



Tecnológico de Monterrey

Instituto Tecnológico y de Estudios Superiores de Monterrey

Procesamiento de Imágenes Médicas para el Diagnóstico (Grupo 101)

Reporte 3. Melanoma vs. Ketosis Seborreica

Profesor

Dr. José Gerardo Tamez Peña

Equipo CT:

Diego De La Barrera Martínez	A01197739
Alexa María de León Durán	A01382990
Juan Luis Flores Sánchez	A01383088
Azul Sofía Moctezuma Enriquez	A01562585

17 de febrero de 2023

Reporte 3. Melanoma vs. Ketosis Seborreica

1. Introducción

En esta práctica analizaremos una imagen para poder identificar si se trata de un melanoma o de queratosis. Como sabemos el melanoma y la queratosis seborreicas, tienen una apariencia similar, lo que ocasiona que no se puedan distinguir fácilmente una de la otra. La distinción de estas es muy importante debido a que la queratosis seborreica es un tumor epidérmico común, benigno, para el cual no es necesario un diagnóstico, por otro lado, el melanoma es una neoplasia cutánea con alta tasa de mortalidad que sí requiere de un diagnóstico y tratamiento.

2. Objetivos

Evaluar y reportar el desempeño de la correcta clasificación de Melanoma usando y comparando dos métodos de Machine Learning,

3. Marco teórico

Machine Learning es un proceso iterativo que consiste en varias etapas para desarrollar modelos de aprendizaje automáticos efectivos. Una de las aplicaciones más comunes es la segmentación de imágenes, que implica la separación de diferentes partes o regiones en una imagen. A continuación, se presenta una descripción general de las etapas de la metodología de Machine Learning:

- Recopilación de datos: Se recopilan datos relevantes para el problema que se va a resolver.
- Preprocesamiento de datos: Se limpian y transforman los datos para que sean adecuados para el modelado.
- Selección de características: Se seleccionan las características más importantes para el problema.
- Selección del modelo: Se selecciona el modelo de aprendizaje automático adecuado para el problema.
- Entrenamiento del modelo: Se entrena el modelo utilizando los datos de entrenamiento.
- Evaluación del modelo: Se evalúa el rendimiento del modelo utilizando los datos de prueba.
- Ajuste de modelo: Se ajusta el modelo para mejorar su rendimiento.
- Uso del modelo: Se utiliza el modelo para hacer predicciones o clasificaciones.

4. Metodología

Para la elaboración de segmentación de las lesiones de piel fue necesario realizar los siguientes pasos:

1. Crear elementos estructurales
2. Carga y procesamiento de imágenes a Matlab
3. Extracción de características de lesión

4. Extracción de características de control
5. Segmentación y evaluación
6. Análisis de segmentación de imágenes

Para la elaboración de un modelo de diagnóstico por computadora con base en herramientas y análisis de *machine learning* fue necesario realizar los siguientes pasos:

1. Carga de datasets de excel
2. Definir canales
3. Identificación de características
4. Obtención de mapa de calor entre las variables
5. Realizar ajuste de modelo
6. Obtener características principales del modelo
7. Obtención de K-nearest neighbour a través de las clases
8. Gráficas de resultados

5. Resultados

Durante las clases se proporcionaron imágenes para realizar un preprocesamiento y entrenamiento en Matlab de la segmentación por medio de los códigos proporcionados en clase, gracias a ellos se pudieron obtener las segmentaciones y evaluaciones de todas las imágenes, anexando dos de referencia (Figura 1 y 2):

