

Ingeniería de la calidad

Práctica II:

Clasificación de defectos en la piel

Iván Vega Gutiérrez

Centro de Investigación en Matemáticas A.C.
Unidad Aguascalientes
E-mail: `ivan.vega@cimat.mx`

I. Introducción

Dentro de las aplicaciones de clasificación de texturas para el aseguramiento de la calidad se encuentra la clasificación de defectos en la piel. Esta es una tarea que ha sido estudiada en [1], [2] y [3], donde se proponen diferentes enfoques para automatizar la clasificación de los defectos en la piel, que van desde métodos de aprendizaje automático hasta implementación de redes neuronales. En este trabajo se pretende hacer dos clasificaciones. Primero hacer una clasificación binaria para detectar si la imagen corresponde a una piel con defecto o no y en una segunda clasificación detectar el tipo de defecto en la piel, los cuales son cinco: folding marks, grain off, growth marks, loose grains y pinhole. Al final se analizan y comparan diferentes métodos implementados para ver la efectividad de cada clasificación.

II. Métodos

Para poder realizar la clasificación binaria y multiclase se siguió el esquema que se muestra en la Figura 1, el cual consiste en cuatro etapas:

1. Preparación de los datos.
2. Extracción de las características.
3. Selección y entrenamiento del modelo.
4. Evaluación del modelo.

La base de datos que se utilizó está formada por 6 carpetas con 600 imágenes cada una con una resolución de 227 x 227 píxeles. Una carpeta corresponde a imágenes que no tienen defecto en la piel y las cinco restantes corresponden a un tipo de defecto en la piel:

- Folding marks
- Grain off
- Growth marks

- Loose grains
- Pinhole

Debido a que todas las carpetas contienen el mismo número imágenes, la base de datos para la clasificación multiclase está balanceada por lo que se optó incluir la métrica de accuracy. Por otro lado, para la clasificación binaria se tendría una base desbalanceada si se tomaran en cuenta todas las imágenes de la base de datos original, por lo tanto, se eligieron aleatoriamente 120 imágenes de cada tipo de defecto de piel para conformar 600 imágenes de defectos en la piel.

A partir de los resultados obtenidos de la práctica anterior, se decidió trabajar únicamente con las imágenes en escala de grises para reducir el costo computacional, ya que los resultados al utilizar los tres canales de rgb fueron muy similares para la tarea de clasificación de texturas.

Con respecto a la etapa de extracción de características, primero se obtuvieron cuatro matrices de co-ocurrencia para cada imagen (con $\theta = \{0, 45, 90, 135\}$). Después, para cada matriz se obtuvieron las siguientes características:

1. Contraste
2. Disimilitud
3. Homogeneidad
4. ASM (angular second momment)
5. Energía
6. Correlación

Finalmente se concatenaron las 6 características por cada matriz. Así, cada imagen fue representada por un vector de dimensión 24. Tanto las matrices como el vector de características fueron normalizados.

III. Resultados

Los modelos que se utilizaron para realizar la clasificación fueron:

- Regresión logística.
- Máquina de soporte vectorial.
- Árbol de decisión.
- Bosque aleatorio.
- Método de Naive Bayes.
- K vecinos más cercanos.

La métrica que se utilizó para hacer la evaluación de los modelos de clasificación multiclase fue macro average sobre las métricas: precision, recall y f1-score, esto debido a la forma en la que se construyeron los datos, ya que las clases son balanceadas. Asimismo, tanto para la clasificación binaria como multiclase, se incluyeron las matrices de confusión para poder visualizar más detalladamente que tan buena fue la clasificación con respecto a cada clase.

