

Identificación de patrones sísmicos en registros 2D mediante máquinas de soporte vectorial (SVM)

Identifying Patterns in 2D Seismic Records Using Support Vector Machines (SVM)*

Recibido: 9 de junio de 2015 • Aceptado: 30 de julio de 2015

Para citar este artículo: J. Rodríguez, S. Cárdenas y J. Fonseca «Identificación de patrones sísmicos en registros 2D mediante máquinas de soporte vectorial (SVM)», *Ingenium*, vol. 17, n.º 33, pp. 68-79, julio de 2015.



Julián Rodríguez Acevedo**
Sandra Cárdenas Buitrago***
Juan Fonseca Montoya****

Resumen

La característica conocida como Brigh spot en los registros sísmicos determina una alta probabilidad de encontrar “trampas” de hidrocarburos. Este trabajo presenta el resultado de la clasificación de registros sísmicos que poseen esta característica empleando una máquina de soporte vectorial con kernel polinómico de tercer orden como clasificador automático. Se “ensambla” una matriz de diez atributos característicos de cada registro sísmico y empleando el análisis de componentes principales para reducir la dimensionalidad de la información sísmica, se obtienen dos características que explican el 99 % de la varianza con lo cual se consigue un reconocimiento promedio de registros con característica Brigh spot del 95.65 %.

* Artículo de investigación, producto derivado del Proyecto de investigación: «Identificación de patrones sísmicos en registros 2D mediante máquinas de soporte vectorial» del Grupo Solsytec. Realizado en el año 2014.

** M. Sc. en Automatización Industrial de la Universidad Nacional. Profesor del Programa de Ingeniería Mecatrónica de la Universidad de San Buenaventura, Sede Bogotá. E-mail: jrodriguez@usbbog.edu.co,

*** M. Sc. en Ingeniería, Área Electrónica y de Computadores de la Universidad de Los Andes, Profesora de Ingeniería de la Universidad de San Buenaventura, Sede Bogotá. E-mail: smcardenas@usbbog.edu.co

**** M. Sc. en Ingeniería, Área Electrónica y de Computadores de la Universidad de Los Andes, Gerente Técnico de SAExploration SAS. Investigador del Grupo Solsytec. E-mail: jdfonseca@saexploration.com

Palabras clave

Patrones sísmicos, Brighth spot, PCA (Análisis de componentes principales), SVM (Máquinas de soporte vectorial).

Abstract

The Bright Spot feature, known in seismic registers, determines a high probability of finding hydrocarbon traps. This work presents the classification results of seismic registers encountering this feature, using a vector support machine with a third-order polynomic kernel as an automatic classifier. A ten characteristic attribute matrix of each seismic register is assembled and with the use of Principal Component Analysis for dimensionality reduction of seismic information, two features are obtained, which explain the 99 % of the variance. With this method, a 95.65 % of average recognition of registers with Bright spot feature is achieved.

Keywords

Seismic patterns, Brighth spot , PCA (Principal Component Analysis), SVM (Support Vector Machines).

I. Introducción

Actualmente los sistemas basados en inteligencia artificial están presentes en muchas áreas del conocimiento, que incluyen entre otras el reconocimiento de imágenes diagnósticas en medicina [1], sistemas automáticos de clasificación a nivel industrial [2] y desarrollo de interfaces humano-máquina basados en reconocimiento de expresiones faciales [3]. En la exploración sísmica para la industria de hidrocarburos, se han realizado numerosos esfuerzos para desarrollar sistemas capaces de interpretar los registros sísmicos en busca de materiales de interés, en [4] se diseña un sistema basado en redes neuronales para reconocer características específicas de trazas estatigráficas trabajándolas como imágenes y calculando sobre ellas los momentos de Hu, para determinar características sobre imágenes con diferentes niveles de ruido, en [5] se emplea de nuevo una red neuronal y se construye un árbol de decisión para la interpretación de los registros sísmicos.

