RELACIÓN DE LA ESCALA DE INTENSIDAD DE MERCALLI Y LA INFORMACIÓN INSTRUMENTAL COMO UNA TAREA DE CLASIFICACIÓN DE PATRONES

Jorge E. Hurtado[1], Daniel Bedoya R.[2]

Recibido: 15/04/2008
Aceptado: 24/11/2008

RESUMEN

A pesar de los progresos ocurridos en la instrumentación sísmica, la valoración de vulnerabilidad sísmica y el daño con índices cualitativos, tal como los proporcionados por Intensidad de Mercalli Modificada (IMM), siguen siendo altamente favorables y útiles para los propósitos prácticos. Para vincular las medidas cualitativas de acción del terremoto y sus efectos, es habitualmente aplicada la regresión estadística. En este artículo, se adopta un planteamiento diferente, el cual consiste en expresar la Intensidad de Mercalli, como una clase en vez de un valor numérico. Una herramienta de clasificación estadística moderna, conocida como máquina de vectores de soporte, se usa para clasificar la información instrumental con el fin de evaluar la intensidad de Mercalli correspondiente. Se muestra que el método da resultados satisfactorios con respecto a las altas incertidumbres y a la medida del daño sísmico cualitativo.

Palabras clave: Intensidad de Mercalli, daño estructural, aprendizaje estadístico, máquinas de vectores de soporte, reconocimiento de modelo.

1 Universidad Nacional de Colombia. Apartado 127, Manizales, Colombia. e-mail: idea@nevado.manizales.unal.edu.co
2 Universidad de Medellín. Medellín, Colombia. Apartado 1983 Medellín, Colombia. e-mail: dabedoya@udem.edu.co

THE RELATIONSHIPS OF MERCALLI INTENSITY TO INSTRUMENTAL INFORMATION AS A PATTERN CLASSIFICATION TASK

ABSTRACT

Despite the progress occurred in seismic instrumentation, the assessment of seismic vulnerability and damage qualitative indexes, such as that provided by Mercalli intensity is highly valuable and useful for practical purposes. In order to link the qualitative measures of earthquake action and its effects, statistical regression is commonly applied. in the paper, a different approach is adopted. It consisten in regarding the Mercalli intensity as a class rather that a numerical value. A modern statistical classification tool known as Support Vector Machine is used for classifying the la instrumental information in order to assess the corresponding Mercalli intensity. It is shown that the method gives satisfactory results with regardto the high uncertainties linked to such a qualitative seismic damage measure.

Keywords: Mercalli Intensity, earthquake damage, statistical learning, support vector machines, pattern recognition.

1. INTRODUCCIÓN

La complejidad de la acción de los terremotos y sus efectos dañinos sobre las estructuras, requiere la adopción de varias perspectivas para su entendimiento. Desde el principio de la ingeniería de terremotos se ha entendido que la instrumentación sísmica no es suficiente para describir tal complejidad de los fenómenos como lo es el daño urbano y regional. La mayoría de la escalas de intensidad en la actualidad representan una descripción subjetiva de la respuesta humana al movimiento y a la descripción asociada al daño de los edificios. Por esta razón, las medidas cualitativas de la acción de los terremotos son desarrolladas sobre la base de la observación de los efectos.

Algunas de las escalas propuestas son bien conocidas. En muchos campos se utiliza la escala de Intensidad de Mercalli Modificada (IMM). Su conocimiento es útil para propósitos de prevención de desastres. La intensidad de Mercalli es habitualmente evaluada después de la ocurrencia de un terremoto importante, en escalas regionales o urbanas. También se estima sobre la base de información histórica de terremotos ocurridos en el pasado.

Los avances en sismología e ingeniería de instrumentación ofrecen la posibilidad de relacionar la intensidad de Mercalli con datos instrumentales como un medio para el entendimiento y la evaluación destructiva de los terremotos. Tales relaciones habitualmente han sido llevadas a cabo usando técnicas de regresión estadística convencionales. El paso inicial en esta dirección relaciona la intensidad *I* con los datos macrosísmicos. Por ejemplo, una relación obtenida para México es (Esteva 1976)

$$I = 1.45M - 5.7\log_{10} R + 7.9 \qquad (1)$$

Donde: M es la magnitud del terremoto y R es la distancia epicentral. La ventaja de la instrumentación de movimientos fuertes es permitir la construcción de relaciones en términos de la aceleración pico del suelo, tal como se muestra en la ecuación 2. Esta expresión es ampliamente usada en el oeste de los Estados Unidos (Trifunac 1975).

$$\log_{10} A = 0.01 + 0.30I \qquad (2)$$

Donde: A es la aceleración horizontal pico del suelo.

Relaciones similares se han desarrollado para la escala de intensidades sísmica de Japón (JMA), utilizando modelos de regresión univariante y multivariante (Yamazaki 2002).

$$I_{JMA} = 0.63 + 1.81\log_{10} A \qquad (3)$$

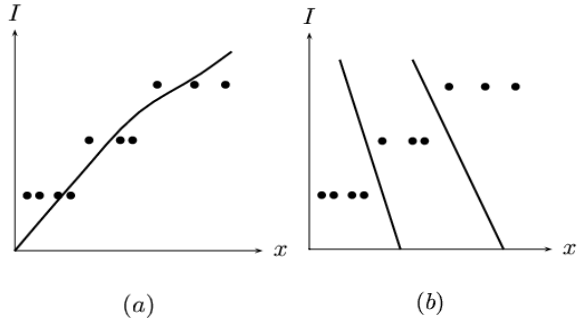
El acercamiento a la estimación del riesgo sísmico dependerá de la experiencia y el conocimiento basado en la certeza del conjunto de datos disponibles, independientemente de su tamaño. Una regresión más sofisticada usando un análisis de componentes principales (Iyengar 1983) ha sido revisada, debido a que los autores relacionan la intensidad de Mercalli con un determinado número de variables instrumentales: doce variables. Esto permite más exactitud en la descripción de la IMM en términos de la información instrumental. La base de datos proporcionada por estos autores será utilizada en este trabajo para demostrar las capacidades del enfoque propuesto.

