

Taller No.7 Clasificador de Rostro con KNN

Niño Rivera María Carolina, Cubides Ortiz Gabriel Andrés
{*maria.nino02*}@*usa.edu.co*, {*gabriel.cubides01*}@*usa.edu.co*
Profesor: Rudas Jorge

Resumen—En este artículo se detalla el proceso de clasificación de rostros femeninos y masculinos a través del algoritmo KNN. La estrategia para esta clasificación corresponde en una redimensión de las imágenes a un tamaño de 50*50 píxeles para luego ser enviado cada foto como descriptor, como etiquetas se utilizan las mismas encontradas en el conjunto de datos. Finalmente se prueban los resultados mediante la matriz de confusión, la curva ROC y el porcentaje de precisión.

Palabras clave—Detección de rostros, clasificación, KNN.

I. INTRODUCCIÓN

El reconocimiento de rostros en imágenes digitales, tiene como finalidad lograr que una máquina determine sin equívocos la identidad de uno o más rostros dentro de una imagen [1]. Hoy en día en el área de visión por computador, el reconocimiento de imágenes en especial de rostros está en auge, debido a las grandes posibilidades que esto trae, como es en el campo de la seguridad, además en campos como minería de datos y reconocimiento de patrones. Entre las aplicaciones que más se utiliza el reconocimiento de datos están [2]:

1. Vigilancia. Buscar a un individuo perseguido por la justicia. Localizar a las personas en ámbitos públicos, aduanas, aeropuertos. Vigilancia doméstica, quien entra puede ser conocido o desconocido: detección de intrusos.
2. Videoconferencia. Tiene que ver con localizar la imagen del individuo en una secuencia de web cam para poder hacer un seguimiento. También saber si está o no está, y quién es.
3. Interfaces hombre-máquina. Seguridad de log in en una computadora.
4. Detección de expresiones faciales (se considera un subcampo del reconocimiento de caras en general): interfaces inteligentes, detectores de cansancio para conductores, aplicaciones médicas.
5. Control de acceso. Tarjetas de identificación, detección de fraudes (usurpar identidad), etc. Ejemplo: detectar carnet falso.
6. Base de datos multimedia. El incremento de imágenes en el world wide web sugiere cada vez más la existencia de aplicaciones CBIR de recuperación de imágenes. En sistemas con bibliotecas digitales de terabytes de video y audio la clasificación de las imágenes cumple un papel fundamental.

El presente documento corresponde a un informe de práctica de laboratorio de Inteligencia Artificial presentado en la Universidad Sergio Arboleda durante el periodo 2022-2.

II. CLASIFICACIÓN KNN

El algoritmo de las K vecinos más cercanos o K-Nearest Neighbors (KNN) es un algoritmo de Machine Learning que pertenece a los algoritmos de aprendizaje supervisado simples y fáciles de aplicar que pueden ser utilizados para resolver problemas de clasificación y de regresión [3]. En aprendizaje supervisado, un algoritmo recibe un conjunto de datos que están etiquetados con los valores de salida correspondientes sobre los que puede entrenarse y definir un modelo de predicción. Este algoritmo podrá luego ser utilizado sobre datos nuevos para predecir sus valores de salida correspondientes [3].

K-Nearest-Neighbor es un algoritmo basado en instancia de tipo supervisado de Machine Learning. Puede usarse para clasificar nuevas muestras (valores discretos) o para predecir (regresión, valores continuos). Como pros tiene sobre todo que es sencillo de aprender e implementar. Tiene como contras que utiliza todo el dataset para entrenar “cada punto” y por eso requiere de uso de mucha memoria y recursos de procesamiento (CPU). Por estas razones kNN tiende a funcionar mejor en conjuntos de datos pequeños y sin una cantidad enorme de características [4].

III. TALLER

Para el taller de clasificación de rostros por medio del algoritmo de KNN se utilizó el conjunto de datos libre de Kaggle [5] el cual es un conjunto de rostros de hombres y mujeres compuesto por 5418 imágenes en formatos “jpeg”, “jpg” y “png”, se encontraron varias imágenes repetidas, además las dimensiones de las imágenes encontradas son diferentes. Con un total de 2698 rostros de mujeres y 2720 en el conjunto de los hombres. El primer paso para realizar este taller consistió en la tarea de reducir las imágenes debido a que la metodología KNN revisa píxel por píxel y en el conjunto de datos se tiene algunas con tamaños superiores a los 3500 píxeles. Además para homogenizar los formatos a utilizarse, para el caso del taller se utilizó el formato “jpeg”. La medida de redimensión para el taller fue de 50*50 píxeles. El algoritmo KNN no utiliza características sino que toma la imagen como elemento de entrada, debido a esto el conjunto de datos se leyó de manera intercalada entre las imágenes de mujeres y hombres, se le adicionó las etiquetas “0” para mujer y “1” para hombre, tal cuál se encontraban etiquetadas los nombres en el conjunto de datos original. Luego de ejecutarse varias veces el mejor resultado sin llegar a un sobre ajuste se encontró al dividir el conjunto de datos en un porcentaje 80 % y 20 %.