En la actualidad la exploración sísmica terrestre se enfoca en áreas de difícil acceso y requiere equipos sofisticados y de gran capacidad computacional para el procesamiento de los datos sísmicos y la construcción de una imagen del subsuelo con base en las trazas sísmicas. Estas trazas sísmicas son obtenidas mediante las señales registradas por sensores luego de una detonación controlada en el subsuelo, estas señales contienen entonces datos estatigráficos, que serán interpretados por un experto. Este trabajo pretende desarrollar un sistema que identifique los datos estatigráficos con mejor probabilidad de contener información acerca de la existencia de hidrocarburos, entrenando una máquina de soporte vectorial previo cálculo de algunos atributos de interés en las señales registradas.

El artículo está organizado de la siguiente manera: En la Sección 2, se hace una explicación acerca de los atributos calculados. En la Sección 3, se muestran los resultados de clasificación de los registros tras entrenar la máquina de soporte vectorial. Finalmente, las conclusiones son presentadas en la Sección 4.

II. Metodología

Con el objeto de determinar algunas características que permitan identificar patrones para posterior entrenamiento del clasificador automático, es necesario calcular algunos atributos de las señales sísmicas. Esta sección describe la forma en que se calculan cada uno de los atributos empleados, así como el proceso y resultado de aplicar el análisis de componentes principales a la información obtenida.

2.1. Cálculo de atributos

Se define un atributo como alguna característica medida u obtenida experimentalmente. Para un conjunto de trazas sísmicas se obtuvieron los siguientes atributos descritos en [6]:

Amplitud instantánea, Amplitud RMS, Amplitud absoluta promedio, Amplitud máxima pico, Amplitud pico promedio, Amplitud Valle promedio, varianza en amplitud, fase instantánea, frecuencia instantánea, derivada de la envolvente.

Antes de hacer la descripción de cada uno de los atributos mencionados, se define la transformada de Hilbert, paso necesario para el cálculo de los mismos:

La transformada de Hilbert básicamente separa las señales en su parte real e imaginaria y se define como:

$$\hat{s}(t) = (t * s)(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\frac{s(\tau)}{t-\tau} \right) d\tau \quad \text{Equ. 1}$$

Donde, corresponde a la convolución de la señal real y la señal $\frac{1}{\pi t}$. A continuación se describe cada uno de los atributos tenidos en cuenta:

2.1.1 Amplitud instantánea

Este atributo proporciona información relacionada con contrastes de impedancia acústica. Cuando se presentan cambios en esta amplitud se puede relacionar con presencia de fluidos. La amplitud instantánea se puede calcular de acuerdo a la ecuación 2

$$A(t) = |F(t)| = \sqrt{(r^2(t) + i^2(t))} \quad \text{Equ. 2}$$

Donde corresponde a la parte real de la traza y es la parte imaginaria. En la figura 1 se ilustran algunas amplitudes de trazas sísmicas.

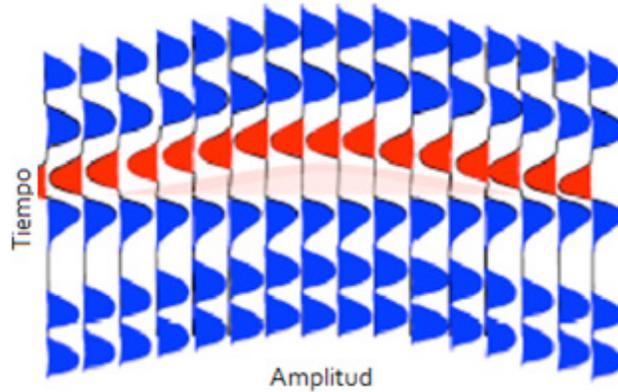


Figura 1. Amplitudes Instantáneas

2.1.2 Amplitud RMS

Se calcula como la raíz cuadrada del promedio de los cuadrados de las amplitudes.

$$Arms(t) = \sqrt{\frac{1}{N} \sum_{i=1}^n a_i^2} \quad \text{Equ. 3}$$

Donde, a_i = amplitudes
 N = número de muestras

2.1.3 Amplitud absoluta promedio

Es el promedio aritmético de los valores absolutos de las amplitudes.