Además de los métodos de regresión clásico lineal y de la componente principal, la regresión también puede ser hecha por medio de otras técnicas tales como redes neuronales, las cuales han demostrado ser útiles en los análisis de regresión en Ingeniería de Terremotos (Dowla 1995, Emami 1996, Hurtado 2001) y campos afines (Yagawa 1996, Hurtado 2001, 2002). En general, algunos tipos de redes neuronales, a pesar de ser popularmente considerados como técnicas de inteligencia artificial, pueden ser legítimamente considerados como un medio para la implementación adaptativa, esto es, una función de regresión dependiente de la muestra que ha localizado un soporte efectivo.

Como consecuencia, las redes neuronales están siendo examinadas en la estructura de la teoría del aprendizaje estadístico (TAE) como una novedosa herramienta para resolver las tareas de la estadística tradicional (Vapnik 1998, Anthony 1999).

Actualmente, la regresión neuronal y los patrones de clasificación son otras de estas tareas para las cuales la redes neuronales han demostrado ser altamente útiles (Bishop 1995) y en las que se investiga de forma intensiva para ser aplicadas en TAE. En este caso, el propósito no es obtener una relación funcional entre los datos de entrada y la salida, sino ubicar los datos de entrada en una clase, la que se distingue de otras clases por una etiqueta. Esta es una consideración que puede ser adoptada en el caso de la intensidad sísmica de Mercalli, debido a que su asignación cualitativa fue hecha por un experto de acuerdo con las observaciones. Se debe destacar la naturaleza cualitativa de la escala IMM, la cual, habitualmente se escribe con números romanos. Esta naturaleza podría también justificarse relacionando la información instrumental a través de técnicas de lógica difusa (Lee 1996).

En resumen, la IMM puede ser relacionada con información instrumental a través de varias técnicas de regresión: paramétricas, no paramétricas, neuronales y difusas. También, la intensidad puede ser considerada como una clase en el sentido estadístico. La diferencia entre las dos técnicas es ilustrada en la figura 1.



Fuente: elababoración propia.

Figura 1. Enfoque estadístico para la valorar la IMM: (a) Por regresión; (b) Por clasificación.

El enfoque de la regresión es probado para atacar una función continua con datos aleatorios, la cual, debido a la naturaleza discretade la IMM, se extiende en la forma indicada en la figura 1a. La regresión lineal clásica difiere de la regresión de la componente principal en que la primera minimiza el promedio de la distancia vertical sobre la función de regresión, mientras la segunda minimiza la distancia ortogonal. Independientemente del enfoque de regresión, cuando se desea obtener una función continua con una variable discreta, esta no se ajusta perfectamente bien a la naturaleza del problema.

Por esta razón, una clase de tarea de clasificación, que redondea la estimación, siempre debe realizarse usando la función de regresión. De lo contrario, el enfoque de clasificación mostrado en la figura 1b es aparentemente más natural para este problema, debido a que la naturaleza cualitativa de la IMM no es forzada. En este enfoque, el énfasis queda directamente en la construcción de las reglas de decisión; en cambio, en la estimación de una función continua, se hace antes de aplicar la decisión del redondeo automatizado. El enfoque de clasificación es más general. En realidad, no requiere valores numéricos de la intensidad de los terremotos para aplicarse un etiquetado por ejemplo A, B, C, etc., para que el método considere la IMM como una variable numérica discreta; es una convención.

La clasificación de información instrumental de acuerdo con los niveles de intensidad puede ser buscado después con métodos de patrones de organización, tales como la discriminación clásica bayesiana, clasificación en árboles, redes neuronales (Ripley 1996) y máquinas de vector soporte (Vapnik 2000). Después de algunas pruebas realizadas con estas técnicas, se halló que el último método ofrece la mejor tasa de cambio de clasificación, y por ello se ha adoptado. Esta superioridad tiene raíces en los principios legítimos sobre los cuales los métodos de vector soporte están basados, como se resume a continuación. A diferencia de las redes neuronales, las máquinas

de vector soporte no han sido desarrolladas con el fin de imitar los procedimientos de aprendizaje del cerebro, pero sí directamente sobre la base de los principios del aprendizaje estadístico (Vapnik 1998). A pesar de que su ecuación de clasificación final se parece a la de las redes neuronales, su principal diferencia respecto a estas es que las reglas de clasificación pueden ser expresadas en términos de algunas muestras cerca a los límites entre clases. Por el contrario, en redes neuronales minimizar el error implica un compromiso sobre todos los patrones de entrenamiento.

En la siguiente sección presentamos un estado del conocimiento sobre los métodos de vector soporte, aplicados al problema de reconocimiento de patrones. Luego se realiza laclasificación de la información sísmica instrumental, de acuerdo con los niveles de IMM, usando la base de datos de la referencia de Iyenger (Iyenger 1983). La exactitud de la valoración es discutida. El artículo finaliza con algunas conclusiones.