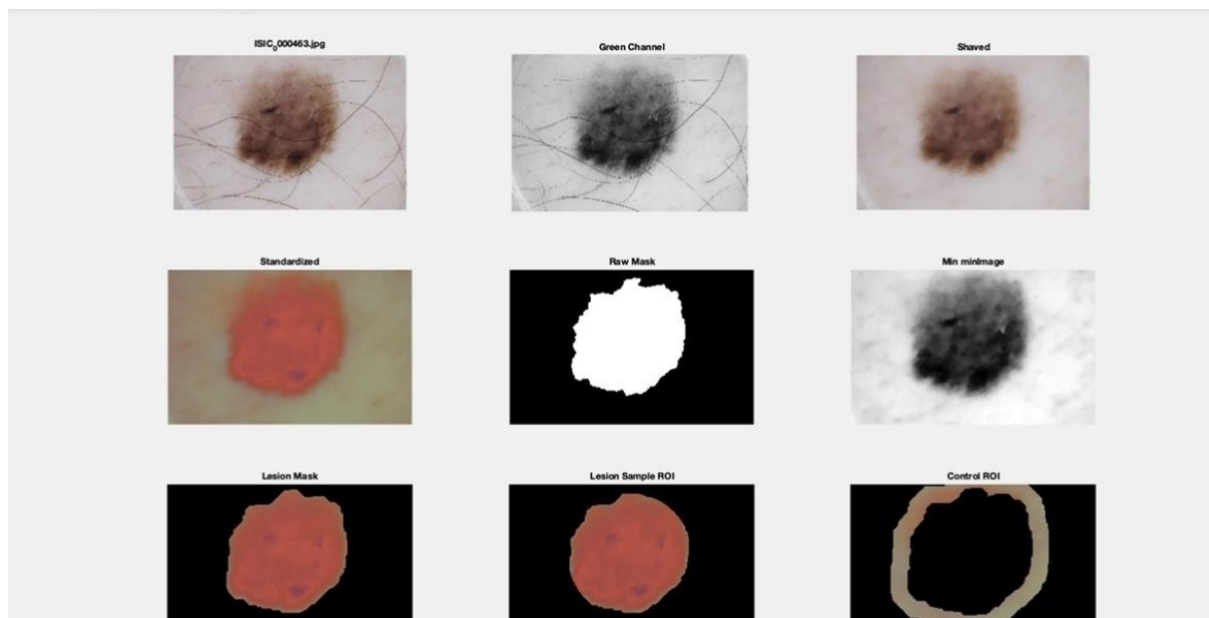
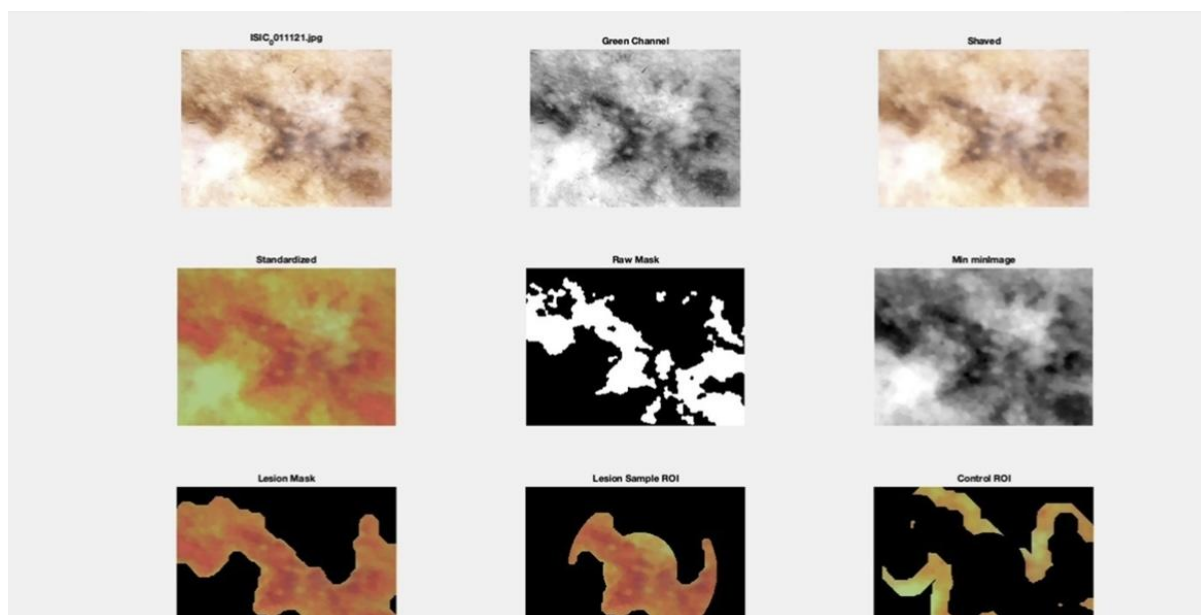
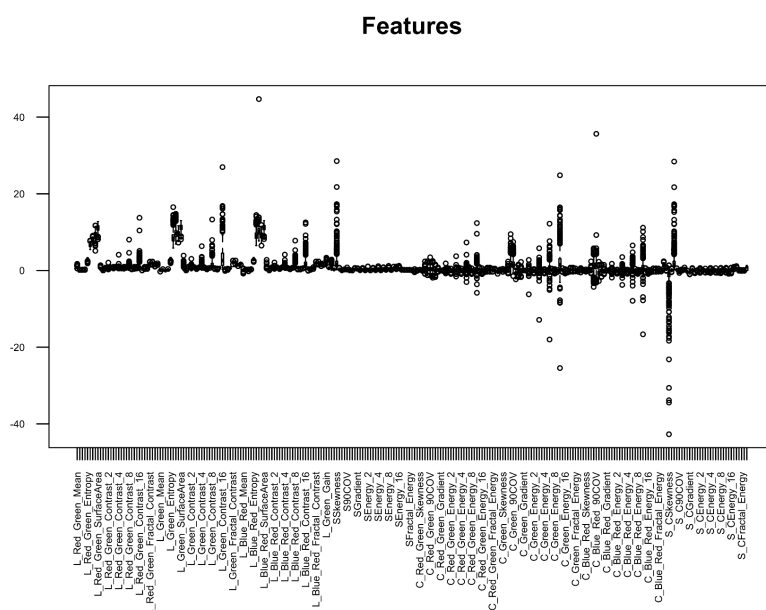


