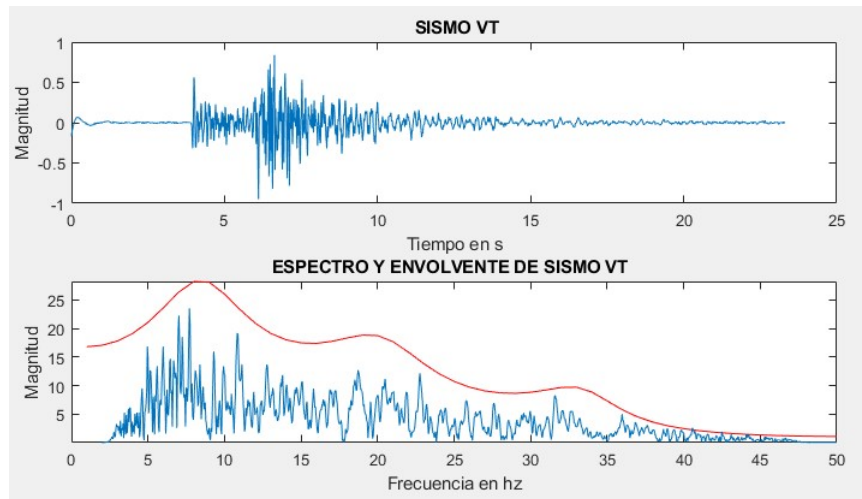
La Tabla 5.2.1 indica el nombre de cada traza, el porcentaje de traslapeo entre cada vector, las unidades extraídas tanto de ruido como de sismos y el número total de unidades que conforma cada traza.

La selección de las muestras que contiene segmentos de ruido y sismo se hizo a partir de 5 trazas de diferentes eventos sísmicos, en donde cada traza fue segmentada a manera vectorial cuyos vectores contienen 300 datos y un determinado porcentaje de traslapeo entre cada uno de estos. Los resultados obtenidos se muestran en la Tabla 5.2.1.

5.2.2. Muestras de Sismos tipo VT y LP

Para la elección de las mejores muestras en la fase de entrenamiento, se tuvo en cuenta el grado de similitud entre cada una mediante su representación espectral y envolvente (Figura 5.2.2 y Figura 5.2.1).

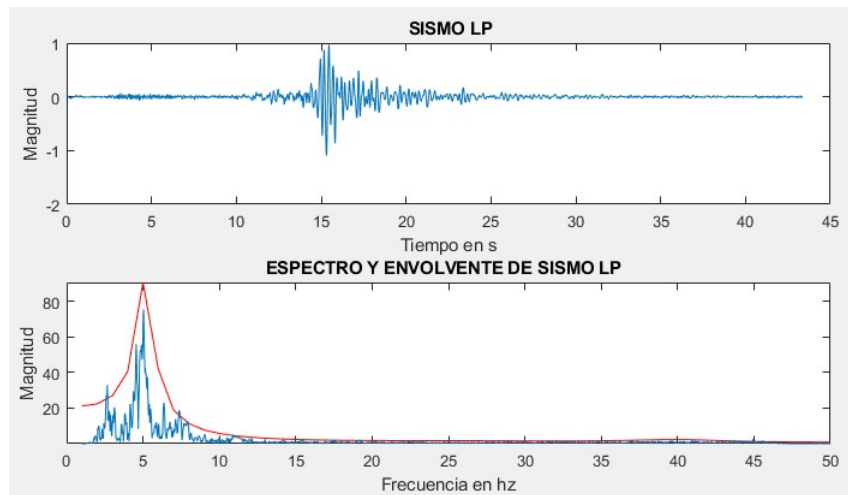
Figura 5.2.1: Representación espectral y envolvente de un sismo de tipo VT.



Fuente: (Ésta investigación)

La Figura 5.2.1 y Figura5.2.2 muestran la representación de un sismo tipo VT y LP con su respectivo espectro y envolvente. Nótese las evidentes diferencias de las señales en el dominio de la frecuencia.

Figura 5.2.2: Representación espectral y envolvente de un sismo de tipo LP.



Fuente: (Ésta investigación)

Las muestras seleccionadas fueron de 1046 trazas de sismos de tipo LP y 713 trazas para sismo tipo VT, éstas señales provienen del monitoreo del VG durante los periodos comprendidos entre el 2006 y 2009.

Se destinaron 500 muestras de sismos VT y 500 muestras de sismos LP en la fase de entrenamiento, las trazas restantes fueron útiles en la validación de resultados.

5.3. Generación de las matrices de entrenamiento

5.3.1. Clasificación de unidades de ruido vs unidades de Sismo

Para el entrenamiento de las RNA se crearon 9 matrices, éstas matrices incluyen determinadas unidades de ruido y sismo. Cada matriz se construyó con vectores LPC de diferentes órdenes, 5, 6 y 7. Los vectores LPC que conforman la matriz, provienen de fragmentos de trazas cuyos vectores contienen 300 datos, el número total de vectores asociados a una traza varía con respecto al porcentaje de traslapo que hay entre cada vector.

La Tabla 5.3.1 indica el nombre de la matriz de entrenamiento, el orden de predicción asociado, el número de ejemplos de los segmentos de sismos y ruido que se utilizó respectivamente y finalmente el número total de ejemplos que conforma cada matriz.