2. ESTADO DEL CONOCIMIENTO SOBRE CLASIFICADORES DE VECTORES DE SOPORTE

Como se ha dicho antes, existen varios métodos estadísticos para realizar la clasificación dentro de grupos de muestras. Un requerimiento habitual en la solución de problemas estadísticos es la necesidad de tener una gran población de muestras disponibles con anterioridad para poder reducir la incertidumbre sobre la predicción. En nuestro caso, sin embargo, es raro poseer abundantes intensidades de Mercalli, por lo que se hace necesario aplicar un método que sea óptimo para pequeñas poblaciones. Por esta razón, es conveniente tener como recurso los métodos de reconocimiento de patrones desarrollados en la estructura de la teoría del aprendizaje estadístico (Vapnik 1998, Anthony 1999), debido a que simplemente apuntan a resolver problemas estadísticos con pequeños tamaños de muestras.

Para las tareas de reconocimiento de patrones, el principal instrumento elaborado de la teoría del aprendizaje estadístico es el llamado clasificador o máquina de vector de soporte (MVS). A pesar de que existen algunos otros métodos de clasificación y variantes de la técnica básica desarrollados en esta estructura en los años recientes (Smola 2000, Muller 2001), los CVS clásicos presentados por Vapnik (Vapnik 1998), serán usados aquí. El resto de esta sección está dedicada a un resumen de este método.

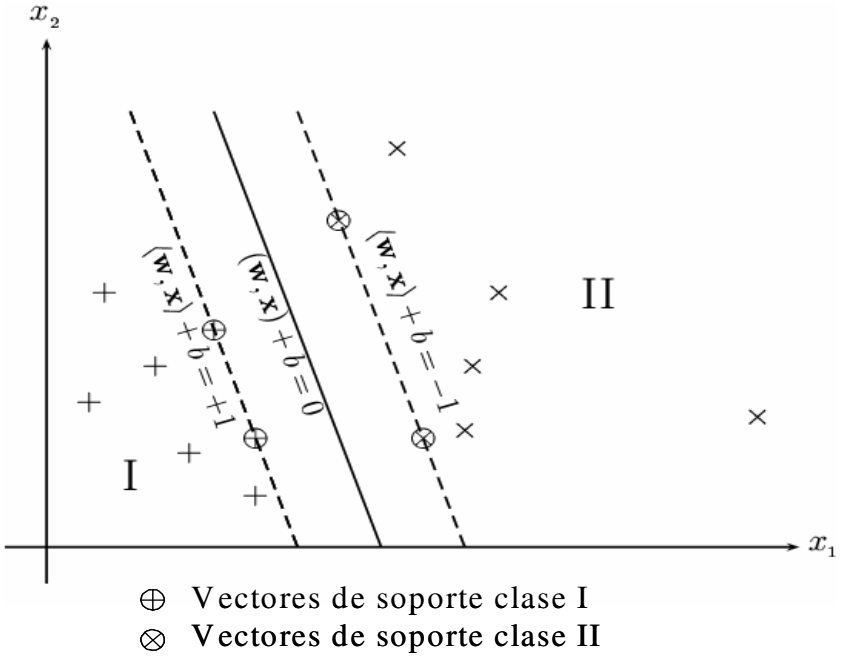
Es necesario empezar el resumen de MVS, a través del problema de clases separables linealmente, es decir, que pueden ser separadas por un hiperplano (figura 2).

Sean los patrones $\{x_1, x_2, ..., x_l\}$ y denotemos las clases para las cuales ellos pertenecen como $c_i \in\{-1, 1\}$, $i = 1, 2, ..., l$.Se busca un hiperplano de separación de la forma

$$f(x)=\langle w \,.\, x\rangle + b = 0 \tag{4}$$

Donde: w es un vector de parámetros que define el vector normal al hiperplano, b es el conocido intercepto del análisis de regresión clásico y <.,.> es el vector normal del producto interno. La condición impuesta para este hiperplano es maximizar la distancia para el patrón dado; esto permite la mejor confianza sobre la clasificación que realiza. Así, la optimización del problema

$$\underset{w,b}{máx}\ \underset{i}{mín}\left\{\|x - x_i\| : \langle w \,.\, x\rangle + b = 0,\ i = 1, 2,...,l\right\} \tag{5}$$



Fuente: elaboración propia.

Figura 2. Hiperplano óptimo de separación

La definición de las clases a través de un signo facilita la formulación de este problema de optimización. Sea γ_i el margen para cada patrón , el cual es expresado como:

$$\gamma_i = c_i\left(\langle w \cdot x\rangle + b\right) \tag{6}$$

Es evidente que una margen positiva siempre indica clasificación derecha, no obstante el signo de la clase. En este punto es importante normalizar los parámetros del hiperplano por $\frac{1}{\|w\|}$ Con esta operación los puntos cerca al hiperplano satisfacen $\left\| \langle w \cdot x\rangle + b \right\| = 1$, como se muestra en la figura 2, y la margen óptima se vuelve

$$\gamma_i = \frac{2}{\|w\|} \tag{7}$$

Como puede ser fácilmente demostrado. Ahora el problema de optimización es el siguiente:

Minimice $\Omega(w) = \frac{\|w\|^2}{2}$

Sujeto a $c_i\left(\langle w \cdot x\rangle + b\right) \geq 1,\ i = 1, 2, ..., l$

Esta restricción en el problema de optimización puede ser reformada como un problema no restringido a través de los multiplicadores de Lagrange $\alpha_i \geq 0$ como

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l} \alpha_i \left[c_i\left(\langle w.x_i\rangle + b\right) - 1\right] \tag{8}$$

