

COMPARACIÓN DE MODELOS DE CLASIFICACIÓN AUTOMÁTICA DE PATRONES TEXTURALES DE MINERALES PRESENTES EN LOS CARBONES COLOMBIANOS

COMPARISON OF MODELS OF AUTOMATIC CLASSIFICATION OF TEXTURAL PATTERNS OF MINERAL PRESENTS IN COLOMBIAN COALS

JAIME LÓPEZ CARVAJAL

Institución Universitaria de Envigado. jalop@iue.edu.co

JOHN WILLIAN BRANCH BEDOYA

Universidad Nacional de Colombia, Sede Medellín. jwbranch@unalmed.edu.co

Recibido para revisar 3 de Mayo de 2004, aceptado 1 de Junio de 2004, versión final 1 de Agosto de 2004

RESUMEN: Esta investigación presenta algunos resultados obtenidos bajo diferentes modelos de clasificación de patrones texturales de minerales presentes en imágenes digitales. El conjunto de datos utilizado está caracterizado por un tamaño pequeño y presencia de ruido. Los modelos implementados fueron el clasificador Bayesiano, red neuronal (2-5-1), máquina de soporte de vectores (SVM), árbol de decisión y 3-vecinos más cercanos. Los resultados obtenidos luego de la validación cruzada demostraron que el modelo Bayesiano (84%) arrojó la mejor capacidad predictiva, debido principalmente a su robustez frente al ruido. La red neuronal (68%) y la SVM (67%) dieron resultados alentadores, que posiblemente puedan mejorarse al incrementar el tamaño del conjunto de datos; mientras el árbol de decisión (55%) y el k-vecinos (54%) no parecen ser adecuados para este problema, dado su sensibilidad al ruido.

PALABRAS CLAVE: Clasificación, imágenes digitales, textura, minerales del carbón

ABSTRACT: The automatic classification of objects is a very interesting approach under several problem domains. This paper outlines some results obtained under different classification models to categorize textural patterns of minerals using real digital images. The data set used was characterized by a small size and noise presence. The implemented models were the Bayesian classifier, Neural Network (2-5-1), Support Vector Machine, decision tree and 3-nearest neighbours. The results after applying crossed validation show that the Bayesian model (84%) proved better predictive capacity than the others, mainly due to its noise robustness behaviour. The Neuronal Network (68%) and the SVM (67%) gave promising results, because they could be improved increasing the data amount used, while the decision tree (55%) and k-NN (54%) did not seem to be adequate for this problem, because of their sensibility to noise.

KEY WORDS: Classification, Digital Images, texture, coal minerals

1. INTRODUCCIÓN

El reconocimiento de patrones se define como la ciencia que estudia cómo se forman los patrones, cómo se puede observar el ambiente que los rodea, cómo se aprende a distinguir patrones interesantes del resto, y cómo se pueden tomar decisiones razonables y confiables sobre sus categorías [1]. Ross citado por Jain [2] retoma las palabras de Herbert Simon, quien menciona que el reconocimiento de patrones es útil para muchas tareas humanas que involucran tomar decisiones; y si son patrones de buena calidad, las decisiones serán mejores. De acuerdo a Morton y Smith [3], un patrón es un arreglo peculiar de elementos estructurales al cual se le puede dar una etiqueta o nombre; por ejemplo una huella digital, una textura, un rostro o una señal.

Esta investigación desarrolló la clasificación supervisada de patrones texturales que caracterizan a dos clases de minerales presentes en los carbones. La textura como tal, es un término vagamente definido en la literatura científica, y muchos autores usualmente usan diferentes definiciones según su área científica. Para el contexto de la presente investigación, la textura se entiende como un atributo que representa la disposición espacial de los niveles de gris de los píxeles dentro de una región de la imagen. Cada patrón está representado en términos de 5 atributos, ocupando un punto en ese espacio 5-dimensional; y cuyo objetivo es separar los patrones texturales pertenecientes a cada categoría, es decir, ocupar regiones diferentes y compactas en ese espacio [4].

La organización de este documento es la siguiente: en la Sección 2 se define el problema de la clasificación automática de minerales; la Sección 3, describe la metodología seguida para el alcance de los resultados; la Sección 4 enseña los resultados alcanzados; la Sección 5 contiene la discusión generada a partir de los resultados; y finalmente, la Sección 6, enumera las conclusiones de la investigación.