Tabla 5.3.1: Datos de las matrices de entrenamiento para la discriminación de segmentos de ruido(**R**) y sismo(**S**)

Nombre Matriz	Orden LPC	# Ejemplos S	# Ejemplos R	# Total de ejemplos
ENTRENA1	5	42	210	252
ENTRENA2	5	84	420	504
ENTRENA3	5	126	630	756
ENTRENA4	6	42	210	252
ENTRENA5	6	84	420	504
ENTRENA6	6	126	630	756
ENTRENA7	7	42	210	252
ENTRENA8	7	84	420	504
ENTRENA9	7	126	630	756

Fuente: (Ésta investigación)

5.3.2. Clasificación de sismos VT vs. sismos LP

En este procedimiento todas las trazas que conforman las matrices de entrenamiento fueron sometidas a un proceso digital de señales, en dicho proceso se eliminó el offset, se normalizo las unidades, se extrajo frecuencias menores a 1hz mediante el filtro de “*Butterword*” y se obtuvo las propiedades características generales.

Para el entrenamiento de todas las anteriores matrices, se utilizó un mapa auto organizativo de dimensión 10x10 con un total de 100 neuronas.

Para la fase de entrenamiento de las RNA se crearon 35 matrices, éstas matrices se construyeron a partir de los coeficientes de prediccion asociados a trazas completas de sismos de tipo VT y LP, el número de ejemplos de entrenamiento para cada matriz oscila entre 100 y 500, donde el orden de LPC asociado a cada traza esta comprendido en el rango entre 4 y 10.

La Tabla 5.3.2 y la Tabla 5.3.3 muestran las características con las que se generaron las matrices de entrenamiento.

Tabla 5.3.2: Configuración de las matrices de entrenamiento para la clasificación entre LP y VT

Nombre Matriz	Orden LPC	# Ejemplos VT	# Ejemplos LP
Matriz1	4	100	100
Matriz2	5	100	100
Matriz3	6	100	100
Matriz4	7	100	100
Matriz5	8	100	100
Matriz6	9	100	100
Matriz7	10	100	100
Matriz8	4	200	200
Matriz9	5	200	200
Matriz10	6	200	200
Matriz11	7	200	200
Matriz12	8	200	200
Matriz13	9	200	200
Matriz14	10	200	200
Matriz15	4	300	300
Matriz16	5	300	300
Matriz17	6	300	300
Matriz18	7	300	300
Matriz19	8	300	300
Matriz20	9	300	300
Matriz21	10	300	300
Matriz22	4	400	400
Matriz23	5	400	400
Matriz24	6	400	400
Matriz25	7	400	400

Fuente: (Ésta investigación)

Tabla 5.3.3: Configuración de las matrices de entrenamiento para la clasificación entre LP y VT

Nombre Matriz	Orden LPC	# Ejemplos VT	# Ejemplos LP
Matriz26	8	400	400
Matriz27	9	400	400
Matriz28	10	400	400
Matriz29	4	500	500
Matriz30	5	500	500
Matriz31	6	500	500
Matriz32	7	500	500
Matriz33	8	500	500
Matriz34	9	500	500
Matriz35	10	500	500

Fuente: (Ésta investigación)

5.4. Validación de resultados

5.4.1. Unidades de sismo vs unidades de ruido

Para el proceso de validación de resultados se utilizó 50 unidades de sismo y 50 unidades de ruido con un total de 100 elementos, éstas unidades no se utilizaron en la etapa de entrenamiento. Para medir directamente la efectividad de cada RNA se cuantificó el número de aciertos.

La tabla 5.4.1 muestra el nombre de la RNA, el nombre de la matriz de entrenamiento asociada a cada RNA, el orden de predicción, el porcentaje de aciertos en la clasificación de unidades de ruido, el porcentaje de aciertos en la clasificación de unidades de sismo y la efectividad porcentual .

Tabla 5.4.1: Resultados de la ejecución de las RNA para unidades de sismo (**S**) y de ruido (**R**).

Nombre RNA	Matriz	LPC	(%)Aciertos R	(%)Aciertos S	(%) Efectividad
CLASIFICA1	ENTRENA1	5	100	100	100
CLASIFICA2	ENTRENA2	5	100	100	100
CLASIFICA3	ENTRENA3	5	100	100	100
CLASIFICA4	ENTRENA4	6	100	100	100
CLASIFICA5	ENTRENA5	6	100	100	100
CLASIFICA6	ENTRENA6	6	100	100	100
CLASIFICA7	ENTRENA7	7	100	100	100
CLASIFICA8	ENTRENA8	7	100	100	100
CLASIFICA9	ENTRENA9	7	100	100	100

Fuente: (Ésta investigación)

La cantidad de aciertos obtenidos por todas las RNA fue de 100/100 del total de los ejemplos evaluados, este resultado se debe a la calidad de las muestras durante el entrenamiento.

La representación espectral para cada segmento juega un papel importante para la distinción de patrones durante la fase de entrenamiento. Mientras las bandas espectrales en los segmentos de sismos se concentran en altas frecuencias (mayores a los 20 Hz), las de los segmentos de ruido se encuentran en frecuencias menores a los 10 Hz

5.4.2. Clasificación de sismos tipo VT vs sismos tipo LP

Para evaluar la efectividad de cada red neuronal se incluyeron 100 trazas de sismos tipo VT y 100 trazas de sismos tipo LP con un total de 200 muestras, las anteriores señales no fueron incluidas en la fase de entrenamiento.

El rango de aciertos está comprendido entre el 77 % y el 97 % en el mejor de los casos.

La RNA cuyo porcentaje de efectividad es del 97 % tiene una dimensión de (10x10) en el mapa SOM, se entrenó con la matriz asociada a un orden de predicción igual a 6, dicha matriz se denomina Matriz17 cuyo número total de ejemplos es de 600, 300 de los cuales corresponden a trazas de sismos tipo VT y 300 a sismos de tipo LP.

A continuación, se muestran los resultados obtenidos para cada matriz cuyos elementos de entrenamiento están incluidos en iguales proporciones.