El problema es hallar los puntos de equilibrio necesarios para minimizar la función de pérdida con respecto a los parámetros del hiperplano (*w, b*) mientras se maximiza con respecto a los multiplicadores de Lagrange. La solución de este problema es:

$$w = \sum_{i=1}^{l} \alpha_i c_i x_i \tag{9}$$

Tomando en cuenta la certeza de los multiplicadores de Lagrange la parte anterior de la ecuación 9 tiene un importante significado: el

vector de pesos puede ser expandido únicamente en términos de los patrones que tienen un multiplicador de Lagrange positivo, mientras el resto no es necesario. Tal patrón especial define el hiperplano que soporta los vectores. En realidad ellos tienen la propiedad especial de quedar simplemente adelante de la margen definida anteriormente. Esto puede ser fácilmente demostrado por medio de la bien conocida condición complementaria Karush-Kuhn-Tucker de la teoría de optimización (Kall 1995), la cual en este caso lleva

$$\alpha_i \left[c_i \left(\langle w . x \rangle + b \right) - 1 \right] = 0, \; i = 1, 2, ..., l \tag{10}$$

Esta condición implica que la solución para los pesos del hiperplano puede ser presentada en la forma

$$w = \sum_{j \in SV} \alpha_j c_j x_j \tag{11}$$

Donde SV es el conjunto de vectores soporte. El umbral *b* puede ser entonces calculado después por medio de la ecuación (10).

Para complementar la solución del problema, es conveniente expresar el problema de optimización en términos de los multiplicadores de Lagrange (variables duales). Reemplazando la ecuación (9) dentro de la ecuación (8). El resultado es Maximice

$$W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i c_i \langle x_i . x_j \rangle \alpha_j c_j \tag{12}$$

sujeto a

$$\alpha_i \geq 0$$

$$\sum_{i=1}^{l} \alpha_i c_i = 0 \tag{13}$$

La solución de este problema de optimización cuadrática debe ser sustituida dentro de la ecuación (10) y (11) para obtener los valores de los parámetros del hiperplano. La función de clasificación es entonces:

$$g(x) = \text{sgn}(f(x)) = \text{sgn}\left(\sum_{i \in SV} \alpha_i c_i \left(\langle x_i . x \rangle + b \right) \right) \tag{14}$$

Todo lo dicho anteriormente concierne a la separación de clases por medio de un hiperplano. Este es el caso de clasificación que nosotros enfrentamos en este artículo; el límite es altamente no lineal, como será mostrado a continuación; esto es necesario para entender la anterior formulación. Esto puede ser hecho a través de la generalización de la última ecuación a

$$g(x) = \text{sgn}(f(x)) = \text{sgn}\left(\sum_{i \in SV} \alpha_i c_i \left(\langle \phi(x_i) . \phi(x) \rangle + b \right) \right) \tag{15}$$

Donde la función $\phi.(\cdot)$ realiza un muestreo desde el espacio de patrones al llamado espacio característico. El muestreo implica una proyección sobre un espacio de alta dimensión para el cual la separación de clases se vuelve más fácil. En realidad, un teorema de la teoría del aprendizaje de clasificación usando funciones lineales indica que la probabilidad correcta de clasificación incrementa si la relación del número de muestras n a la dimensión d del espacio de entrada decrece (Vapnik 1998, Fine 1999). Si el número de muestras es limitado, un camino conveniente para mejorar la capacidad de clasificación es aumentar la dimensionalidad a través de un muestreo no lineal, hecho precediendo la ecuación. Note, sin embargo, que la función $\phi.(\cdot)$ que aparece en la ecuación (15) es únicamente la representación del producto interno. Consecuentemente, se puede usar el hecho de Kernels

$$K(x, z) = \langle \phi(x) . \phi(z) \rangle \tag{16}$$

en cambio de la función no lineal. Aquí, el producto interno es definido en un espacio de Hilbert. Con esta modificación la ecuación (15) se vuelve

$$g(x) = \text{sgn}(f(x)) = \text{sgn}\left(\sum_{i \in SV} \alpha_i c_i \left(K(x_i, x) + b \right) \right) \tag{17}$$

Esto significa que el muestreo actual $x \mapsto \phi(x)$ no necesita, explícitamente, serconocido.

En el campo del aprendizaje estadístico se han demostrado las significantes ventajas para el aprendizaje desde las muestras, si los Kernels derivados pueden ser flexibles, esto es si ellos son parametrizados por las muestras dadas (Cherkassky 1998). Esto está garantizado en la ecuación (17) por la presencia de x_i como una variable independiente dentro de Kernel. Tal flexibilidad explica los éxitos en las aplicaciones en redes neuronales para tareas de aprendizaje en estructuras o mecánica de suelos como se citó en la introducción. Este éxito, junto con el parecido de la ecuación (15) para aquellos muestreos con redes neuronales, justifica el uso común de redes neuronales con Kernel para clasificadores de vector soporte. En este artículo, se hace uso de la función base de radial de Kernel dada por

$$K(x,z)=\exp\left(-\frac{\|x-z\|}{2\sigma^2}\right) \qquad (18)$$

Al final de esta sección un importante tema que es habitual para funciones de regresión y clasificación será tratado. Esta es la habilidad de generalización de la función, por ejemplo, su habilidad para ofrecer bajos errores cuando es usada para valoración con datos no empleados en la fase de entrenamiento. Para los propósitos del presente artículo, el problema de generalización es crucial para el uso práctico de los clasificadores cuando reciben nueva información instrumental. Sin embargo, un resumen sobre la teoría de generalización podría requerir varias páginas.

