

Clasificación de espectros estelares mediante DeepLearning

Ignacio Loayza
Judytza Tapia

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

Clasificación del espectro de estrellas

El esquema de harvard de clasificación estelar divide el espectro en siete categorías principales de acuerdo a la temperatura de la estrella:

- O, B, A, F, G, K, M

“O” representa la mayor temperatura ($T > 30000$ K) y M la de menor (2200-3700 K)

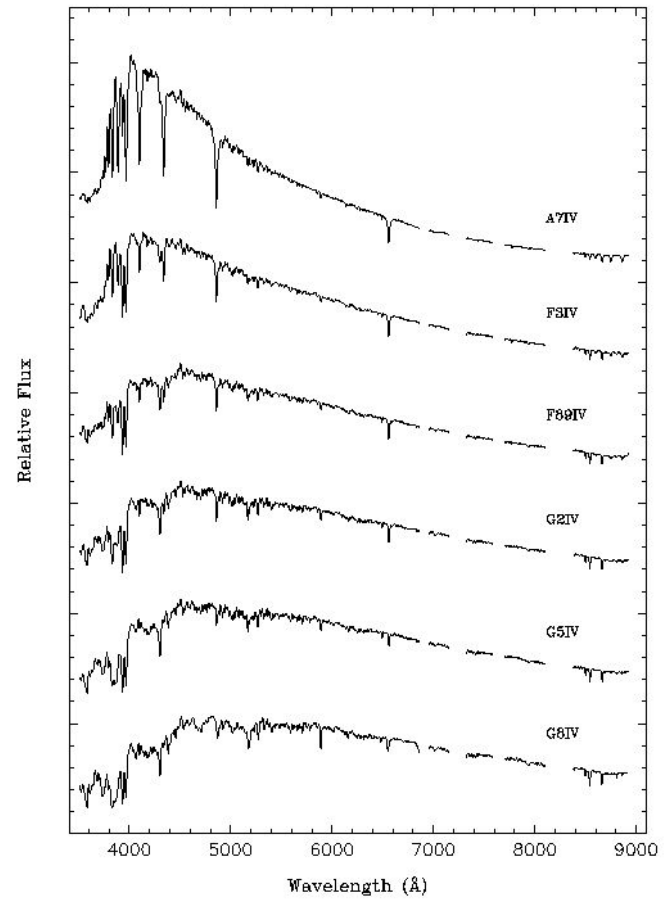
Cada clase contiene diez subclases (identificadas con números del 0 a 9) indicando decrementos en temperatura dentro de la clase.

Morgan & Keenan complementan el esquema posteriormente con clases de luminosidad:

- Ia, Ib, II, III, IV, V

La clasificación de acuerdo al esquema MK es convencionalmente realizada mediante criterio experto a través de la inspección visual de cada espectro en particular.

Un esquema de clasificación completo comprende tanto la clase del esquema harvard como la adición de Morgan & Keenan, por ejemplo: O9II, A0V



Solución – Conjuntos de datos

El objetivo principal es lograr clasificar los espectros de estrellas según el esquema Harvard sobre conjuntos grandes como LAMOST.

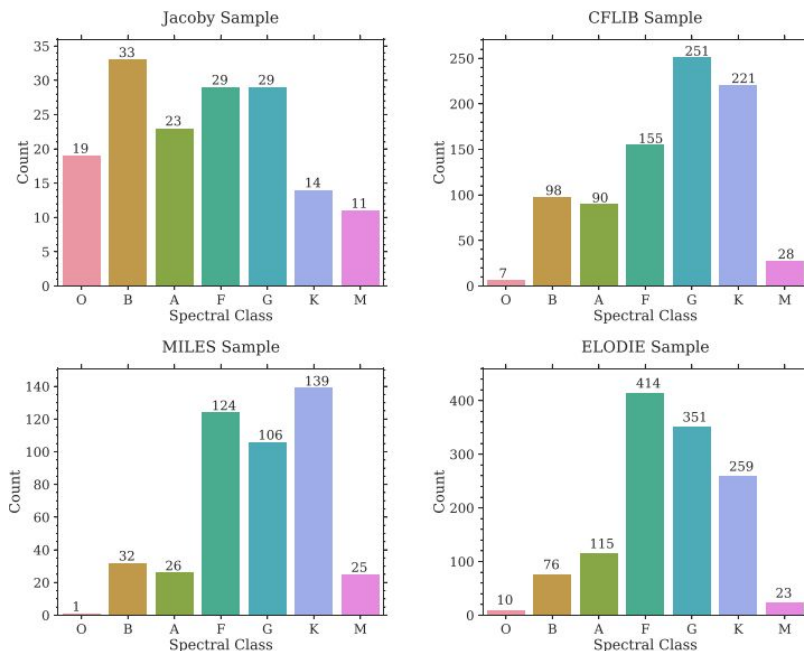
Para este estudio se evaluará sobre data de **LAMOST DR5** la cual contiene las siguientes clasificaciones ya realizadas sobre los espectros medidos:

