

# ***Data Science. Aplicación a Imágenes Satelitales***

## ***Seminario de Investigación***

*Prof. Elvira Martínez de Icaya Gomez*

*Prof. M. Estíbaliz Martínez Izquierdo*

*Ingeniera Mayra Martínez Mena*

***Máster Universitario en Ciencias y Tecnologías de la Computación***



**Curso 2017-18**

**25 de Abril 2018**



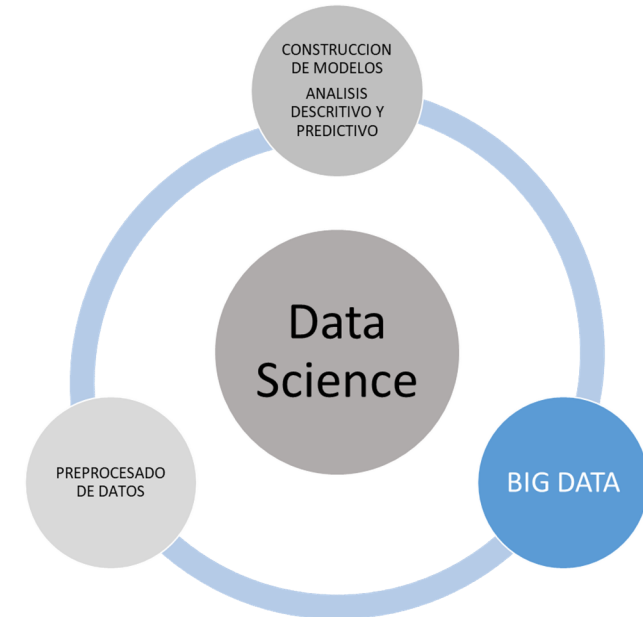
# Data Science

- En 2006, Clive Humby (UK), matemático:

*Data is the new oil*



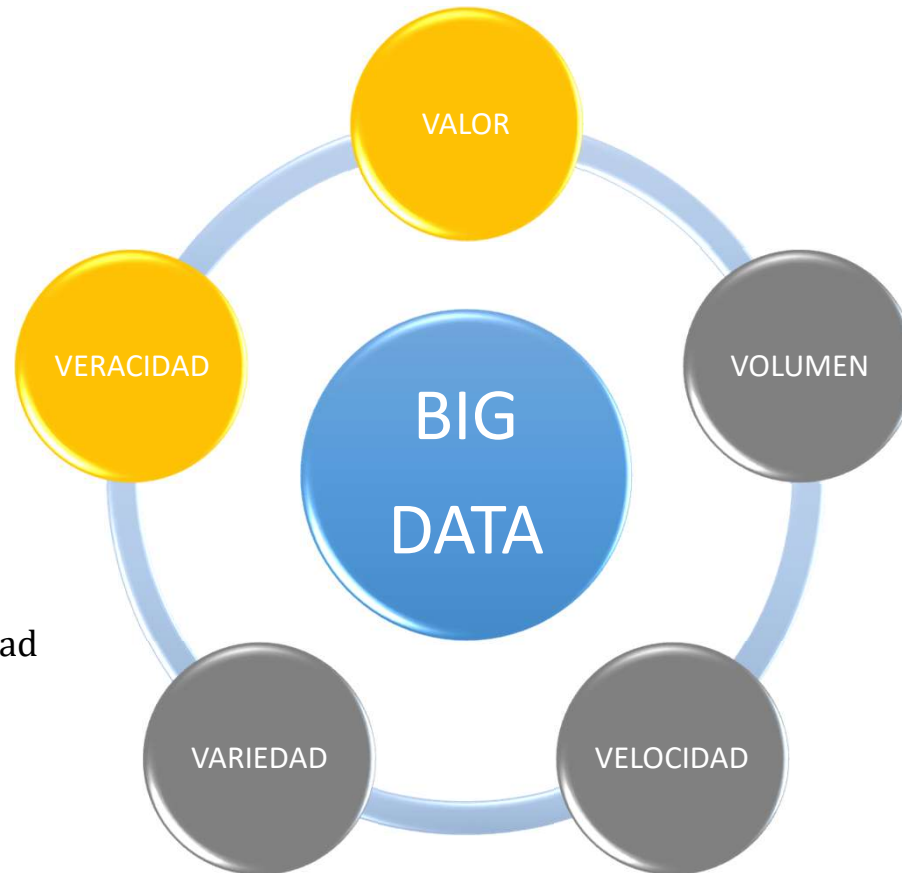
- *Data Science* es un campo multidisciplinario. Es la intersección entre los siguientes dominios:
  - Conocimiento del negocio (plan de investigación)
  - Aprendizaje (automático)
  - Programación



