



Departamento de Estadística, Universidad Nacional

---

# Implementación de modelos ANN para clasificación de objetos astrofísicos identificados por SDSS ( Galaxias, Estrellas y Quasars)

Jose Alejandro Ramirez  
Oscar Julian Layton

Junio 19, 2020



- 1 SDSS (Sloan Digital Sky survey)
- 2 Descripción de la información
- 3 Implementación de los modelos ANN
- 4 Entrenamiento y validación. Modelo 1
- 5 Entrenamiento y validación. Modelo 2
- 6 Entrenamiento y validación. Modelo 3
- 7 Conclusiones



# SDSS (Sloan Digital Sky survey)

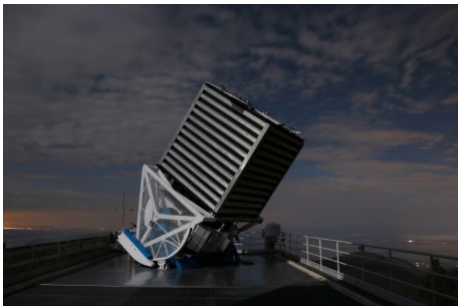


Figure: Sistema SDSS. Patrick Gaulme

SDSS (Sloan Digital Sky Survey) es un proyecto de investigación espacial que emplea un sistema fotométrico con un espejo primario de 2.5 m de diámetro para capturar y procesar imágenes de objetos astrofísicos.



# Image 1 SDSS



Figure: Tomado de [4]. Messier 81 (Bode's Galaxy). Esta galaxia se encuentra a unos 12 millones de años luz de distancia en la constelación de la Osa Mayor. SDSS



## Image 2 SDSS



Figure: Tomado de [4]. Cluster de estrellas m13. SDSS



## Image 3 SDSS



Figure: Tomado de [4]. Dos galaxias espirales una al lado de la otra. La galaxia espiral M51 y su compañera más débil. SDSS

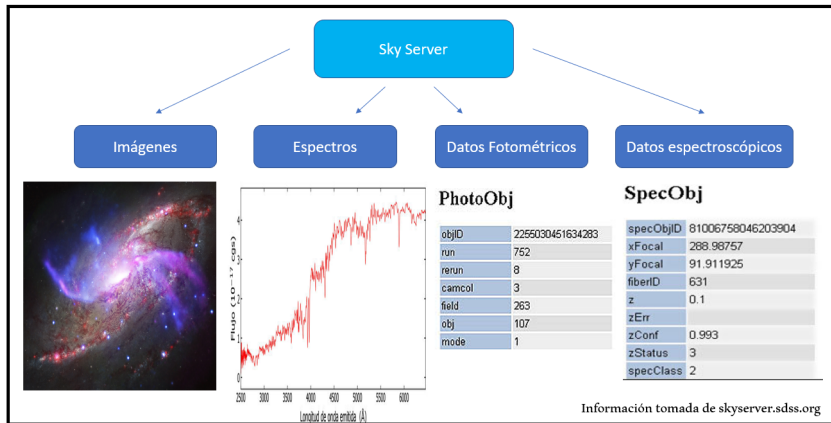


Figure: Sky Server data. SDSS



La fotometría y la espectroscopia son ramas de la óptica que se ocupan de la medición y el análisis de las variables físicas asociadas a la luz emitida por una fuente.

- Ascensión y declinación recta: coordenadas astronómicas de un objeto sobre la esfera celeste.
- Flujo luminoso total: cantidad de luz emitida por una fuente de luz, sus unidades en el SI es lúmen.
- Intensidad luminosa: cantidad de flujo luminoso que emite una fuente por unidad de ángulo sólido. Su unidad de medida es candela.
- Desplazamiento al rojo: representa un cambio en la longitud de onda de la radiación por el efecto de una fuente en movimiento.