Figura 1. Segmentación precisa de la imagen 0000463.



Como se puede observar, aunque haya un preprocesamiento previo, al subir una imagen de la carpeta “clean”, el código puede llegar a tener fallas en la segmentación.

El gráfico (Figura 3) representa un conjunto de características dentro del modelo de aprendizaje automático y cómo se relacionan entre sí. Se genera utilizando técnicas de análisis de los componentes principales para reducir la dimensión de los datos. Luego, los datos se representan en el plano cartesiano, donde cada punto representa una observación y las variables se representan como vectores. Gracias a este gráfico podemos visualizar cómo se agrupan las observaciones en función o las relaciones entre las variables, además de permitirnos identificar si existen variables altamente correlacionadas, lo que puede afectar el rendimiento del modelo.



En la figura 4 Cada fila y columna representa una variable, y cada celda de la matriz representa la relación entre esas variables. La intensidad del color de cada celda indica el valor numérico de la relación correspondiente, con colores que indican los valores entre -2 y 2. Esta representación gráfica nos permite hacer una visualización más fácil de los conjuntos de datos, ya que se está trabajando con un número alto de información, además podemos detectar patrones y tenencia dentro de los datos para el análisis correspondiente.

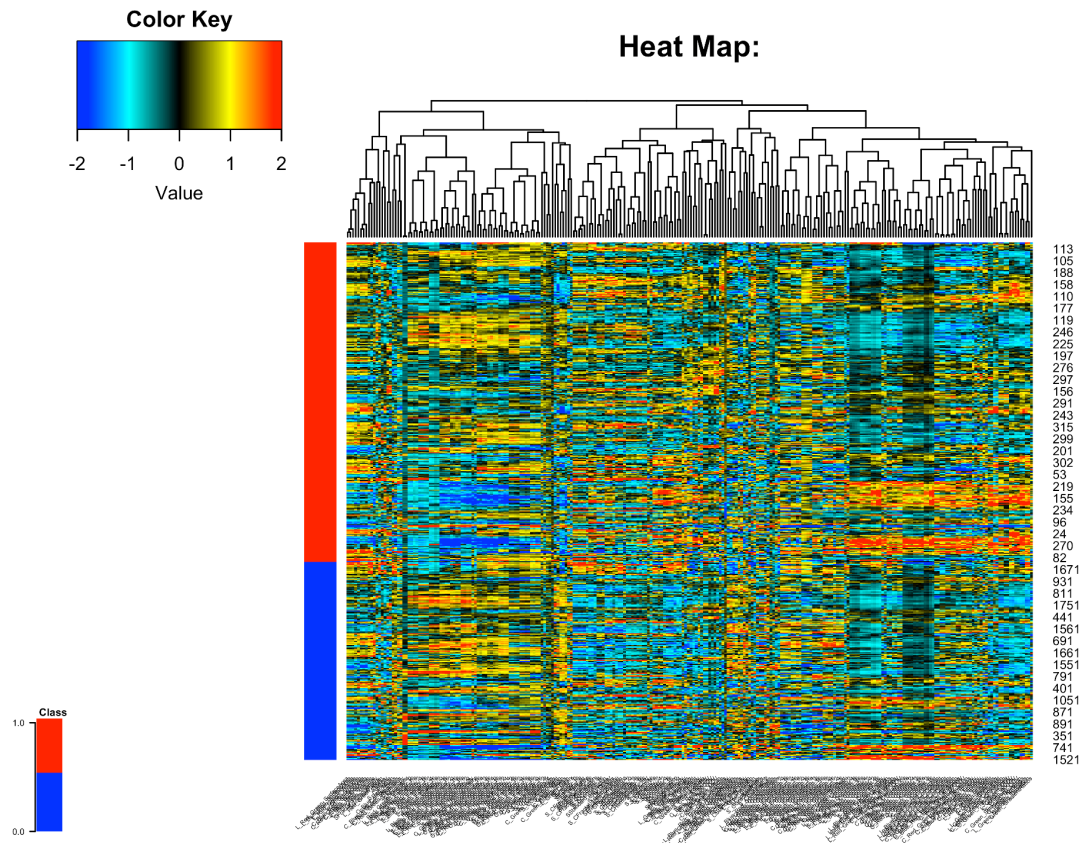


Figura 4. Mapa de calor entre variables

Después de ajustar el modelo de aprendizaje automático podemos generar el gráfico de las características principales (Figura 5), el cuál nos permite visualizar las características que más contribuyen a la predicción del resultado del modelo. Las características se encuentran en orden de acuerdo a su importancia relativa, que se calcula a través de los métodos de selección.