Involucra métodos científicos, procesos y sistemas para extraer conocimiento o un mejor entendimiento de datos en sus diferentes formas

# BIG DATA

Como concepto es definido alrededor de cinco aspectos sobre los datos:



El volumen, la variedad y la velocidad se refieren a procesos de:

- ✓ **Generación**
- ✓ **Forma de captura**
- ✓ **Almacenamiento**

de los datos

El valor y la veracidad están relacionados con características de:

- ✓ **Calidad**
- ✓ **Utilidad**

de los datos



**SMART DATA**

# SMART DATA

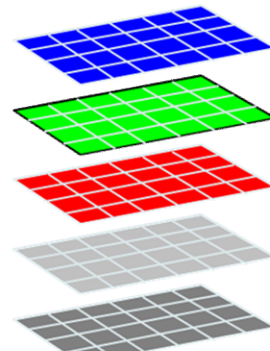
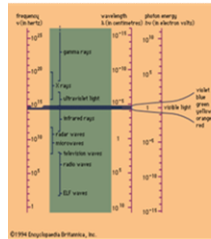
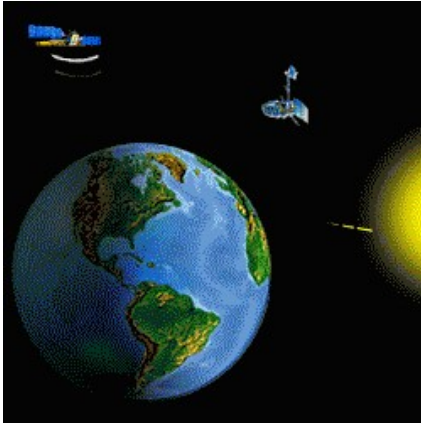
Ayudan a tomar decisiones inteligentes



Dos casos de estudio en Teledetección:

1. Reducción de datos mediante calculo de índices espectrales
2. Clasificación con Lógica Difusa

# DATOS SATELITALES



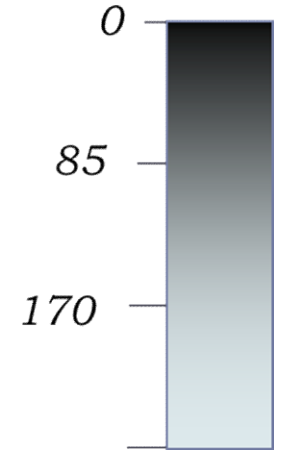
Espectral



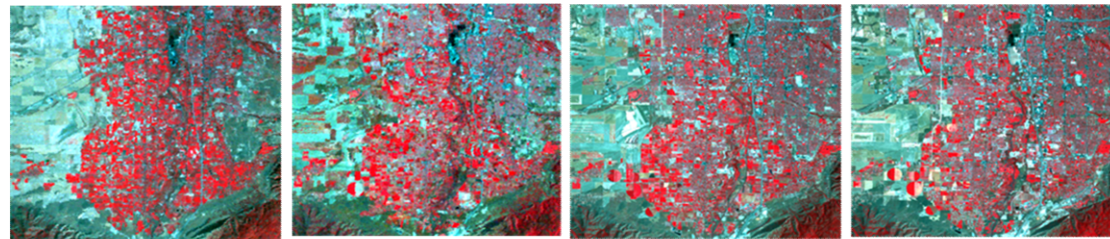
Espacial



Radiométrica



Temporal



# Reducción de datos mediante calculo de índices espectrales

Primer caso de estudio



Curso 2017-18

25 de Abril 2018



# Fuentes de Información Satelital



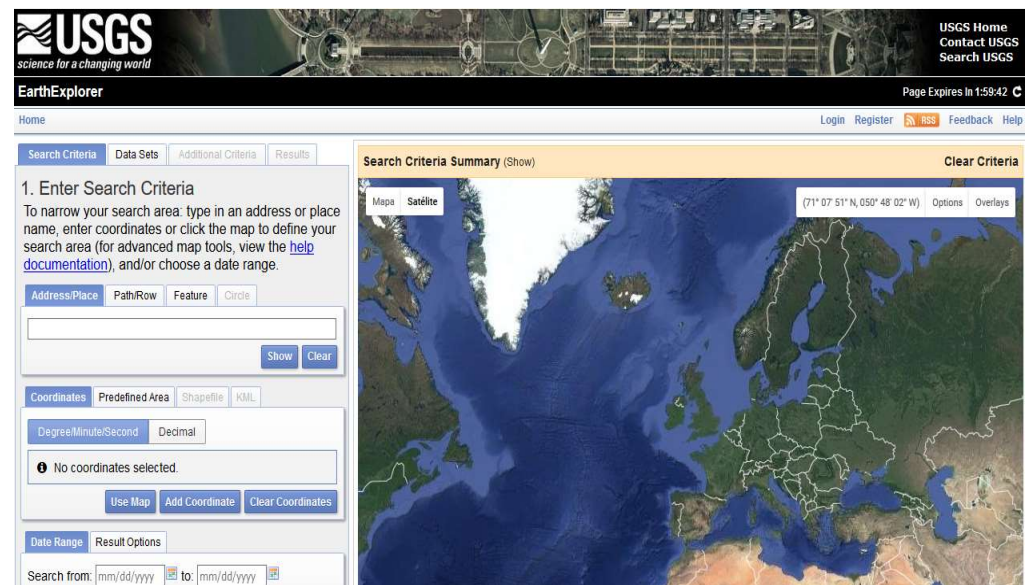
# Programas de Observación de la Tierra

## Programa Copérnico

## USGS



<https://scihub.copernicus.eu/dhus/>

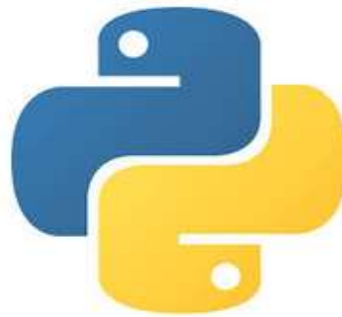


<https://earthexplorer.usgs.gov/>



# Lenguajes para análisis de datos

- ✓ Muchos lenguajes que ofrecen las capacidades para ejecutar operaciones de análisis de datos
- ✓ Están emergiendo con fuerza: R, Python, MATLAB.



# Pasos a aplicar a imágenes satelitales



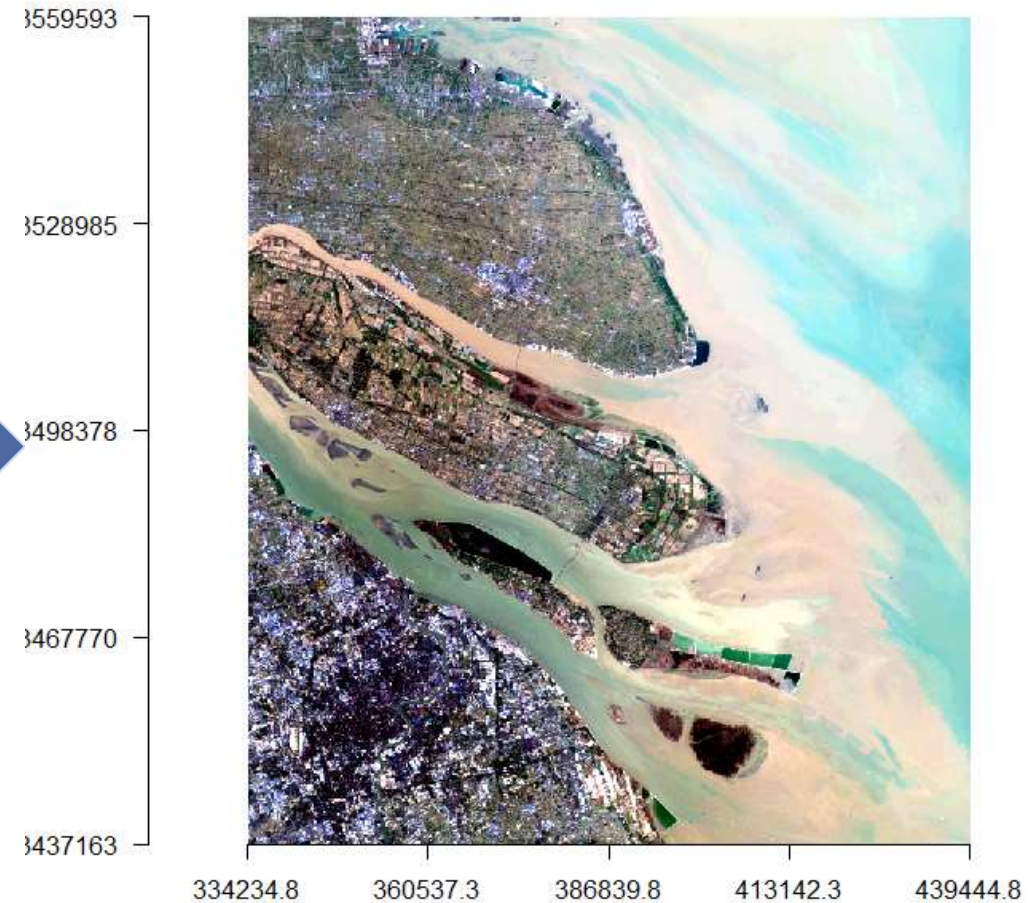
# Importación y Exportación de Datos

```
Blue<-"C:\Blue.TIF"  
Green<-"C:\Green.TIF"  
Red<-"C:\Red.TIF"  
NIR<-"C:\NIR.TIF"  
Swir1<-"C:\Swir1.TIF"  
Swir2<-"C:\Swir2.TIF"
```

```
s<-brick("C:\stack_2017.TIF")
```