2. CLASIFICACIÓN

La clasificación es una tarea básica en el análisis de datos y en el reconocimiento de patrones. Los fundamentos teóricos presentados a continuación están basados en la teoría de decisión clásica de Wald, citado por [1].

Una instancia (una textura para nuestro caso) es un vector de características $X = x$ de alguna población X , tal que x pertenece a X , es el espacio muestral. La pertenencia de clase verdadera es $\omega \in \Omega = \{\omega_1, \omega_2, \dots, \omega_M\}$, siendo M el número total de clases ($M = 2$ para nuestro caso). Este es el verdadero estado de la naturaleza. Se asume que las parejas (X, ω) fueron tomadas independientemente de la distribución de probabilidad conjunta $p(x, \omega)$ sobre $X \times \Omega$ [5]. Dada una observación (vector de atributos), $X = x$, se quiere tomar la decisión sobre la verdadera clase de la instancia. Esto es lo que se conoce como problema de clasificación, predecir la verdadera clase ω para una instancia $X = x$. El espacio de acción Y es el espacio de decisiones permisibles para nuestro problema de clasificación. Un elemento de Y se denomina y . Este espacio de acción generalmente puede verse como una extensión del espacio de clases, donde $Y = \{\Omega, \text{'valor extremo'}, \text{'valor dudoso'}, \dots\}$, tal como lo define [6]. Para la presente investigación, se asume que el espacio de acción es igual al espacio de clases, es decir $Y = \Omega$.

3. DEFINICIÓN DEL PROBLEMA

La clasificación de minerales presentes en imágenes digitales de carbones es una tarea que suelen realizar manualmente los expertos petrógrafos para su correcta caracterización y la determinación del uso potencial de los carbones. Esta tarea es una labor dispendiosa que requiere de personal capacitado y experimentado, y por ende costoso, dedicado una gran cantidad de tiempo a la observación de las imágenes para discriminar manualmente todos los minerales que encuentra en cada una de las imágenes observadas. Este es un proceso altamente subjetivo debido a la influencia de algunos

factores intrínsecos como son: el conocimiento del experto, su estado de ánimo, la cantidad de horas dedicadas, el agotamiento de la visión, la calidad de las imágenes y el estado de los equipos.

Todo lo anterior, tiene una alta repercusión en la calidad de la clasificación de los minerales, lo que se refleja en un tiempo de procesamiento muy lento, una precisión de la clasificación no muy alta, lo cual significa una pérdida económica significativa para las empresas que explotan o comercializan el carbón, así como unos costos muy altos para el laboratorio encargado de esta función.

Por lo tanto, se pretende enfrentar este problema utilizando un atributo que sea lo más representativo y diferenciador entre los minerales presentes, de tal forma, que se pudieran conformar patrones que fueran clasificados automáticamente por un modelo que incorpore el conocimiento que tienen los expertos, pero sin la subjetividad, la lentitud e imprecisión que este personal humano involucra durante su actividad.

De acuerdo a investigaciones previas [7-8], se pudo concluir que la textura representa un patrón que tiene un potencial muy alto para la clasificación automática de los minerales presentes en las imágenes digitales de

carbones, superando a otros atributos como la forma o el color.

Dado que las muestras de carbones provienen de distintos lugares del país, aumentando la presencia de variaciones o alteraciones de su conformación típica, se requiere implementar un modelo que permita su correcta clasificación. Por tal motivo, se investigó el comportamiento de diferentes métodos de clasificación supervisada, para obtener un método para la realización automática de este proceso, que permita reducir el porcentaje de errores, reducir los costos asociados y optimizar la obtención de los resultados.

4. DESARROLLO EXPERIMENTAL

4.1. IMÁGENES DIGITALES

El banco de datos de imágenes está compuesto por aproximadamente 400 imágenes que fueron usadas para extraer características de los minerales y la posterior implementación de los diferentes clasificadores ensayados. Las condiciones de la captura pueden verse en la Tabla 1. Los minerales en las imágenes digitales de carbones fueron identificados con apoyo de petrógrafos expertos.

