



Búsqueda y caracterización de exoplanetas mediante técnicas de IA

Propuesta de trabajo

Eduardo Ojero. ESAC, Madrid

Ana Ulla, Universidad de Vigo

Minia Manteiga, Universidad de A Coruña

1.- Técnicas de IA

Machine Learning

- No supervisadas → Agrupamiento **redes SOM, PCA, K-means..**
→ con etiquetas: clasificación
- Supervisadas → Clasificación, parametrización **RNAs (incluyendo Deep Learning), SVM, Random-Trees,...**

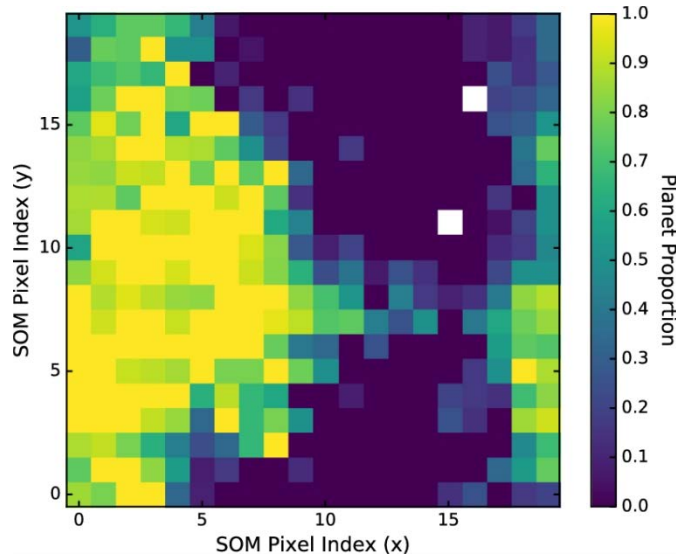


Figure 3. Proportion of disposed KOIs within each SOM pixel which are true planets
Transit shapes and self-organizing maps as a tool for ranking planetary candidates: application to *Kepler* and *K2*
D. J. Armstrong D. Pollacco A. Santerne. MNRAS 465, 2634, **2017**

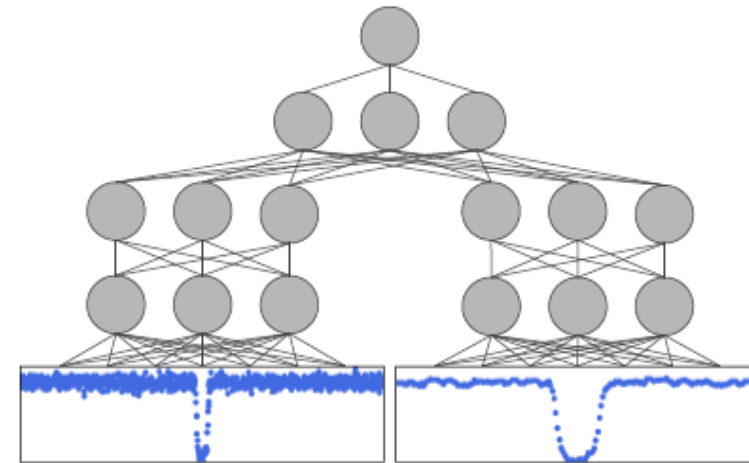


FIG. 4.— Fully connected neural network architecture for classifying light curves, with both global and local input views.

IDENTIFYING EXOPLANETS WITH DEEP LEARNING: A FIVE PLANET RESONANT CHAIN AROUND KEPLER-80 AND AN EIGHTH PLANET AROUND KEPLER-90
Christopher J. Shallue & Andrew Vanderburg, ApJ 155:94, **2018**