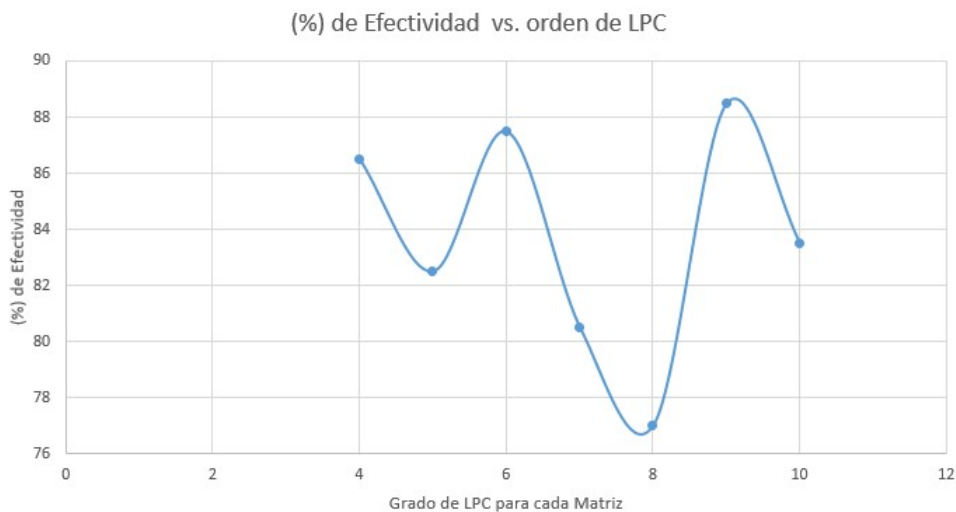
Tabla 5.4.2: Resultados de la ejecución de las RNA entrenadas con 100 trazas de sismos LP y 100 trazas de sismos VT

Nombre RNA	Matriz	LPC	(%) Aciertos VT	(%) Aciertos LP	(%) Efectividad
RNA1	Matriz1	4	99	74	86.5
RNA2	Matriz2	5	99	66	82.5
RNA3	Matriz3	6	95	80	87.5
RNA4	Matriz4	7	97	64	80.5
RNA5	Matriz5	8	95	59	77
RNA6	Matriz6	9	95	82	88.5
RNA7	Matriz7	10	84	83	83.5

Fuente: (Ésta investigación)

La Tabla 5.4.2 muestra el nombre de la RNA, el nombre de matriz con la que fue entrenada, el orden de LPC asociado, el porcentaje de aciertos que obtuvo en la clasificación de sismo tipo VT y LP, y finalmente el porcentaje efectividad de cada RNA (Esta descripción también es válida para las posteriores tablas). Se obtuvo un promedio porcentual total de 83.7 % en la efectividad a diferentes órdenes de predicción lineal, todos éstos están asociados a una matriz de 100 ejemplos de sismo VT y 100 de sismo LP.

Figura 5.4.1: Efectividad vs orden LPC para un conjunto de matrices entrenadas con 100 trazas LP y 100 trazas VT.



Fuente: (Ésta investigación)

la Figura 5.4.1 indica que al orden de 9 en la predicción lineal se alcanza un porcentaje en la efectividad del 88.5 % , es interesante mencionar el comportamiento que describen las RNA entre los órdenes de 8 y 9, ya que apartir de estos se encuentran el peor y mejor resultado respectivamente.

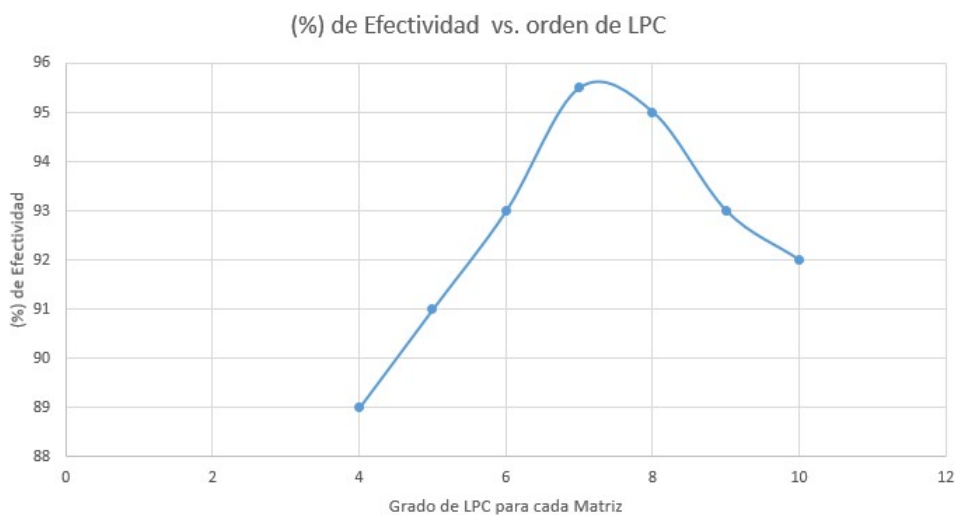
Tabla 5.4.3: Resultados de la ejecución de las RNA entrenadas con 200 trazas de sismos LP y 200 trazas de sismos VT

Nombre RNA	Matriz	LPC	(%) Aciertos VT	(%) Aciertos LP	(%) Efectividad
RNA8	Matriz8	4	99	79	89
RNA9	Matriz9	5	96	86	91
RNA10	Matriz10	6	97	89	93
RNA11	Matriz11	7	95	96	95.5
RNA12	Matriz12	8	93	97	95
RNA13	Matriz13	9	90	96	93
RNA14	Matriz14	10	86	98	92

Fuente: (Ésta investigación)

En la Tabla 5.4.3 se obtuvo un promedio porcentual total de 92.6 % en la efectividad a diferentes órdenes de predicción lineal, todos éstos estan asociados a una matriz de 200 ejemplos de sismo VT y 200 de sismo LP.