Tabla 1. Condiciones de la Captura de imágenes digitales
Table 1. Conditions of the Apprehension of digital images

CONDICION	VALOR
Lámpara del microscopio:	7 amperios
Sensibilidad:	400
Tamaño de imagen:	640 x 480 píxeles
Tiempo de exposición (3):	0.5, 5 y 1/20 de segundo
Tipo de captura:	manual
Objetivo:	32x oel 0.65
Luz	Reflejada standard
Polarizador:	no
Calibración:	no
Formato:	.jpg

4.2. MATRICES DE COOCURRENCIA

Son descriptores texturales ampliamente conocidos, dados a conocer por Haralick et al [9] y que han demostrado sus bondades en tareas relacionadas con la segmentación, clasificación y búsqueda de imágenes.

Estas matrices usualmente se configuran de acuerdo al problema investigado. En este caso, se seleccionaron 8 ángulos (0, 45, 90, 135 grados y sus respectivos ángulos opuestos) como resultado del desarrollo experimental, dado que los patrones texturales utilizados no presentan una regularidad bien definida; por lo cual, se requiere una cobertura completa del patrón textural (invarianza rotacional), que lo brindan la cantidad y dirección de los ángulos seleccionados.

4.3. CUANTIZACIÓN

Es un proceso que permite reducir la cantidad de niveles de gris en las imágenes evitando una pérdida significativa de información. Además, busca mejorar la validez estadística de las matrices de coocurrencia y mejorar el desempeño del algoritmo implementado. De acuerdo al análisis experimental desarrollado, aplicando una discriminación lineal a los patrones texturales para diferentes grados de cuantización (8,16 y 32 niveles de gris), se encontró que el nivel de cuantización más apropiado fue de 16 niveles de gris para las imágenes de minerales utilizadas, tal como se puede observar en la Figura 1.

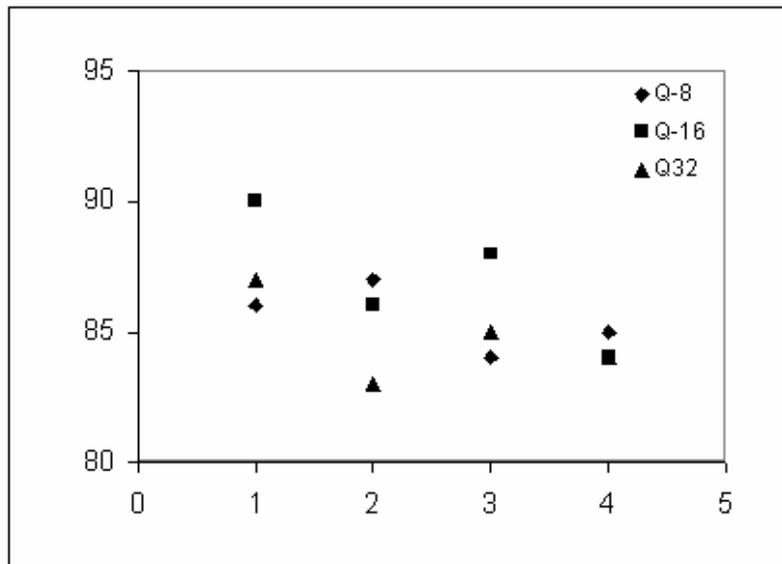


Figura 1. Porcentaje de discriminación obtenido para cada nivel de cuantización Q8-Q16-Q32 (Eje horizontal: distancias entre píxeles 1-4)

Figure 1. Percentage of discrimination obtained for every level of quantization Q8-Q16-Q32 (Horizontal axis: distances between pixels 1-4)

4.4. ATRIBUTOS TEXTURALES

Inicialmente se extrajeron 12 atributos texturales de cada patrón, los cuales fueron sometidos a un proceso de selección de los mejores atributos, utilizando diferentes técnicas de minería de datos contenidas en la herramienta Weka [10]. Este análisis

permitió la selección de los cinco atributos texturales más representativos [11,12], para este caso investigado, como fueron contraste, entropía, disimilaridad, varianza y prominencia, cuyas respectivas definiciones se encuentran en la Tabla 2.