IV. RESULTADOS

El algoritmo presentó unos resultados buenos ya que la precisión encontrada fue de un 88 % de precisión aproximada en los datos de prueba y un 93 % en los datos de entrenamiento. Las métricas de desempeño puede verse mejor en la tabla I. Al revisar la matriz de confusión se puede observar que de 549 mujeres clasificó de forma correcta 480 y en el caso de los hombres de 531 casos clasificó correctamente 471, tal como se muestran en la figura 1

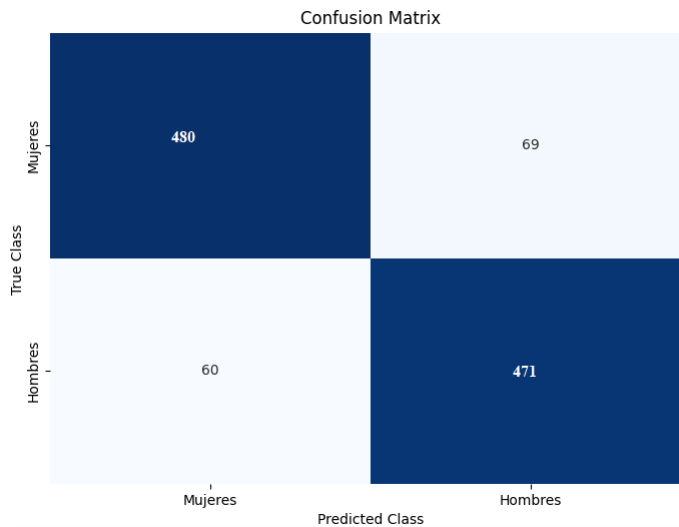


Figura 1. Resultados Matriz de Confusión

La siguiente es la imagen donde detecta los puntos encontrados en el algoritmo de KNN



Figura 2. Imagen donde se resaltan los puntos encontrados en KNN

Adicionalmente se realizó la gráfica ROC para evaluar la calidad de salida del clasificador. Las curvas ROC suelen presentar una tasa de verdaderos positivos en el eje Y y una tasa de falsos positivos en el eje X. Esto significa que la esquina superior izquierda de la gráfica es el punto “ideal”: una tasa de falsos positivos de cero y una tasa de verdaderos positivos de uno. Esto no es muy realista, pero significa que un área bajo la curva (AUC) más grande suele ser mejor [6].

La “inclinación” de las curvas ROC también es importante, ya que es ideal para maximizar la tasa de verdaderos positivos y minimizar la tasa de falsos positivos. Las curvas ROC se utilizan normalmente en la clasificación binaria para estudiar la salida de un clasificador. Para extender la curva ROC y el área ROC a la clasificación de etiquetas múltiples, es necesario binarizar la salida. Se puede dibujar una curva ROC por etiqueta, pero también se puede dibujar una curva ROC

Resultados métricas de desempeño				
Mujeres	Precisión 0.87	Recall 0.89	F1-Score 0.88	Support 540
Hombres	0.89	0.87	0.88	540
Accuracy			0.88	1080
Macro avg	0.88	0.88	0.88	1080
Weighted avg	0.88	0.88	0.88	1080

Tabla I
MÉTRICAS DE DESEMPEÑO

considerando cada elemento de la matriz del indicador de la etiqueta como una predicción binaria (micropromedio) [6]. En los resultados obtenidos en la curva ROC se demuestra que esta cerca al punto ideal, lo cual demuestra una cercanía con el valor de precisión obtenido. La gráfica 3 muestra los resultados obtenidos. El AUC significa área bajo la curva (ROC) en inglés. Generalmente, cuanto mayor es la puntuación AUC, mejor es el rendimiento de un clasificador binario para una tarea de clasificación dada, en el caso del taller el porcentaje AUC obtenido es de 88.06 %.

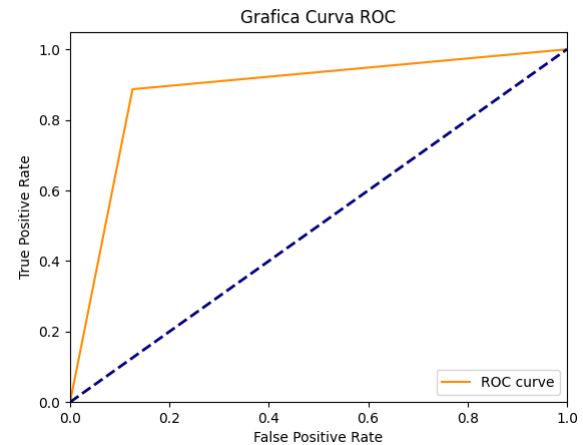


Figura 3. Resultados Grafica Curva ROC

IV-A. Preguntas

¿Qué pasa si clasificamos rostros de hombres y mujeres sin descriptores y utilizando knn? Al tratarse de una clasificación supervisada en donde se le asignan las etiquetas a las imágenes se logra hacer una clasificación bastante buena a pesar de no utilizar los descriptores, en este caso utiliza como descriptor toda la imagen y extrae las características píxel a píxel, debido a esto no es posible utilizar imágenes muy grandes, ya que tomaría demasiado proceso computacional para terminarse.

