

Guillermo Ortiz Macías
216787558
guillermo.ortiz7875@alumnos.udg.mx
Maestría en Ciencia de los Datos
Programación 2

Challenge 3: Análisis de métodos de Machine Learning para el problema de tesis.



Este challenge trata de repasar las distintas técnicas de estadística y de machine learning que se estuvieron viendo a través del semestre e investigar cómo estas técnicas podrían ayudar a resolver mi problema de tesis.

Mi tesis trata sobre el diagnóstico temprano de las retinopatías pigmentarias mediante análisis de imágenes del fondo del ojo, utilizando técnicas de inteligencia artificial y ciencia de datos. La retinopatía pigmentaria es una patología ocular hereditaria que se caracteriza por la pérdida lenta y progresiva de la visión, destruyendo de manera gradual las células de la retina.

Durante el semestre estas fueron las técnicas estadísticas y de machine learning que se estudiaron:

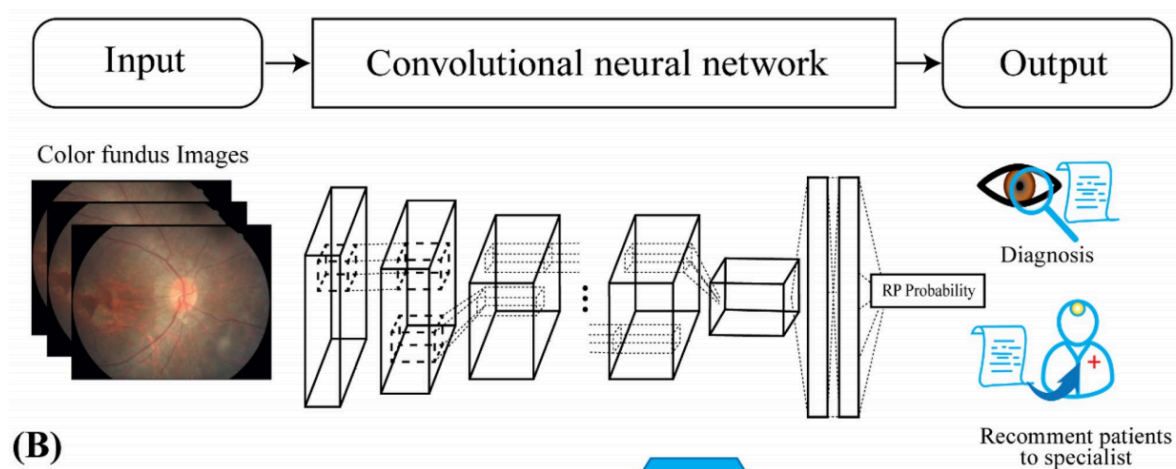
- Regresión lineal
- Regresión logística
- Árboles de decisión
- Random Forest
- Máquinas de soporte vectorial (SVM)
- Redes Neuronales
- k-Vecinos más cercanos (k-NN)
- Gradient Boosting (XGBoost, LightGBM)
- Clústering por K-means
- Clústering jerárquico
- DBSCAN
- Análisis de componentes principales (PCA)
- Autoencoders
- Mapas Autoorganizados (SOM)
- Reducción de dimensionalidades con t-SNE

Entre todas estas distintas técnicas, investigué cuáles son las que me más podrían ayudar para mi problema de tesis, y estas fueron las más prometedoras:

Redes Neuronales

Al investigar sobre el uso de las redes neuronales para la detección temprana de enfermedades por medio de imágenes, y específicamente patologías relacionadas al fondo del ojo, encontré que las redes neuronales han sido utilizadas en varios trabajos de investigación que han intentado resolver este mismo problema y problemas similares.

En la investigación de [Chen et al., 2021] se aplicaron 3 modelos de *deep learning* y de redes neuronales convolucionales llamados **Inception V3**, **Inception Resnet V2** y **Xception**. En la investigación se entrenó una red neuronal profunda para cada uno de estos tres modelos y se encontró que Xception es el que tiene mejores resultados con un área bajo la curva ROC del 96.89%:



La investigación también comenta que, aunque los resultados son buenos, para los estándares de deep learning la cantidad de imágenes que se utilizaron es poca (1284 para entrenamiento y 386 para validación), por lo que se necesitan otras investigaciones con conjuntos de datos más grandes.