$$Absprom(t) = \frac{1}{N} \sum_{i=1}^n |a_i| \quad \text{Equ. 4}$$

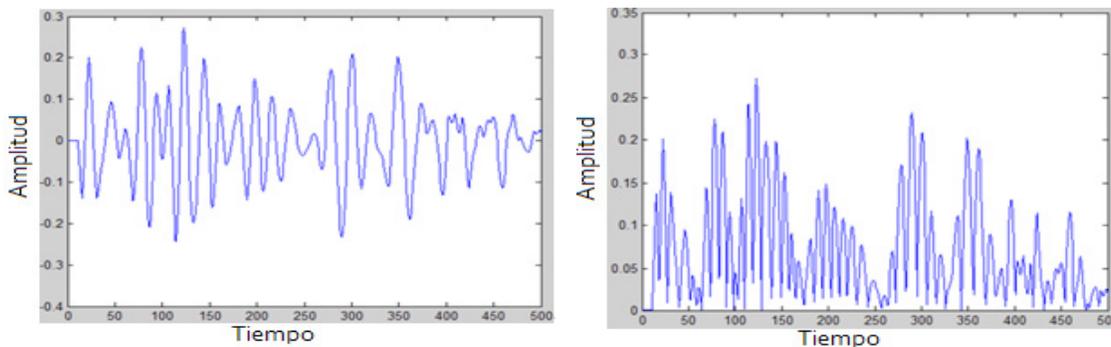


Figura 2. En la izquierda, la traza original. En la derecha, los valores absolutos.

2.1.4 Amplitud máxima pico

Máxima amplitud de las trazas $f(t)$ en un mismo instante de tiempo.

$$Amp(t) = \max(f(t)) \quad \text{Equ. 5}$$

2.1.5 Amplitud pico promedio

Se calcula con la sumatoria de todos los valores positivos de la traza $a_i > 0$ y la sumatoria es dividida entre el número de muestras con amplitud positiva. En la figura 1 se muestran en azul las amplitudes positivas

$$App(t) = \frac{1}{N} \sum_{i=1}^n a_i, a_i > 0 \quad \text{Equ. 6}$$

2.1.6 Amplitud valle promedio

Este atributo es calculado como el promedio de todos los valores negativos de amplitud $a_i < 0$.

$$Avalleprom(t) = \frac{1}{N} \sum_{i=1}^n a_i, a_i < 0 \quad \text{Equ. 7}$$

En la figura 1 se muestran en rojo las amplitudes negativas

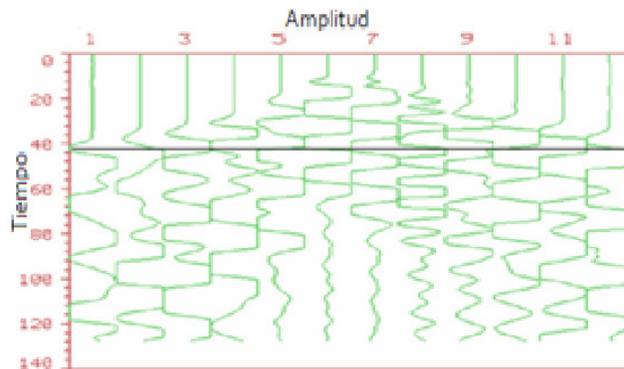


Figura 3. Amplitudes Instantáneas

2.1.7 Varianza en amplitud

Para cada traza, este proceso calcula primero el promedio de todas las muestras x_i . Luego sustrae el promedio a cada muestra y calcula el valor de la diferencia al cuadrado. La salida es la suma de todas estas variables divididas por el número de muestras N .

$$V(t) = \frac{1}{N} \sum_{i=1}^n (x_i - x)^2 \quad \text{Equ. 8}$$

2.1.8 Fase instantánea

Importante por brindar información para el seguimiento de rasgos estratigráficos, se calcula determinando la posición angular del fasor en los instantes muestreados.

$$phi(t) = atan\left(\frac{i(t)}{r(t)}\right) \quad \text{Equ. 9}$$

Donde $i(t)$ =parte imaginaria de la señal y $r(t)$ =parte real de la señal

2.1.9 Frecuencia instantánea

Relevante por estar asociada a efectos de la propagación de la onda, la frecuencia se calcula como la derivada de la fase $\phi(t)$ de la señal, con lo que se relaciona la velocidad de cambio de la fase en función del tiempo.