La habilidad de generalización de los clasificadores de vector soporte está controlada por la llamada dimensión Vapnik-Chervonenkis (VC) de una clase de función. Esto es definido como el máximo número de dicotomías que pueden ser implementados en un espacio dado por una función $f(x, w)$ de una clase definida por un conjunto de parámetros $w \in \Omega$. Por ejemplo, las líneas rectas en el espacio pueden ser separadas en tres muestras dentro de cualquier 2^3 posibilidades; sin embargo, por ningún medio las líneas rectas pueden realizar 2^4 posibilidades de manera discriminante de cuatro muestras dentro de dos grupos. Esto significa que la dimensión VC de la clase de líneas rectas en el plano $d = 2$ es $d + 1 = 3$. En general, la dimensión VC de un hiperplano en un espacio d dimensional es $d + 1$. La dimensión VC ha sido calculada por la mayoría de los modelos usados por los métodos estadísticos flexibles (Burges 1998).

Uno de los principales logros de la teoría del aprendizaje estadístico es la derivación de muchas distribuciones de límites libres para la generalización del error de un modelo de clasificación, regresión y problemas de estimación de densidad, sobre la base de de dos componentes: el riesgo empírico (el cual es solamente determinado por las muestras) y la dimensión VC (la cual no tiene probabilidad pero sí una definición geométrica). El riesgo empírico es el que corresponde a las muestras de entrenamiento y está dado por

$$R_e(w,n)=\frac{1}{n}\sum_{i=1}^{n}\Gamma(c_i = g(x_i,w)) \qquad (19)$$

Donde $(x_1, c_1), (x_2, c_2), \ldots, (x_n, c_n)$, son los pares de muestras y sus correspondientes clases y $\Gamma(\cdot)$ es el indicador de la función:

$$\Gamma(z)=\begin{Bmatrix}1,\ si\ z\geq 0\\ 0,\ si\ z<0\end{Bmatrix} \qquad (20)$$

Para los problemas de clasificación binario, una expresión general para la distribución de límites libres es

$$R_e(w)\leq R_e(w,n)+\frac{\varepsilon}{2}\left(1+\left[1+\frac{4R_e(w,n)}{\varepsilon}\right]^{\frac{1}{2}}\right) \qquad (21)$$

donde

$$\varepsilon=\frac{\alpha\nu\left[\ln(\beta\eta/\nu)+1\right]-\ln(\eta/4)}{\eta} \qquad (22)$$

En esta ecuación v es la dimensión VC y los parámetros α y β están en los rangos $0 < \alpha < 4$

y $0 < \beta < 2$.

La ecuación (21) tiene varias leyendas. La primera indica si el modelo tiene una dimensión VC infinita (el cual es el caso de por ejemplo $g(x, w) = \text{sgn}[\sin(w\,x)]$), entonces el riesgo es ilimitado. Por otra parte, el riesgo tiene un límite que no depende sobre la medida de la probabilidad de las muestras, porque ni el riesgo empírico ni la dimensión VC requieren su conocimiento.

Esto significa que si un método de clasificación aplica la minimizacióndel límite buscando la dimensión VC bajo control, no es necesaria la suposición de la estructura probabilística de los datos de entrenamiento, en contraste con el método clásico de discriminación bayesiano. La estimación de la estructura de probabilidad está más involucrada que el objetivo de la clasificación; es evidente que haciendo la dimensión VC, la solución controlando la variable es un problema más complicado, cuando se evita un paso del intermedio. Note en este contraste la suposición del error de la estructura de probabilidad, que es un paso común en la regresión lineal clásica para la valoración del intervalo de confianza estimado (Sen, 1990).

Otro significado de la ecuación (21), que es importante para una buena valoración de la escala de intensidades IMM, con escasa información, es la posibilidad de minimizar el riesgo cuando existen pocas muestras de entrenamiento. En realidad, si n (que es la variable de control en la estadística clásica), es grande, entonces ambos, el riesgo empírico y ε, serán pequeños, así el limite será también reducido. Sin embargo, si n es pequeño, entonces la reducción del error de entrenamiento es todavía posible, haciendo la dimensión VC la variable de control. Esta es la esencia del llamado principio de minimización del riesgo estructural (Vapnik, 1998), que es materia para el presente artículo. En esencia, el método del vector soporte aplica este principio directamente, para que bajo la generalización del error, pueda ser obtenido incluso con pequeños tamaños de muestras. Esto se demuestra en la siguiente sección.

3. Clasificación de los datos sísmicos con niveles IMM

Los métodos y la teoría anteriormente presentados, fueron aplicados a la base de datos de la referencia (Iyenger 1983), que contiene información instrumental y niveles IMM. Los datos comprenden 92 terremotos e información sobre 12 variables, como aparece en la tabla 1.

