

Artículo de investigación

Clasificación de Espectros Astrofísicos usando Algoritmos de Aprendizaje Profundo Pre-entrenados

Classification of Astrophysical Spectra using Deep Learning Pre-Trained Algorithms

Vanesa Lara Cruz¹ ✉ y Luz Ángela García ¹ ✉

¹Universidad ECCI, Cra. 19 No. 49-20, Bogotá, Colombia, Código Postal 111311

Recepción: 13-jun-2023 **Aceptado:** 15-abril-2024 **Publicado:** 23-julio-2024

Cómo citar: Lara Cruz, V., & García Peñaloza, L. Ángela. (2024). Clasificación de Espectros Astrofísicos usando Algoritmos de Aprendizaje Profundo Pre-entrenados. *Ciencia en Desarrollo*, 15(2). <https://doi.org/10.19053/01217488.v15.n2.2024.15965>

Resumen

La implementación de la espectroscopia como herramienta en la clasificación de objetos astrofísicos es una técnica que cada día se vuelve de mayor utilidad para la astronomía moderna. Actualmente, hay una serie de telescopios espaciales y terrestres cartografiando el cielo cada noche, incrementando el volumen de datos de forma exponencial en el tiempo. Esto lleva al límite los esfuerzos de los astrónomos para clasificar eficientemente los espectros de los objetos observados. En este trabajo, se implementan los algoritmos pre-entrenados: INCEPTION V3, RESNET 50 y MNIST, que se basan en Visión Computacional y Redes Neuronales Convolucionales (CV y CNN, por sus siglas en inglés). En este trabajo realizamos una comparación del rendimiento en la clasificación de espectros tomados de la campaña Sloan Digital Sky Survey (SDSS), Data Release (DR12) del espectrógrafo Baryon Oscillation Spectroscopic Survey (BOSS) con una muestra de 300000 espectros. En el proceso de aprendizaje se ajustaron los hiperparámetros asociados a la entrada de cada modelo, y cada algoritmo se valoró implementando las métricas de exactitud, precisión y sensibilidad. Los resultados de la clasificación demuestran que el mejor clasificador para espectros de cuásares es RESNET 50 con un rendimiento de más de 60 % con respecto a su desempeño con galaxias y estrellas, además ratifican que la implementación de algoritmos de Aprendizaje Profundo para la clasificación de espectros astrofísicos es una estrategia que debe incorporarse a las rutinas de trabajo de los nuevos *surveys* espectroscópicos con el fin de optimizar el tiempo y la calidad durante el proceso de clasificación manual que hacen los astrónomos.

Palabras Clave: Espectros Astronómicos, Cartografiado de galaxias, Redes Neuronales Convolucionales, Astroinformática.

Abstract

The implementation of spectroscopy as a tool in the classification of astrophysical objects is a technique that is becoming increasingly useful for modern astronomy. Currently, a collection of space and ground-based telescopes map the sky every night, exponentially increasing data volume over time. The growing number of datasets pushes astronomers to efficiently classify the spectra of observed objects. In this work, we implement the pre-trained algorithms: INCEPTION V3, RESNET 50 and MNIST, that are based on algorithms of Computer Vision (CV) and Convolutional Neural Networks (CNN). We compare the performance of the classifiers applied on spectra taken from the Sloan Digital Sky Survey (SDSS) in its Data Release DR12 with the Baryon Oscillation Spectroscopic Survey (BOSS) spectrograph with a sample of 300000 spectra. In the learning process, the hyperparameters associated with each model were adjusted, and each model was evaluated with accuracy, precision, and sensitivity metrics. The exercise results show that the best classifier for quasar spectra is RESNET 50 with a performance of more than 60 % accuracy. In addition, a low loss rate was obtained in the case of star classification with the INCEPTION V3 model. Furthermore, this investigation confirms that this procedure should be introduced as a regular step in the pipeline of ongoing and upcoming spectroscopy surveys in order to optimize time and quality during the manual classification process carried out by astronomers.

Keywords: Astronomical Spectra, Galaxy survey, Convolutional Neural Network, Astroinformatics.

1. Introducción

La cantidad de galaxias observadas aumenta exponencialmente en el tiempo con la llegada de nuevos *surveys* espectroscópicos que permiten estudiar la distribución de la estructura a gran escala: 2dfGRS¹, 6dfGS², SDSS³ al igual que BOSS⁴ (Baryon Oscillation Spectroscopic Survey; [1]) y su versión extendida (eBOSS), 4-meter Multi Object Spectroscopy Telescope (4MOST⁵ [2]), DESI⁶ (Dark Energy Spectroscopic Instrument; [3, 4], en el futuro Euclid⁷, el Observatorio Vera Rubin (o LSST⁸) y el telescopio espacial Nancy Grace Roman⁹ (o WFIRST) son campañas que están en pleno desarrollo, con su entrada en funcionamiento, nuevos retos se imponen para la clasificación y caracterización de este tipo de espectros. Por ejemplo, en las primeras campañas de BOSS clasificaban los espectros visualmente, lo que implicó para las siguientes campañas corregir errores humanos aleatorios e invertir mucho tiempo para completar correctamente esta tarea.

Haciendo uso de la espectroscopía, que permite estudiar en detalle la trayectoria de la luz emitida por diferentes objetos astrofísicos y su interacción con el medio intergaláctico, en diferentes longitudes de onda. Esta técnica se ha representado a través de los años demostrando que cuando la luz de un objeto pasa a través de un prisma o un dispositivo especial, se divide en diferentes colores, como un arcoíris. Cada color corresponde a una longitud de onda específica. Estos colores nos dan información valiosa sobre la estructura y propiedades de los objetos, como su temperatura, composición química y movimiento propio. Es por eso que en esta investigación la técnica es un principio para clasificar los espectros de cada objeto. Así como lo señala Missey y Hanson [5], la espectroscopía se utiliza en muchos campos puesto que es una herramienta poderosa para estudiar la naturaleza de los diferentes objetos.

