
COMBINACIÓN FIJA DE CLASIFICADORES PARA LA DISCRIMINACIÓN DE SEÑALES SÍSMICAS VOLCÁNICAS

FIXED COMBINING OF CLASSIFIERS FOR DISCRIMINATING SEISMIC VOLCANIC SIGNALS

Cristian Andrés Chu Salgado¹, Mauricio Orozco Alzate² y John Makario Londoño Bonilla³.

1. Ingeniero Electrónico, Departamento de Ingeniería Eléctrica, Electrónica y Computación, Universidad Nacional de Colombia Sede Manizales

2. Doctor en Ingeniería - Automática, Profesor asistente en dedicación exclusiva, Departamento de Informática y Computación, Universidad Nacional de Colombia Sede Manizales

*3. Doctor en Geofísica, Profesional Especializado (Sismología), Observatorio Vulcanológico y Sismológico de Manizales, Instituto Colombiano de Geología y Minería - INGEOMINAS
morozcoa@bt.unal.edu.co*

Recibido para evaluación: 5 de Marzo de 2010 / Aceptación: 7 de Mayo de 2010 / Recibida versión final: 23 de Mayo de 2010

RESUMEN

Los eventos sísmicos tienen una gran variedad de clases, las cuales se manifiestan en mayor o menor frecuencia dependiendo del volcán en estudio. En este artículo se presenta un método de clasificación de eventos sísmicos del volcán Nevado del Ruiz mediante reglas fijas de combinación de clasificadores, buscando de esta manera explotar la multiplicidad de datos de los registros del Observatorio Vulcanológico y Sismológico de Manizales, provenientes de las diversas estaciones de monitoreo. El objetivo es aprovechar los múltiples conjuntos de entrenamiento que se pueden obtener a partir de eventos que fueron registrados simultáneamente en varias estaciones, lo cual garantiza la independencia de datos y, por lo tanto, el aumento del desempeño de clasificadores combinados bajo las reglas de producto y suma.

PALABRAS CLAVES: Clasificador base, combinación de clasificadores, fusión, selección, reglas fijas, Nevado del Ruiz.

ABSTRACT

Seismic events have a large variety of classes, which are more or less observed depending on the volcano under study. In this paper, a classification method of seismic events at Nevado del Ruiz volcano by using fixed rules for combining classifiers is presented. It is aimed at exploiting data multiplicity in the registers taken by the Volcanological and Seismological Observatory at Manizales, which are recorded at different monitoring stations. The main objective is taking advantage of multiple training sets which are obtained from events that were registered simultaneously at several stations; this ensures data independence and, therefore, an increase in the performance of classifiers combined with product and sum rules.

KEY WORDS: Base classifiers, classifier combining, fusion, selection, fixed rules, Nevado del Ruiz.

1. INTRODUCCIÓN

El Observatorio Vulcanológico y Sismológico de Manizales (OVSM) del INGEOMINAS monitorea permanentemente la actividad del volcán Nevado del Ruiz mediante estaciones distribuidas estratégicamente, en la cuales se han registrado una gran cantidad de datos, tanto analógicos como digitales. Las señales digitales constituyen una gran base de datos donde cada evento sísmico contiene los registros que proveen las estaciones donde fue detectado; por lo tanto, se tienen datos múltiples e independientes de un mismo evento.

Resolver el problema de la clasificación de eventos sísmicos es fundamental en el desarrollo de estudios que pretenden descubrir la interacción entre los sismos volcánicos y los procesos volcánicos. En la base de datos del OVSM del INGEOMINAS se encuentran eventos de once clases: avalanchas (AV), sismos distantes (DS), sismos híbridos (HB), sismos de hielo (IC), sismos de largo período (LP), sismos regionales (RE), relámpagos (RY), sismos tectónicos locales (TL), tornillos (TO), tremores (TR) y sismos volcano-tectónicos (VT). Sin embargo, cuando se está monitoreando un volcán específico, los sismólogos usan su propia clasificación con una descripción detallada para cada subtipo de evento (Zobin, 2003).

Los trabajos realizados en clasificación automática de señales sísmicas se han enfocado principalmente en la aplicación de redes neuronales artificiales (ANN). En (Langer et al., 2006) se propone el uso de una ANN para clasificar cinco clases de eventos sísmicos: VT, RE, LP, HB y derrumbes (ROC), tomando como características amplitudes y momentos estadísticos. En (Scarpetta et al., 2005) se emplean ANN para clasificar eventos VT de otras señales (explosiones submarinas, truenos) a partir de características espectrales y parámetros de amplitud. Algunos otros estudios similares usando redes neuronales se encuentran en (Romeo et al., 1995; Ezin et al., 2002; Del Pezzo et al., 2003; Avossa et al., 2003; Benbrahim et al., 2005; Esposito et al., 2006). A nivel exploratorio, también se han aplicado técnicas de clasificación basadas en disimilitudes (Orozco-Alzate et al., 2006) para la identificación automática de eventos LP y VT del volcán Nevado del Ruiz. Además, en diversos estudios se han empleado modelos ocultos de Markov (Ohrnberger, 2001; Alasonati et al., 2006; Gutiérrez et al., 2006; Benítez et al., 2007) y redes bayesianas (Riggelsen et al., 2007) pero todos ellos basados únicamente en señales provenientes de una única estación de registro.