A0, A0III, A1IV, A1V, A2IV, A2V, A3IV, A3V, A5, A5V, A6IV, A6V, A7, A7III, A7IV, A7V, A8III, A9, A9V, B, B6, B9, F0, F2, F3, F4, F5, F6, F7, F8, F9, G0, G1, G2, G3, G4, G5, G6, G7, G8, G9, K0, K1, K2, K3, K4, K5, K7, K9, M1, M2, M3, M4, M5, M6, M7, M8, M9

Como conjunto de entrenamiento se tiene considerado usar datos ya etiquetados presentes en LAMOST junto con alguna/s de las siguientes librerías espectrales:

Librería	N° espectros
JHC Atlas	161
ELODIE3.1	1962
CFLIB	1273
MILES	985

Solución – Distribución de clases en las librerías espectrales consideradas



Solución – Arquitecturas a probar

Modelos basados en DL

- 1D CNN (**Baseline**)
- 1D VGG16
- 1D ResNet
- 1D Densenet

Trabajos recientes¹ muestran que:

- Arquitecturas convolucionales 1D son eficaces en resolver el problema de clasificación de espectros estelares.
- Capas densas simples pueden aprender atributos relevantes en para la clasificación de espectros estelares.

Modelos de clasificación de ML

Directamente sobre el espectro

- SVM (**Baseline**)
- Random Forest
- XGBoost

Sobre un vector de características obtenido de las activaciones internas de la arquitectura 1D CNN.

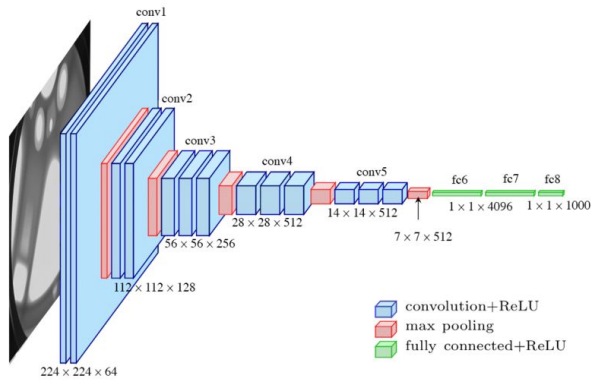
- SVM
- XGBoost
- Random Forest

¹ Application of convolutional neural networks for stellar spectral classification
.- Sharma et al (2020)..

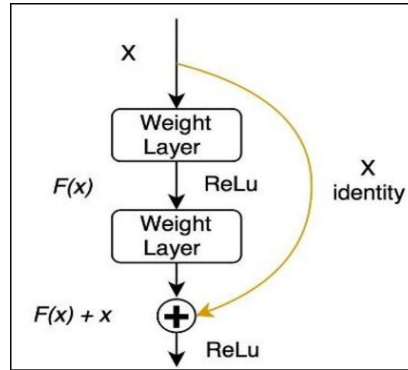
Automated Classification of Massive Spectra Based on Enhanced Multi-Scale
Coded Convolutional Neural Network.- Jian et al (2020)

Solución – Arquitecturas a probar

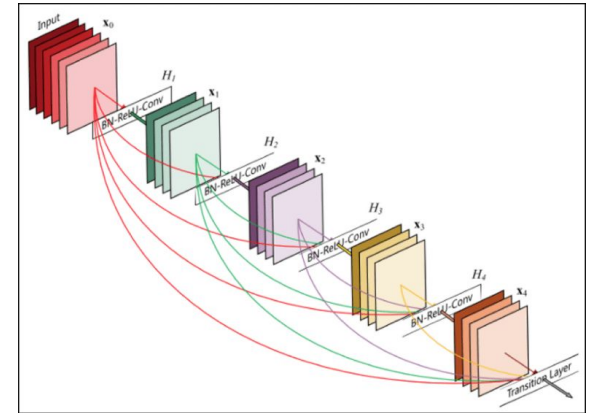
VGG16



Resnet



Densenet



Evaluación de la solución

Este es un problema de clasificación multiclase, por lo que será evaluado con las métricas adecuadas para el problema:

- F1 Score
- Recall
- Precision
- ROC/AUC

Las arquitecturas anteriores serán modificadas de forma de presentar en su última capa la cantidad de clases que se desee predecir, aún está en discusión hasta qué nivel de detalle en cuanto a las clases disponibles (esquema Harvard y/o MK) se abordará en este proyecto.

Fin