$$frec(t) = \left(d \frac{\phi(t)}{dt} \right) \quad \text{Equ. 10}$$

2.1.10 Derivada de la envolvente

Relacionada con la energía de la traza sísmica, se calcula como la derivada de la amplitud instantánea

$$dE(t) = \left(d \frac{A(t)}{dt} \right) \quad \text{Equ. 11}$$

2.2 Análisis de componentes principales

Se calcularon los atributos descritos en la sección anterior a un conjunto de 25 imágenes de registros sísmicos con características de probabilidad de existencia de hidrocarburos (característica Bright spot) y 25 con menor probabilidad de existencia de información de interés. La figura 4 ilustra una imagen de muestra de cada uno de los dos grupos empleados.

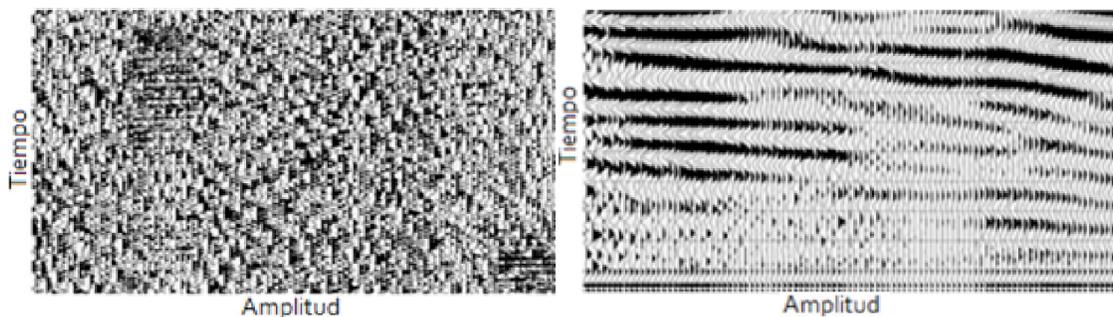


Figura 4. Izquierda: traza sin características de interés. Derecha: traza con probable información de interés

Con estos 50 registros sísmicos se “ensambla” la Tabla 1 de tamaño correspondiente a las 50 muestras y los 10 atributos calculados para cada una de las imágenes procesadas.

| Muestra | A. Instantánea | A. RMS | A. Absoluta promedio | ... | Derivada de la envolvente |
|---------|----------------|--------|----------------------|-----|---------------------------|
| 1 | 0,5 | 17,89 | 713,2 | ... | -1,02x10 ⁻⁵ |
| ... | ... | ... | ... | ... | ... |
| 50 | 0,53 | 34,67 | 2,5x10 ⁻³ | ... | 4,67x10 ⁻⁵ |

Tabla 1. Muestra de las características calculadas

Con el objetivo de reducir a la menor cantidad posible los datos para el entrenamiento del clasificador automático, lo cual reduce el tiempo y los recursos computacionales

requeridos, se empleó un análisis de componentes principales (PCA) [7] a los vectores conformados por las 10 características de la tabla 1.

El análisis de componentes principales consiste en realizar una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos, en el cual la varianza de mayor tamaño del conjunto de datos es capturada en el primer eje (llamado el Primer Componente Principal), la segunda varianza más grande es el segundo eje, y así sucesivamente, de esta manera obtenemos un vector de características con un mínimo de redundancia en la información.

El proceso resumido del PCA se describe a continuación:

1. Cálculo de la matriz de covarianza *cov*.
2. Cálculo de los eigenvectores (*eigenvect*) y eigenvalores (*eigenval*) de la matriz de covarianza.
3. Seleccionar los componentes y armar un vector de características.
4. Organizar los eigenvectores en orden de acuerdo a los eigenvalores correspondientes.