Figura 5.4.2: Efectividad vs orden LPC para un conjunto de matrices entrenadas con 200 trazas LP y 200 trazas VT.



Fuente: (Ésta investigación)

La Figura 5.4.2 muestra que en el orden 7 de predicción lineal se alcanza una efectividad del 95.5 %, siendo la mayor para este caso. La curva presenta un comportamiento ascendente entre los ordenes 4 y 7, posteriormente a éstos, la efectividad cae paulatinamente.

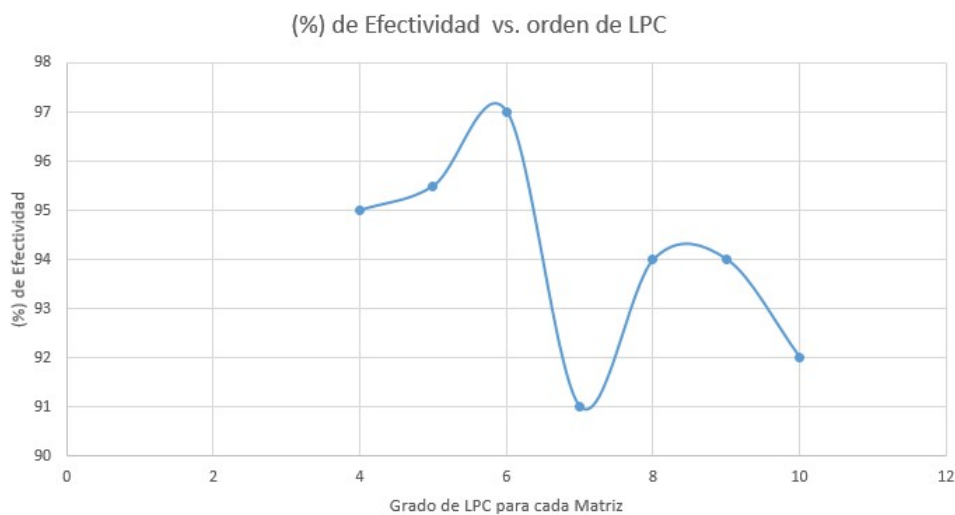
Tabla 5.4.4: Resultados de la ejecución de las RNA entrenadas con 300 trazas de sismos LP y 300 trazas de sismos VT

Nombre RNA	Matriz	LPC	(%) Aciertos VT	(%) Aciertos LP	(%) Efectividad
RNA15	Matriz15	4	92	98	95
RNA16	Matriz16	5	99	92	95.5
RNA17	Matriz17	6	98	96	97
RNA18	Matriz18	7	87	95	91
RNA19	Matriz19	8	98	90	94
RNA20	Matriz20	9	91	97	94
RNA21	Matriz21	10	86	98	92

Fuente: (Ésta investigación)

En la Tabla 5.4.4 se obtuvo un promedio porcentual total de 94.1 % en la efectividad a diferentes órdenes de predicción lineal, todos éstos están asociados a una matriz de 300 ejemplos de sismo VT y 300 de sismo LP.

Figura 5.4.3: Efectividad vs orden LPC para un conjunto de matrices entrenadas con 300 trazas LP y 300 trazas VT.



Fuente: (Ésta investigación)

La Figura 5.4.3 muestra un comportamiento inestable en la efectividad con diferentes órdenes de LPC, se observa que en el orden 6 la RNA alcanza su mejor desempeño, pero en el orden 7, la curva decae drásticamente.

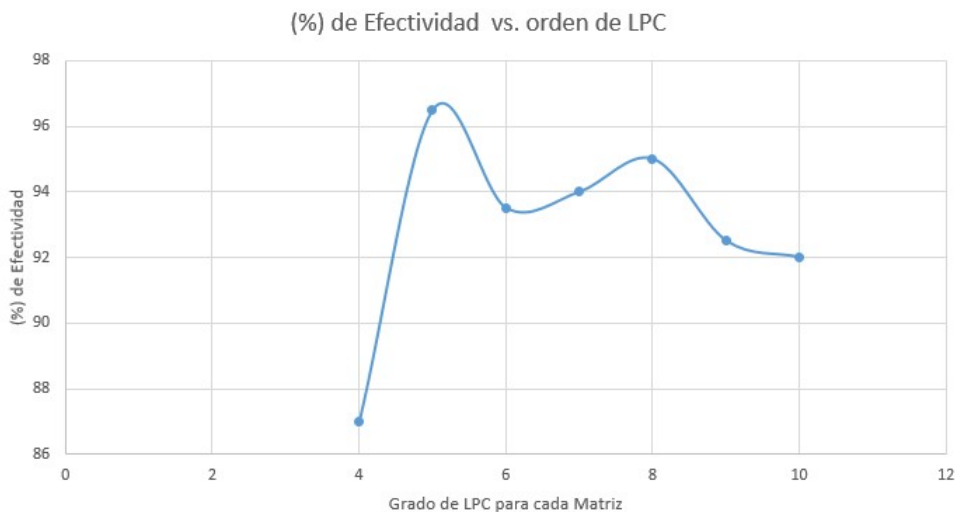
Tabla 5.4.5: Resultados de la ejecución de las RNA entrenadas con 400 trazas de sismos LP y 400 trazas de sismos VT

Nombre RNA	Matriz	LPC	(%) Aciertos VT	(%) Aciertos LP	(%) Efectividad
RNA22	Matriz22	4	83	91	87
RNA23	Matriz23	5	97	96	96.5
RNA24	Matriz24	6	97	90	93.5
RNA25	Matriz25	7	95	93	94
RNA26	Matriz26	8	95	95	95
RNA27	Matriz27	9	91	94	92.5
RNA28	Matriz28	10	88	96	92

Fuente: (Ésta investigación)

En la Tabla 5.4.5 se obtuvo un promedio porcentual total de 92.9 % en la efectividad a diferentes órdenes de predicción lineal, todos éstos están asociados a una matriz de 400 ejemplos de sismo VT y 400 de sismo LP.