Varios estudios han demostrado que los ensambles de clasificadores superan considerablemente el desempeño de los clasificadores individuales cuando operan en un escenario donde se presentan datos múltiples e independientes (Kittler et al., 1998). Partiendo de este hecho, en este estudio se clasifican dos clases de eventos con una alta frecuencia de aparición, LP y VT (Lesage, 2009), empleando reglas fijas de combinación de clasificadores. Los primeros se producen por efecto de actividad de fluidos dentro del volcán, mientras que los segundos son producidos por microfracturamientos internos (Trombley, 2006).

Algunas veces se hace difícil la clasificación o diferenciación entre sismos LP y VT puesto que, usualmente al viajar por el medio (la corteza terrestre), las señales sísmicas se modifican, y las formas, firmas y frecuencias de los sismos VT y LP pueden confundirse. En consecuencia, una señal LP puede lucir o aparentar ser una señal VT y viceversa. De ahí que se requiera del uso de métodos numéricos y técnicas de estadística computacional que permitan explorar y analizar en mayor detalle las señales sísmicas, permitiendo una diferenciación que a simple vista sería muy complicada.

Los datos fueron extraídos de las estaciones de registro denominadas Alfombrales (ALF), BIS y Tolda Fría (TOL), seleccionando sólo aquellos eventos que fueron registrados simultáneamente en las tres estaciones.

La organización de este artículo es la siguiente: en la sección 2 se presenta una revisión general de la combinación de clasificadores, describiendo métodos y estrategias; en la sección 3 se exponen los resultados experimentales; por último, en las secciones 4 y 5, se presentan las conclusiones y el trabajo futuro respectivamente.

2. COMBINACIÓN DE CLASIFICADORES

La combinación de clasificadores se puede sustentar a partir de sus ventajas estadísticas, computacionales y representacionales (Kuncheva, 2004). En lugar de usar un único clasificador, se pueden tomar varios clasificadores base y combinarlos de tal manera que se puedan aprovechar las ventajas de cada uno. Las técnicas de combinación de clasificadores están orientadas, especialmente, a problemas de clasificación que estén enmarcados en alguna de las siguientes situaciones (Jain et al., 2000):

- Tener la posibilidad de entrenar varios clasificadores.
- Tener varios conjuntos de entrenamiento, siendo necesario que se garantice la independencia de los datos.
- Cuando se tienen varios clasificadores y cada uno tiene un desempeño elevado en determinadas regiones del espacio de características.

2.1. Arquitecturas de combinación

La arquitectura de un combinador depende del problema que se esté tratando. De esta manera se tiene la arquitectura en pila, la paralela y la secuencial (o en cascada).

En la arquitectura en pila, todos los clasificadores base se entrenan con datos que provienen de un único espacio de características (sólo un conjunto de entrenamiento) y se combinan sus salidas. Por otro lado, en la arquitectura paralela se combinan clasificadores que han sido entrenados en diferentes espacios de características.

Por último se tiene la arquitectura secuencial, donde inicialmente se encuentra activo sólo un clasificador y los demás permanecen pasivos. A medida que se desarrolla el proceso de clasificación, cada clasificador opera directamente sobre la salida del clasificador anterior.

2.2. Fusión y selección

Las principales estrategias de combinación son la fusión y la selección de clasificadores (Kuncheva, 2004). En la fusión, todos los clasificadores base conocen el espacio de características en su totalidad, mientras que en la selección cada clasificador base conoce una parte del espacio de características y es responsable por los objetos en dicha región.

Como es de esperarse, la etiqueta asignada a un objeto en una combinación basada en selección está dada por la decisión que tome un sólo clasificador base, a diferencia de la fusión, donde la etiqueta a asignar es el resultado de alguna regla que involucre a todos los clasificadores base, pudiendo ser un promedio, un voto mayoritario, entre otros.

Adicionalmente, es importante diferenciar los métodos que generalmente se manejan al momento de iniciar el diseño de un clasificador combinado, ya que de esto depende lo que se desea optimizar. La primera clase de métodos, los no generativos, son aquellos que buscan optimizar el combinador a partir de un conjunto de clasificadores base predeterminados. La segunda clase, los generativos, se enfocan en diseñar y optimizar diferentes clasificadores base a partir de un combinador fijo (Valentini y Masulli, 2002).

2.3. Reglas fijas de selección y fusión de casificadores

Antes de mencionar las reglas de combinación que se emplean con más frecuencia, es importante tener en cuenta que las salidas de un clasificador pueden ser de diferentes tipos, y en consecuencia las reglas de combinación se eligen de acuerdo a las salidas de los clasificadores base. Las salidas de un clasificador pueden ser de nivel abstracto, de rango o de medida (Kuncheva, 2004; Jain et al., 2000).