Un efecto cosmológico importante a tener en cuenta para tratar las líneas espectrales es su desplazamiento con respecto a la longitud de onda en reposo, conocido como *redshift* (corrimiento al rojo en español). Hay algunas líneas anchas que absorben buena parte del flujo intrínseco emitido por el objeto astrofísico, y por tanto, hacen difícil la clasificación automática de los espectros [8]. En el caso de los cuásares podrían presentarse en líneas de absorción anchas, -BAL por sus siglas en inglés; [6, 7]- o DLA, Damped Lyman Alpha Systems [9]. Algoritmos como QuasarNet [10], SQUEZE (Spectroscopic QUasar Extractor and redshift (z) Estimator) [11, 12] se diseñaron rutinas de clasificación de espectros de cuásares con resultados muy favorables, dado su enfoque en identificar aquellas líneas que resultan problemáticas para el clasificador (ver por ejemplo el tratamiento mostrado en [13]). Estos trabajos se fundamentan en la aplicación de técnicas *Convo-*

lutional Neural Networks (CNN) y *Computer Vision* (CV) para la clasificación de objetos astrofísicos que son especialmente útiles si se trabaja, el espectro en formato de imagen como entrada del algoritmo. Dado que para los algoritmos es más sencillo identificar matrices que componen una imagen.

En este trabajo evaluamos el rendimiento de los modelos pre-entrenados para clasificar espectros astrofísicos: INCEPTION V3 [14], la RESNET 50 [15] y MNIST [16], tomados a partir de los resultados de la competencia ILSVRC 14 (ImageNet Large Scale Visual Recognition Challenge) [17] que utiliza el principio del córtex del ojo humano para identificar patrones de espectros. Este proyecto considera tres métricas de evaluación y diferentes configuraciones de los hiperparámetros del modelo para la consecución de la evaluación.

2. Metodología

En el desarrollo de esta investigación exploratoria se revisa el rendimiento de cada modelo de aprendizaje profundo, aprovechando el entrenamiento previo de cada algoritmo con su respectivo conjunto de datos y sintonizando algunos parámetros para ajustar la entrada a cada modelo como se muestra en la Figura 1.

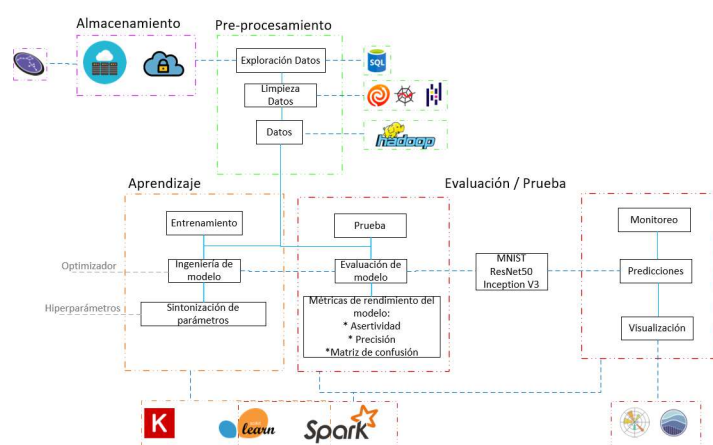


Figura 1: Diagrama representativo de la metodología implementada en este trabajo con cada uno de los algoritmos.

2.1. Selección de datos

Los espectros utilizados en este trabajo corresponden a datos tomados del espectrógrafo BOSS, en dos versiones: v5_7_0 y v5_7_2¹⁰ que pertenecen al catálogo DR12 de la campaña SDSS. El catálogo seleccionado tiene cerca de 2 millones de objetos astrofísicos. Aplicando la técnica de Muestreo Aleatorio Estratificado (MAE) se recolectó un conjunto de datos de 300000 espectros emitidos por estrellas, cuásares y

¹<http://www.2dfgrs.net>

²<http://www.6dfgs.net>

³<https://www.sdss.org/dr12/>

⁴<https://www.sdss4.org/surveys/boss/>

⁵<https://www.4most.eu/cms/home/>

⁶<https://www.desi.lbl.gov>

⁷<https://www.euclid-ec.org>

⁸<https://www.lsst.org>

⁹<https://www.stsci.edu/roman>

¹⁰<https://data.sdss.org/sas/dr12/boss/spectro/redux/images/>

galaxias, con un redshift entre 0 y 3.5, como se muestra en la Figura 2.

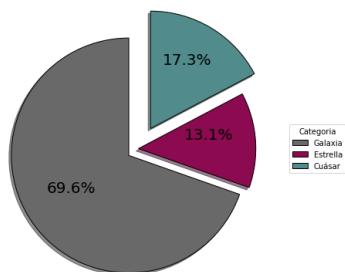


Figura 2: Diagrama circular distribución de clases en la muestra del conjunto de datos completo que corresponde de 300000 espectros.

Una vez seleccionada la muestra de espectros que se convierten en el conjunto de datos definitivo para aplicar las pruebas, se obtiene una distribución de la cantidad de objetos a través de histogramas como se muestran a continuación en las Figuras 3, 4, 5, correspondientes estrellas, cuásares y galaxias, respectivamente.

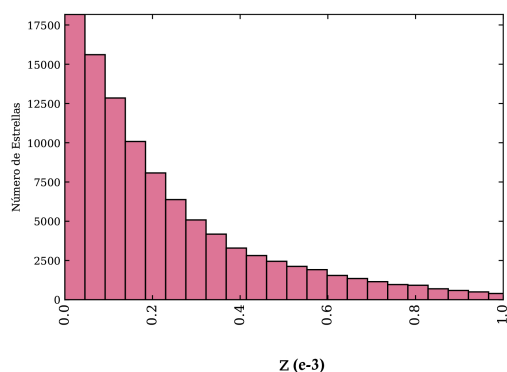


Figura 3: Distribución de estrellas en la muestra como función del redshift. La tendencia de creciente indica que la mayor cantidad de estrellas observadas se encuentran en el vecindario local, dentro o en cercanías a nuestra galaxia, la Vía Láctea, y su brillo es cada vez menos detectado a mayor distancia

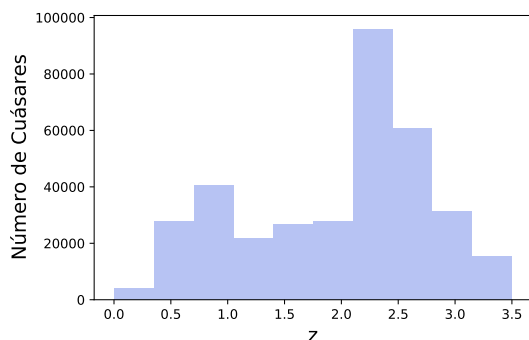


Figura 4: Histograma de cuásares en nuestra muestra con respecto al z observado. El pico de la distribución se encuentra en $z \sim 2.4$ (cuásares con bosque de Lyman α) de gran interés en cosmología.