Tabla 2. Atributos texturales seleccionados, con su definición matemática
Table 2. Selected textural attributes, with its mathematical definition

Atributo textural	Formula
Entropia	$\sum_i \sum_j P_a(i,j) \log P_a(i,j)$
Contraste	$\sum_i \sum_j (i-j)^2 P_a(i,j)$
Disimilaridad	$\sum_i \sum_j P_a(i,j) i-j $
Varianza	$\sum_i \sum_j P_a(i,j) (i-\mu)^2$
Prominencia	$\sum_i \sum_j (i+j - \mu_x - \mu_y)^2 P_a(i,j)$

4.5. TAMAÑO DE LA TEXTURA

Se escogió un tamaño para la imagen de la textura 32 x 32 píxeles, motivado principalmente por el tamaño de los minerales en las imágenes (carbonato y pirita); dado que un tamaño superior incorporaría ruido, lo cual afectaría el comportamiento de los clasificadores ensayados [13]. En Figura 2, se ilustran los dos patrones correspondientes a los minerales estudiados en la presente investigación, como fueron el carbonato y la pirita.

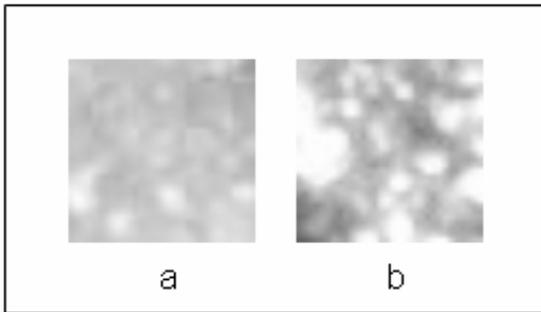


Figura 2. Patrones texturales utilizados:
a) carbonato y b) pirita
Figure 2. Textural patterns used:
a) carbonate and b) pyrite

4.6. DISCRETIZACIÓN

Para la discretización, se utilizó el método de rangos iguales [14], también conocido como *binning*, el cual transforma los datos originales de tipo continuo, a datos discretos

ubicados en diez categorías previamente definidas.

4.7. INDUCCIÓN DEL CONOCIMIENTO DE EXPERTOS

También conocido como proceso de elicitación, para el cual se diseñaron Tablas que contienen 10 imágenes texturales que corresponden a cada una de las categorías resultantes luego de la discretización [15]. Este procedimiento se llevo a cabo para la determinación de los parámetros de la distribución de probabilidad Dirichlet del clasificador Bayesiano Informativo.

5. RESULTADOS

La métrica de evaluación de los modelos utilizada para la presente investigación fue el rendimiento predictivo (o el error de clasificación), técnica útil para estimar la confiabilidad de cada modelo con relación al conjunto de datos [16].

Los resultados arrojados por la presente investigación se encuentran sintetizados en la Tabla 3. Allí se encuentran las Tablas de contingencia de cada uno de los modelos de clasificación ensayados, en donde en su diagonal principal se encuentra el porcentaje de instancias clasificadas correctamente (*rendimiento predictivo*), y en su diagonal secundaria las clasificadas de forma incorrecta (*error de clasificación*), de acuerdo a la siguiente formula:

$$\text{Rendimiento del clasificador (\%)} = \Sigma(\text{diagonal principal})$$

Una vez obtenido el rendimiento del clasificador, se procede a calcular su respectivo error de clasificación, utilizando la siguiente fórmula:

$$\text{Error del Clasificador (\%)} = 100 - \text{Rendimiento del clasificador}$$

Para ejemplificar este cálculo, a continuación se ilustra el caso del clasificador *Naive Bayes*, su rendimiento predictivo y su error de clasificación se obtuvieron de la siguiente forma:

$$\begin{aligned} \text{Rendimiento} &= \Sigma(24 + 45) = 69 \% \\ \text{Error} &= 100 - 69 = 31 \% \end{aligned}$$

Tabla 3. Tablas de contingencia para los modelos de clasificación (T1 corresponde a la textura carbonato, y T2 a la textura pirita)

Table 3. Tables of contingency to classification models (T1 texture of carbonate, T2 texture of pyrite)

Naive Bayes			Bayes Informativo		
	<i>T1</i>	<i>T2</i>		<i>T1</i>	<i>T2</i>
<i>T1</i>	24	23	<i>T1</i>	51	5
<i>T2</i>	8	45	<i>T2</i>	9	35
error = 31%			error = 14%		
k-vecinos			Arbol de decision		
	<i>T1</i>	<i>T2</i>		<i>T1</i>	<i>T2</i>
<i>T1</i>	23	24	<i>T1</i>	5	50
<i>T2</i>	25	28	<i>T2</i>	5	40
error = 49%			error = 55%		
Red Neuronal			SVM		
	<i>T1</i>	<i>T2</i>		<i>T1</i>	<i>T2</i>
<i>T1</i>	27	25	<i>T1</i>	18	31
<i>T2</i>	10	38	<i>T2</i>	2	49
error = 35%			error = 33%		