En el nivel abstracto, cada clasificador produce una etiqueta para cada objeto. No hay información acerca de la certeza de las etiquetas asignadas ni alguna otra etiqueta recomendada. En el nivel de rango, para un objeto determinado el clasificador categoriza todas las clases de acuerdo al nivel de certeza que tiene dicho objeto de pertenecer a cada una de las clases. La etiqueta del objeto estaría dada por la clase que esté más alta en el ranking. Por último, en el nivel de medida, cada clasificador produce un vector c -dimensional $[d_{i,1}(\mathbf{x}), \dots, d_{i,c}(\mathbf{x})]$ ($c =$ número de clases). El valor $d_{i,j}(\mathbf{x})$ representa el soporte de decisión del clasificador i para la hipótesis de que \mathbf{x} pertenece a la clase ω_j .

Las reglas fijas se derivan de las reglas del producto y suma (Kittler et al., 1998). A partir de estas reglas, y considerando que un conjunto de clasificadores $D(\mathbf{x})$, con c clases y L clasificadores base, se describe por la matriz en (1), se deducen las reglas fijas de fusión y selección para determinar la certeza o soporte de decisión $\mu_j(\mathbf{x})$ de un objeto \mathbf{x} de pertenecer a la clase ω_j .

$$D(\mathbf{x}) = \begin{bmatrix} d_{1,1}(\mathbf{x}) & \cdots & d_{1,j}(\mathbf{x}) & \cdots & d_{1,c}(\mathbf{x}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{i,1}(\mathbf{x}) & \cdots & d_{i,j}(\mathbf{x}) & \cdots & d_{i,c}(\mathbf{x}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{L,1}(\mathbf{x}) & \cdots & d_{L,j}(\mathbf{x}) & \cdots & d_{L,c}(\mathbf{x}) \end{bmatrix} \quad (1)$$

En las reglas de selección se busca elegir un único clasificador base, de acuerdo al criterio máximo, mínimo o mediana. Es necesario que las salidas de los clasificadores base sean medibles, y por lo tanto $\mu_j(\mathbf{x})$ también será medible. En la Tabla 1 se describen las reglas de selección de clasificadores.

Tabla 1. Reglas de selección de clasificadores

Regla de selección	Descripción
Máximo	$\mu_j(\mathbf{x}) = \max_i \{d_{i,j}(\mathbf{x})\}$
Mínimo	$\mu_j(\mathbf{x}) = \min_i \{d_{i,j}(\mathbf{x})\}$
Mediana	$\mu_j(\mathbf{x}) = \text{med}_i \{d_{i,j}(\mathbf{x})\}$

Las reglas de fusión fijas no requieren de parámetros extra de entrenamiento; para el caso de fusión de valores continuos o medibles, se tiene la regla de la media, el producto y la media generalizada; ver (2) - (4).

- Suma (media):

$$\mu_j(\mathbf{x}) = \frac{1}{L} \sum_{i=1}^L d_{i,j}(\mathbf{x}) \quad (2)$$

- Producto:

$$\mu_j(\mathbf{x}) = \prod_{i=1}^L d_{i,j}(\mathbf{x}) \quad (3)$$

- Media generalizada:

$$\mu_j(\mathbf{x}, \alpha) = \left(\frac{1}{L} \sum_{i=1}^L d_{i,j}(\mathbf{x})^\alpha \right)^{1/\alpha} \quad (4)$$

Donde,

$$\alpha = \begin{cases} -\infty, & \text{mínimo} \\ \infty, & \text{máximo} \\ -1, & \text{media armónica} \\ 0, & \text{media geométrica} \\ 1, & \text{media aritmética} \end{cases}$$

La regla de combinación de salidas abstractas más empleada es el voto mayoritario (5), donde la etiqueta de un objeto x se asigna por democracia entre todos los clasificadores base. Asumiendo c clases y D_i clasificadores base, donde $i=1, \dots, L$, se obtiene el vector de etiquetas $[d_{i,1}, \dots, d_{i,c}]^T$, siendo $d_{i,j}=1$ si el clasificador D_i etiqueta x de clase ω_j .

$$\text{Asignar } x \rightarrow \omega_k \text{ si: } \sum_{i=1}^L d_{i,k} = \max_{1 \leq j \leq c} \sum_{i=1}^L d_{i,j} \quad (5)$$

3. RESULTADOS

Después de seleccionar los eventos sísmicos de clases LP y VT que fueron registrados simultáneamente en la estación Alfombrales (ALF), Bis (BIS) y Tolda Fría (TOL), se obtuvieron tres conjuntos de datos diferentes, uno para cada estación, garantizando de esta manera la independencia de los datos en los tres espacios de características. En total fueron seleccionados 350 eventos, cuyas señales en el dominio del tiempo fueron normalizadas a una media $\mu = 0$ y varianza $\sigma^2=1$. Todos los experimentos fueron realizados con el toolbox de Matlab PRTools¹.

Las diferencias en el contenido espectral se suelen usar en la discriminación visual de los diferentes tipos de sismos volcánicos (Zobin, 2003). En concordancia con ello, para este estudio se eligió la transformación al dominio de la frecuencia como método de representación, aplicando la transformada rápida de Fourier (FFT) de 2000 puntos a todos los eventos. El número de puntos para calcular la FFT se eligió con base en la longitud de la señal sísmica más corta; a pesar que la práctica usual es escoger una resolución diádica, es decir, equivalente a una potencia de 2. De esta manera, cada conjunto de datos contiene 350 eventos, cada uno con 2000 características. En la Figura 1 se muestra la representación en frecuencia de un evento de clase VT registrado simultáneamente en ALF, BIS y TOL.