Figura 5.4.4: Efectividad vs orden LPC para un conjunto de matrices entrenadas con 400 trazas LP y 400 trazas VT.



Fuente: (Ésta investigación)

La Figura 5.4.4 muestra un ascenso en la curva entre los órdenes 4 y 5 de LPC donde la efectividad adquiere sus valores máximo y mínimo respectivamente, posteriormente la curva describe un comportamiento oscilante entre los siguientes órdenes de predicción.

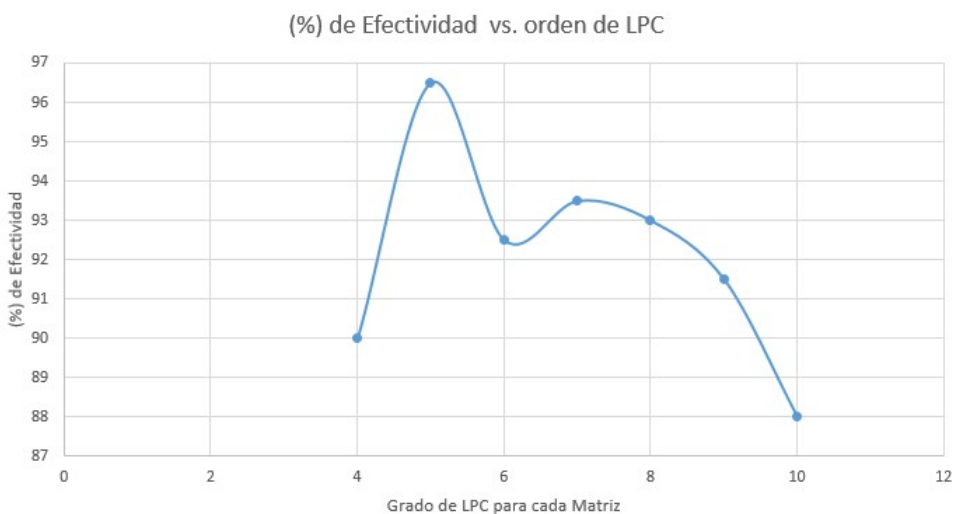
Tabla 5.4.6: Resultados de la ejecución de las RNA entrenadas con 500 trazas de sismos LP y 500 trazas de sismos VT

Nombre RNA	Matriz	LPC	(%) Aciertos VT	(%) Aciertos LP	(%) Efectividad
RNA29	Matriz29	4	90	90	90
RNA30	Matriz30	5	97	96	96.5
RNA31	Matriz31	6	97	88	92.5
RNA32	Matriz32	7	95	92	93.5
RNA33	Matriz33	8	91	95	93
RNA34	Matriz34	9	90	93	91.5
RNA35	Matriz35	10	83	93	88

Fuente: (Ésta investigación)

En la Tabla 5.4.6 se obtuvo un promedio porcentual total de 92.1 % en la efectividad a diferentes órdenes de predicción lineal, todos éstos están asociados a una matriz de 500 ejemplos de sismo VT y 500 de sismo LP.

Figura 5.4.5: Efectividad vs orden LPC para un conjunto de matrices entrenadas con 500 trazas LP y 500 trazas VT.



Fuente: (Ésta investigación)

La figura 5.4.5 muestra que al orden 5 de LPC la RNA alcanza la mayor efectividad igual al 96.5 %, se destaca también el descenso de la curva a partir del orden de 7 en la predicción lineal.

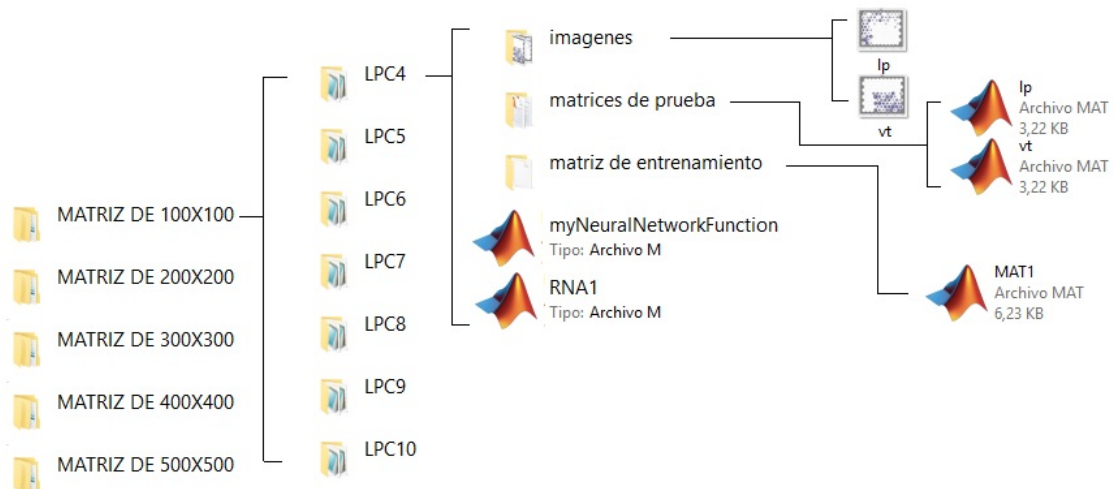
5.5. Almacenamiento y operación de las RNA

Para la correcta ejecución de las diferentes RNA, se recomienda usar una versión igual o superior a la R2017a de MATLAB, mencionando que en la elaboración de las distintas rutinas computacionales se incluyeron comandos solo aceptados en dichas versiones, todo con el objetivo de evitar posibles errores.