# La esfera celeste

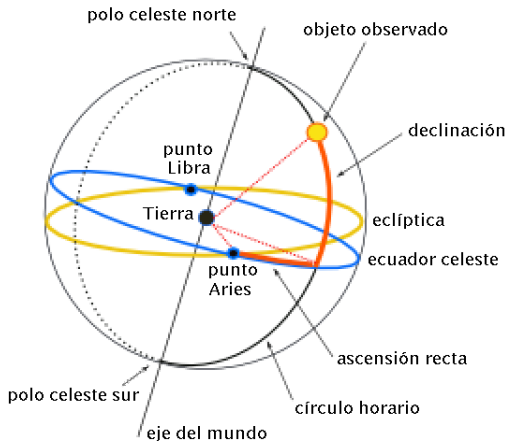


Figure: Tomado de [4]. Sistema de coordenadas ecuatoriales



**Los sistemas de adquisición de datos fotométricos, se clasifican segun el ancho de banda.**

Banda ancha:  $\text{FWHM} > 500\text{\AA}$

Banda intermedia:  $\text{FWHM} \ 100\text{-}300\text{\AA}$

Banda angosta:  $\text{FWHM} \ 10\text{-}100\text{\AA}$

FWHM: Anchura a media altura, dada por la diferencia entre los dos valores extremos de la variable independiente en los que la variable dependiente es igual a la mitad de su valor máximo.



## Sistemas fotométricos de banda ancha

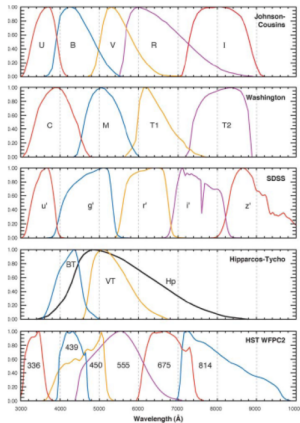


Figure: Tomado de [3]. Principales sistemas fotométricos



# Sistema fotométrico Sloan ('u', 'g', 'r', 'i', 'z')

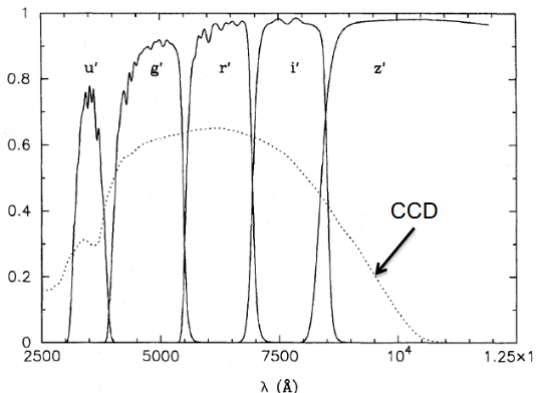


Figure: Tomado de [3]. Curva de transmitancia ('u', 'g', 'r', 'i', 'z')



# Descripción de la información

- El telescopio toma las señales electromagnéticas provenientes de los objetos celestes utilizando un sistema fotométrico de cinco filtros.
- Las señales se procesan por medio de dispositivos de carga acoplada como imagenes (CCD - charge coupled device ).
- Las imágenes, se procesan generando listas de datos fotometricos de objetos observados con varias variables.



# Descripción de la información

Este conjunto de datos está formado por 10.000 registros de observaciones asociadas a objetos astrofísicos. Cada observación se describe mediante 17 columnas de características y 1 columna de etiqueta que identifica la observación como una estrella, una galaxia o un Quásar.

Table: Variables del conjunto de datos

<b>objid</b> Identificador de objeto	<b>ra</b> Ascensión recta (banda r)	<b>dec</b> declinación J2000 (banda r)	<b>u</b> Ajuste de magnitud 1
<b>g</b> Ajuste de magnitud 2	<b>r</b> Ajuste de magnitud 3	<b>i</b> Ajuste de magnitud 4	<b>z</b> Ajuste de magnitud 5
<b>run</b> Número de ejecución 1	<b>camcol</b> columna de cámara	<b>field</b> número de campo	<b>specobjid</b> Identificador de objeto
<b>class</b> clasificación	<b>redshif</b> Desplazamiento al rojo	<b>plate</b> número de placa	<b>mjd</b> MJD de observación
<b>fiberid</b> ID de fibra			