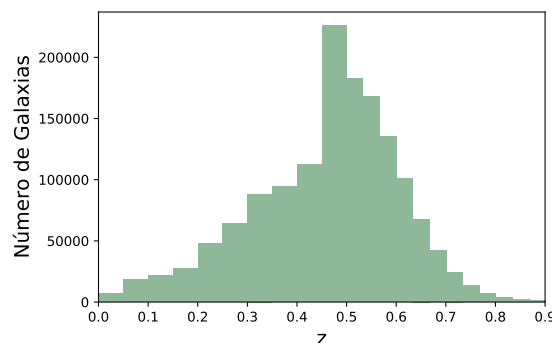


Figura 5: Histograma de galaxias como función de z: la distribución indica una alta cantidad de galaxias en un redshift de 0.5, consistente con los resultados reportados por las diferentes campañas de SDSS. Esta distribución revela que nuestra técnica de selección aleatoria cumple con las condiciones de un muestreo estadístico exitoso

2.2. Variables

Durante la implementación de los algoritmos para la clasificación se ajustaron algunos parámetros e hiperparámetros, principalmente con el objetivo de preparar las imágenes de los espectros y variables asociadas al rendimiento de cada modelo.

Algunos de los hiperparámetros relevantes ajustados en cada modelo son:

- Kernel:** es una función que permite transformar los datos de entrada en un espacio de características de mayor dimensión. Esta transformación ayuda a abordar problemas en los que los datos no son linealmente separables en su forma original. En lugar de realizar explícitamente esta transformación, el *kernel* realiza cálculos internos basados en productos escalares en el espacio de características de mayor dimensión. Esto nos permite trabajar con los datos en este espacio ampliado sin tener que conocer las coordenadas exactas de ese espacio.
- Stride:** es un parámetro que determina la cantidad de filas o columnas que mueve el kernel durante el procesamiento de cada entrada de datos a la red neuronal. Dado que el tamaño de las imágenes cambian para la entrada de cada algoritmo, es necesario ajustar este parámetro y estandarizar para todos los algoritmos, manteniéndolo en la $3 \times 3 \times 1$.
- Épocas:** se refiere a un paso completo a través de todo el conjunto de datos de entrenamiento durante la fase de entrenamiento de un modelo. Durante una época, el modelo procesa secuencialmente cada ejemplo de entrenamiento y ajusta sus parámetros internos en base a los resultados obtenidos. En cada época, el modelo toma cada imagen de entrenamiento, la procesa a través de su arquitectura y realiza ajustes en sus parámetros internos para reducir el error y mejorar su capacidad de clasificación. Una vez que todas las imágenes del conjunto de entrenamiento han sido procesadas, se considera que ha pasado una época

completa. Las épocas pueden incrementar o disminuir dependiendo de los resultados obtenidos en cada iteración, como se muestra en la sección **Resultados**.

- **Clase:** corresponde a cada uno de los tipos de datos que cada algoritmo clasifica, para este trabajo son espectros de: cúasares, estrellas y galaxias.
- **Capa inicial:** corresponde al tamaño de las entradas, siendo estas las imágenes de los espectros, entonces el pixel de ancho \times pixel largo \times el pixel de profundidad. El último pixel corresponde al color de las entradas, 1 para imágenes en escala de grises y 3 para imágenes a color RGB (Red, Blue, Green).

Los parámetros que se nombran a continuación se mantienen acorde al algoritmo seleccionado:

- **Función de activación:** es una puerta de salida que regula la salida de una neurona o una capa de neuronas dependiendo de la entrada recibida. Entonces si una neurona se activa y determina cuánta información pasa a través de ella hacia las capas posteriores. Dado que la última convolución clasifica cada clase, esta simplifica el resultado de la capa anterior para convertirla en la entrada de la siguiente.
- **Weight:** representa un nodo o una unión entre cada capa de la red neuronal, también como peso. Se utiliza en todos los algoritmos matemáticos *backpropagation*¹¹ para ajustar los valores de los *kernels* para entrenar y la preponderancia de las capas.

2.3. Procedimiento

Se realizó una exploración de la capacidad de aprendizaje que tienen los modelos INCEPTION V3, RESNET 50 y MNIST, haciendo una recolección de espectros, seguido del pre-procesamiento de los mismos, luego de la selección de los algoritmos, pasando a la fase de entrenamiento de cada modelo y finalizando con la evaluación de la clasificación, así:

Pre-procesamiento

Finalizada la selección de los 300000 espectros para la etapa de pre-procesamiento, fue necesario eliminar el ruido de la señal, ya que los datos espectroscópicos contenían ruido aleatorio, esto se logró aplicando el filtro de media móvil. Una vez eliminado el ruido se grafican los espectros como en la Figura 6, siendo cada espectro un ejemplo de lo que es la entrada en cada algoritmo.

Algoritmos

Los algoritmos seleccionados tienen un componente de CV y CNN, aunque cada uno de estos modelos tiene una arquitectura-lógica distinta, hay algunas capas en común que determinan las etapas esenciales en un proceso como

el aprendizaje profundo, que son: que son el entrenamiento y la prueba, que se explican en la Figura 7. Los modelos seleccionados son: INCEPTION V3, RESNET 50 y MNIST, cada uno de ellos previamente entrenado para otros objetivos de ciencia. En la Tabla 1, se muestra su arquitectura y parametrización original.