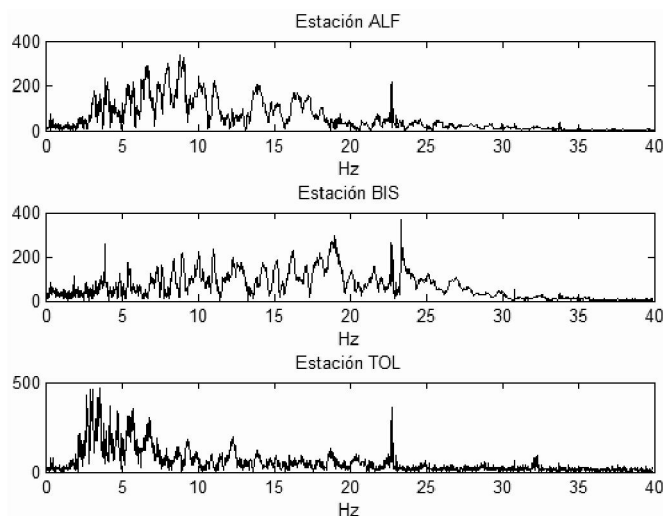


Figura 1. Evento VT registrado en ALF, BIS y TOL.

1. <http://www.prttools.org>

Posteriormente, para reducir el costo computacional y evitar la maldición de la dimensión (Jain et al., 2000), se seleccionaron 100 características, tomando como criterio de evaluación la distancia de Mahalanobis y como método de optimización el SFS (selección secuencial hacia adelante) (Van der Heijden et al., 2004). Las características seleccionadas corresponden entonces a valores de la energía de la señal en 100 frecuencias particulares. Con el fin de visualizar la separabilidad y similitud de las clases, se realizó un escalamiento multi-dimensional (MDS). Esta técnica estadística permite visualizar datos de un espacio n-dimensional, siendo posible determinar de manera cualitativa la separabilidad de las clases. En las Figuras 2, 3 y 4 se muestran los resultados del MDS aplicado a los datos de cada una de las estaciones estudiadas. Los ejes coordenados -etiquetados como Característica 1, Característica 2 y Característica 3- corresponden a combinaciones no lineales de las coordenadas originales, calculadas de manera que se preserven las distancias entre los objetos tan bien como sea posible.

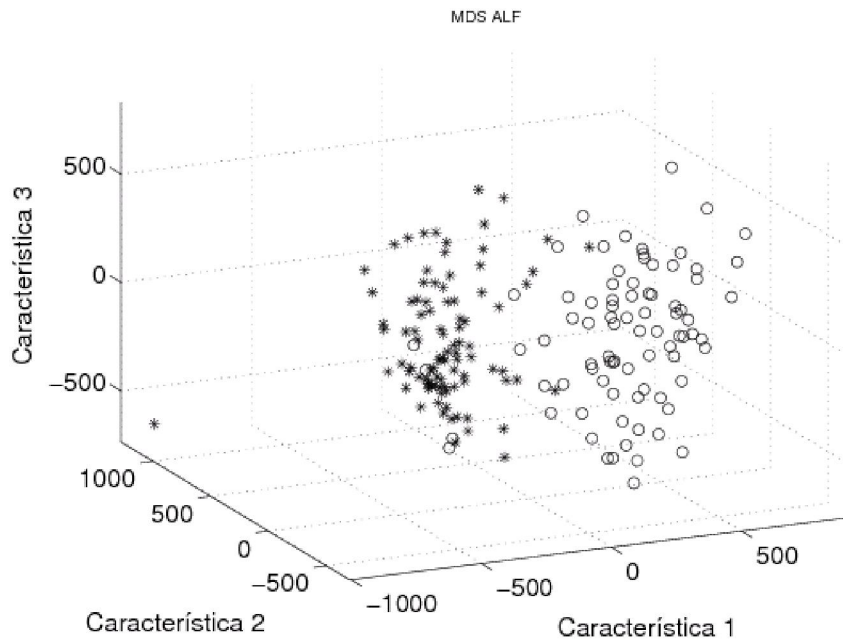


Figura 2. Escalamiento multi-dimensional (datos estación ALF).

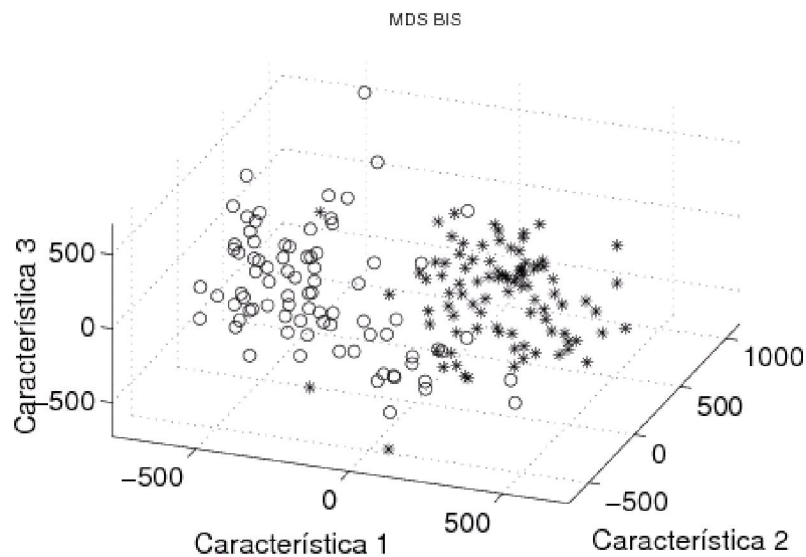


Figura 3. Escalamiento multi-dimensional (datos estación BIS).