# Descripción de la información

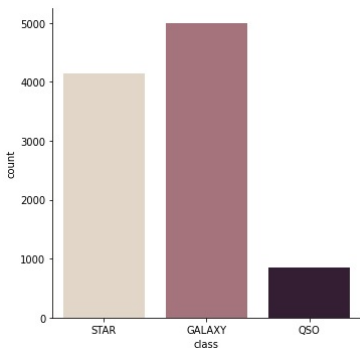


Figure: Histrograma clasificación



# Descripción de la información

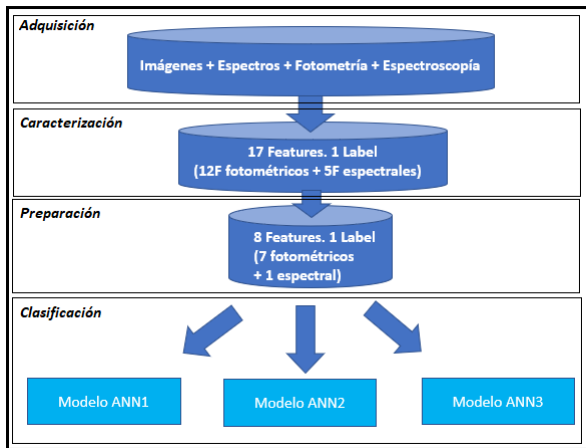


Figure: Tratamiento de la data





# Implementacion de los modelos ANN

Se implementaron tres modelos, en el primero se consideran las variables asociadas a la posición del objeto y el corrimiento al rojo, en el segundo las variables asociadas a los filtros del sistema fotométrico y en el tercero una combinación de features usados en los modelos anteriores. Las estructuras diferentes de redes neuronales artificiales secuenciales permiten clasificar objetos astrofísicos observados por el SDSS en Galaxias, Estrellas y Quasars.

- Input: Máximo 7 variables fotométricas (ra, dec, u g, r, i,z) y una variable espectral (desplazamiento al rojo).
- Labels: variable "class", para 3 objetos astrofisicos (Galaxias, Estrellas y Quasars)
- Regularizador: Dropout
- Funciones de activación: Relu (capa de entrada y ocultas) y Softmax (capa de salida)



# Estructura del modelo 1

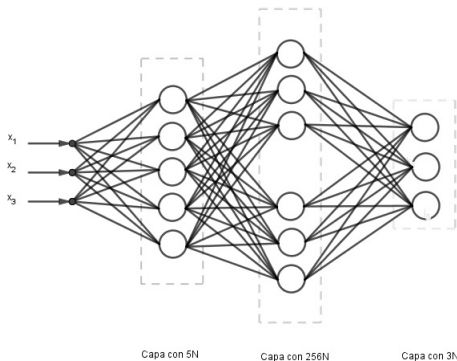


Figure: Red neuronal artificial. 1 capa oculta.



# Estructura del modelo 2

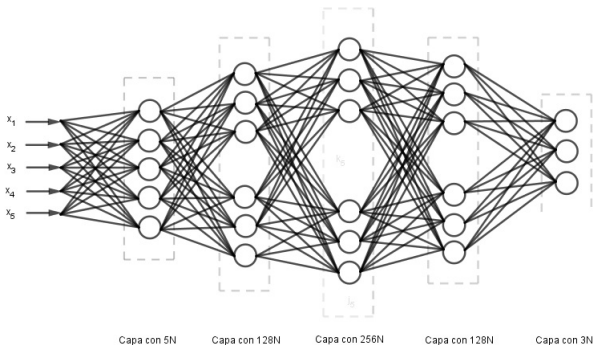


Figure: Red neuronal artificial. 3 capas ocultas.