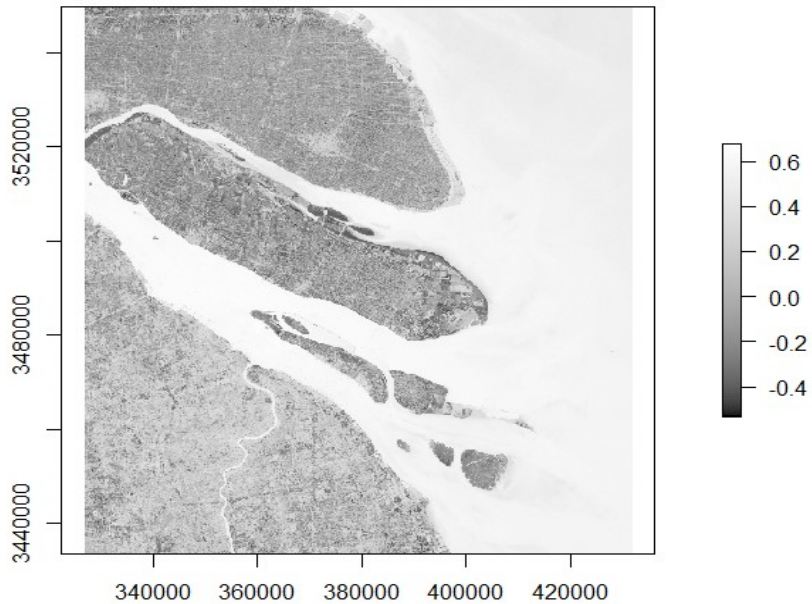
```
plotRGB(s, r = 3, g = 2, b = 1, stretch = "lin",  
main = "Landsat 8 True Color Composite")
```