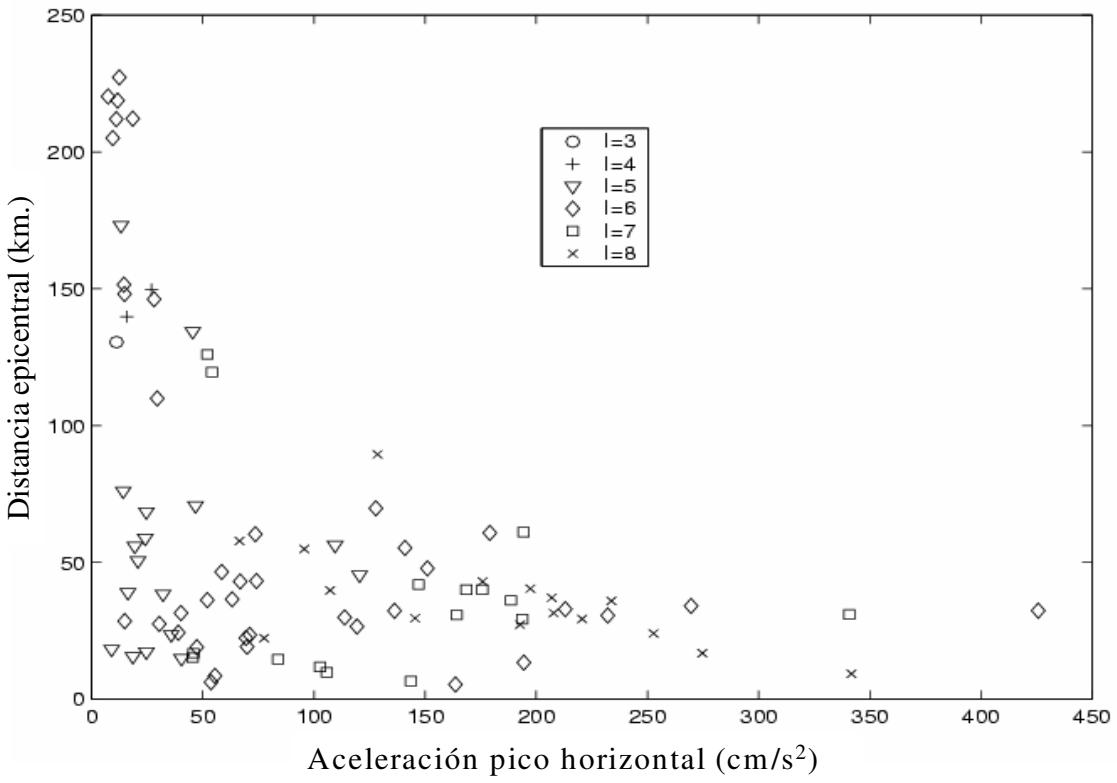
Tabla 1. Variables instrumentales de la base de datos. (Iyengar 1983)

N°	Variable
1	Magnitud
2	Duración
3	Aceleración pico horizontal (A)
4	Velocidad pico horizontal
5	Desplazamiento pico horizontal
6	Tiempo de la aceleración pico horizontal
7	Relación de la aceleración pico horizontal de la componente ortogonal con (A)
8	Relación de la aceleración pico vertical con (A)
9	Distancia epicentral
10	Condición del suelo
11	Máxima pseudo velocidad relativa espectral
12	Número de cruces por cero de la componente horizontal dominante (ZCR)

Fuente: elababoración propia.

El número de variables básicas fue la principal razón para seleccionar esta base de datos, debido a que el problema de clasificación de Mercalli es de hecho complejo. Una razón adicional es que la base de datos es homogénea y los registros pertenecen a alguna región sísmica (occidente de Estados Unidos), que es una importante consideración cuando se usa el enfoque de regresión o clasificación para datos sismológicos.

La figura 3 es un gráfico de la intensidad de Mercalli con respecto a dos variables que son importantes de determinar de acuerdo con Iyengar (1983). Se puede observar que las clases están lejos de ser perfectamente separables, esto es, la discriminación del problema es altamente no lineal. Este hecho valora la proyección no lineal que aporta el método del vector soporte descrito anteriormente.



Fuente: elaboración propia.

Figura 3. Intensidad de Mercalli con respecto a la aceleración pico horizontal y la distancia epicentral.

El enfoque de clasificación fue aplicado como sigue. Primero, un CVS fue calculado para decidir si el correspondiente vector de datos pertenece a I ≤ 4 o no. El vector en la clase I ≤ 4 no fue filtrado afuera pero presentado, sin embargo, como otro CVS para la siguiente pregunta (I ≤ 5 o no). Y así hasta el dilema final (I ≤ 8 o no). En otras palabras, un conjunto de cinco CVS fue entrenado de tal manera que los resultados de un clasificador no fueran tomados en cuenta para seleccionar la población de entrenamiento siguiente. Este criterio fue adoptado sobre la siguiente base: (a) La intensidad de Mercalli es un etiquetado cualitativo asignado por un experto y puede estar sujeta a errores; de aquí la importancia de incluir para los efectos un posible vector que pertenezca a bajas clases sobre los cálculos de un clasificador de altas clases; (b) el filtrado de los datos de entrenamiento debe dejar una población muy pequeña para el entrenamiento del clasificador de altas intensidades.

Una materia de preocupación que se levanta de esta anulación de filtrarse es que el riesgo de inconsistencias es abierto. Por una inconsistencia puede ser entendida la situación cuando un clasificador indica que el vector correspondiente a una intensidad es menor que cuatro, pero el siguiente clasificador indica que pertenece a una intensidad mayor que seis. Si la condición de ser menor o igual al valor de referencia es etiquetado con -1 y siendo mayor que con +1, la consistencia de clasificación con cinco CVS debe tomar la forma tal como [+1+1+1-1-1], que significa que el vector debe ser asignado a I = 7. En la aplicación del procedimiento descrito no fue encontrada ninguna inconsistencia. Este resultado es positivo y es debido a la robustez del método CVS, que es una consecuencia de la definición del clasificador en términos del vector cerca al límite.

En orden a usar el conjunto de clasificadores cuando una nueva información es suministrada, es importante probar el conjunto con casos no presentados en la fase de entrenamiento. Los cálculos preliminares mostraron que todos los clasificadores del conjunto requieren cerca de 60 vectores soporte. El número de vectores soporte es menor que las muestras de entrenamiento; esto fue decidido para entrenar todas las máquinas de vector soporte con las primeras 80 muestras, y reservar 12 para las pruebas. Para las máquinas de vector soporte fue usado un Kernel de base radial con parámetros σ = 0.0005. La secuencia mínima del método de optimización (Platt 1999), fue usada para resolver la restricción del problema de optimización puesto por el vector soporte de entrenamiento. Usando esta técnica, cada entrenamiento de un clasificador toma menos de cinco segundos en un computador personal normal.