Para el cálculo de los vectores característicos se ensambló una matriz con 10 columnas (una por cada atributo calculado) y 50 filas (50 registros sísmicos) y del cálculo de PCA en esta matriz se tomaron 2 componentes principales, los cuales explican el 99 % de la varianza como se observa en la tabla 2.

| Componente | Porcentaje de explicación de la varianza |
|------------|--|
| 1 | 98,01 % |
| 2 | 1,04 % |
| Total | 99,05 % |

Tabla 2. Explicación varianza PCA

Los coeficientes encontrados de PCA correspondientes a los vectores propios de mayor magnitud se presentan en la tabla 3.

| N.º | V. propio 1 | V. propio 2 |
|-----|-------------------------|--------------------------|
| 1 | $3,14 \times 10^{-5}$ | -0.0016 |
| 2 | 0.0064 | -0.0446 |
| 3 | 0.7071 | $-9,95 \times 10^{-4}$ |
| 4 | $1,055 \times 10^{-5}$ | 0.0029 |
| 5 | 0.7071 | -0.0010 |
| 6 | $-2,28 \times 10^{-21}$ | $-9,017 \times 10^{-19}$ |
| 7 | 0.0017 | 0.99 |

| N.º | V. propio 1 | V. propio 2 |
|-----|------------------------|------------------------|
| 8 | $2,65 \times 10^{-10}$ | $1,99 \times 10^{-8}$ |
| 9 | $-5,63 \times 10^{-7}$ | $-5,10 \times 10^{-4}$ |
| 10 | $-8,13 \times 10^{-9}$ | $3,56 \times 10^{-7}$ |

Tabla 3. Vectores propios

De manera que las dos nuevas características son combinación lineal de las más relevantes, de acuerdo a la ecuación 12 que genera una matriz de 50 muestras y 2 características.

$$\text{NuevasCaract} = \text{vecpropios}'x \text{características}' \quad \text{Equ. 12}$$

Una máquina de soporte vectorial SVM (por sus siglas en inglés) es un tipo de clasificador que consiste en encontrar un hiperplano o conjunto de hiperplanos de alta dimensionalidad que separe de forma óptima puntos pertenecientes a dos o más clases [8]. En este sentido, se busca aprovechar las características de las SVM de una clase [9], capaces de clasificar ruido presente en las trazas sísmicas y permitir su eliminación sin afectar la integridad de la información, mediante la implementación de un kernel (núcleos) Gauss que entre los núcleos disponibles es el más utilizado [10].

Las SVM que se basa en la teoría de la dimensión VC de aprendizaje estadístico y el principio de minimización del riesgo estructural, siendo esta una medida estadística de la capacidad de clasificación de los algoritmos y entendiéndola como la cardinalidad del mayor conjunto de datos que el algoritmo puede clasificar o separar. Se presentan como un algoritmo robusto con gran capacidad de adaptación, optimización global y en general un buen desempeño en este tipo de tareas.

Los registros sísmicos suelen estar en grupos de trazas de múltiples canales, para este caso contamos con 7000 canales y diferentes disparos, lo que permite la redundancia de datos y la división en conjuntos de entrenamiento y prueba, necesario en el proceso de diseño y evaluación del funcionamiento del algoritmo de clasificación.

III. Clasificación y resultados

Esta sección presenta los pasos seguidos en el proceso de clasificación de la información, los datos a clasificar luego del análisis de componentes principales, así como la descripción del clasificador y los resultados obtenidos.

En primera instancia, se grafican las dos características seleccionadas por el análisis de componentes principales (figura 5) en esta se observan las características correspondientes a registros sísmicos con y sin bright spot, es decir, aquellos que tienen mayor probabilidad de representar presencia de hidrocarburos frente a los de menor probabilidad. Además, se observan características que no son separables linealmente, lo que evidencia la necesidad de implementar una máquina de soporte vectorial con kernel polinómico para su clasificación.