Landsat 8 True Color Composite

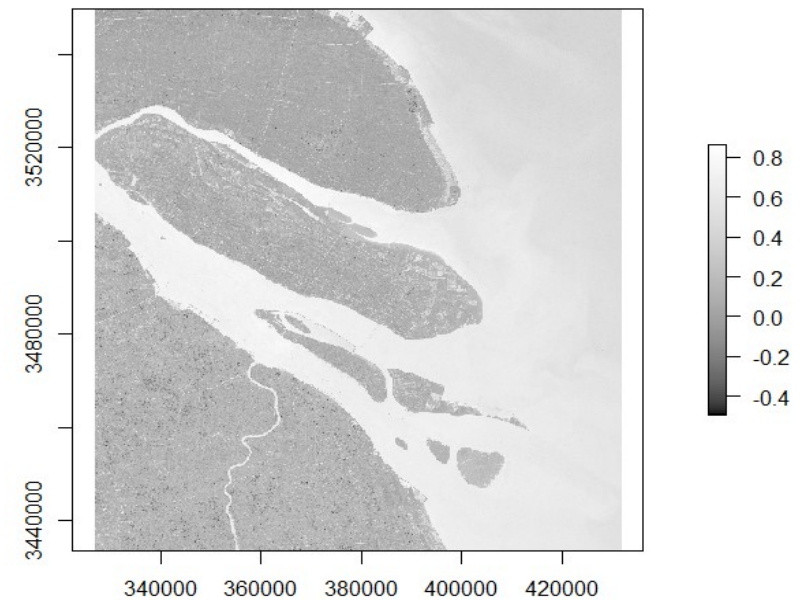


# Indices Espectrales

NDVI



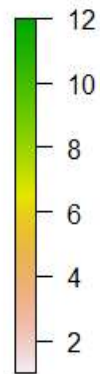
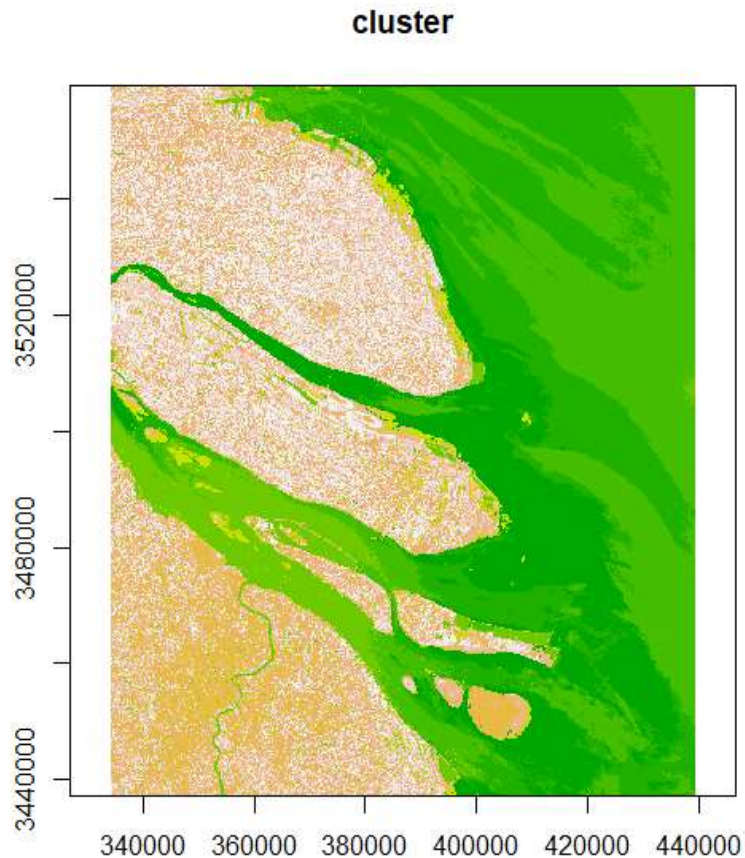
NDWI



```
band3 = raster(i, layer=3)
band4 = raster(i, layer=4)
NDVI=(band3-band4)/(band3+band4)
plot(NDVI, col=grayscale_colors, main='NDVI')
```

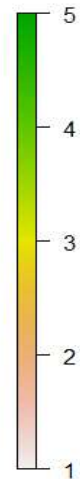
```
band3 = raster(i, layer=3)
band6 = raster(i, layer=6)
NDWI=(band3-band6)/(band3+band6)
plot(NDWI, col=grayscale_colors, main='NDWI')
```