5.5.1. RNA en la clasificación de VT y LP

La manera en la que se organizaron los distintos archivos para cada RNA en la carpeta “*CLASIFICACIÓN VT Y LP*” se representa en la Figura 5.5.1

Figura 5.5.1: Forma en la que se almacenaron cada una de las RNA para la clasificación de VT y LP.



Fuente: (Ésta investigación)

La Figura 5.5.1 muestra el contenido de la carpeta “MATRIZ 100X100” (siendo análoga para las demás carpetas), este nombre hace referencia a la cantidad de unidades de cada tipo

usadas durante el entrenamiento, en este caso corresponde a 100 unidades de sismo tipo VT y 100 unidades de sismo tipo LP, en el interior de ésta carpeta se incluyen 7 subcarpetas, una de ellas llamada “LPC4” haciendo alusión al orden de LPC (en este caso es de 4) asociado a cada ejemplo.

La carpeta LPC4 contiene 5 elementos, entre ellos 3 carpetas que incluyen las imágenes de las regiones activadas del mapa SOM, las matrices de prueba usadas para la identificación de éstas regiones, la matriz de entrenamiento utilizada para cada RNA y finalmente 2 ejecutables, uno denominado “RNA1” y la función de ajuste de los parámetros para cada RNA (“myNeuralNetworkFunction”).

Antes de ejecutar el script “RNA1” se requiere especificar 2 variables que contienen las direcciones para los archivos, una de ellas denominada “ruta” y la siguiente “ruta1” (Figura 5.5.2).

Figura 5.5.2: Segmento del script de una RNA donde se indican un ejemplo sobre las variables “ruta” y “ruta1” .

```
1 - clear
2 - clc
3
4 - cd
5 - %RUTA DONDE SE ENCUENTRAN LOS SISMOS QUE SE DESEAN CLASIFICAR
6 - ruta='C:\Users\ASUS\Desktop\30 vt';
7 - %RUTA Y NOMBRE DEL ARCHIVO .xlsx DONDE SE ENCUENTRAN LOS RESULTADOS
8 - ruta1='C:\Users\ASUS\Desktop\Resultados1.xlsx';
9 - %XXXXXXXXXXXXXXXXXXXXXXXXX ORDEN LPC XXXXXXXXXXXXXXXXXXXXXXXX
10 - orden_lpc=4
11 - %XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
12 - folder=cd(ruta);
13 - if strcmp(folder,cd)==1;
14 - else
15 -     folder=cd(ruta);
16 - end
17 - archivos=ls(folder);
18 - wq=size(archivos);
19 - tamaño=wq(1,1);
```

Fuente: (Ésta investigación)

La variable “ruta” pide la ubicación de la carpeta en cuyo contenido se encuentran las trazas que deseamos clasificar, y la variable “ruta1” pide la ubicación junto al nombre del archivo con extensión .xlsx en el que se encuentran todos los resultados obtenidos.

Posteriormente al asignar las anteriores variables, ejecutamos el script “RNA1”, luego, al terminar la ejecución se genera el archivo de resultados cuyo nombre y ubicación es el especificado en “ruta1” con extensión .xls .

Al abrir el archivo en la dirección especificada encontramos información tal como el nombre de la traza del sismo, el tipo al que pertenece según la clasificación de la RNA1, y un valor

de etiquetado (Figura 5.5.3).

Figura 5.5.3: Ejemplo de los resultados que se obtienen al ejecutar una RNA.

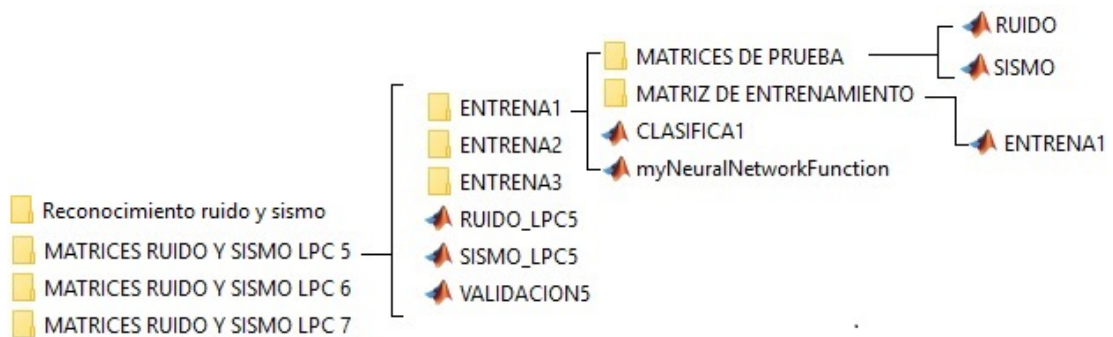
	A	B	C
1	Nombre del sismo	Tipo de sismo	VT=1,LP=2
2	F_200601-03-224-01S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
3	F_200601-03-225-75S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
4	F_200601-04-091-75S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
5	F_200601-04-110-02S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
6	F_200601-05-031-54S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	VT	1
7	F_200601-05-031-74S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	VT	1
8	F_200601-05-125-85S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
9	F_200601-05-145-20S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
10	F_200601-05-160-35S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
11	F_200601-05-165-45S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	LP	2
12	F_200602-12-070-53S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	VT	1
13	F_200602-16-225-12S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	VT	1
14	F_200602-20-091-62S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	VT	1
15	F_200602-26-085-43S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	VT	1
16	F_200602-26-100-95S.OVP_18_1_CUVZ_SB_Z_SAC.asc.txt	VT	1

Fuente: (Ésta investigación)

5.5.2. RNA en la clasificación de sismo y ruido

La Figura 5.5.4 muestra cómo se organizaron los archivos correspondientes a cada RNA en la carpeta “*CLASIFICACIÓN RUIDO VS SISMO*” .