Es importante destacar que los resultados consignados en las Tablas de contingencia se obtuvieron a partir de una validación cruzada de 10 corridas (conocido como *10-fold cross validation*) para obtener un estimado del error de cada uno de ellos. Cada uno de los algoritmos de los modelos de clasificación se implementaron bajo su configuración convencional (sin optimizarlos), y sometidos a la misma cantidad de datos tanto para su entrenamiento (100 texturas) como para su evaluación (300 texturas).

Se procuro abarcar los diferentes paradigmas conocidos, como fueron los basados en instancias, los conexionistas y estadísticos [17]; los cuales incluyen desde los Bayesianos hasta los más modernos como las máquinas de soporte de vectores (SVM). En la Tabla 3 se registran las Tablas de

contingencia (también conocidas como Tablas de confusión) para cada uno de los modelos ensayados.

6. DISCUSIÓN

Los resultados obtenidos en esta investigación indican que el modelo de clasificación Bayesiano Informativo arroja el mejor rendimiento (menor error). Esto es producto de sus cualidades al enfrentar el problema de clasificación cuando el tamaño de muestra es pequeño y el ruido de los datos es alto, y además se valió de la incorporación del conocimiento y dominio de personal experto en el dominio (petrógrafos), a través de los parámetros de la distribución ajustada.

En este sentido, se suele argumentar que un tamaño de muestra pequeño generalmente no contiene toda la información requerida para generar un modelo confiable, pero en muchos dominios de aplicación los datos no son abundantes, por lo cual se debe trabajar con la información disponible, así no cumplan con todos los requerimientos deseables (muchas instancias, bajo ruido, baja variación intraclase, alta variación interclase, etc.).

En la Figura 3, se observa como los modelos como la red neuronal y la SVM vieron afectada su capacidad predictiva por el sobreentrenamiento a unos datos con ruido, lo cual se refleja en un bajo rendimiento ante patrones nuevos, es decir, aquellos que no se utilizaron para generar el modelo. Como estos modelos presentan un bajo sesgo; favoreciendo el reconocimiento de relaciones no lineales en los datos, generó un bajo desempeño del modelo en este caso particular, posiblemente causado por la alta presencia de ruido.

Para el caso de los modelos de clasificación basados en instancias como los k-vecinos y el árbol de decisión, los rendimientos obtenidos no han sido los mejores, presumiblemente por su carácter de predicción local, lo que ha degradado su funcionamiento al enfrentarse con un conjunto de datos no tan representativo, como sería lo ideal.

En la Figura 3, se ilustra como los modelos con fuerza representacional más alta, como la Red Neuronal y la Máquina de Soporte de Vectores (SVM), arrojaron un rendimiento inferior al esperado para este tipo de modelos; siendo el ruido de los datos empleados, un posible causante de este comportamiento. Por su parte, el modelo Bayes Informativo, que se apoyo en la información de expertos y no en los datos, para el ajuste de sus parámetros, alcanzo el mejor comportamiento ante este conjunto de prueba ruidoso.

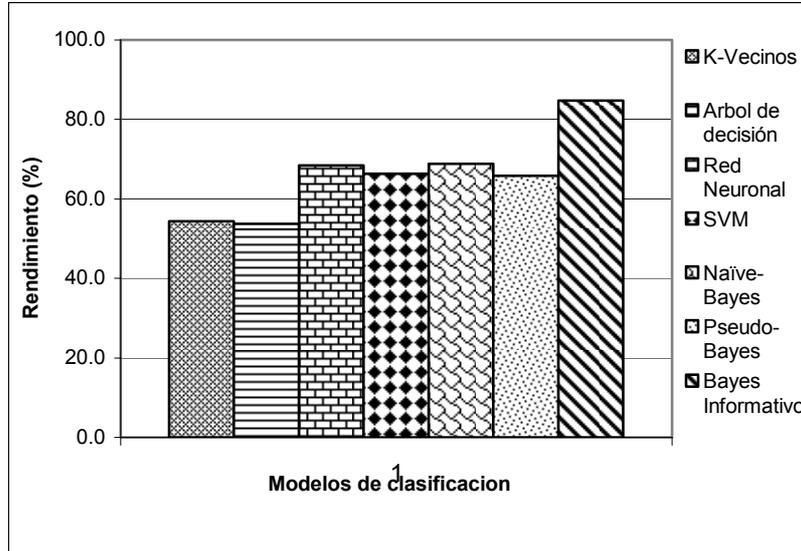


Figura 3. Rendimiento de los diferentes modelos de clasificación
Figure 3. Results of the different models of classification