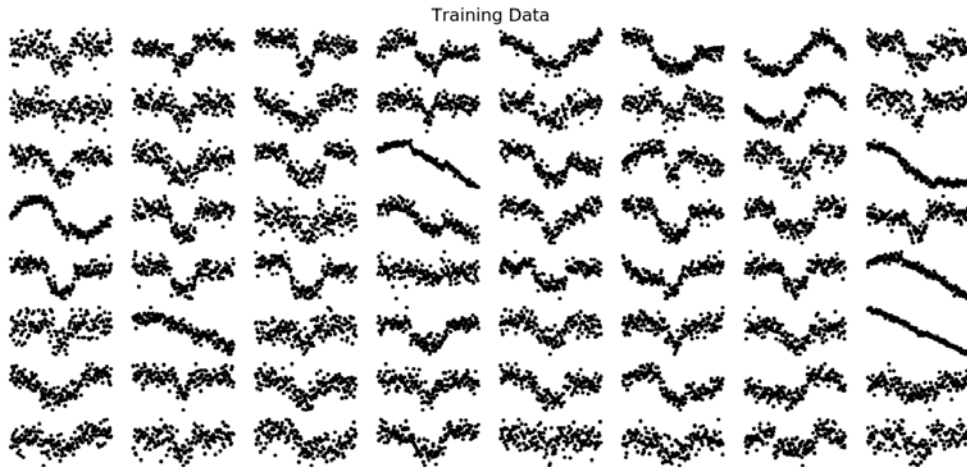
2.- Búsqueda de exoplanetas mediante IA

Desarrollo incipiente 2015-2018, necesidad de unificar criterios para poder estudiar poblaciones de exoplanetas

Datos de entrada para la búsqueda de exoplanetas: [Series fotométricas](#), Espectro de potencias, Curvas de velocidad radial

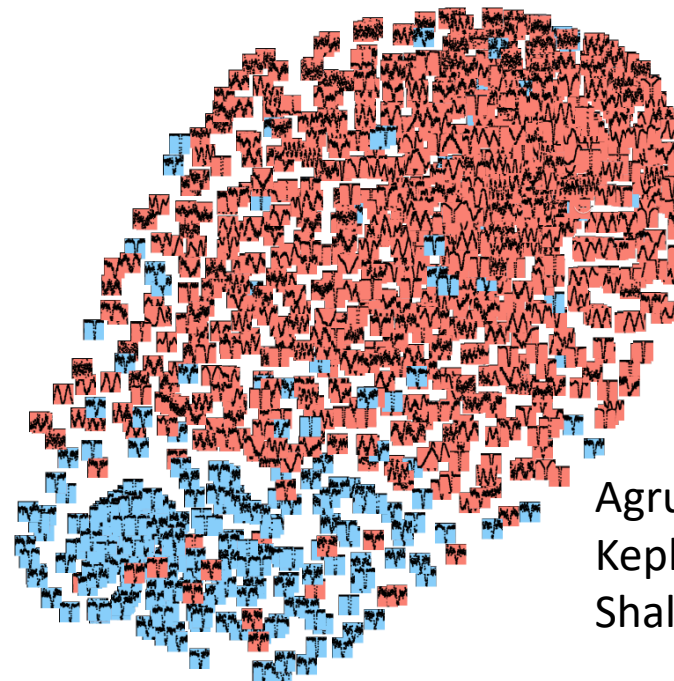
Entrenamiento

- Curvas de luz sintéticas (Pearson et al. 2017), TCEs o **KOIs**, datos modificados con píxeles simulados (Shallue & Vanderburg, 2018)
- Problemas: efectos sistemáticos instrumentales, variabilidad estelar, tránsitos debidos a binarias
- Preprocesado: detrending (pipelines automáticas) + **phase folding** → asumir un periodo, remuestrear la curva de luz resultante, eliminar bajas frecuencias, escalar las curvas



Curvas sintéticas.

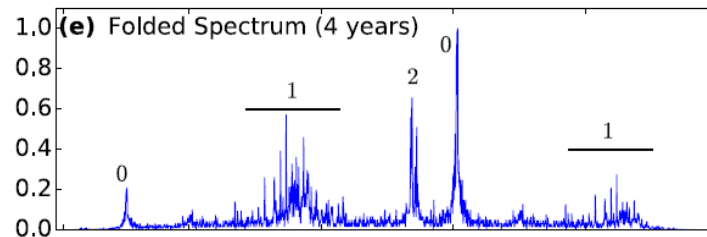
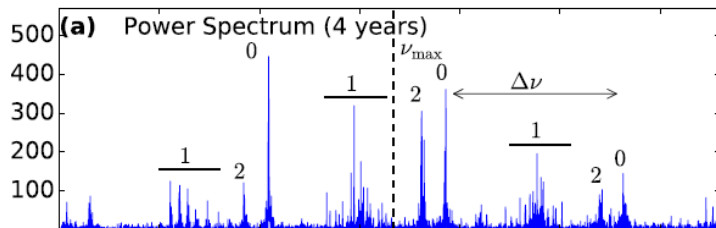
Pearson, Palafox & Griffith, MNRAS 474, 2018