Figura 5.5.4: Forma en la que se almacenaron cada una de las RNA para la clasificación de ruido y sismo.



Fuente: (Ésta investigación)

La Figura 5.5.4 muestra los archivos contenidos en la carpeta “MATRICES RUIDO Y SISMO LPC 5” (Siendo análogo para las demás carpetas) que representa el orden de LPC asociado a cada matriz (en este caso de orden 5), en su interior se destacan 6 archivos correspondientes a 3 carpetas y 3 matrices con extensión .mat, estas carpetas contienen las matrices de entrenamiento asociadas a determinadas unidades de ruido y sismo de acuerdo a la Tabla 5.3.1

De las matrices “RUIDO_LPC5.mat” y “SISMO_LPC5.mat” se extrajo los ejemplos que se utilizaron en la fase de entrenamiento y validación de los resultados, la matriz “VALIDACION5” contiene las unidades destinadas únicamente para evaluar la efectividad de cada RNA asociada un orden de LPC en concreto. La carpeta “ENTRENA1” contiene las matrices de prueba usadas para identificar las regiones que se activan en el mapa SOM, la matriz de entrenamiento asociada, el ejecutable “CLASIFICA1” y la función donde se encuentran establecidos los parámetros de la RNA.

En la carpeta “*Reconocimiento de ruido y sismo*” encontramos el ejecutable “*clasificador1click.m*” el cual muestra de una manera más intuitiva el reconocimiento de segmentos de ruido y sismo en una señal, solo basta llamar el archivo cuya traza este previamente cargada en *Current Folder* de MATLAB, posterior a la ejecución, se procede a seleccionar un punto sobre la gráfica que se despliega, el programa toma un segmento de 300 datos a partir de dicho punto, calcula los coeficientes de predicción asociados y muestra en la ventana de *Command Window* a que grupo pertenece el segmento seleccionado.

Capítulo 6

Discusión

El uso de mapas auto organizados es una de las técnicas más aplicadas debido a su versatilidad y confiabilidad, dado que el aprendizaje, contrariamente al supervisado no está sujeto a aprender únicamente lo que el supervisor les enseña si no, más bien, el propio sistema se encarga de extraer características en los datos de entrada y por ende evita posibles errores en la clasificación debido a una mala elección de las muestras durante el entrenamiento por parte del usuario.

A pesar de que, durante el desarrollo de este trabajo se obtuvo buenos resultados, la técnica “clustering” mostro ser muy susceptible a la calidad de los datos de entrenamiento, esto debido a que uno de los inconvenientes de la técnica radica en que el sistema debe ubicar el centro de los “clusters”.

El método de representación de señales se enfocó exclusivamente en el análisis espectral, siendo esta una alternativa entre muchas más, debido a esta razón el método puede ser mejorado mediante otras formas de representación en diferentes dominios, inclusive el del tiempo, todo esto con el objetivo de incluir más variedad de características que diferencien un conjunto de otro, por lo tanto, la clasificación se realizara con mayor precisión.

Existen algunos problemas tales como la reducción de dimensionalidad donde los datos de entrada deben de ser agrupados en un subespacio con una dimensionalidad más baja que la de los datos, es decir, los datos son representados en la dimensión predeterminada del mapa SOM que en su defecto tiene una representación bidimensional, esto puede incurrir a una pérdida de la información mediante la representación en el mapa SOM.

La discriminación de los segmentos de ruido y sismo se realizó satisfactoriamente, alcanzando un porcentaje en los aciertos del 100 % en las pruebas realizadas, estos resultados pueden explicarse basándose en las notorias diferencias de sus envolventes sobre cada espectro, siendo bastante similares entre los del mismo tipo y notablemente diferente entre los pertenecientes a otro grupo, esto facilitó el reconocimiento de los diferentes clústers asociados a cada segmento y por lo tanto su clasificación.

Con base a los resultados obtenidos, no fue necesario re-entrenar las redes neuronales en esta ocasión, esto debido a la correcta discriminación de los cluster desde el primer entrenamiento, además, el tamaño en los mapas autorizados se redujo ya que la presencia de regiones bien definidas por los clusters lo permitió, destacando una correcta clasificación incluso con un mapa auto organizado de tamaño 4x4 con un total de solamente 16 neuronas.

Para la clasificación entre los sismos de tipo VT y LP fue necesario contemplar varios parámetros que según los resultados obtenidos tienen mucha relevancia en cuanto a la efectividad de las RNA se refiere.

Un parámetro determinante en la efectividad de las RNA es la cantidad de ejemplos usados durante el entrenamiento, es decir, a medida que el número total de ejemplos aumenta la eficiencia también lo hace, esto queda en evidencia en los resultados, mencionando que se obtuvo la menor eficiencia (equivalente a 77 %) en una RNA entrenada con 100 unidades de LP y 100 unidades de VT y una de las máximas eficiencias equivalentes al 96.5 % en una RNA entrenadas con 500 unidades de sismos LP y 500 unidades de sismos VT