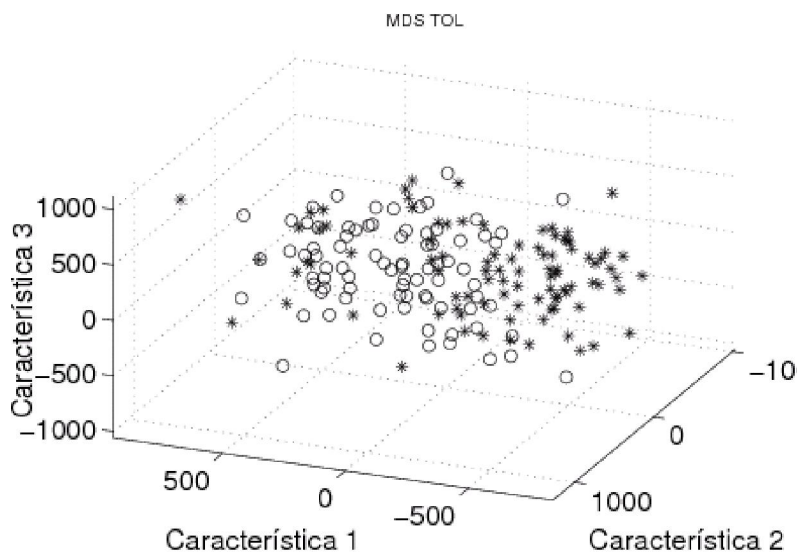


Figura 4. Escalamiento multi-dimensional (datos estación TOL).

3.1. Selección y entrenamiento de clasificadores base

Los clasificadores pre-seleccionados para operar como clasificadores base fueron el análisis discriminante lineal (LDA o Bayes-Normal-1) con regularización $r = 0.02$ (Friedman, 1989), el de media más cercana (NMC) y el clasificador de los cinco vecinos más cercanos (5-NN). El valor de regularización se escogió de acuerdo con la sugerencia general de hacerlo igual, en la práctica, a 0.05, 0.01 o menos (Palka y Duin, 2005). En las Figuras 5, 6 y 7 se muestran las curvas de aprendizaje para los clasificadores entrenados en los tres conjuntos de datos. El eje horizontal corresponde a diversos tamaños para el conjunto de entrenamiento.

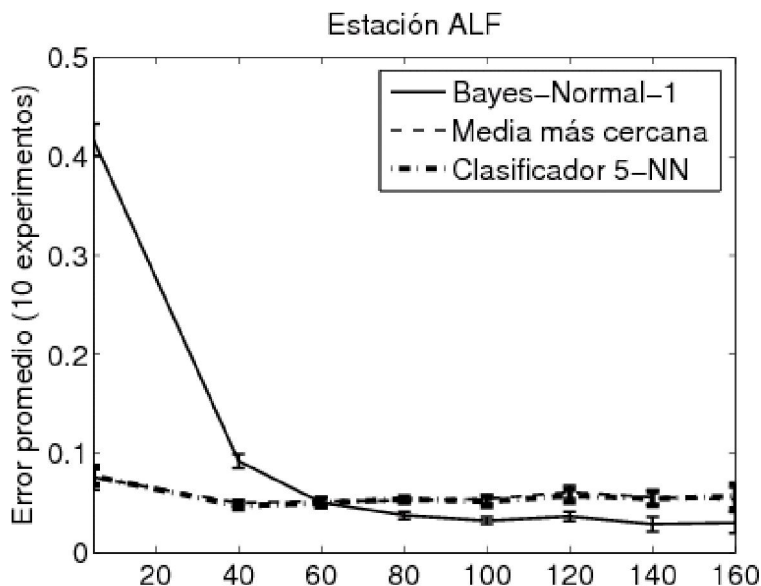


Figura 5. Curva de aprendizaje (datos estación ALF).