Agrupamiento 2D de curvas de luz
Kepler TCEs, con y sin tránsitos
Shallue & Vanderburg, ApJ 155, 2018

2.- Búsqueda de exoplanetas mediante IA

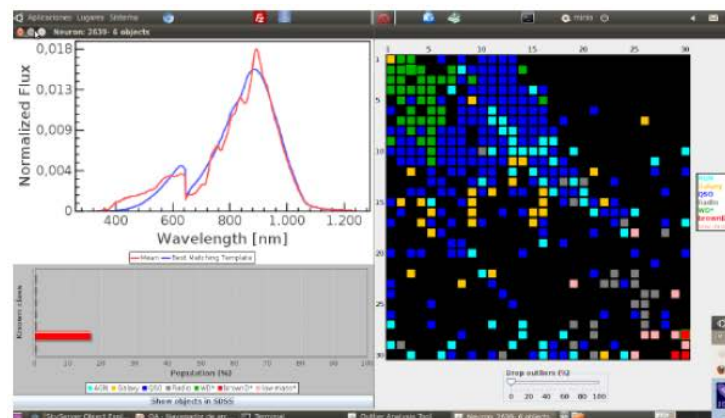
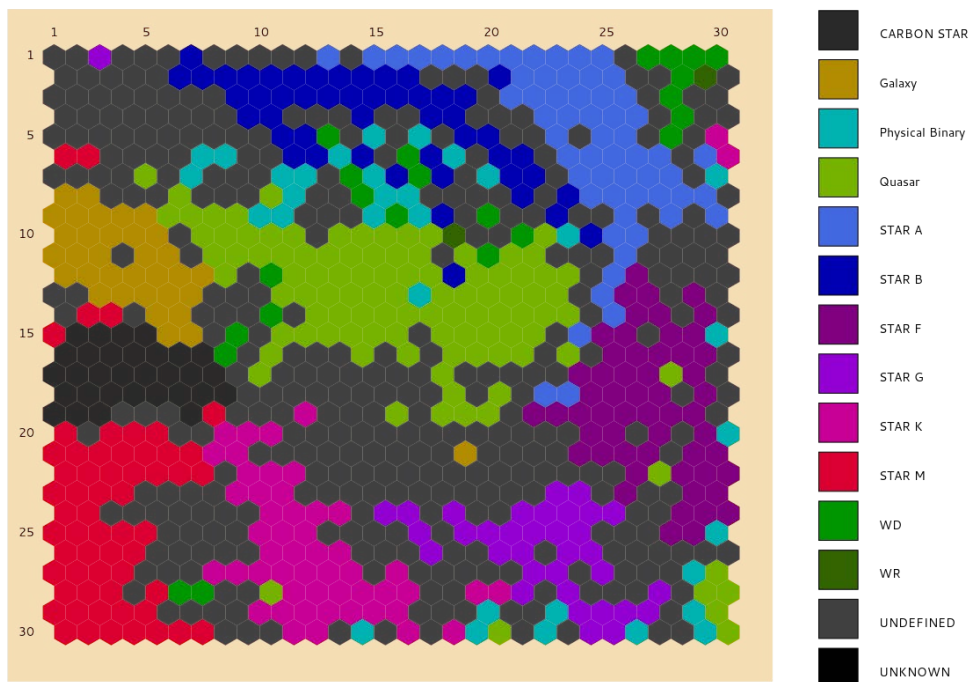
- **Evaluación del funcionamiento**, figuras de mérito Accuracy, Precisión, Recall, Matriz de confusión
- **Arquitectura de las redes**
 - Supervisadas: Redes profundas completamente conectadas con retropropagación del error, **Redes convolucionales**
Capa de salida única: **redes binarias**: transit planet SI o NO
 - No supervisadas: **Mapas SOM**: agrupa curvas de luz parecidas, separa tránsitos de binarias y de planetas
- **Variantes:**
 - No asumir periodos, remuestrear en todos los periodos de interés. Ventaja de técnicas automáticas
 - Dominio de Fourier: Usar los espectros de potencia. Buenos resultados separando tipos de gigantes (Hon et al. 2018)
 - Utilizar curvas sintéticas para poder parametrizar los tránsitos