Para el clasificador Informativo, el cual presento un aumento considerable en su precisión predictiva que parece ser el resultado de la incorporación del

conocimiento inducido del experto, lo que confirma las bondades del procedimiento seguido. En este punto, es importante destacar que en cada dominio particular, la

tarea de un clasificador es decidir o apoyar una decisión sobre patrones futuros, es decir, de aquellos que no se conoce su categoría con anticipación, por lo cual, se considera muy importante generar un modelo de clasificación que ofrezca un alto rendimiento predictivo, como fue el caso del clasificador Bayesiano informativo (84%), como se confirma en la Tabla 4.

El hecho de enfrentar un problema con datos reales, tal como se hizo en la presente

investigación, supuso un reto superior, dada la posibilidad de encontrarse con dificultades no estimadas previamente o que no se pueden modificar en el transcurso de la investigación, como fue el caso de la captura de las imágenes, dado que se escogió un formato de compresión con pérdida, lo cual posiblemente afectó los patrones texturales de los minerales estudiados, pero que habrá que confirmarla precisamente en investigaciones posteriores.

Tabla 4. Tasa de error para los modelos de clasificación ensayados, utilizando una validación cruzada de 10 corridas

Table 4. Error rate for the tried classification models, using crossed validation of 10 tests

METODO	Entrenamiento	Prueba	Diferencia
	% error (d.e)	% error (d.e)	
Naïve Bayes	75,3 (5,58)	68,9 (3,07)	6,3
Bayes Informativo	63,8 (6,15)	84,7 (2,50)	-21,0
k-vecinos	91,6 (1,84)	54,4 (2,84)	37,2
Arbol de decision	91,9 (2,78)	53,7 (1,89)	38,2
Red Neuronal	97,3 (1,84)	68,5 (3,17)	28,8
Maquina de Soporte de Vectores	84,9 (1,49)	66,4 (3,63)	18,5

El comportamiento de los modelos de clasificación como la red neuronal y la SVM se pueden considerar como promisorios, porque aunque sus rendimientos predictivos no fueron muy altos, con el incremento del tamaño de muestra su comportamiento puede mejorar significativamente, al recoger información de un conjunto de datos mas significativo, y no sufrir el sobreentrenamiento a conjuntos ruidosos, tal como sucedió en la presente investigación. Además, según [6], la SVM ha demostrado un buen comportamiento en un dominio amplio de problemas, lo que apoya la hipótesis del tamaño de muestra enunciado anteriormente.

Para el caso de los modelos k-vecinos y árbol de decisión, dada su característica de predicción local, no parecen susceptibles de mejora con un aumento del tamaño del conjunto de datos. Aunque posiblemente con la modificación del número de vecinos en el caso del primero, y de la aplicación de podas en el segundo, se podría mejorar su rendimiento predictivo. En la Figura 4, se observa como la validación cruzada, arroja rendimientos muy similares para las diferentes corridas del modelo, lo que permite establecer que su rendimiento dificilmente puede ser mejorado para este dominio particular.

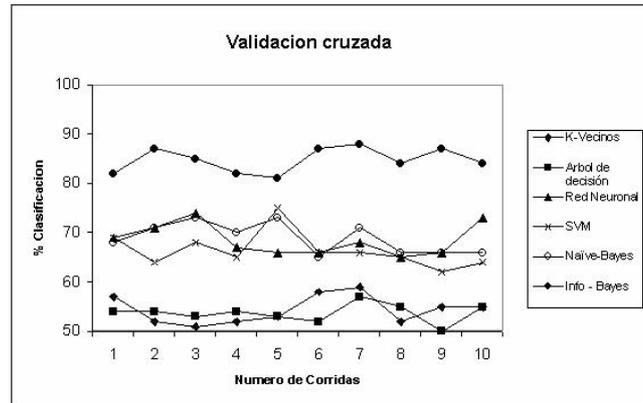


Figura 4. Clasificaciones obtenidas por aprobación cruzada para cada modelo
Figure 4. Classifications obtained by crossed validation for every model

De lo anterior, es importante mencionar que los modelos de clasificación que no arrojaron buen comportamiento bajo este dominio, no se pueden considerar de potencia inferior a los demás; porque según [16], ningún modelo podrá superar a los demás en todos los ámbitos, porque su comportamiento está íntimamente ligado al concepto buscado y a las características de los datos, como se confirmó para el dominio de los minerales presentes en las imágenes digitales de carbones.