La tabla 2 muestra alguna información sobre la fase de entrenamiento.

Tabla 2. Información de los clasificadores

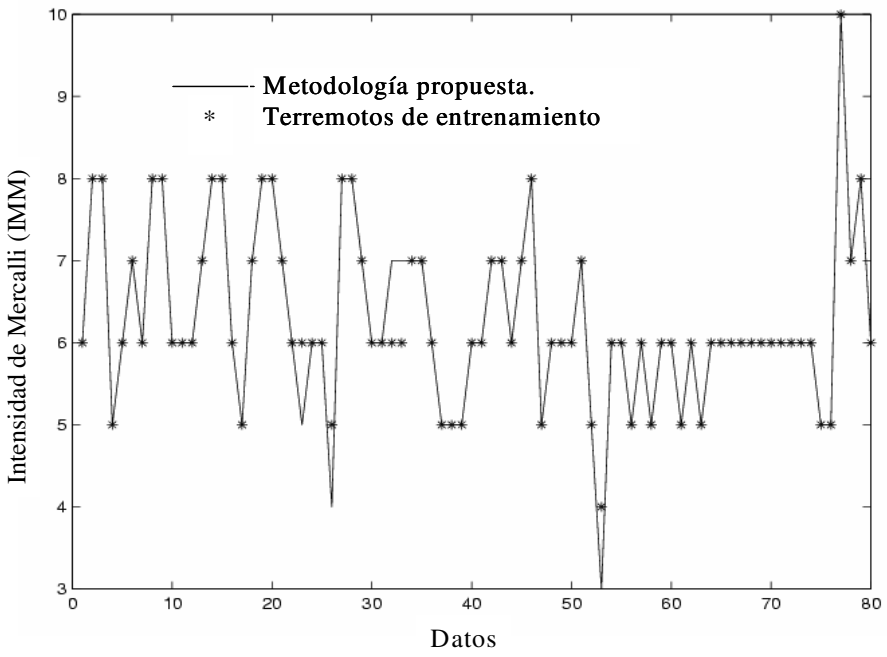
Intensidad (IMM)	No. de vectores soporte (SV)	No. de clasificadores
4	58	0
5	56	1
6	62	3
7	61	1
8	60	0

Fuente: elaboración propia.

Note que el número de vectores soporte en todos los casos es muy alto, denotando la complejidad del problema de clasificación. El número de errores de entrenamiento para cada clasificador es, no obstante, bajo. Note que estos errores no son exclusivamente debidos al método como tal, pero sí son inherentes a algunos problemas de clasificación, donde las clases no son perfectamente separables, como es ilustrado por la figura 3. También, la presencia de errores en este caso se debe a la subjetividad asignada a la IMM.

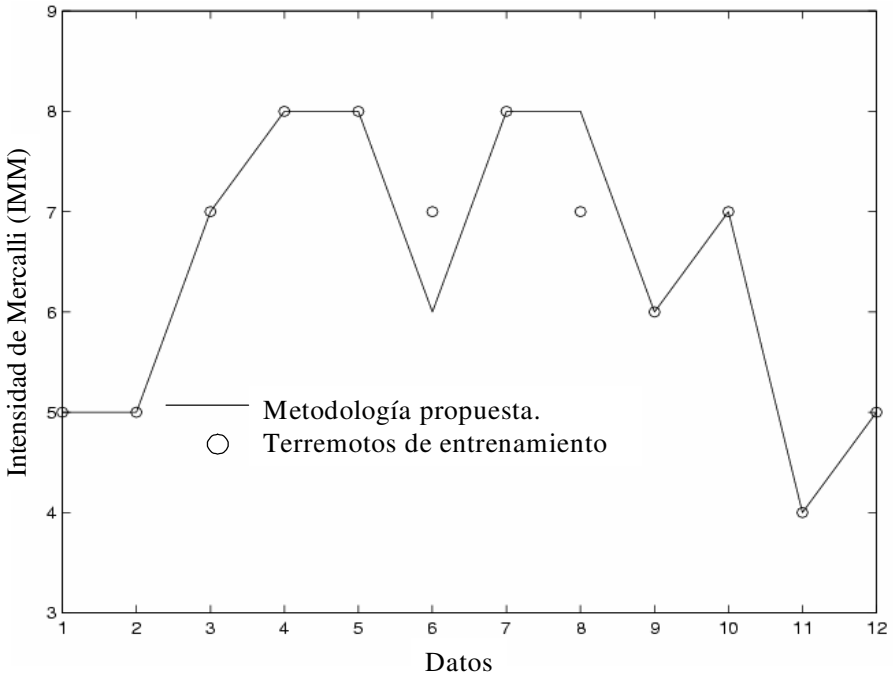
La figura 4, muestra el cálculo de las intensidades con el procedimiento descrito anteriormente.

La comparación con los valores actuales es muy buena, como puede verse. Sin embargo, la comparación para los casos no usados en la fase de entrenamiento desplegada en la figura 5 es más relevante. En este caso, sólo dos de los doce terremotos no fueron bien clasificados. En resumen, en la clasificación de las dos figuras, hay un total de siete errores, y todos los errores están cerca de un ocho por ciento, que puede ser considerado bajo con respecto a la alta no linealidad del problema del daño sísmico y las sombras subjetivas al asignar la IMM. También, se debe notar que en ambas figuras en ningún caso el vector de clasificación difiere de más de dos niveles de intensidad IMM, asignados por un experto.