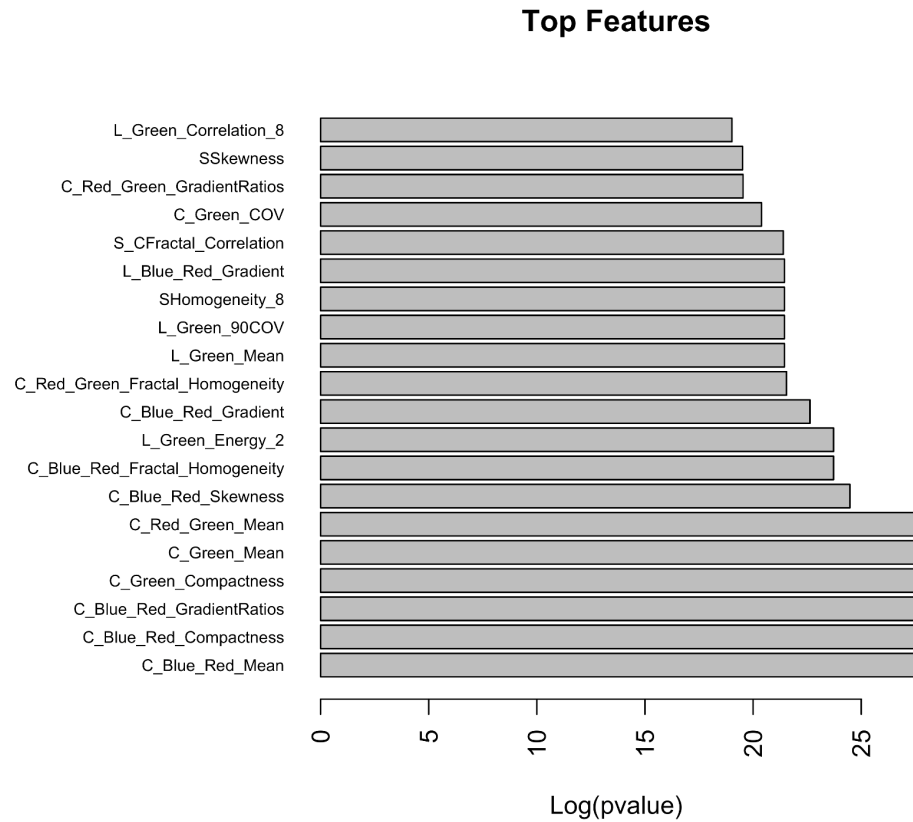


Figura 5. Visualización de características principales