En tareas de *Machine Learning*, es usual dividir el conjunto de datos en tres *sets* que corresponden a etapas: Entrenamiento, Aprendizaje y Validación. Dado que los algoritmos implementados en esta investigación son pre-entrenados, no se realizó la tercera fase pues se tomaron en cuenta los hiperparámetros que cada algoritmo tenía previo con el entrenamiento correspondiente a cada conjunto de datos.

Aprendizaje

En esta fase de construcción y entrenamiento de los modelos de aprendizaje profundo, el modelo aprende a hacer predicciones y tomar decisiones a partir de los datos de entrada. El proceso se basa en la minimización de una función de costo o pérdida del modelo y los valores reales. También se realizan iteraciones en las que cada modelo ajusta gradualmente sus parámetros e hiperparámetros para reducir el error y mejorar su rendimiento. El éxito del proceso de aprendizaje depende en gran medida de la calidad y cantidad de datos disponibles. Cuanto más sean representativos los datos y mayor sea su diversidad, mejor es la capacidad de clasificación de cada modelo.

Cada algoritmo se entrenó con una proporción de datos diferente en relación a la muestra, evaluadas de la siguiente manera:

- Evaluación 1: con un 60 % de aprendizaje y 40 % de prueba.
- Evaluación 2: con un 70 % de aprendizaje y 30 % de prueba.
- Evaluación 3: con un 80 % de aprendizaje y 20 % de prueba.

Los resultados corresponden a las pruebas ejecutadas con la proporción de la Evaluación 3, porque cada modelo utiliza el conjunto de entrenamiento para ajustar sus parámetros mediante el algoritmo de *backpropagation*, esto significa que la red aprende a partir de sus errores comparando las salidas con las etiquetas de los datos del entrenamiento.

Tabla 1. Parametrización inicial de cada uno de los algoritmos

Algoritmo	Inception V3	RESNET 50	MNIST
Tipo de dataset	Imagenet	Imagenet	MNIST
Tamaño de dataset	1.2E-6	1.2E-6	60E-3
Clase	1000	1000	10
Entrada (pixeles)	224 \times 224 \times 3	224 \times 224 \times 3	28 \times 28 \times 1
Ciclo de corrida	42	50	15
ILSVRC	2015	2016	N.A.
Épocas	140 - 200	64	10 - 25
AcceptanceRate(%)	92.3	74.9	96.3
ErrorRate(%)	7.7	25.1	3.7

Evaluación

En la fase de prueba se evalúa el rendimiento de los tres algoritmos bajo los siguientes casos:

¹¹<http://neuralnetworksanddeeplearning.com/chap2.html>

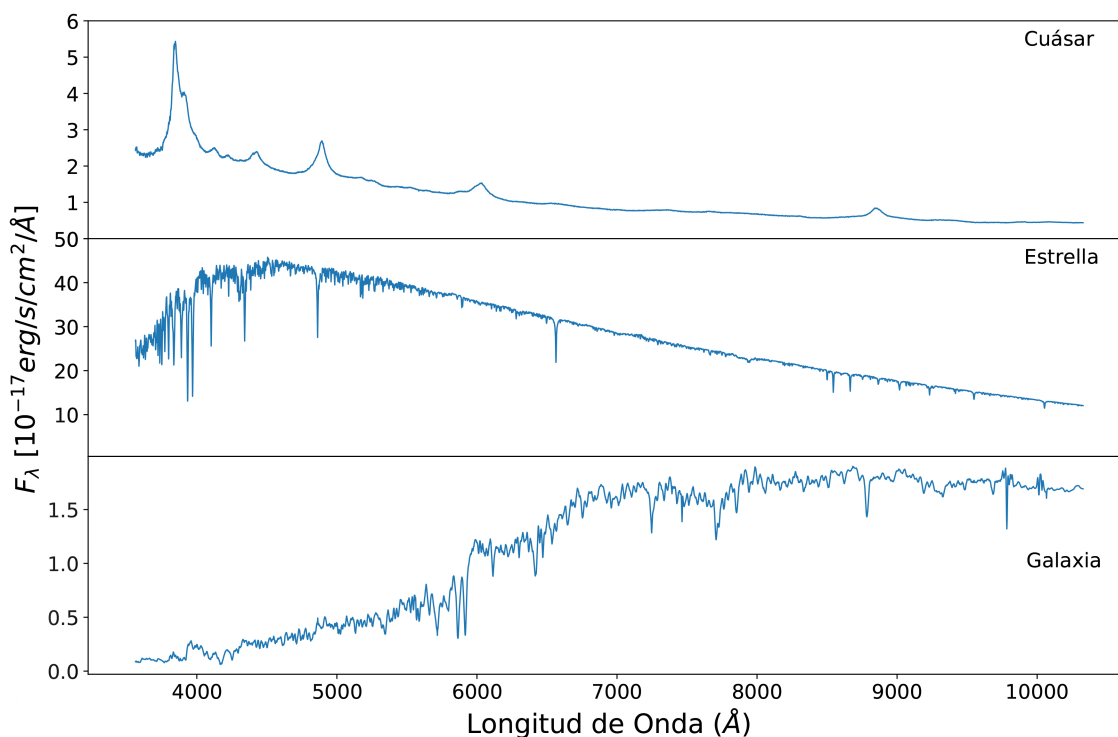


Figura 6: Ejemplos de espectros que se tomaron como entrada para cada uno de los algoritmos. El primer panel corresponde al espectro de un cúasar. El panel central es el espectro de una estrella y finalmente el tercer panel muestra el espectro de una galaxia. Estos rasgos característicos son los que los algoritmos intentan reconocer para realizar la clasificación.

- Reducción de pérdida por número de épocas.
- Proporción en conjunto de datos con métricas de evaluación: exactitud, precisión y sensibilidad.