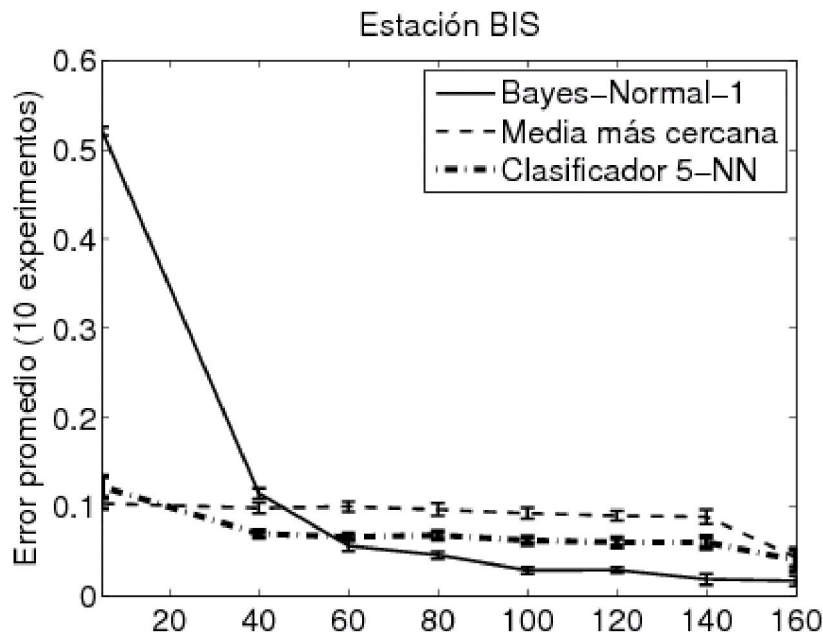


Figura 6. Curva de aprendizaje (datos estación BIS).

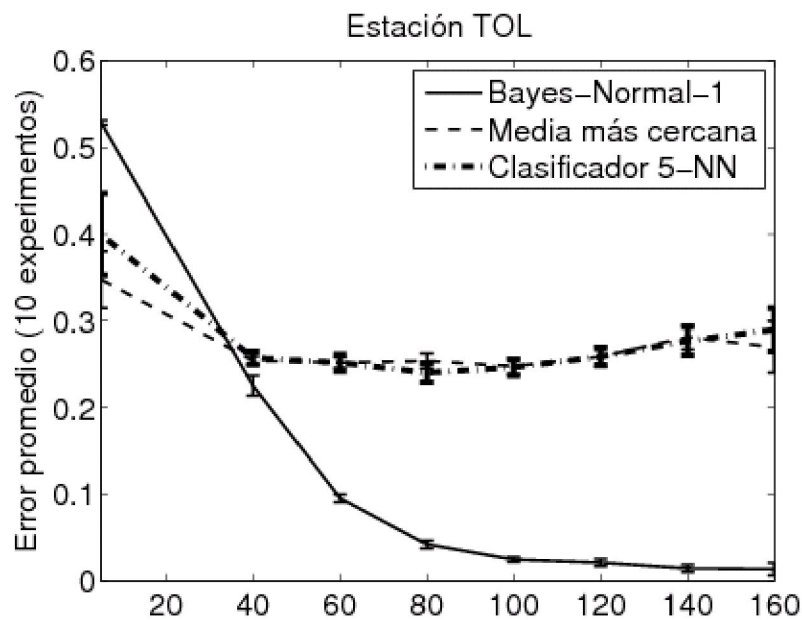


Figura 7. Curva de aprendizaje (datos estación TOL).

Como los eventos están balanceados (164 de clase LP y 186 de clase VT), se optó por particionar los conjuntos de datos de manera equitativa para generar los subconjuntos de entrenamiento y de prueba, 175 para entrenamiento y 175 para prueba. Los clasificadores base pre-seleccionados fueron entrenados en cada uno de los subconjuntos de entrenamiento (uno por cada estación). Como se esperaba, según las curvas de aprendizaje, el clasificador con mejor desempeño fue el LDA con regularización $r=0.02$. En las Tablas 2, 3 y 4 se muestran los resultados de validación cruzada con dos particiones, 175-175, y veinte repeticiones².

2. Se usó la función *crossval* del toolbox *PRTools*

Tabla 2. Validación cruzada - Datos estación ALF

Clasificador	Error	σ
LDA	0.0374	± 0.0081
NMC	0.0499	± 0.0022
5-NN	0.0474	± 0.0036

Tabla 3. Validación cruzada - Datos estación BIS

Clasificador	Error	σ
LDA	0.0349	± 0.0082
NMC	0.0894	± 0.0050
5-NN	0.0607	± 0.0057

Tabla 4. Validación cruzada - Datos estación TOL

Clasificador	Error	σ
LDA	0.0363	± 0.0095
NMC	0.2593	± 0.0091
5-NN	0.2371	± 0.0133

Es particularmente interesante notar el desempeño del clasificador LDA para la estación TOL (Tabla 4). La diferencia de desempeño del clasificador LDA respecto a los otros clasificadores -en esta estación- podría explicarse como una mejor capacidad de generalización ante datos contaminados o ruidosos. En contraste, los clasificadores NMC y 5-NN se ven más afectados por las perturbaciones puesto que sus fronteras de decisión se construyen con base en un criterio de distancia en lugar de basarse en criterios de densidad de probabilidad. Desde el punto de vista físico, este hecho puede atribuirse al fenómeno denominado efecto de sitio (Londoño, 1996), el cual es un parámetro que modifica las señales sísmicas debido a la geología local donde se encuentra la estación sísmica que registra la señal. El efecto de sitio puede ser removido con técnicas de procesamiento de la señal, o bien podría usarse un clasificador insensible al mismo como parece ser en este caso.