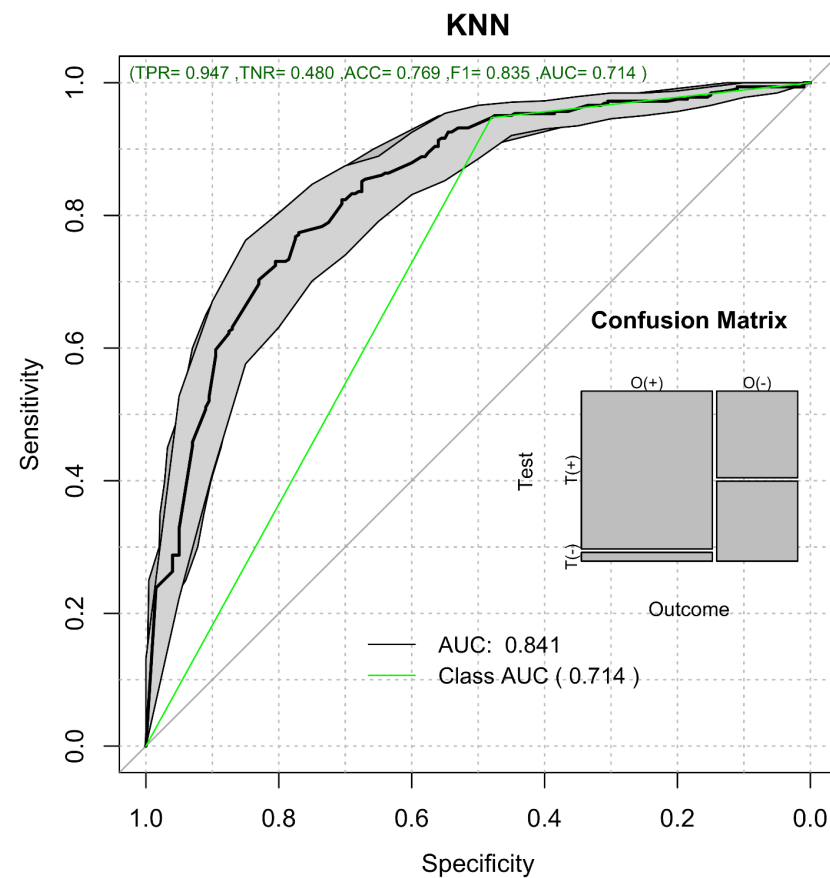
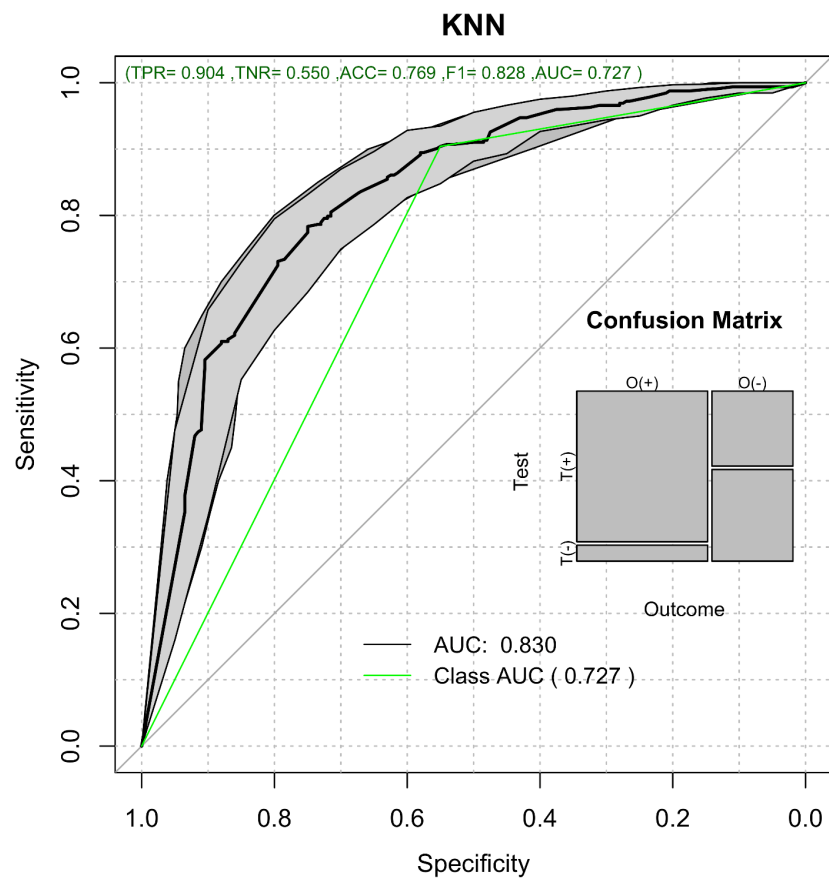