Autoencoders

En la investigación de [Radha, K, Karuna, Y., 2025], se menciona que la retinopatía pigmentaria es una patología rara cuando se compara con otras patologías oculares, por lo que la cantidad de datos es limitado, esto incluye a los datasets que se pueden encontrar sobre imágenes del fondo del ojo de ojos que presentan esta patología.

Como describí en la sección de redes neuronales, una opción de modelo es utilizar una red neuronal convolucional y estas necesitan de datasets grandes. Por esta razón es posible que en algún momento de mi tesis vaya a tener que generar imágenes sintéticas de fondos de ojo con retinopatía pigmentaria. Una forma de generar estas imágenes es a través de autoencoders.

En su investigación [Radha, K, Karuna, Y., 2025] hablan sobre el diagnóstico de retinopatías diabéticas usando deep learning y sobre cómo utilizaron autoencoders variacionales (VAE) y redes generativas adversarias (GAN) para generar imágenes sintéticas realistas del fondo del ojo.

Máquinas de soporte vectorial (SVM)

En la investigación de [Pin, W. En, F. Peng, W., 2021], hacen una comparación de la métrica entre un modelo de tipo máquina de soporte vectorial y una red neuronal convolucional para la clasificación de imágenes de distintos conjuntos de datos. Buscan comparar el rendimiento de estos dos modelos para:

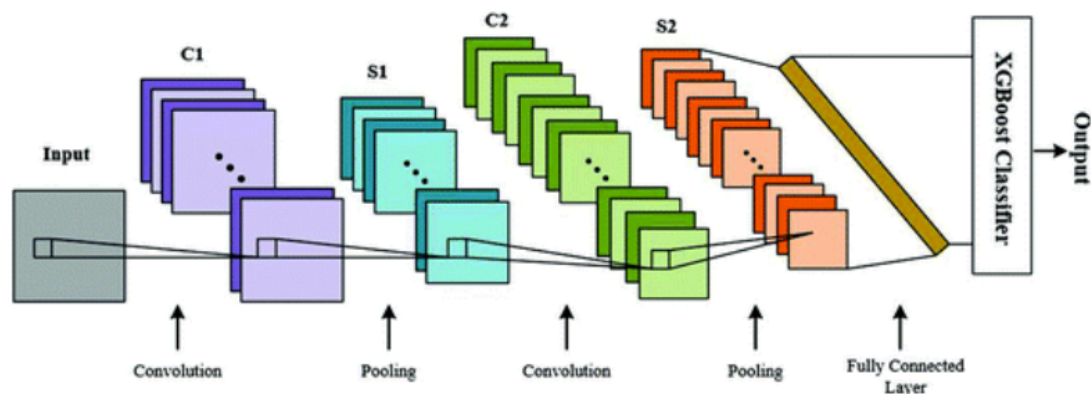
- Conjuntos de datos grandes (10,000 imágenes)
- Conjuntos de datos pequeños (1,000 imágenes)
- Distintos tamaños de imágenes: 64x64, 128x128 y 256x256

Las conclusiones a las que llega la investigación es que mientras más imágenes se tenga en el conjunto de datos y mientras más grandes sean las dimensiones de las imágenes, las CNN tienden a tener un mejor rendimiento a comparación de las SVM; pero con un conjunto de datos menor y con imágenes con menores dimensiones una SVM puede dar también buenos resultados (SVM tiene exactitud de 0.86 y el CNN de 0.83).

Esto me da la idea de que, con un conjunto de datos con un número reducido de imágenes, podría empezar a hacer un primer modelo usando una máquina de soporte vectorial y, conforme se vayan obteniendo más imágenes en el conjunto de datos, pasar a utilizar una red neuronal convolucional.

