

# DESCUBRIMIENTO DE CONOCIMIENTO EN BASES DE DATOS

## Clasificación supervisada [Caso de uso I]

- Estoy interesado en el tipo de objetos T
- Tengo una serie de ejemplos (espectros) de objetos T
- Tengo miles o millones de casos (espectros) de entre los cuales quiero entresacar los de tipo T

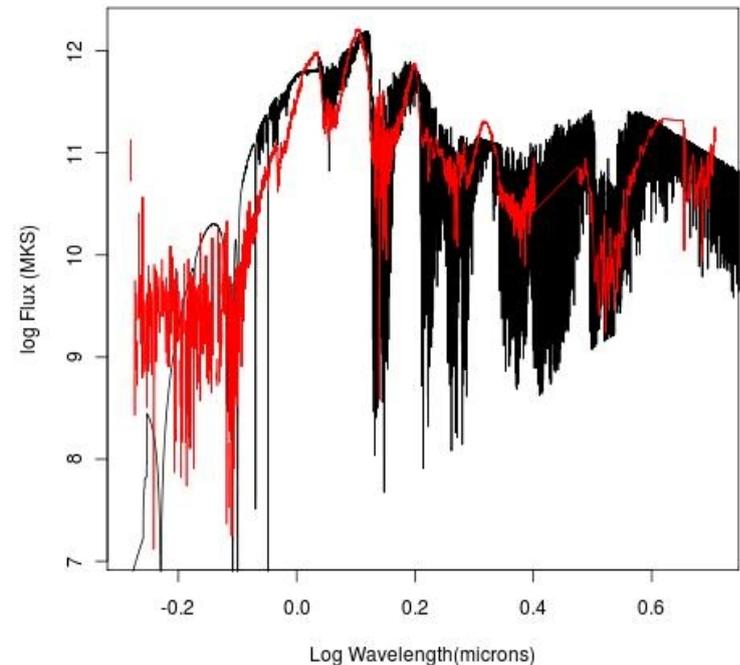
+ Reducción de la dimensionalidad (¡los espectros son muy largos!)

+ Evaluación de los modelos

## Clasificación supervisada [Caso de uso II: regresión]

- Quiero caracterizar físicamente los objetos de tipo T a partir de su espectro
- Tengo calibrados los ejemplos (espectros, índices de color) conocidos
- Reducción de la dimensionalidad y Evaluación de los modelos

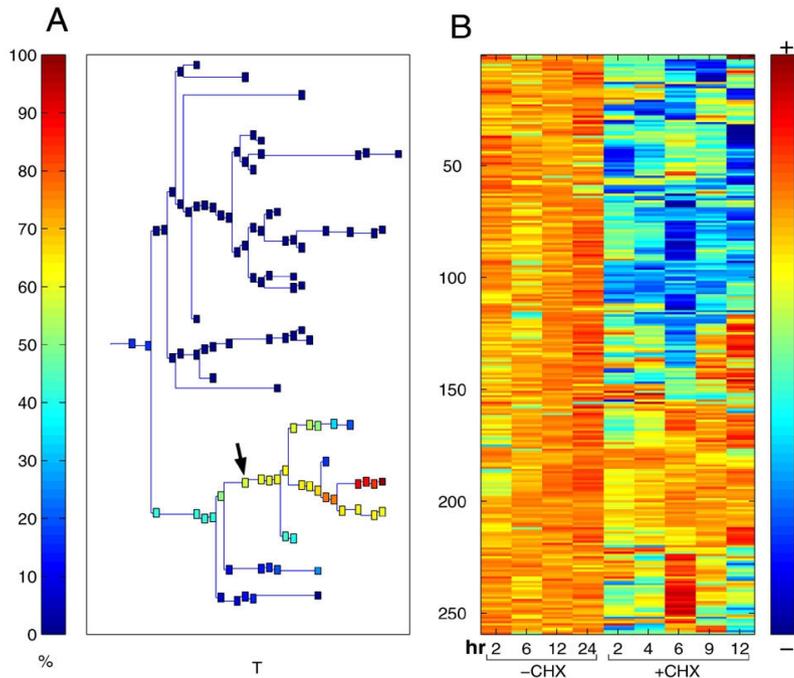
Quizás cometo un error muy pequeño en sistemas improbables...



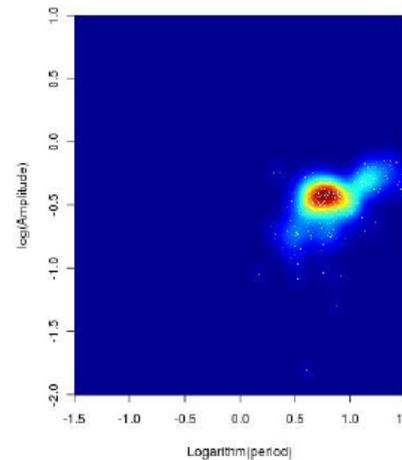
# DESCUBRIMIENTO DE CONOCIMIENTO EN BASES DE DATOS

## Agrupamiento o clustering [Caso de uso III]

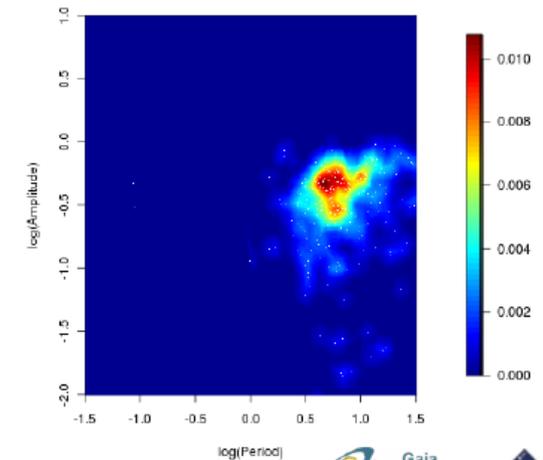
- Tengo una muestra de objetos cuya distribución es conocida a la resolución pre GAIA
- Confiamos en la fuerza (ley asintótica para muestras grandes)
- Realizamos análisis multivariante para identificar grupos, jerarquías,...



Hipparcos characterization Of Milky Way Cepheids



CU7 Characterization of Milky Way Cepheids



Si descubrimos un nuevo tipo de objeto...

# No exactamente Minería de Datos...

Selección de modelos [Caso de uso IV]:

- Tengo dos o más modelos (por ejemplo de complejidad creciente)
- Quiero saber qué parámetros del modelo pueden explicar las observaciones
- Quiero saber si está justificado seleccionar el modelo con más parámetros (navaja de Occam)
- Quiero conocer intervalos de confianza para los parámetros
- Quiero hacer contraste de hipótesis