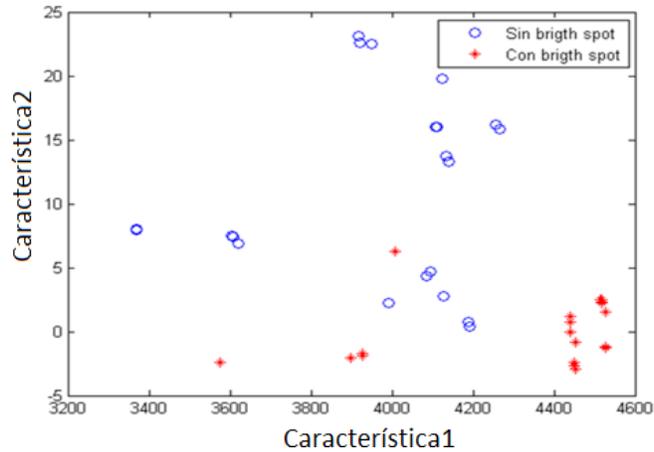


Figura 5. Datos de Entrenamiento

Posteriormente, se establecen los datos de entrenamiento y verificación, utilizando los graficados en la figura 5 como matriz de entrenamiento de la máquina de soporte vectorial, y los graficados en la figura 6 como matriz de verificación del clasificador.

El proceso de entrenamiento de la SVM corresponde al algoritmo SMO presentado en [11], y consta de los pasos enunciados a continuación: $y_i \forall y_j$

1. Se empleó la función kernel $\varphi(x) = (x_1 x_2) \rightarrow (x_1^2, x_1 x_2, x_2^2)$.
2. Construcción de la matriz de kernel: $H_{i,j} = y_i y_j K(x_i x_j)$. Esta es conocida como la matriz Hessiana.
Como tenemos 44 datos de entrenamiento, esta matriz es $H_{44 \times 44}$
3. Inicializar los multiplicadores de LaGrange: $\alpha_i = 0$ y los gradientes $\nabla_i = -1, 1 \leq i \leq n$ (En nuestro caso $n = 44$)
4. Hallar el vector que maximiza el gradiente:

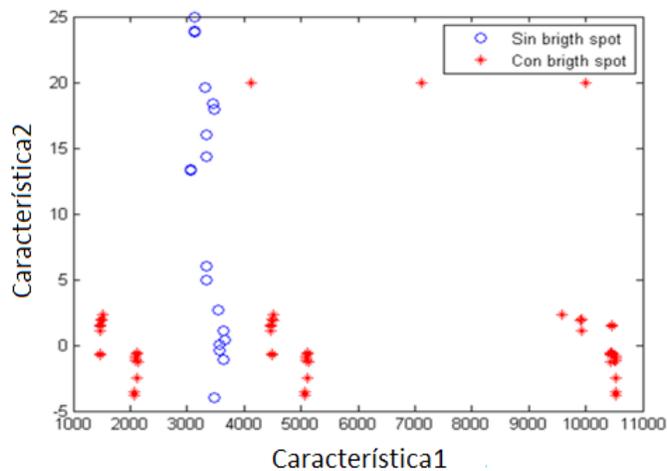


Figura 6. Datos de prueba, $\max[-y_i, \nabla_k]$, donde $\nabla_k = \nabla_k + ((H_{ik} \nabla_{\alpha_i}) + (H_{jk} \nabla_{\alpha_j}))$. ∇_{α_i} . Representa la variación en el valor de α_i , además $1 \leq k \leq n$.

5. Leer la columna i de la matriz Hessiana H , que correspondería a un vector de dimensión $1 \times n$.

6. Cálculo de la variación en gradiente asociado con cada vector candidato k :

$$dif \nabla_k = \nabla_{max} + \nabla_k$$

7. Cálculo del coeficientes cuadrático $Q_{coi,k}$ que está asociado a los vectores i, j :

$$Q_{coi,k} = H_{ii} + H_{kk} + (2y_i H_{ik})$$

8. Mínimo de la variación de la función objetivo Δ_{fk}

$$\Delta_{fk} = \min \left[-\frac{(dif \nabla_{max})^2}{Q_{coi,k}} \right]$$

Si $Q_{coi,k} \geq 0$, de lo contrario se hace $Q_{coi,k} = \tau$, Dónde τ se asocia a un valor muy pequeño del orden de 10^{-6} que garantice un valor grande de y eventualmente se descarte.