Gradient boosting:

En los artículos [Azupur, J.,] [Pang L et al, 2019] tienen la idea de combinar una red neuronal convolucional con un clasificador de tipo eXtreme Gradient Boosting para mejorar el rendimiento del modelo cuando funciona únicamente con una CNN. La idea se basa en extraer características de calidad usando la CNN y alimentar el XGBoost con estas características para hacer la clasificación de imágenes, siendo el XGBoost la capa final de la red neuronal:



Análisis de componentes principales (PCA)

En la investigación de [Ji, M., Yuyu, Y., 2019] se dice que PCA puede ayudar a reducir las dimensionalidades de las características extraídas de modelos de aprendizaje profundo que utilizan imágenes, como son las redes convolucionales. En el paper se describe un proceso que realiza tres cosas:

1. Extraer características de imágenes utilizando redes neuronales convolucionales profundas.
2. Utilizar PCA para reducir las dimensiones de las características extraídas por la CNN.
3. Optimización del algoritmo de PCA para procesar la información altamente dimensional extraída.

Las características principales de una imagen para hacer una clasificación pueden ser un histograma, texturas, formas de contornos, etc, una sola imagen puede tener cientos de distintas características importantes para su clasificación, PCA puede ayudar a conservar las características principales haciendo reducción de las dimensiones de la imagen, manteniendo componentes con mucha varianza y quitando aquellos que no tengan tanta [Ji, M., Yuyu, Y., 2019].

Para deep learning la reducción de dimensionalidades de PCA se puede ayudar con la **entropía de información**. Se utiliza primero la entropía de información para quitar características no útiles, para después realizar la reducción de dimensionalidad. Mientras mayor sea la entropía de información de un dato, mayor información de ese dato se debería mantener, y viceversa: si un dato tiene poca entropía de información, entonces ese dato contiene poca información, por lo que se podría quitar ese dato a la hora de hacer reducción de dimensionalidad. Cuando se aplica entropía de la información se define, utilizando un umbral, si un dato se debe mantener o no dependiendo de su magnitud de entropía de información [Ji, M., Yuyu, Y., 2019].

Siguiente el mismo paper [Ji, M., Yuyu, Y., 2019] se aplica un algoritmo mejorado de PCA de esta forma

- Se extraen las características de las imágenes utilizando aprendizaje profundo (CNN) y se transforman en matrices.
- Se calcula la entropía de información de cada característica y se rechazan aquellas que no tengan la entropía suficiente.
- Se aplica el algoritmo de PCA utilizando solo las características que pasaron el filtro de entropía. Esto logra mejores resultados de reducción de dimensionalidad, utilizando menores recursos de cómputo y de tiempo.

Una ventaja de utilizar este algoritmo mejorado de PCA está en el consumo de memoria, en el paper se compara que sin el algoritmo mejorado el uso de memoria fue de 6133 MB mientras que con el algoritmo mejorado apenas llega a los 1000MB.

También hay una gran mejoría en el tiempo de procesamiento: sin la mejora del algoritmo se toma 1300s, mientras que con la mejora son solo 100s.

J. Ma, Y. Yuan/J. Vis. Commun. Image R. 63 (2019) 102578

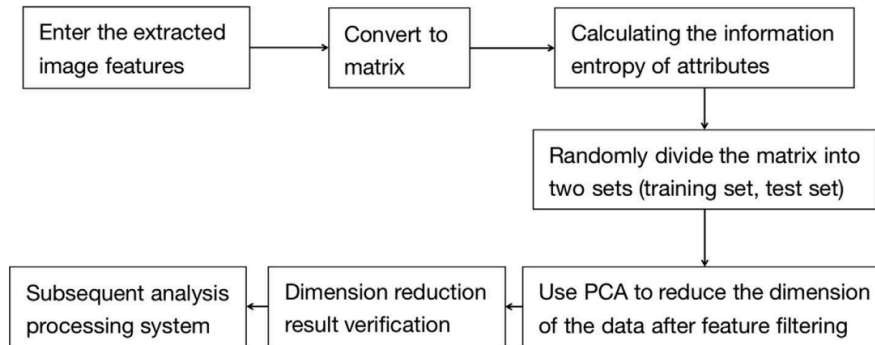


Fig. 1. Flow chart of dimension reduction processing of this method.

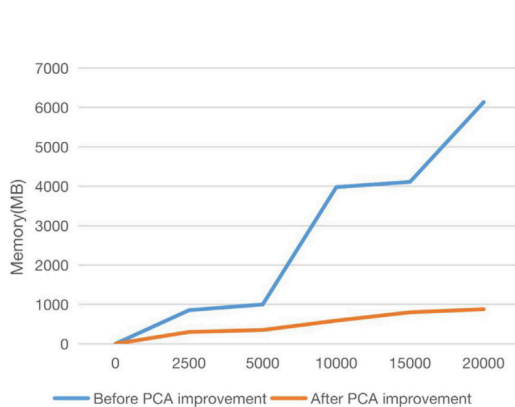


Fig. 4. Memory usage comparison of PCA algorithm before and after improvement.