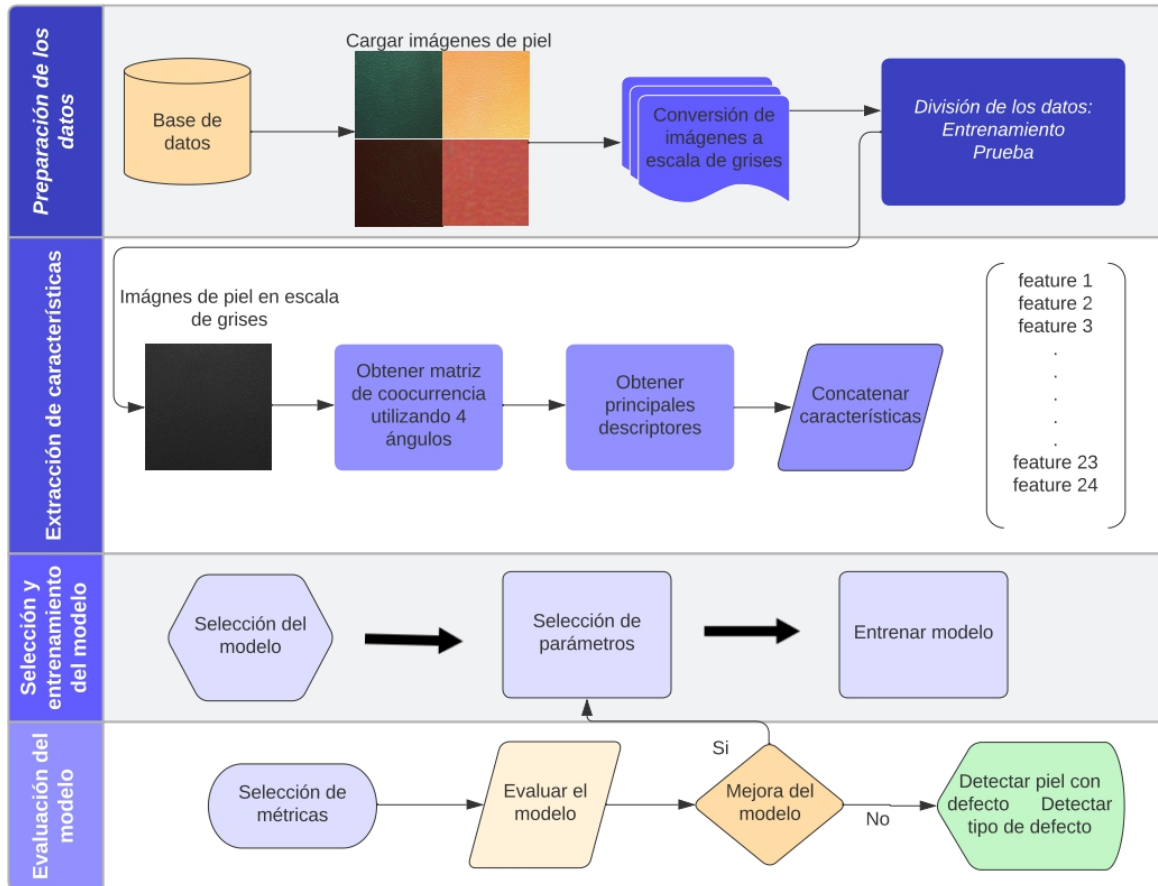


Figura 1: Pipeline

Método	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.69	0.69	0.69	0.69
SVM	0.75	0.73	0.74	0.74
Decision Trees	0.8	0.8	0.8	0.8
Random Forest	0.84	0.84	0.84	0.84
Naive Bayes	0.75	0.73	0.73	0.74
K-Nearest Neighbor	0.78	0.78	0.78	0.78

Tabla 1: Métricas de evaluación para la clasificación binaria

En las tablas 1 y 2 se muestran las métricas del rendimiento para la clasificación binaria y multi-clase, respectivamente. Se puede observar que para la clasificación binaria el bosque aleatorio es el modelo que mejor desempeño tiene, mientras que para la clasificación multiclase los mejores métodos fueron árbol de decisión, bosque aleatorio y k vecinos más cercanos.

Por otro lado, en la figura 2 se observa en general que se logra una ligera mejora al clasificar piel sin defecto. Con respecto a la clasificación multiclase, en la figura 3 se nota que en general, todos los modelos logran una buena clasificación para detectar los defectos: folding marks, growth marks y loose grains. Un resultado interesante fue que al utilizar el clasificador SVM fue incapaz de detectar el defecto pinhole.

Método	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.38	0.44	0.35	0.42
SVM	0.3	0.44	0.33	0.42
Decision Trees	0.54	0.55	0.54	0.54
Random Forest	0.53	0.55	0.54	0.55
Naive Bayes	0.5	0.47	0.41	0.48
K-Nearest Neighbor	0.54	0.56	0.54	0.55

Tabla 2: Métricas de evaluación para la clasificación multiclase

IV. Discusión

A partir de los resultados obtenidos podemos concluir que los modelos implementados en general son buenos para predecir si una imagen tiene algún defecto o no, sin embargo, al tratar de hacer la clasificación con respecto al tipo de defecto en la piel, el problema se vuelve más complejo y esto se ve reflejado en las métricas de evaluación.