En la fase de evaluación de cada algoritmo se tomaron las distribuciones de las evaluaciones 1, 2 y 3 con proporciones de 60/40, 70/30 y 80/20, respectivamente. Con cada evaluación se realizaron actualizaciones de ponderaciones con 20, 30 y 40 épocas. Finalmente, se aplicaron las métricas de evaluación a los modelos iterados con 40 épocas. Esta ruta de evaluación para los algoritmos se realiza con el fin de determinar cuál es el comportamiento de cada uno clasificando los espectros. En la fase de evaluación cada modelo se somete a un conjunto de datos por separado que no se utilizó en el entrenamiento, estos datos de prueba son independientes de los datos utilizados para entrenar cada modelo y proporcionan una medida objetiva del rendimiento.

La clasificación de espectros realizados por los algoritmos se cuantifica con las siguientes categorías:

- Si el espectro corresponde a un objeto determinado y el algoritmo lo clasifica como tal, es un Verdadero Positivo (VP).
- Si el espectro no corresponde al objeto determinado y el algoritmo lo clasifica como cualquiera de las otras clases, es un Verdadero Negativo (VN).
- Si el espectro no corresponde a un objeto determinado y el algoritmo lo clasifica en una clase incorrecta, es

un Falso Negativo (FN).

- Si el espectro no corresponde al objeto determinado y el algoritmo lo clasifica como tal, es un Falso Positivo (FP).

Teniendo en cuenta los criterios para la cuantificación, a continuación se definen las métricas de evaluación implementadas:

Exactitud

Esta métrica de evaluación definida en la ecuación (1) busca identificar la frecuencia con la que el algoritmo se equivoca, es decir, reconoce los valores verdadero positivo y verdadero negativo, comparándolos con el total para así determinar, esto arrojaría el resultado de las predicciones correctas, de acuerdo a las proporciones de las clases en el conjunto de datos.

$$Exactitud = \frac{VP + VN}{(VP + FP + FN + VN)}. \quad (1)$$

Precisión

La métrica de precisión determina cuantos valores falsos positivos y verdaderos positivos está clasificando el de una categoría, pero que en realidad corresponde a otra distinta. Adicionalmente, con esta métrica es posible determinar si hay un desequilibrio en las proporciones del conjunto de datos, porque al clasificar espectros de objetos que no corresponden a la clase correcta, se incrementa el margen de error del modelo.

$$\text{Exactitud} = \frac{VP}{(VP + FP)}. \quad (2)$$

Sensibilidad

Esta métrica es también conocida como *Recall*, busca identificar los espectros clasificados correctamente como positivos dentro del conjunto de positivos verdaderos, tal como se muestra en la ecuación (3):

$$\text{Sensibilidad} = \frac{VP}{(VP + FN)}. \quad (3)$$

3. Resultados y discusión

Tomando los 300000 espectros para la clasificación con los algoritmos INCEPTION V3, RESNET 50 y MNIST, los resultados se interpretan partiendo de margen de error y el mejor rendimiento para cada clase.

El margen de error es una variable significativa para determinar la asertividad de la clasificación de los espectros, como lo muestra la Tabla 1 en el valor *Error rate*, por ello se evaluó cada algoritmo hasta 40 iteraciones, como se muestra en la Figura 7. Además, se ejecutaron las iteraciones con base en la distribución del conjunto de datos correspondiente a la Evaluación 3, tomando el 80 % del conjunto de espectros para la fase de entrenamiento y el 20 % restante para prueba.

Los resultados muestran una considerable variación en el porcentaje de pérdida desde cuando los algoritmos se evalúan con 20 épocas hasta cuando se evalúa con 40, progresivamente. Teóricamente, el porcentaje de pérdida debería ser inversamente proporcional al número de épocas, ya que a medida que el número de épocas incrementa el porcentaje de pérdida debe reducirse, como se puede ver en la Figura 7.

Para la evaluación con 20 épocas los resultados muestran que la cantidad de iteraciones no es suficiente para obtener una clasificación confiable, puesto que muestra un leve incremento del error. Asimismo, RESNET 50 es el algoritmo que mayor error arroja cuando se evalúa con el número mínimo de épocas en la fase de entrenamiento. No obstante, la diferencia entre la etapa de prueba y entrenamiento del mismo no presenta diferencias significativas. Es por esto que se evaluaron con más épocas, con el fin de identificar el número de iteraciones con el que cada modelo podría obtener su punto mínimo de error.

En la evaluación con 40 épocas el comportamiento de los algoritmos muestra fluctuaciones moderadas aunque intermitentes, porque en la medida en que crece el número de épocas de evaluación logra regularizar y reducir la tasa de pérdida. Esto demuestra un progreso en el aprendizaje de los modelos, pues que clasifica cada espectro en la clase correcta. Un elemento destacable en el comportamiento de los tres modelos de algoritmo es a causa del número de épocas, dado que cuando superan la iteración 10 presentan una fluctuación más acentuada en la fase de prueba que en la fase de entrenamiento, esto es a causa del progreso en el aprendizaje que cada modelo va teniendo en la medida que las iteraciones incrementan.

El rendimiento de cada modelo se valora considerando los resultados de las métricas exactitud, precisión y sensibilidad como se muestran en las Figuras 10, 9 y 10, por medio de matrices de confusión, que arrojan la cantidad de veces que cada modelo clasifica bien o no un determinado espectro. Además, los resultados de los algoritmos se normalizan con el objetivo de escalar los datos sin distorsionar las diferencias en los intervalos o perder información. En las Figuras 8, 9 y 10 se muestran los resultados por medio de las matrices de confusión, con las clases E, C y G (estrella, cuásar y galaxia, respectivamente). Los resultados están normalizados en un rango entre 0 y 1, indicando el porcentaje de éxito del algoritmo aplicado a cada clase con 40 iteraciones.