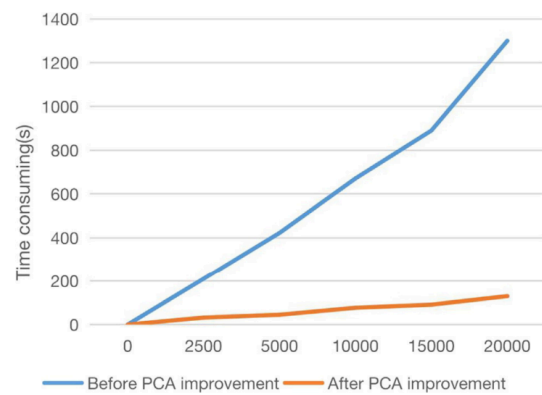


Fig. 5. Time-consuming comparison of PCA algorithm before and after improvement.

Y lo mejor de todo es que que esta mejora del algoritmo de PCA no afecta en la precisión del modelo, sino que lo mejora:

Table 1

Performance comparison of classification accuracy before and after improvement of PCA algorithm.

Image	Before the PCA algorithm is improved		After the PCA algorithm is improved	
	Feature dimension	Classification accuracy (%)	Feature dimension	Classification accuracy (%)
1	3692	88.3	923	91.3
2	4325	89.6	1109	90.6
3	2917	87.3	514	89.3
4	3977	86.7	1003	90.2

Conclusiones

Con toda esta investigación realizada pienso que lo más correcto para mi proyecto de tesis es primero hacer un modelo utilizando una máquina de vector soporte (SVM) pues pienso que al principio no voy a tener muchas imágenes del fondo de ojo disponibles.

A la vez, podría empezar a investigar sobre las redes neuronales y sobre todo las redes neuronales convolucionales para entender la matemática detrás de ellas y cómo implementarlas.

Los autoencoders los podría utilizar para generar imágenes sintéticas y así tener más imágenes para hacer el entrenamiento de los modelos.

Y podría combinar las redes neuronales convolucionales con una capa final de tipo clasificación de gradient boosting, teniendo una capa intermedia de PCA donde se haga reducción de dimensionalidad para facilitar así el proceso de clasificación.

Referencias

[Chen et al., 2021] Chen, TC., Lim, W.S., Wang, V.Y. et al. Artificial Intelligence–Assisted Early Detection of Retinitis Pigmentosa — the Most Common Inherited Retinal Degeneration. J Digit Imaging 34, 948–958 (2021). <https://link.springer.com/article/10.1007/s10278-021-00479-6>

[Radha, K, Karuna, Y., 2025] Radha, K., Karuna, Y. Latent space autoencoder generative adversarial model for retinal image synthesis and vessel segmentation. BMC Med Imaging 25, 149 (2025). <https://bmcmmedimaging.biomedcentral.com/articles/10.1186/s12880-025-01694-1>

[Pin, W. En, F. Peng, W., 2021] Pin Wang, En Fan, Peng Wang. Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. Pattern Recognition Letters, Volume 141, Pages 61-67 (2021). <https://www.sciencedirect.com/science/article/abs/pii/S0167865520302981>

[Azupur, J.,] CNNs and Gradient Boosting for Image Classification. <https://jonaac.github.io/works/deepxgboost.html>

Pang L, Wang J, Zhao L, Wang C, Zhan H. A Novel Protein Subcellular Localization Method With CNN-XGBoost Model for Alzheimer's Disease. Front Genet. 2019 Jan 18;9:751. doi: 10.3389/fgene.2018.00751. PMID: 30713552; PMCID: PMC6345701. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6345701/>

[Ji, M., Yuyu, Y., 2019] Ji Ma, Yuyu Yuan, Dimension reduction of image deep feature using PCA, Journal of Visual Communication and Image Representation, Volume 63, 2019.
<https://sci-hub.se/10.1016/j.jvcir.2019.102578>