El incremento de los ejemplos de entrenamiento tiene mucho sentido, ya que las RNA entrenadas con una mayor variedad de ejemplos permite una gran versatilidad en la clasificación de nueva información. Como mencionamos anteriormente este método es muy susceptible a la “calidad” de los elementos de entrenamiento, esto se considera una limitación que debe tenerse en cuenta. Las muestras con más con mayor grado de similitud definen la parte central de cada clúster, mientras que las demás tienden a separarse cada vez más hasta tener una posición prácticamente arbitraria en el mapa SOM, una manera de contrarrestar este efecto es afinar el número de ejemplos cuyos grados de similitud entre si sea alto, esto con el objetivo de obtener regiones muy específicas en el mapa, otra posibilidad es re-entrenar cada RNA hasta que se observe una distinción relativamente clara entre cada clúster, esto cuando es posible, además en algunas ocasiones es conveniente incrementar el tamaño del mapa SOM pero solamente es recomendable cuando el número total de ejemplos usados en la fase de

entrenamiento excede en gran medida el número total de neuronas presentes en el mapa. Las RNA que mostraron mejores resultados son las que en organizaron cada clúster lo más simétricamente posible, en el proceso de etiquetado de dichas regiones, en algunos casos concretos fue necesario incrementar el área alrededor de cada clúster, esto más que todo en las RNA con órdenes de LPC grandes y con relativamente pocos ejemplos de entrenamiento. Las RNA que surgen como producto del estudio de los mapas SOM tienen la capacidad de clasificar ya sea sismos de tipo VT y de tipo LP o segmentos de ruido y sismo, por lo tanto, las señales asociadas a fenómenos externos o relacionadas con tipos de sismicidad diferentes a VT y LP serán erróneamente clasificadas, dado que este tipo de señales no fueron incluidas en el presente trabajo.

Capítulo 7

Conclusiones

La metodología implementada para la clasificación de sismos volcánicos proporciona un alto grado de confiabilidad dentro de las condiciones necesarias para su aplicación.

La efectividad en la discriminación para determinadas unidades de ruido y sismo fue del 100 % independientemente al grado de predicción asociado a cada unidad, no obstante, al existir gran variedad de eventos sísmicos los resultados pueden diferir al deseado, esto teniendo en cuenta que el ruido filtrado en cualquier señal puede tener considerables variaciones con respecto a diversos factores externos que no se tuvieron en cuenta.

La efectividad de las RNA en la clasificación de trazas de eventos de tipo VT y LP varía desde un 77 % hasta superar el 95 %, alcanzando un 97 % en el mejor de los casos.

La variación de estos resultados son consecuencia directa de la calidad de los ejemplos de entrenamiento, el grado de LPC asociado y el número total de unidades que se incluye en cada matriz de entrenamiento.

Se obtuvo resultados bastante buenos entre los órdenes 5,6,7 de predicción lineal, esto infiere que es la mejor alternativa para representar una señal en términos de sus LPC, la eficiencia del método en estos órdenes de predicción lineal radica en que se logra recopilar la suficiente información que permite su correcta discriminación y no incluir en detalles que de cierta manera la complicaría.

Archivos adjuntos

1. La carpeta “**CLASIFICACIÓN RUIDO VS SISMO**” contiene todas las RNA para la clasificación de segmentos de sismo y ruido, la forma en la que se organizaron los archivos se muestran en la Figura 5.5.4.
2. La carpeta “**CLASIFICACION VT VS LP**” contiene todas las RNA dedicadas a la clasificación de sismicidad de tipo VT y LP, la Figura 5.5.1 muestra como se organizaron todos sus archivos.
3. La carpeta “**TRAZAS VT Y LP**” contiene las trazas de algunos ejemplos de sismicidad de tipo VT y LP.
4. El archivo “**Tesis**” contiene el documento final del trabajo de grado con extensión “.pdf”.

Bibliografía

- [1] O. E. CADENA IBARRA. *Detección y clasificación automática de registros sísmicos en el Observatorio Vulcanológico y Sismológico de Pasto utilizando redes neuronales artificiales*, 2011.
- [2] G. CORTÉS MORENO et al. Reconocimiento de señales sismo-volcánicas mediante canales específicos basados en modelos ocultos de markov. 2016.
- [3] P. P. CRUZ and A. HERRERA. *Inteligencia artificial con aplicaciones a la ingeniería*. Alfaomega, 2010.
- [4] A. D. FIROOZABADI, F. SEGUEL, I. SOTO, D. GUEVARA, F. HUENUPAN, M. CURILEM, and L. FRANCO. *Evaluation of Llaima volcano activities for localization and classification of LP, VT and TR events*. *Journal of Electrical Engineering*, 68(5):325–338, 2017.
- [5] C. S. GEOLÓGICO. Observatorio vulcanológico y sismológico de pasto. consultado el 15 de febrero de 2016, 2019.
- [6] L. JAKSON. *Digital filters and signal processing* 2nd edition, 1989.
- [7] H. LANGER, S. FALSAPERLA, and G. THOMPSON. *Application of artificial neural networks for the classification of the seismic transients at Soufriere Hills volcano, Montserrat*. *Geophysical research letters*, 30(21), 2003.
- [8] UNIVERSIDAD DE MADRID. *Los mapas auto-organizados de Kohonen (SOM)*. <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/DM/tema5dm.pdf>, 2019.
- [9] M. MALFANTE, M. DALLA MURA, J. I. MARS, and J.-P. MÉTAXIAN. *Machine learning for automatic classification of volcano-seismic signatures*. In *2017 25th European Signal Processing Conference (EUSIPCO)*, pages 2393–2397, 2017.

- [10] M. MALFANTE, M. DALLA MURA, J.-P. MÉTAXIAN, J. I. MARS, O. MACEDO, and A. INZA. *Machine learning for volcano-seismic signals: Challenges and perspectives. IEEE Signal Processing Magazine*, 35(2):20–30, 2018.
- [11] A. V. OPPENHEIM and R. W. SCHAFER. *Tratamiento de señales en tiempo discreto*. 2000.
- [12] S. G. C. OVSP . *Volcán Galeras-Generalidades*. 2013.
- [13] J. G. PROAKIS, D. S. G. MANOLAKIS, V. SANTALLA DEL RÍO, A. CASTRO, J. LUIS, et al. *Tratamiento digital de señales* . 1998.
- [14] G. ÁLVAREZ ALBA. *Diseño de Redes neuronales artificiales no supervisadas orientadas a la inteligencia de negocio*. 2016.