La Figura 8 muestra los resultados de la prueba con una distribución de datos del 60 % para entrenamiento y el 40 % para prueba. El análisis de estos resultados se puede interpretar que INCEPTION V3 y RESNET 50 clasifican cuásares en un 80 % mejor que las galaxias con la métrica de exactitud. Además, con la métrica de precisión la clasificación se mantiene en un 80 % para la clase de cuásares, aún con la métrica de sensibilidad se reduce al 70 %. Si se mantuviese esta proporción en el conjunto de datos los mejores clasificadores de cuásares son los modelos INCEPTION V3 y RESNET 50.

Los resultados de la Figura 9 nos llevan a proponer una distribución de mayor proporción para la fase de entrenamiento, puesto que entre más información de entrada tiene cada algoritmo, el modelo reduce su porcentaje de pérdida, como lo muestra la Figura 10. Cambiando la proporción del conjunto de datos, al incrementar 10 % en la fase de entrenamiento obtenemos resultados similares a la Evaluación 1 como lo muestra la Figura 9, manteniendo la clasificación de los espectros de cuásar con una bondad del 80 % para los algoritmos RESNET 50 y MNIST; no obstante, INCEPTION V3 reduce en al menos un 15 % de exactitud la clasificación para este mismo tipo de objeto. Además, en la clasificación de espectros de estrellas, ResNet 50 clasifica al 50 % de los espectros de prueba. Por otro lado, para la clasificación de espectros de galaxias ningún modelo muestra un buen desempeño. Esto nos motiva a hacer una tercera evaluación con una distribución para la fase de entrenamiento con un 80 % como se muestra en la Figura 10.

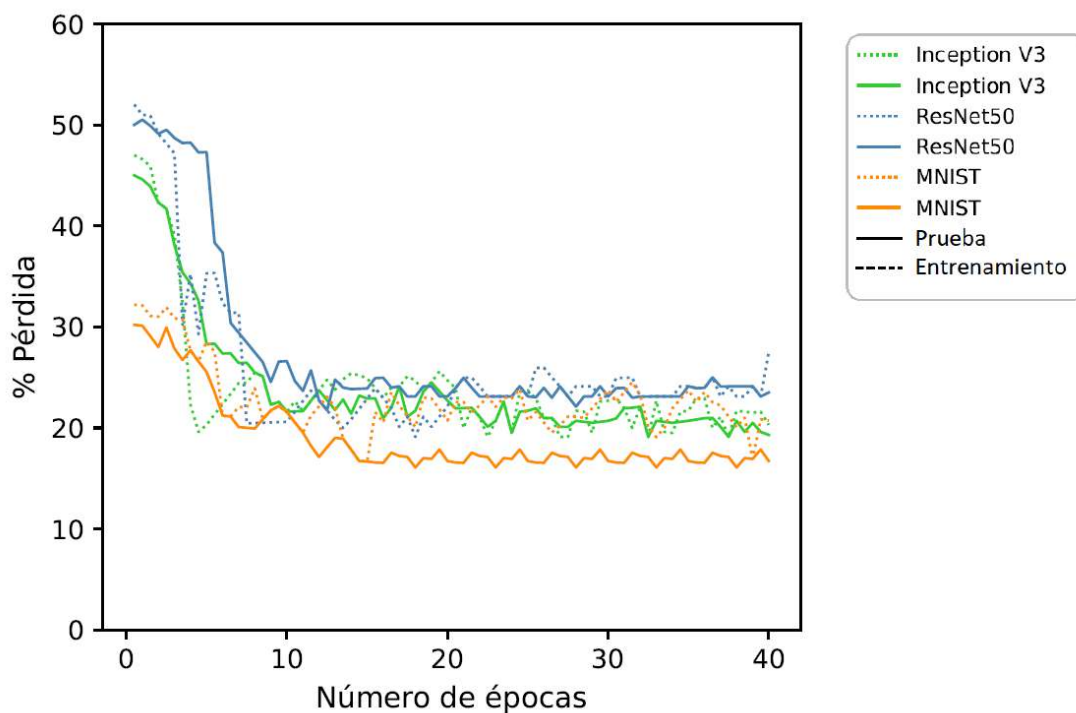


Figura 7: Resultado de las pérdidas por algoritmo evaluado en 40 épocas. Las líneas continuas representan la fase de prueba, mientras que las líneas discontinuas dan cuenta del entrenamiento.

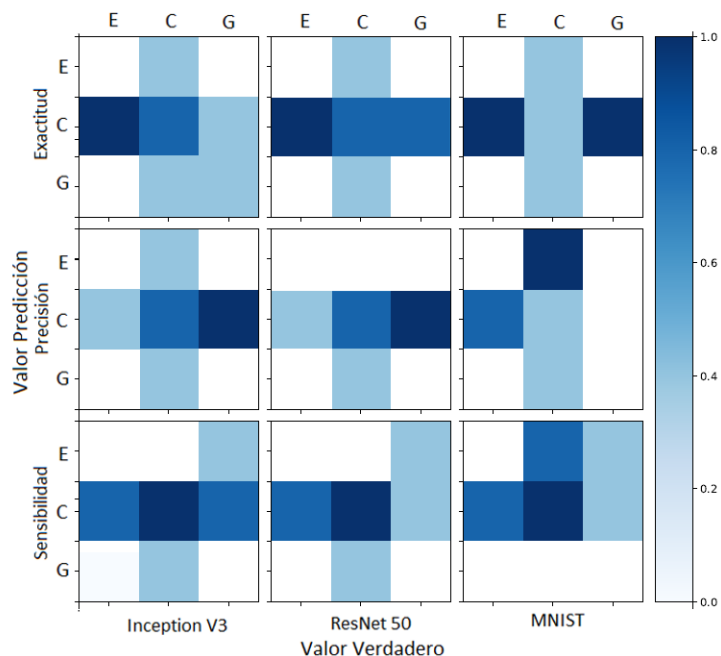


Figura 8: Resultados de balanceo 60/40 de entrenamiento y prueba, respectivamente. En primera fila se evalúa la métrica de exactitud, en la segunda, la de precisión y en la tercera, sensibilidad. De otro lado, la columna a la izquierda muestra los resultados del algoritmo INCEPTION V3, mientras la columna en la mitad, los resultados de RESNET 50 y en la columna a la derecha, MNIST.