# Agrupamiento (Clustering) de niveles digitales. Clasificación no supervisada



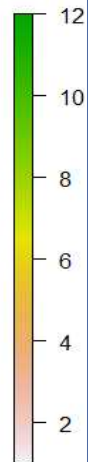
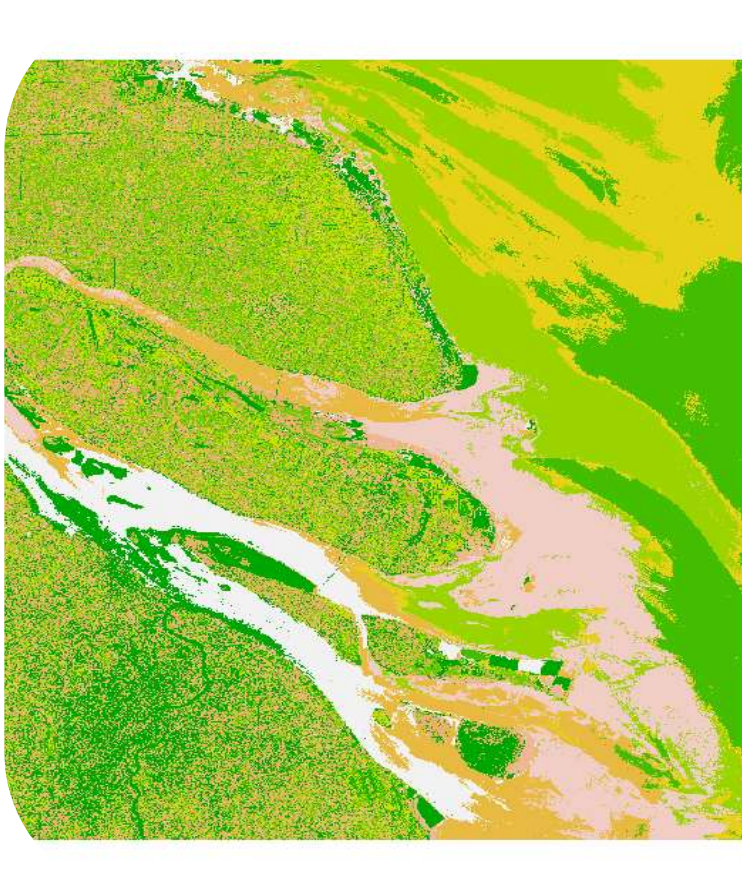
```
clus<-clara(v,12,sample=500, metric="manhattan", pamLike = T)  
clara_raster<-raster(s)  
clara_raster[i]<-clus$clustering  
plot(clara_raster, main="cluster")
```

# Agrupamiento (Clustering) de niveles digitales. Clasificación no supervisada



```
s.kmeans<-kmeans(s[,5,iter.max = 100,nstart = 3])  
kmeansraster<-raster(s)  
kmeansraster[]<-s.kmeans$cluster  
plot(kmeansraster, main="kmeans")
```

# Clasificación Random Forest (Optimización de la clasificación)



```
E_rf <- kmeans(rf_prox, 12, iter.max = 100, nstart = 10)
rf <- randomForest(vx, as.factor(E_rf $ cluster), ntree = 500)
rf_raster <- predict(s, rf)
plot(rf_raster)
```