Como trabajo a futuro se pretende mejorar los resultados obtenidos al optimizar los parámetros de cada modelo de clasificación, utilizando grid search. Asimismo, otro factor que es determinante para obtener una buena clasificación es en la etapa de características, por lo que sería interesante promediar las matrices de coocurrencia y a partir de una única matriz obtener un vector de características de dimensión seis, esto podría influir positivamente en el rendimiento de la clasificación. Otro enfoque que se podría implementar sería hacer un preprocesamiento de las imágenes más elaborado y reducir el ruido que exista en cada imagen para evidenciar las texturas que se presentan, ya que al revisar la base de datos de manera manual se nota algo heterogénea.

Por último, este enfoque de clasificación de texturas podría ser implementado en la industria para detectar fallas en materiales, más aún, identificar que tipo de falla se presenta en algún material, lo que sería de gran utilidad al verificar la calidad de algún producto.

Referencias

- [1] Moganam, P. K., Sathia Seelan, D. A. (2022). Deep learning and machine learning neural network approaches for multi class leather texture defect classification and segmentation. Journal of Leather Science and Engineering, 4(1), 7.
- [2] Moganam, P. K., Seelan, D. A. S. (2020). Perceptron Neural Network Based Machine Learning Approaches for Leather Defect Detection and Classification. Instrumentation, Measures, Metrologies, 19(6)
- [3] Kumar, M. P., Ashok, D. (2020). A multi-level colour thresholding based segmentation approach for improved identification of the defective region in leather surfaces. Engineering Journal, 24(2), 101-108..