Hon et al 2018, MNRAS 476, 2018

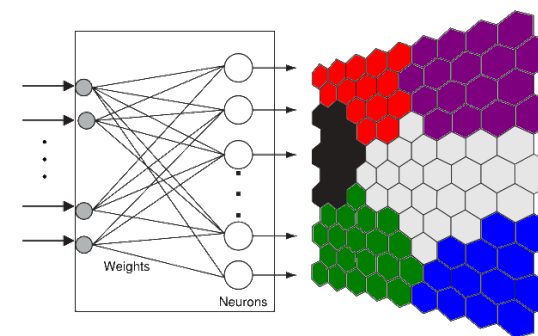
3.- Experiencia y herramientas disponibles en el grupo gallego de Gaia

3-1. GUASOM: herramienta para el desarrollo y análisis de mapas SOM



New BD candidates

Espectrofotometría Gaia BP/RP.
Conjunto de validación DR3 8M fuentes
Validation results – ApsisVal3.1f (Xmatch
SIMBAD labels)



Análisis de Outliers de clasificación del pipeline de SDSS DR4
Fustes et al, A&A 559, 2013

3-2. ANN: parametrización de espectros estelares mediante redes supervisadas

Dominio: espectros Gaia/RVS 8500 Å, R=11500

Parámetros atmosféricos estelares, para estrellas frías e intermedias

Entrenamiento con espectros sintéticos calculados ad-hoc

Datos con baja SNR: descomposición A y D WAVELETS

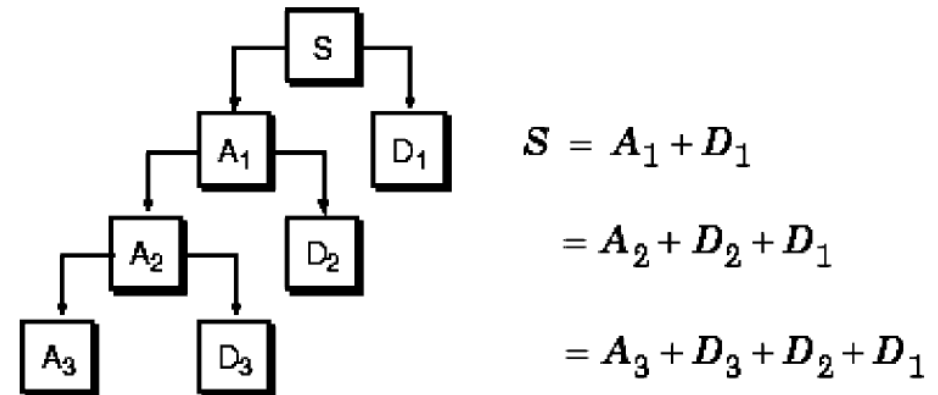