7. CONCLUSIONES

El modelo de clasificación que mejor se comportó fue el clasificador Bayesiano Informativo, el cual arrojó un porcentaje de predicción alto; lo que confirma la robustez del método Bayesiano al incorporar el manejo probabilístico de la incertidumbre, así como la incorporación de valiosa información de expertos.

El modelo de red neuronal y SVM son viables de mejorar, pero requieren mejorar el conjunto de datos, tanto en calidad como en cantidad, para obtener rendimientos predictivos aceptables.

Por el contrario, los modelos k-vecinos y árbol de decisión han demostrado no ser viables para el problema investigado, dada su alta sensibilidad ante conjuntos de entrenamiento ruidosos, lo que resulta en una predicción pobre.

REFERENCIAS

- [1] DUDA RO, HART PE, STORK WC. 2000. Pattern recognition. New York: Wiley.
- [2] JAIN AK, DUIN RPW, MAO J. 2000. Statistical pattern recognition: to review. IEEE. Transactions on Pattern Analysis and Machine Intelligence 22(1):4-37.
- [3] MORTON N, SMITH EP. Pattern recognition engineering. John Wiley & Sons Inc. It USES. 1993. 588 p.
- [4] MATERKA A, STRZELECKI M. 1998. Texture analysis methods: A review. Technical report. University of Lodz. Institute of Electronics, Brussels. 33 p.
- [5] LANGAAS M. Discrimination and classification. Dept. of Mathematical Sciences. Norwegian Institute of Technology. 44 p. 1994.
- [6] RIPLEY B. Pattern recognition and Neural Networks. Cambridge U. Press, 1996.
- [7] BLANDÓN A, RESTREPO A, ORTEGA O, LÓPEZ J. 2003. Reconocimiento de patrones en los factores de forma utilizados para la descripción de liptinitas. VI Congreso Nacional de Ciencia y Tecnología del Carbón. Medellín, Agosto. 4 p.

- [8] LÓPEZ J, CABRERA KR, BRANCH JW. 2003. Análisis cluster de texturas utilizando matrices de coocurrencia. En: Memorias VIII Simposio de Tratamiento de señales, imágenes y visión artificial. p.23-24. UPB. Medellín, Noviembre.
- [9] HARALICK RM, SHANMUGAN K, DINSTEN I. 1973. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics. 3(6):610-620.
- [10] WITTEN IH, FRANK E. WEKA: Machine Learning Software in Java. University of Waikato. New Zeland. 2002
- [11] MANJUNATH BS, MA WY. 1996. Texture features for browsing and retrieval of image data. IEEE Transactions in Pattern Analysis and Machine Intelligence 18(8): 837-841
- [12] TURCERYAN M, JAIN A. Texture analysis. In The Handbook of pattern recognition and computer vision (2nd ed.). por C.H. Chen et al. (Eds.), Chapter 2.1. p.235-276. World Scientific Publishing Co. Singapore. 1993.
- [13] PUIG D, GARCÍA MA. 2001. Determining optimal window size for texture feature extraction methods. IX Spanish Symposium on Pattern Recognition and Image Analysis, Castellon, Spain, 2:237-242. 6 p.
- [14] DOUGHERTY J, KOHAVI R, SAHAMI M. 1995. Supervised and unsupervised discretization of continuous features. In: Proceedings of the 12th International Conference on Machine Learning, pages 194–202. Morgan Kaufmann Publishers, Inc.
- [15] CABRERA, KR. 2003. Comunicación personal.
- [16] HILARIO M, KALOUSIS A. 1999. Characterizing learning models and algorithms for classification. Esprit LTR Project METAL 26.357 CUI - University of Geneva. 16 p.
- [17] MICHIE D, SPIEGELHALTER DJ, TAYLOR CC. Machine learning, neural and statistical classification. Prentice-Hall, 1994.