9. Hallar el valor $\Delta_{i,j}$, que es un indicador de cuánto deben variar los multiplicadores de LaGrange:

$$\Delta_{i,j} = \pm \nabla_i - \nabla_j Q_{coi,k}$$

El signo \pm depende de si $y_i \neq y_j$ y son iguales o no.

10. Actualización de los multiplicadores:

$$\alpha_i = \alpha_i + \delta_{i,i}; \alpha_j = \alpha_j + \delta_{j,j}$$

11. Condición de finalización: Se examina la suma de los gradientes de i y j si este valor es menor que el umbral predeterminado ya tenemos un valor óptimo y el programa finaliza, de lo contrario volvemos al paso 4

Los valores de los calculados aparecen en la tabla 4

| | | | | | | | | | | | | |
|------------|--------|--------|--------|--------|--------|--------|-------|-------|-------|-------|--------|--------|
| α_i | -0,296 | -0,957 | -0,957 | -0,101 | -0,957 | -0,957 | 0,136 | 0,898 | 1,048 | 1,048 | 1,0477 | 0,0468 |
|------------|--------|--------|--------|--------|--------|--------|-------|-------|-------|-------|--------|--------|

Tabla 4. Multiplicadores de LaGrange encontrados

La función de decisión está dada por la ecuación 13:

$$f(x) = \text{sing}(\sum_{i \in SV} \alpha_i (y_i K(x, x_i)) + b) \quad \text{Equ. 13}$$

Y el valor de b , se determina de acuerdo a la ecuación 14:

$$b = \frac{(\sum_{i \in SV} \alpha_i H_{ii})}{N_{SV}} \quad \text{Equ. 14}$$

Para nuestro caso el valor de b es: $b=1,1794$

En la Figura 7 se ilustra la región de separación entre clases obtenida por el SVM empleando un kernel polinómico de 3.º orden, mientras que en la Figura 8 muestra el resultado junto con los datos de verificación en el proceso de clasificación.