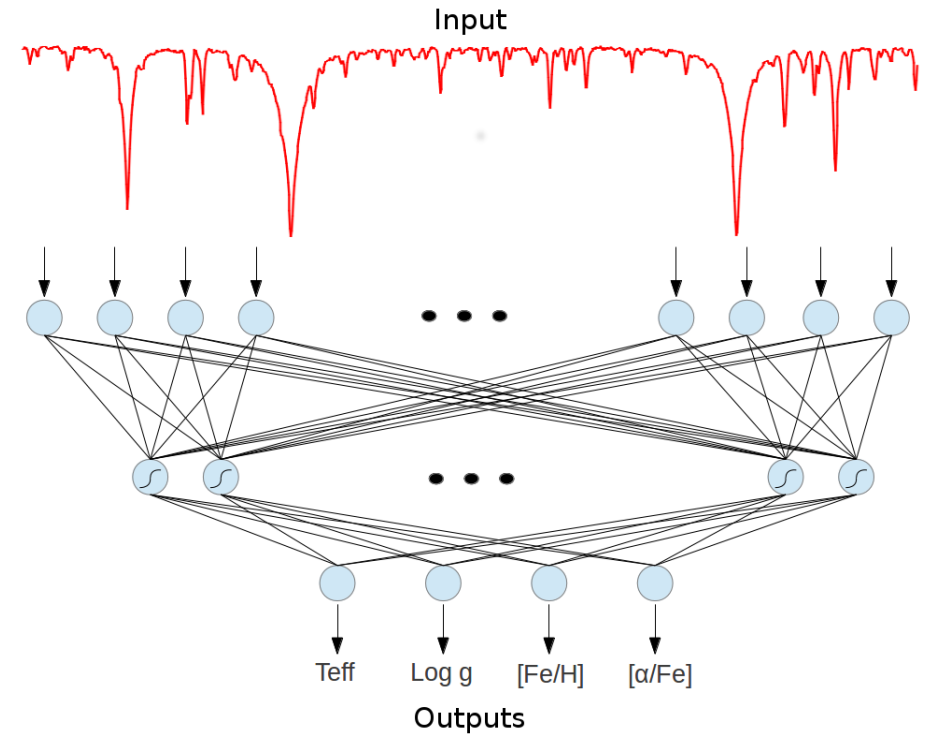
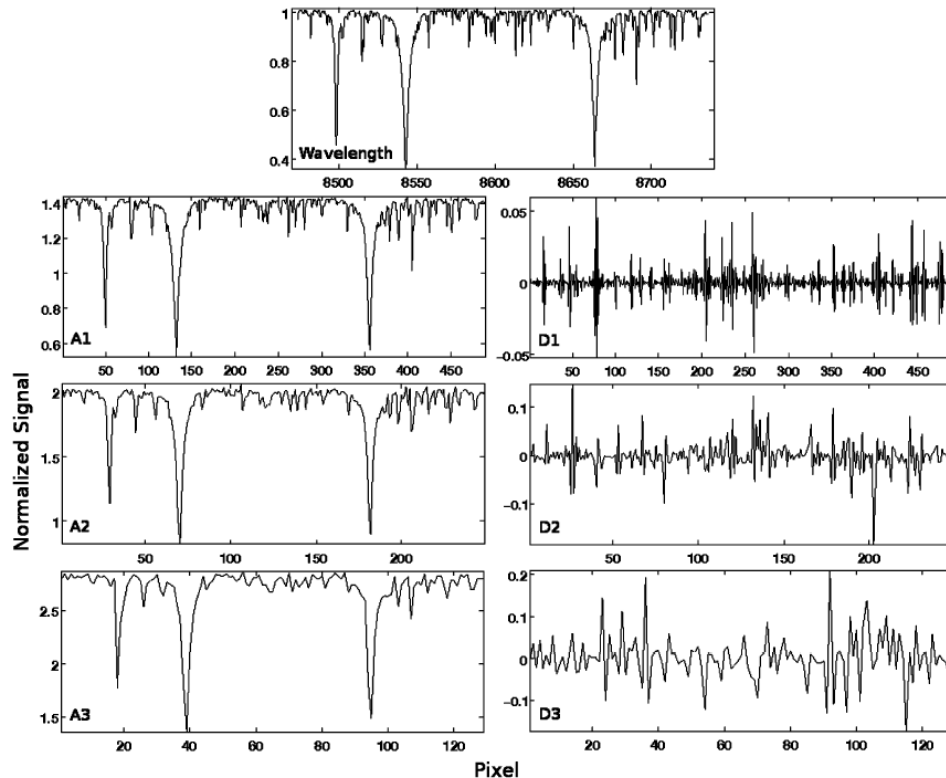


FIG. 1.—Wavelet multilevel analysis.

Manteiga et al., PASP 122, 2010

4.- Propuesta de trabajo y tesis doctoral de Eduardo

4-1.- Comenzar con las curvas de luz de los KOI

a) Preprocesado de las curvas de luz

b) Clustering con GUASOM, interpretación de los prototipos

c) Construir un conjunto de entrenamiento para redes supervisadas

- basado en observaciones

- basado en curvas sintéticas

d) Probar diferentes arquitecturas de redes supervisadas, con varias capas ocultas (Deep Learning) e incluyendo redes convolutivas

e) Probar diferentes dominios: temporal, frecuencias, wavelets

4-2.- Extender el estudio a curvas de luz obtenidas con otros instrumentos, CARMENES