¿Qué pasa si pre procesamos un poco cada imagen? En este caso es necesario realizar un preprocesamiento de reducción de la imagen para que pueda extraerse las características. Al realizar un preprocesamiento para dejar más cuadradas las fotos, adicionó un marco negro y por esta razón los resultados obtenidos fueron peores.

¿Alguna propuesta para mejorar el performance? Identificamos que al utilizar 3 vecinos y no cinco el desempeño era

Resultados métricas de desempeño 3 vecinos				
	Precisión	Recall	F1-Score	Support
Mujeres	0.93	0.95	0.94	540
Hombres	0.95	0.92	0.94	540
Accuracy			0.94	1080
Macro avg	0.94	0.94	0.94	1080
Weighted avg	0.94	0.94	0.94	1080

Tabla II
MÉTRICAS DE DESEMPEÑO

mucho mejor, pasando de 93 % a 98 % en entrenamiento y de 88 % a 94 % en los datos de prueba, debido a que realizaba mejor la distinción de las pequeñas pero importantes diferencias entre los rostros de hombres y mujeres.

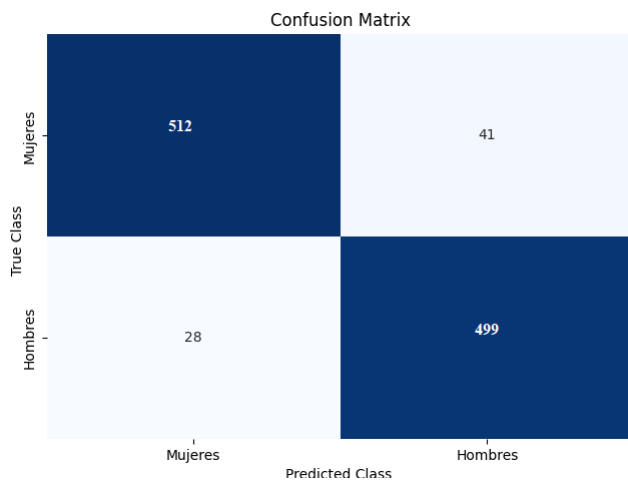


Figura 4. Resultados Matriz de Confusión 3 Vecinos

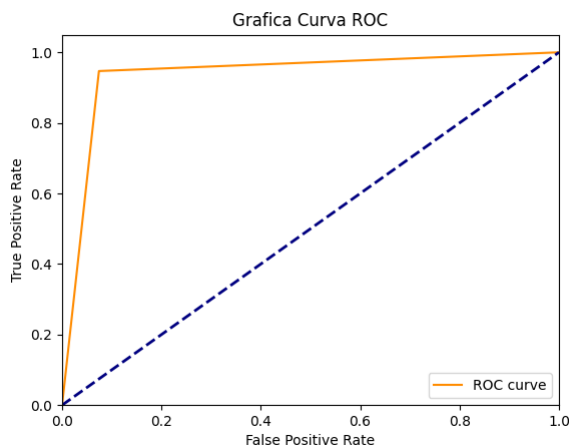


Figura 5. Resultados Grafica Curva ROC 3 vecinos

¿Por qué si funcionaba con el problema de clasificación de números? Funcionaba mejor debido a que las características a examinar no son tantas como en el caso de los rostros.

REFERENCIAS

- [1] Reconocimiento de Rostros en Tiempo Real sobre. **Cardona Lopez, Alexander and Pineda Torres, Franklin**. Mayo 29, 2018.

- [2] Analisis comparativo de metodos basados en subespacios aplicados al reconocimiento de caras. e. **Ferri, Francesc and Armengot, Marcelo**. 2006.
- [3] ¿Qué es el algoritmo KNN? **DataScientest**. 12/28. <https://datascientest.com/es/que-es-el-algoritmo-knn>
- [4] Aprenda Machine Learning **aprendemachinelarning**. 2018/07/10. <https://www.aprendemachinelarning.com/clasificar-con-k-nearest-neighbor-ejemplo-en-python/>
- [5] Male and female faces dataset *Kaggle*, 2021 <https://www.kaggle.com/datasets/ashwngupta3012/male-and-female-faces-dataset?resource=download>. Recuperado el 11 de Marzo de 2023.
- [6] Receiver Operating Characteristic (ROC) *scikit*, 2023 https://scikit-learn.org/1.0/auto_examples/model_selection/plot_roc.html. Recuperado el 11 de Marzo de 2023.