Fuente: elaboración propia

Figura 4. Comparación de la intensidad de Mercalli con el conjunto de entrenamiento.



Fuente: elaboración propia

Figura 5. Comparación de la intensidad de Mercalli con el conjunto de validación.

Los resultados anteriores, junto con la teoría en la que está basado el método de los CVS, sugiere que el enfoque propuesto es fiable para la evaluación de la intensidad sobre la base de información sísmica instrumental.

4. CONCLUSIONES

Se presentó un método para la evaluación del daño sísmico por medio de la Intensidad de Mercalli Modificada, usando información sísmica instrumental. El método está basado con respecto a la intensidad, como una clase en lugar de una función de los datos cuantitativos. El método de clasificación del vector soporte fue seleccionado debido a su legítima fundamentación teórica, ello da una alta eficiencia para clasificar fenómenos complejos no lineales. Los resultados demuestran que este enfoque es una herramienta conveniente para estimar intensidades de terremotos, en regiones sísmicas, con información de movimientos fuertes. También es una importante herramienta, para la valoración de otras clasificaciones del daño por terremotos. Investigaciones en este aspecto están siendo llevadas a cabo por los autores.

AGRADECIMIENTOS

Los autores agradecen el apoyo financiero y logístico, para la realización de la presente investigación, a la Universidad Nacional de Colombia y a la Universidad de Medellín.

REFERENCIAS

ANTHONY, M., BARTLETT, P. L. (1999). Neural Network Learning: Theoretical Foundations. Cambridge University Press, Cambridge.

BISHOP, Ch. 1995. Neural Networks for Pattern Recognition. Oxford, Ed. Clarendon Press.

BURGES, Ch. J. (1998). A Tutorial on Support Vector Machines for pattern Recognition. Knowledge Discovery and Data Mining, 2, 121–167 pp.

CHERKASSKY, V. MULIER, F. (1998). Learning from Data, Conceptos, Teoría and Methods. New York, Ed. John Wiley and Sons, INC.

DOULA, F. U., ROGERS, L. (1995). Solving problems in environmental engineering and geosciences with artificial neural networks, The M. I. T. Press, Cambridge.

EMAMI, S. M. R., IWAO, Y., HARADA, T. (1996). A Method for Prediction of Peak Horizontal Acceleration by Artificial Neural Networks. In Procedings of the Eleventh World Conference on Earthquake Engineering, pp. 1238, Rotterdam.

ESTEVA, L. (1976). Seismicity. In C. Rosenblueth, editors, Seismic Risk Engineering Decisions. 179-225 pp.

FINE, T. (1999). Feedforward Neural Network Methodology, Springer Verlarg, New York.

HURTADO, J. E. (2002). Analysis of One – Dimensional Stochastic Finite Elements Using Neural Networks. Probabilistic Engineering Mechanics, 17, 34–44 pp.

HURTADO, J. E., ALVAREZ, D. A. (2001). Neural – Networks – Based Reliability Analysis: A Comparative Study. Computer Methods in Applied Mechanics and Engineering, 191, 113–132 pp.

HURTADO, J. E., LONDOÑO, J. M. & MEZA, M. A. 2001. On Applicability of Neural Networks for Soil Dynamics Amplification Analysis. Soil Dynamics and Earthquake Engineering, V. 21, 579 – 591

IYENGAR, N., PRODHAN, C. (1983). Classification and rating of strong motion earthquake records. Earthquake Engineering and Structural Dynamics, 11, 415–426 pp.

KALL, P., WALLACE, S. W. (1995). Stochastic Programming, Jhon Wiley and Sons, Chichester.

KARIM, K. R., YAMAZAKI, F. (2002). Correlation of JMA instrumental seismic intensity with strong motion parameters. Earthquake Engineering and Structural Dynamics, 31, 1191–1212 pp.

LEE, G. C. S., LIN. C. T. (1996). Neural Fuzzy Sytems. Prentice Hall, Upper Saddle River, USA.

MÜLLER, K. R., MIKA, S., RÄTSCH, G., TSUDA, K. SCHÖLKOPF, B. (2001). Introduction to Kernel-based Learning Algorithms. In Y. H. Hu and J. N. Hwang, editors, Handbook of Neural Networks Signal Processing, 4.1 - 4.40 pp.

PLATT, J. C. (1999). Fast Training of Support Vector Machines Using Sequential Minimal Optimization, In B. Schölkopf, C. J. C. Burges, and A. Smola, editors, Advances in Kernel Methods, 185 – 208 pp.

RIPLEY, B. D. (1996). Pattern Recognition and Neuronal Networks, Cambridge University Press, Cambridge.

SEN, A., SRIVASTABA. 1990. Regresión Analysis, Springer Verlarg, New York.

SMOLA, A., BARTLETT, P., SCHÖLKOPF, B., SCHUURMANS, (2000). Adavances in Large Margin Classifiers. The M. I. T. Press, Cambridge.

TRIFUNAC, M. D., BRADY, A. G. (1975). On correlation of seismic intensity scales with the peaks of recorded strong sotion. Bulletin Of The Seismological Society Of America. 65, (1) 139–162 pp.

VAPNIK, V. (1998). Statistical Learning Theory. Jhon Wiley and Sons, New York.

VAPNIK, V. (2000). The Nature of Statistical Learning Theory. New York, Ed. Springer Verlag.

YAGAWA, G., OKUDA, H. (1996). Neural networks in computational mechanics. Archives of Computational Methods in Engineering, 3, 435–512 pp.