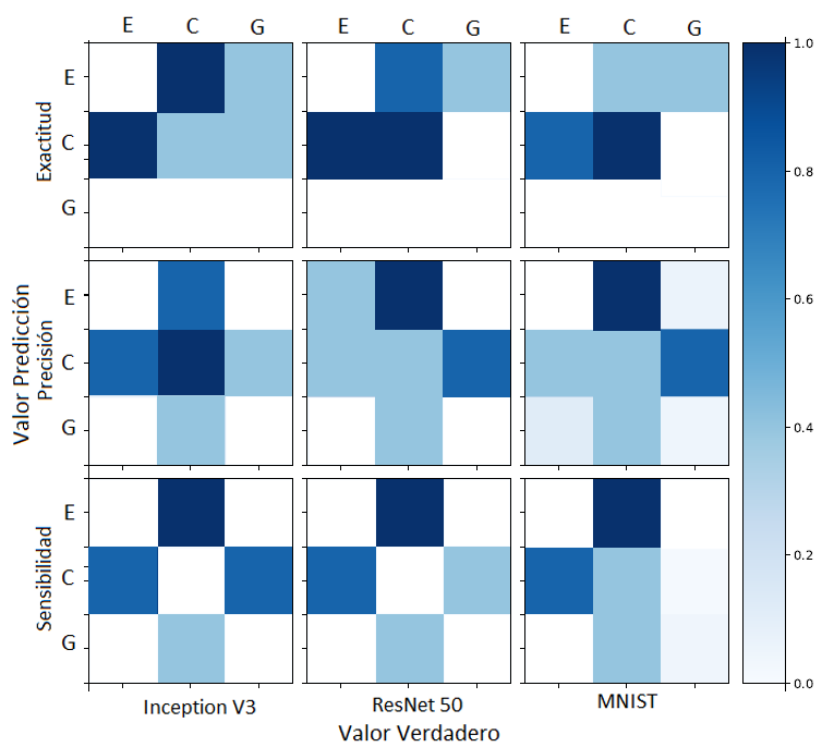


Figura 9: Resultados de balanceo 70/30 de entrenamiento y prueba, respectivamente. De arriba hacia abajo se muestran las métricas de exactitud, precisión y sensibilidad, mientras que de izquierda a derecha se presentan los resultados de los algoritmos INCEPTION V3, RESNET 50 y MNIST.

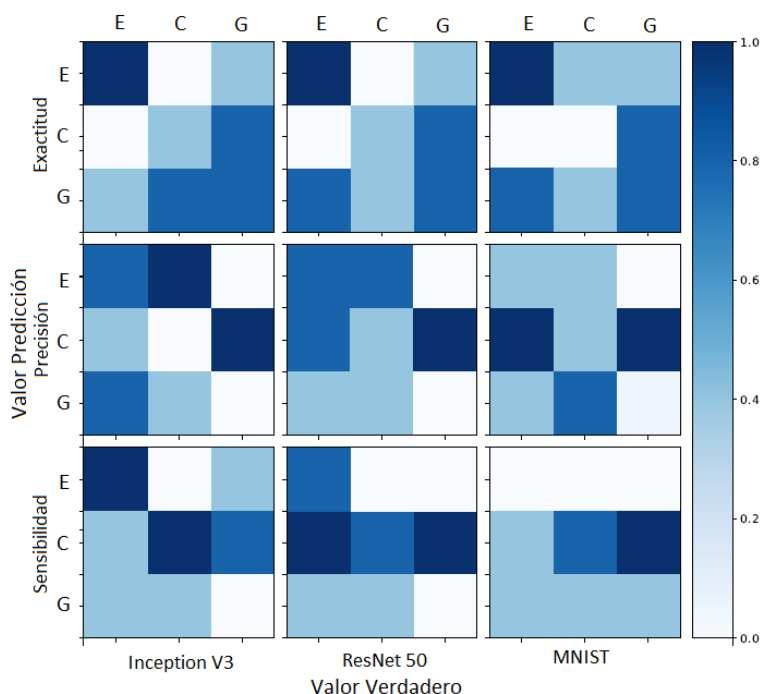


Figura 10: Resultados de balanceo 80/20 de entrenamiento y prueba, respectivamente. En primera fila se evalúa la métrica de exactitud, en la segunda la de precisión y en la tercera sensibilidad. Las columnas muestran los resultados del algoritmo INCEPTION V3, RESNET 50 y MNIST, de izquierda a derecha.

Los espectros de estrellas y cuásares resultan estar mejor clasificados por el algoritmo RESNET 50, bajo la métrica de precisión, mientras que los espectros de galaxias muestran tener una bajo índice de éxito en la clasificación, con un 10 %.

Para esta evaluación (la número 3, con porcentajes de 80/20 en aprendizaje y entrenamiento), con una mayor proporción de datos de aprendizaje muestra una mejoría en la clasificación ante algunas métricas. Por ejemplo, en la métrica de exactitud la clasificación incrementa en al menos 10 % para cada clase en INCEPTION V3 y RESNET 50. Sin embargo, para MNIST la clasificación se reduce para los cuásares. Esto indica que se pueden implementar los modelos de INCEPTION V3 y RESNET 50 para la clasificación de espectros de estrellas, galaxias y cuásares con estas métricas, con excepción MNIST para la clasificación de cuásares. Para estos casos en los que los algoritmos tienen una fase de pre-entrenamiento, es importante disponer al menos el 80 % del conjunto de datos para re-entrenar cada algoritmo. Durante el desarrollo, los algoritmos pre-entrenados no variaron los pesos predefinidos por los desarrolladores, se conservaron los valores para comparar la tasa de pérdida y de aprendizaje en cada modelo.

En relación al balanceo de las clases, distribuimos el conjunto de datos en dos partes, una para el entrenamiento y el otro para prueba. Sin embargo, en aplicaciones de algoritmos para clasificación se suele dividir en tres, siendo esta última una fase de validación. Para este trabajo tuvimos en cuenta únicamente las dos primeras fases puesto que cada algoritmo trae intrínsecamente un aprendizaje previo adquirido por el conjunto de datos original con el que cada uno fue diseñado y entrenado.