Figura 6. Resultados de Análisis «K- Vecinos Más Cercanos»

Se muestran los resultados de los resultados predictivos del algoritmo KNN para 5 y 10 vecinos más cercanos (izquierda y derecha, respectivamente). Este gráfico está complementado con ROC y un área bajo la curva (AUC) lo que nos indica la calidad del modelo de clasificación binaria, como se observa, el modelo con los 10 vecinos más cercanos tiene un AUC mayor (0.841).

6. Conclusiones:

Azul Sofía Moctezuma Enriquez: Durante esta práctica trabajamos temas importantes vistos previamente en clase, como lo es Machine Learning, donde en conjunto puedo concluir que una práctica exitosa de segmentación de imágenes, en este caso de melanoma, se debe tener en cuenta varios factores importantes, como la calidad y cantidad de datos de entrenamiento, la elección del modelo adecuado, la selección de características relevantes y la evaluación rigurosa del modelo. En cuanto a los procesos realizados, fue de suma importancia realizar el preprocesamiento a través de Matlab para la segmentación de cada imagen y así obtener un resultado de Machine Learning exitoso, donde se arrojaron diversas gráficas con las características más importantes para el modelo buscado, además de que nos apoyamos bastante en el AUC de las gráficas KNN para obtener un análisis más preciso de la segmentación realizada.

Diego Alejandro De la Barreda Martínez: Durante la práctica se pudo observar dos métodos en Machine Learning usados para clasificación de Melanoma. Primero vemos la segmentación de imágenes a través de una visualización en Matlab donde con un código nos da diferentes segmentaciones para una imagen de acuerdo a las necesidades que tengamos. Después se hizo un análisis con características donde se pudo identificar si la imagen mostrada era Melanoma o otro tipo de lesión. Los resultados arrojados nos muestran la relación entre las variables, las características más importantes y una curva ROC con área bajo el umbral donde nos muestra la especificidad y sensibilidad del modelo diciéndonos lo acertado que llega a ser.

Alexa María de León Durán: Con esta actividad se puso en práctica los conceptos vistos en clase, específicamente Machine Learning, a través de Matlab y el software R. Para este caso fue fundamental hacer un proceso de obtención de características y un análisis para poder hacer una clasificación de imágenes adecuada y así poder conocer el diagnóstico de las imágenes proporcionadas.

Juan Luis Flores Sánchez: Al finalizar esta práctica, fue posible adquirir nuevos conocimientos, reforzar conceptos y poner en práctica lo aprendido utilizando herramientas de *machine learning* como lo son MATLAB y R, todo esto a partir del proceso de obtención de características, discriminación de las mismas por importancia, generación y evaluación de modelos computacionales que permitan clasificar imágenes de acuerdo a sus características con la finalidad de darle un uso médico de diagnóstico a partir del entrenamiento de los mismos.

7. Referencias

[]. Gavilán, I. G. (2021, June 28). Metodología para Machine Learning (I): CRISP-DM. Ignacio G.R.Gavilán.
<https://ignaciogavilan.com/metodologia-para-machine-learning-i-crisp-dm/>