# Estructura del modelo 3

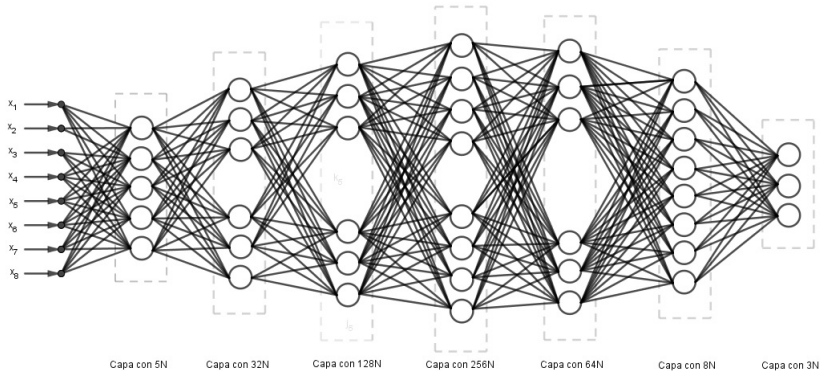
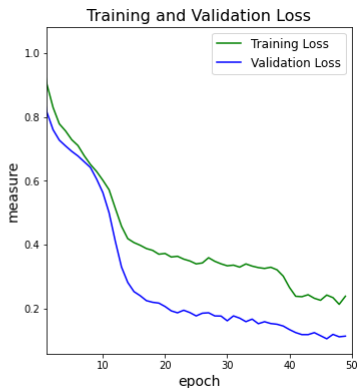
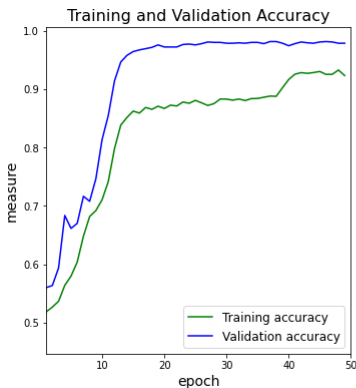


Figure: Red neuronal artificial. 5 capas ocultas.



# Entrenamiento y validación. Modelo 1

## Modelo 1:

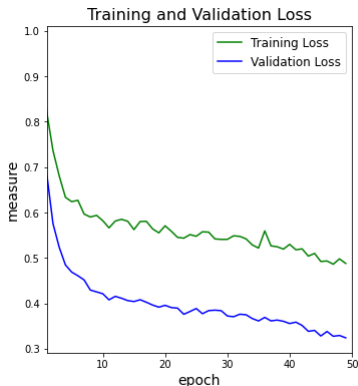
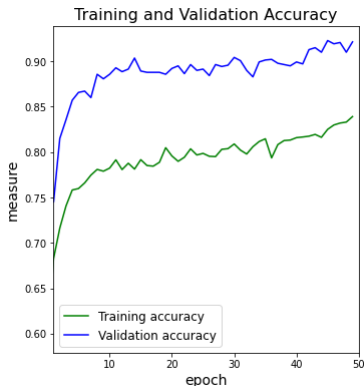


Modelo 1. Accuracy y funcion de perdida para los conjuntos de entrenamiento y validacion.



# Entrenamiento y validación

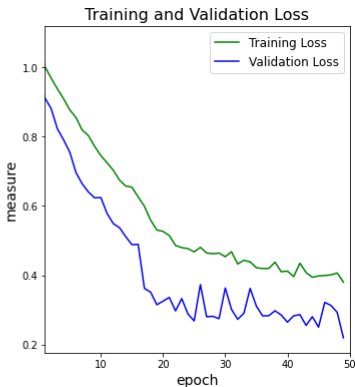
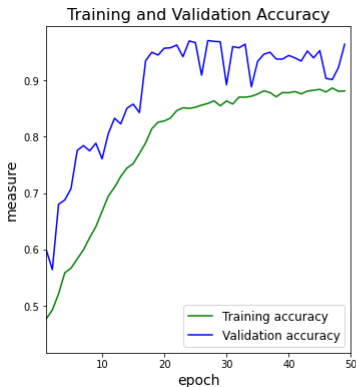
## Modelo 2:



Modelo 2. Accuracy y funcion de perdida para los conjuntos de entrenamiento y validacion.



## Modelo 3:



Modelo 3. Accuracy y funcion de perdida para los conjuntos de entrenamiento y validacion.



- La clasificación de objetos astrofísicos a través de redes neuronales presentó curvas de aprendizaje apropiadas para los tres modelos planteados.
- Las estructuras de las redes empleadas fueron construidas por uno proceso de agregación de capas y contraste continuo con la métrica de ajuste.