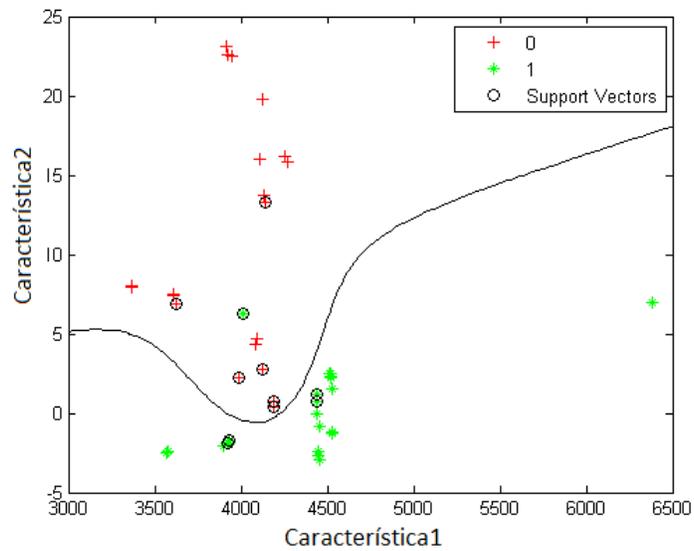


Figura 7. Datos de entrenamiento usados

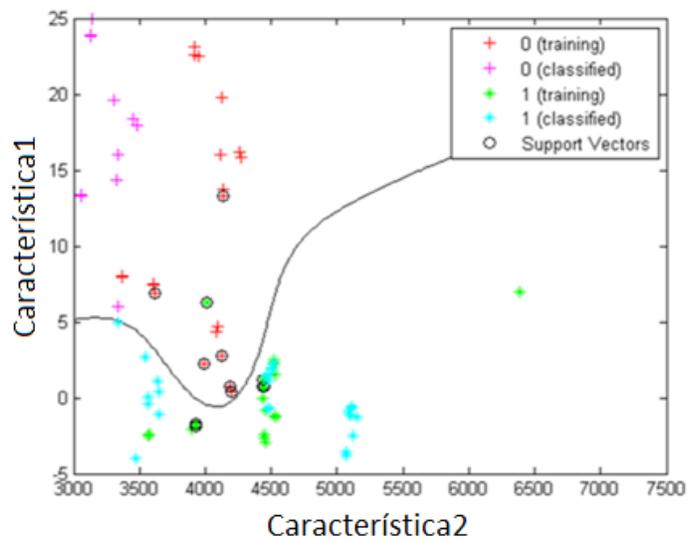


Figura 8. Datos clasificados mediante la SVM implementada

La tabla 5 es la matriz de confusión en porcentajes obtenida con el clasificador, se puede observar que el 95,65% de la información correspondiente a datos sísmicos con característica de bright spot fue clasificada como tal, mientras que de los datos que no tienen esta característica el 61,91 % fue clasificada correctamente, lo que quiere decir que hubo una tasa del 4,35% de falsos positivos en la clasificación.

| | Sin bright spot | Con bright spot |
|-----------------|-----------------|-----------------|
| Sin bright spot | 61,91 % | 38,091 % |
| Con bright spot | 4,35 % | 95,65 % |

Tabla 5. Matriz de confusión (en porcentajes) del clasificador

IV. Conclusiones

Se emplearon 25 datos de registros sísmicos como conjunto de prueba y 25 como datos de validación, luego de calcular 10 atributos y extraer sus componentes principales, se evidenció que para la característica de interés en este trabajo no son relevantes la mayoría de ellos.

Aunque las dos características empleadas en el entrenamiento de la máquina de soporte vectorial no son separables linealmente, se encontró el kernel polinómico adecuado que permite la separación de los grupos en una dimensión mayor.

Para la característica de interés en este trabajo (Bright spot), la cual determina una alta probabilidad de encontrar “trampas” de hidrocarburos, el desempeño de la máquina de soporte vectorial tuvo un importante 95.65% de tasa de reconocimientos; sin embargo, es de interés en un trabajo futuro clasificar otras características junto con la encontrada en este trabajo para incrementar la certeza de existencia de hidrocarburos en la región de donde se captura la información sísmica.

Referencias

- [1] N. E. Ross, C. J. Pritchard, D. M. Rubin y A. G. Dusé, «Automated image processing method for the diagnosis and classification of malaria on thin blood smears», *Medical and Biological Engineering and Computing*, vol. 44, n° 5, p. 427-436, 2006.
- [2] P. Oosterlinck y A. Suetens, «Industrial pattern recognition», de *Pattern Recognition in Practice II*, Elsevier, 2012.
- [3] J. Prieto y F. Rodríguez, «Analyzing the relevance of shape descriptors in automated recognition of facial gestures in 3D images.» de *Three-Dimensional Image Processing (3DIP) and Applications SPIE*, 2013.
- [4] K. Y. Huang, «Neural network for robust recognition of seismic patterns», de *Neural Networks, 2001. Proceedings. IJCNN'01*, 2001.
- [5] K. Y. Chao y Y. Huang, «Seismic pattern recognition using neural network and tree automaton», de *Geoscience and Remote Sensing Symposium. IEEE International*, 2004.
- [6] C. M. Orietta, *Implementación de algoritmos para estudios de factibilidad de atributos sísmicos utilizando datos sintéticos*, Universidad Simón Bolívar, 2005.
- [7] T. K. Hasselman, M. C. Anderson y W. Gan, «Principal components analysis for nonlinear model correlation, updating and uncertainty evaluation», *SPIE proceedings series*, 1998.
- [8] T. Hastie, S. Rosset, R. Tibshirani y J. Zhu, «The entire regularization path for the support vector machine», *The Journal of Machine Learning Research*, vol. vol. 5, p. pp. 1391-1415, 2004.
- [9] W. Z. Zhong, «A new one-class svm based on hidden information», *ScienceDirect. Knowledge-Based Systems*, 2014.
- [10] Y. X. Xu, «Two methods of selecting gaussian kernel parameters for one-class svm», *Science Direct Knowledge Based Systems*, 2014.
- [11] P. Jiménez y L. Rengifo, «Al interior de una máquina de soporte vectorial», *Revista de ciencias Universidad del Valle*, 2010.