3.2. Combinación de clasificadores base

Como clasificador base fue seleccionado el LDA, (6), con regularización $r=0.02$, ya que su desempeño en los tres subconjuntos de entrenamiento es elevado, sobretodo en el subconjunto de la estación TOL, donde su desempeño es notablemente superior con relación al NMC y el 5-NN. La regularización aumenta el desempeño del clasificador porque mejora la estimación de la matriz de covarianza.

$$g_i(\mathbf{x}) = \log(P(\omega_i)) - \frac{1}{2} \mu_i^T \Sigma^{-1} \mu_i + \mu_i^T \Sigma^{-1} \mathbf{x} \quad (6)$$

El soporte de decisión $d_{i,j}(\mathbf{x})$ de un clasificador LDA es numérico y representado por la probabilidad a posteriori $P(\omega_j|\mathbf{x})$. Aprovechando este tipo de salida, los combinadores fijos que se probaron partieron de la regla del producto y de la suma, obteniendo los mejores resultados con la primera.

A partir de (3), se puede deducir (7), que es la regla del producto en términos de las probabilidades a posteriori producidas por cada clasificador base (L: Clasificador base, \mathbf{x} : Medida, c: Clase.)

$$\text{Asignar } x \rightarrow \omega_j \text{ si: } P^{-(L-1)}(\omega_j) \prod_{i=1}^L P(\omega_j|\mathbf{x}) = \max_{1 \leq j \leq c} [P^{-(L-1)}(\omega_j) \prod_{i=1}^L P(\omega_j|\mathbf{x})] \quad (7)$$

El LDA con regularización $r=0.02$ se entrenó en diez subconjuntos de entrenamiento diferentes³ (un subconjunto de entrenamiento tiene a su vez tres subconjuntos, uno para cada estación), y posteriormente fueron combinados bajo la regla (7). Los clasificadores combinados fueron probados en los correspondientes subconjuntos de prueba, arrojando un error promedio de clasificación de 0.51%.

4. CONCLUSIONES

El empleo de múltiples conjuntos de entrenamiento aumenta la capacidad de generalización de un sistema de clasificación, y esto se puede aprovechar mediante el uso de técnicas de combinación de clasificadores. Se pudo evidenciar que el clasificador combinado resultante siempre superó a todos los clasificadores individuales que se entrenaron (5-NN, discriminante lineal, media más cercana).

La selección de características usando la distancia de Mahalanobis como criterio de evaluación fue uno de los puntos clave para que el clasificador LDA tuviera un desempeño alto con los datos de la estación TOL, los cuales son los más problemáticos debido a la poca separabilidad de sus clases. El criterio de selección de características, que es de naturaleza probabilística, significó un mejoramiento inmediato del clasificador LDA.

El combinador se ejecutó bajo la regla del producto, ya que fue la que mostró mejor desempeño en comparación con la regla de la suma. El hecho de la independencia de los datos favoreció la aplicación de estas reglas (producto y suma), ya que parten de la suposición de dicha independencia.

Estas reglas fueron las escogidas para los experimentos, dado que las salidas del clasificador base son probabilísticas y la mejor manera de aprovechar su información es a través de la combinación, en lugar de reglas de selección (máximo, mínimo o mediana). Las reglas de selección de clasificadores son más adecuadas para salidas de nivel de rango.

5. TRABAJO FUTURO

Trabajos posteriores buscarán solucionar el problema multiclase. Se estudiarán otros métodos de representación y se usarán reglas de combinación entrenadas. Además, se compararán diferentes formas de atacar el caso de los datos problemáticos, debidos seguramente a efectos de sitio. Así, una combinación de métodos sísmicos de remoción del efecto de sitio con las técnicas de fusión de clasificadores podrían generar una clasificación muy acertada. También sería importante vincular más estaciones, sobre todo la estación ubicada en el cráter Olleta que sirve de referencia en el OVSM del INGEOMINAS.

AGRADECIMIENTOS

Este trabajo se realizó en el marco del proyecto de investigación "Sistemas múltiples de clasificación para el reconocimiento automático de eventos sísmicos en el complejo volcánico Cerro Machín - Cerro Bravo (código: 8970)" y el Semillero de Investigación en Software para Reconocimiento de Patrones; ambos financiados por la Dirección de Investigación de la Universidad Nacional de Colombia Sede Manizales.

3. El conjunto de datos fue particionado en partes iguales (175-175) para los subconjuntos de entrenamiento y de prueba. Cada subconjunto de entrenamiento tiene su correspondiente conjunto de prueba, garantizando de esta manera que sean disjuntos.