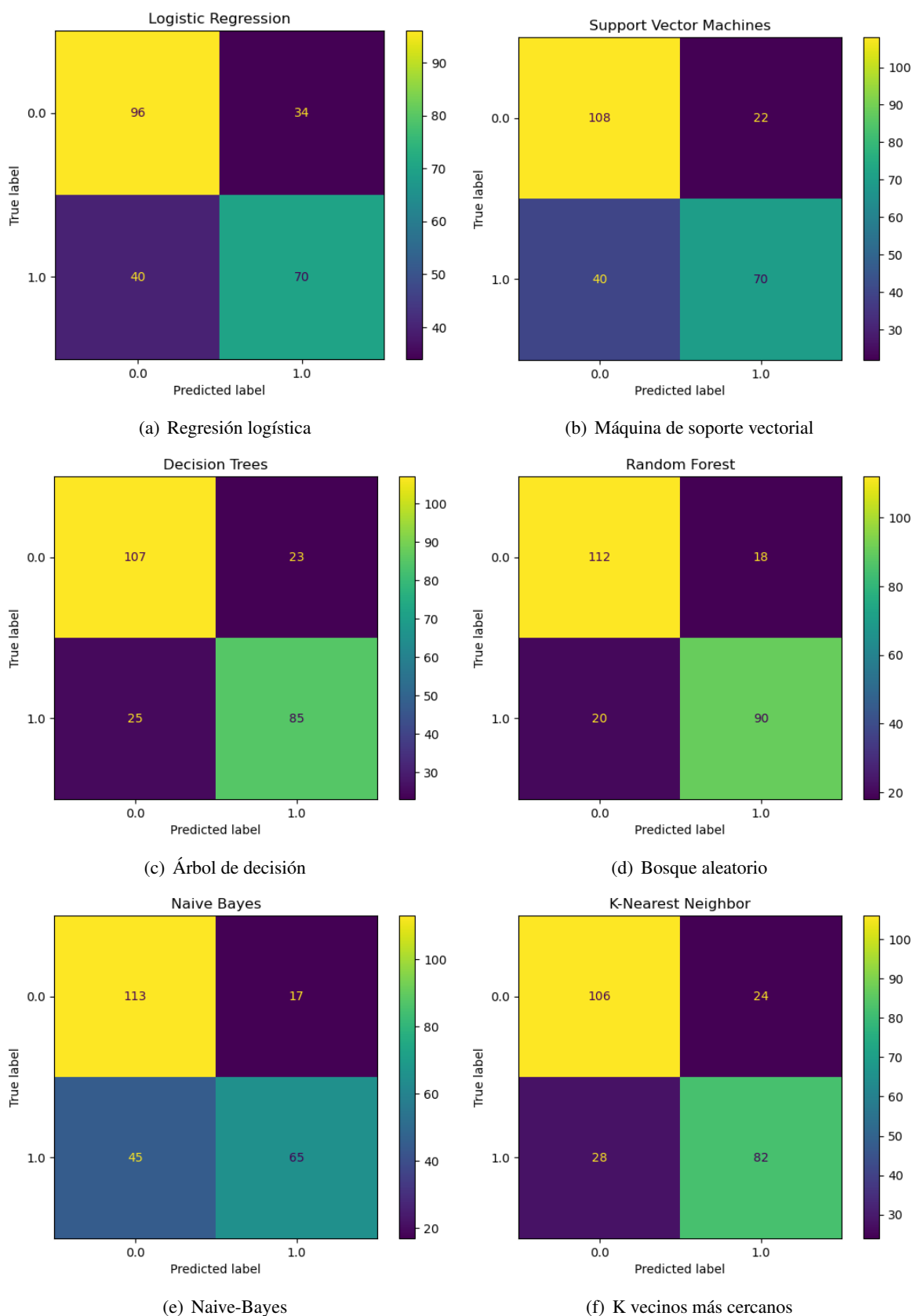


Figura 2: Matriz de confusión para la clasificación binaria

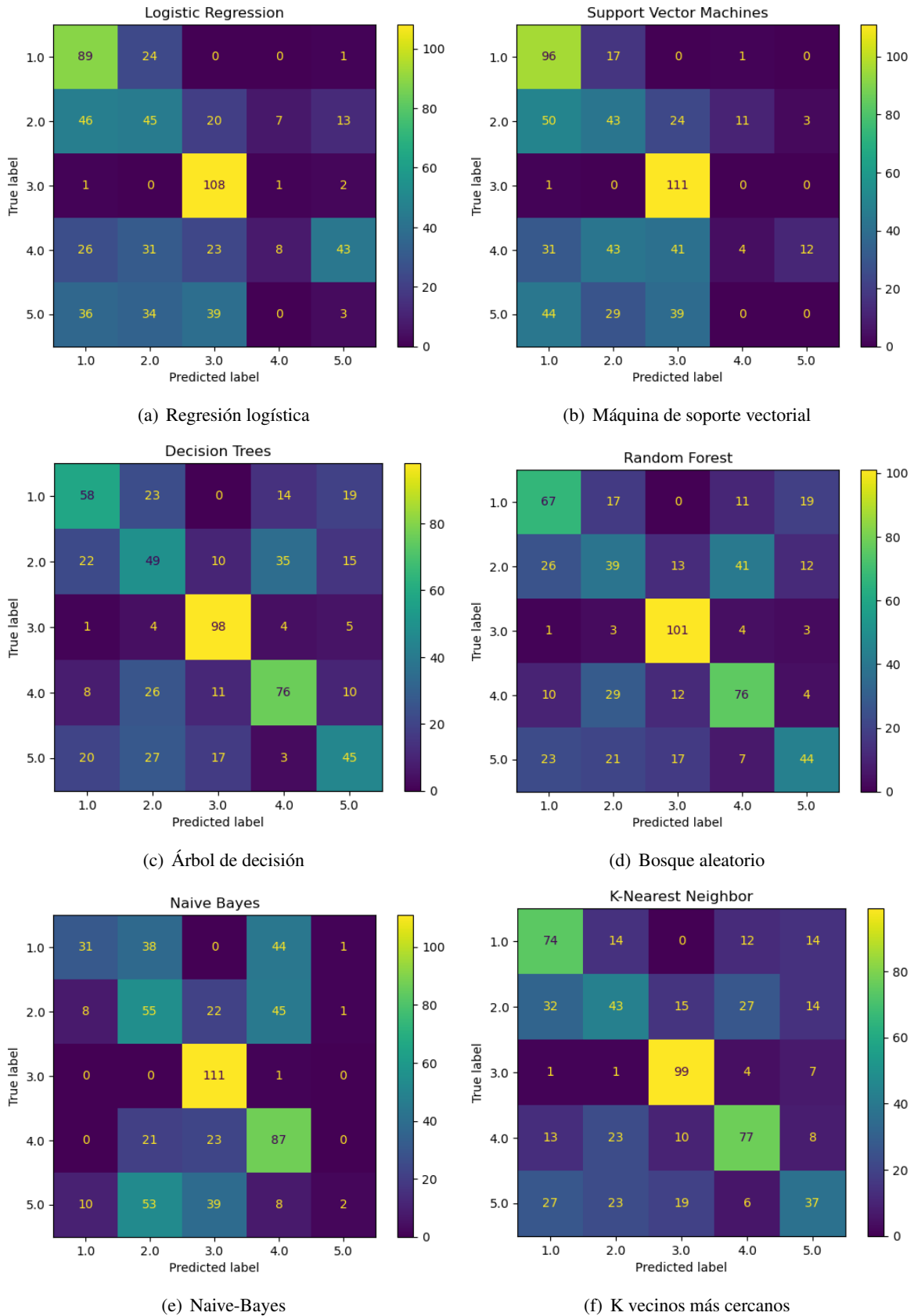


Figura 3: Matriz de confusión para la clasificación multiclase