4. Conclusiones

En este trabajo se pusieron a prueba tres algoritmos de Visión Computacional y Redes Neuronales Convolucionales: INCEPTION V3, RESNET 50 y MNIST para clasificar espectros astrofísicos de cuásares, estrellas y galaxias con una muestra de 300000 elementos pertenecientes a la campaña DR12 del cartografiado SDSS. Para evaluar la calidad de la clasificación se consideraron 3 evaluaciones con distinta proporción entre los conjuntos de entrenamiento y prueba, con porcentajes de 60/40, 70/30 y 80/20, respectivamente.

Las métricas de evaluación implementadas demuestran que en el proceso de clasificación, cada modelo reconoce patrones diferentes, esto indica que bajo las pruebas realizadas difícilmente se puede definir a un algoritmo como el mejor clasificador para todas las clases, puesto que el nivel de exactitud varía dependiendo de cada métrica y de cada clase. Dicho esto, podemos inferir que el modelo INCEPTION V3 es mejor clasificador para espectros de estrellas y galaxias con un 85 % y 75 %, respectivamente. También, el algoritmo RESNET 50 clasifica mejor los espectros de estrellas y galaxias con un 83 % y 73 %, respectivamente. Finalmente el modelo de MNIST presenta una clasificación parcial para las tres clases de espectro con un 78 % para estrellas, 22 % para estrellas y 62 % para galaxias, los anteriores resultados fueron obtenidos con una proporción de datos del 20 % para prueba y el 80 % para entrenamiento.

Además de la variación de los parámetros e hiperparámetros en cada algoritmo, la cantidad de iteraciones que se aplica a cada modelo es fundamental para el proceso de aprendizaje, porque en cada iteración representa un incremento en la tasa de aprendizaje de cada modelo.

A lo largo de nuestro estudio se aplicaron los algoritmos pre-entrenados para responder a la clasificación del grupo de espectros, obteniendo resultados concluyentes que demuestran la aplicación de algoritmos de aprendizaje profundo como una excelente herramienta para automatizar la categorización de datos que se obtienen en los telescopios actuales y por tanto, reducir el tiempo que lleva separar, caracterizar y estudiar estos grandes volúmenes de espectros de alta resolución. Los hallazgos encontrados en esta investigación demuestran el impacto que tiene el *Machine Learning* en el campo de la astrofísica, para lo clasificación de espectros. Incrementando la capacidad de procesamiento eficiente, mejorando la identificación de patrones complejos, permitiendo a los modelos tener adaptabilidad y aprendizaje continuo y contribuye a la automatización del proceso de clasificación y a la reducción de los errores humanos.

En futuros trabajos, se considerará hacer una exploración más exhaustiva de los hiperparámetros de los modelos, así como un entrenamiento propio con la optimización de los pesos de los modelos, adaptados a las necesidades de nuestro problema.

Agradecimientos

Las autoras de este trabajo agradecen al centro de investigación CEINTECCI por garantizar acceso al clúster con el que se computaron estos resultados. V. Lara-Cruz agradece el apoyo prestado en la fase de pre-procesamiento con los datos al equipo técnico SDSS. Las autoras agradecen a la Universidad ECCI, en particular al comité de I+D+i, por su apoyo en la puesta en marcha del proyecto y por el tiempo otorgado.

Declaración de conflicto de interés

Los autores manifiestan no tener conflictos de interés.

Referencias

- [1] Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, APJS, 219, 12. doi:10.1088/0067-0049/219/1/12
- [2] de Jong, R. 2012, vol. 8446. doi:10.1117/12.926239.
- [3] DESICollaboration, Adame, A. G., Aguilar, J., et al. 2023, arXiv:2306.06307. doi:10.48550/arXiv.2306.06307
- [4] DESICollaboration, Adame, A. G., Aguilar, J., et al. 2023, arXiv:2306.06308. doi:10.48550/arXiv.2306.06308
- [5] Massey, P. & Hanson, M. M. 2013, Planets, Stars and Stellar Systems. Volume 2: Astronomical Techniques, Software and Data, 35. doi:10.1007/978-94-007-5618-22

- [6] Hall, P. B., Gunn, J. E., Knapp, G. R., et al. 2002, *Mass Outflow in Active Galactic Nuclei: New Perspectives*, 255, 161. doi:10.48550/arXiv.astro-ph/0107182
- [7] Trump, J. R., Hall, P. B., Reichard, T. A., et al. 2006, *APJS*, 165, 1. doi:10.1086/503834
- [8] García, L. Á., Martini, P., Gonzalez-Morales, A. X., et al. 2023, arXiv:2304.05855. doi:10.48550/arXiv.2304.05855
- [9] Wolfe, A. chaska, J. M., X. Gawiser, 2005, E., & Pro ARAA, 43, 861. doi:10.1146/annurev.astro.42.053102.133950
- [10] Busca, N. & Balland, C. 2018, arXiv:1808.09955. doi:10.48550/arXiv.1808.09955
- [11] Pérez-Ràfols, I., Pieri, M. M., Blomqvist, M., et al. 2020, *MNRAS*, 496, 4931. doi:10.1093/mnras/stz3467
- [12] Pérez-Ràfols, I. & Pieri, M. M. 2020, *MNRAS*, 496, 4941. doi:10.1093/mnras/staa1786
- [13] Guo, Z. & Martini, P. 2019, *APJ*, 879, 72. doi:10.3847/1538-4357/ab2590
- [14] Szegedy, C., Vanhoucke, V., Ioffe, S., et al. 2015, arXiv:1512.00567. doi:10.48550/arXiv.1512.00567
- [15] He, K., Zhang, X., Ren, S., et al. 2015, arXiv:1512.03385. doi:10.48550/arXiv.1512.03385
- [16] Y. Lecun, et al., pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
- [17] Russakovsky, O., arXiv. doi:10.48550/arXiv.1409.0575.