REFERENCIAS

- Alasonati, P., Wassermann, J. and Ohrnberger, M., 2006. Signal classification by wavelet-based hidden Markov models: application to seismic signals of volcanic origin. *Statistics in Volcanology*. Geological Society of London.
- Avossa, C., Giudicepietro, F., Marinaro, M. and Scarpetta, S., 2003. Supervised and unsupervised analysis applied to Strombolian E.Q. 14th Italian Workshop on Neural Nets, WIRN VIETRI 2003, Lecture Notes in Computer Science, vol. 2859, Springer. pp. 173-178.
- Benbrahim, M., Daoudi, A., Benjelloun, K. and Ibenbrahim, A., 2005. Discrimination of seismic signals using artificial neural networks. The Second World Enformatika Conference, WEC'05, Enformatika, Çanakkale, Turquía. pp. 4-7.
- Benítez, M. C., Ramírez, J., Segura, J. C., Ibañez, J. M., Almendros, J., García-Yeguas, A. and Cortés, G., 2007. Continuous HMM-based seismic-event classification at Deception Island, Antarctica. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 1, pp. 138-146.
- Del Pezzo, E., Esposito, A., Giudicepietro, F., Marinaro, M., Martini, M. and Scarpetta, S., 2003. Discrimination of earthquakes and underwater explosions using neural networks. *Bulletin of the Seismological Society of America*, vol. 93, no. 1, pp. 215-223.
- Esposito, A. M., Giudicepietro, F., Scarpetta, S., D'Auria, L., Marinaro, M. and Martini, M., 2006. Automatic discrimination among landslide, explosion-quake, and microtremor seismic signals at Stromboli Volcano using neural networks. *Bulletin of the Seismological Society of America*, vol. 96, no. 4, pp. 1230-1240.
- Ezin, E. C., Giudicepietro, F., Petrosino, S., Scarpetta, S. and Vanacore, A., 2002. Automatic discrimination of earthquakes and false events in seismological recording for volcanic monitoring. WIRN VIETRI 2002: Proceedings of the 13th Italian Workshop on Neural Nets-Revised Papers, Lecture Notes in Computer Science, vol. 2486, pp. 140-145, Londres, UK. Springer-Verlag.
- Friedman, J., 1989. Regularized discriminant analysis. *Journal of the American Statistical Association*, vol. 84, no. 405, pp. 165-175.
- Gutiérrez, L., Ramírez, J., Benítez, C., Ibañez, J., Almendros, J. and García-Yeguas, A., 2006. HMM-based classification of seismic events recorded at Stromboli and Etna volcanoes. *IEEE International Geoscience & Remote Sensing Symposium, IGARSS 2006*, pp. 2765-2768.
- Jain, A. K., Duin, R. P. W. and Mao, J., 2000. Statistical Pattern Recognition: A Review. *IEEE transactions on pattern recognition and machine intelligence*, vol. 22, no. 1, pp. 4-37.
- Kittler, J., Hatef, M. and Duin, R. P. W., 1998. On Combining Classifiers. *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 3, pp. 226-239.
- Kuncheva, L. I., 2004. *Combining Pattern Classifiers. Methods and algorithms*. John Wiley & Sons, New Jersey.
- Langer, H., Falsaperla, S., Powell, T. and Thompson, G., 2006. Automatic classification and a-posteriori analysis of seismic event identification at Soufrière Hills volcano, Montserrat. *Journal of volcanology and geothermal research*, vol. 153, no. 1-2, pp. 1-10.
- Lesage, P., 2009. Interactive Matlab software for the analysis of seismic volcanic signals. *Computers & Geosciences*, vol. 35, no. 10, pp. 2137-2144.
- Londoño, J.M., 1996. Temporal change in coda Q at Nevado Del Ruiz Volcano, Colombia. *Journal of Volcanology and Geothermal Research*, vol. 73, no. 1-2, pp. 129-139.
- Ohrnberger, M., 2001. Continuous Automatic Classification of Seismic Signals of Volcanic Origin at Mt. Merapi, Java, Indonesia. PhD thesis, University of Potsdam, Postdam, Germany.
- Orozco-Alzate, M., García-Ocampo, M. E., Duin, R. P. W. and Castellanos-Domínguez, C.G., 2006. Dissimilarity-

- based classification of seismic volcanic signals at Nevado del Ruiz Volcano. *Earth Sciences Research Journal*, vol. 10, no. 2, pp. 57-65.
- Palka, E. and Duin, R. P. W., 2005. *The Dissimilarity Representation for Pattern Recognition: Foundations and Applications*, vol. 64, serie: *Machine Perception and Artificial Intelligence*. World Scientific, Singapore.
- Riggelsen, C., Ohrnberger, M. and Scherbaum, F., 2007. Dynamic bayesian networks for real-time classification of seismic signals. *PKDD 2007: Proceedings of the 11th European conference on Principles and Practice of Knowledge Discovery in Databases, Lecture Notes in Artificial Intelligence*, Berlín, Heidelberg. Springer-Verlag. vol. 4702, pp. 565-572.
- Romeo, G., Mele, F. and Morelli, A., 1995. Neural networks and discrimination of seismic signals. *Computers & Geosciences*, vol. 21, no. 2, pp. 279-288.
- Scarpetta, S., Giudicepietro, F., Ezin, E. C., Petrosino, S., Del Pezzo, E., Martini, M. and Marinaro, M., 2005. Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks. *Bulletin of the Seismological Society of America*, vol. 95, no. 1, pp. 185-196.
- Trombly, R. B., 2006. *The Forecasting of Volcanic Eruptions*. iUniverse.
- Valentini, G. and Masulli, F., 2002. Ensembles of learning machines. *WIRN VIETRI 2002: Proceedings of the 13th Italian Workshop on Neural Nets-Revised Papers, Lecture Notes in Computer Science*, Londres, UK. Springer-Verlag. vol. 2486, pp. 3-22.
- Van der Heijden, F., Duin, R. P. W., De Ridder, D. and Tax, D. M. J., 2004. *Classification, parameter estimation and state estimation*. John Wiley & Sons, England.
- Zobin, V., 2003. *Introduction to volcanic seismology*. Elsevier, The Netherlands.