# Clasificación con Lógica Difusa

Segundo caso de estudio



Curso 2017-18

25 de Abril 2018

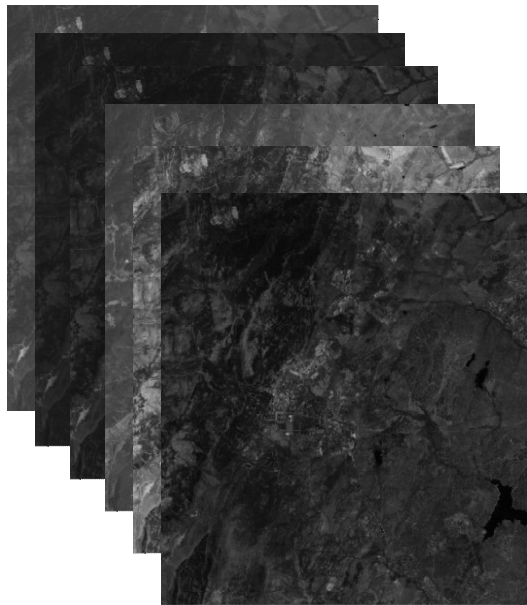




# Clasificación temática de imágenes satelitales

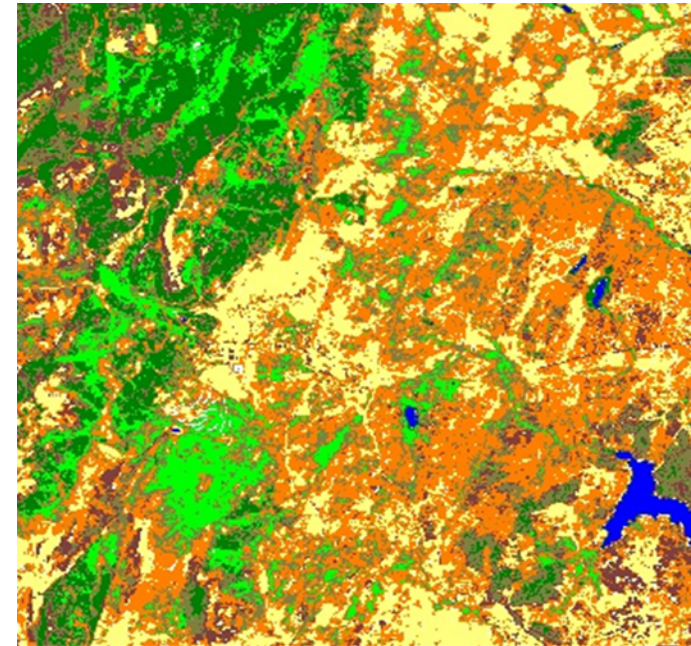
¿Qué es una imagen clasificada?

B1 (B, VIS)  
B2 (G, VIS)  
B3 (R, VIS)  
B4 (NIR)  
B5 (IR)  
B6 (IR)



LANDSAT

Toma de decisiones



	Agua		Pinar
	Bosque		Suelo claro
	Matorral		Suelo oscuro
	Mixto		No clasificados

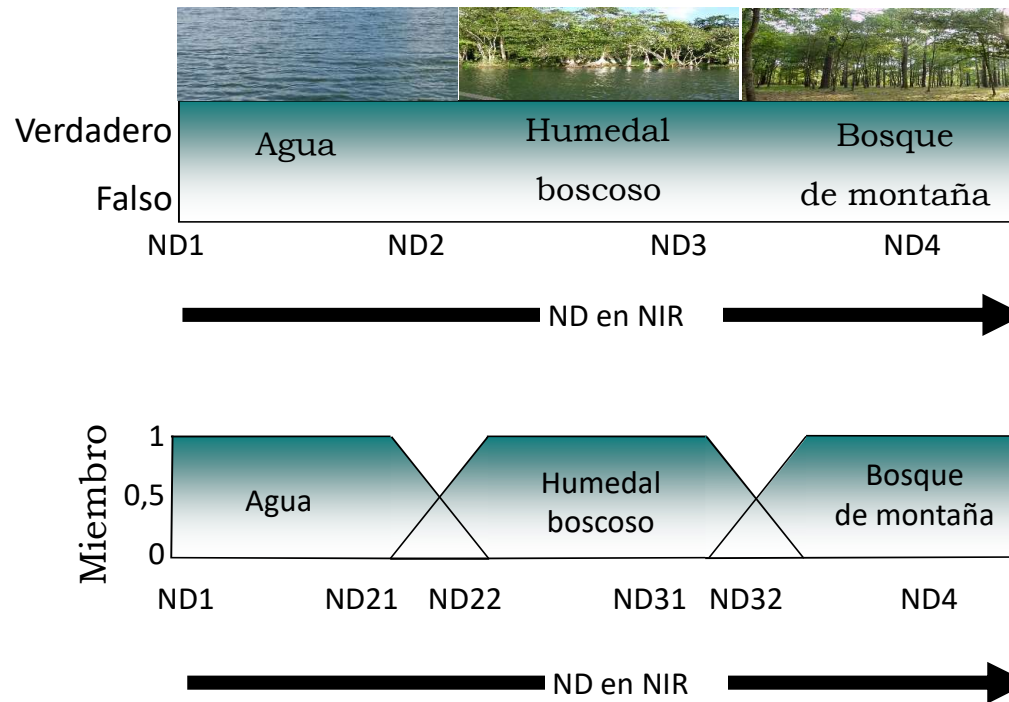
Etiquetas

# Clasificación temática de imágenes satelitales

## I. Entrenamiento → Reglas

## II. Asignación

*La representación lingüística y el poder de procesamiento de la lógica difusa hacen que se considere una herramienta para la inteligencia computacional, para tender un puente simbólico entre la inteligencia natural y la artificial*



Reglas de decisión:

Agua

Humedal boscoso

Bosque de montaña

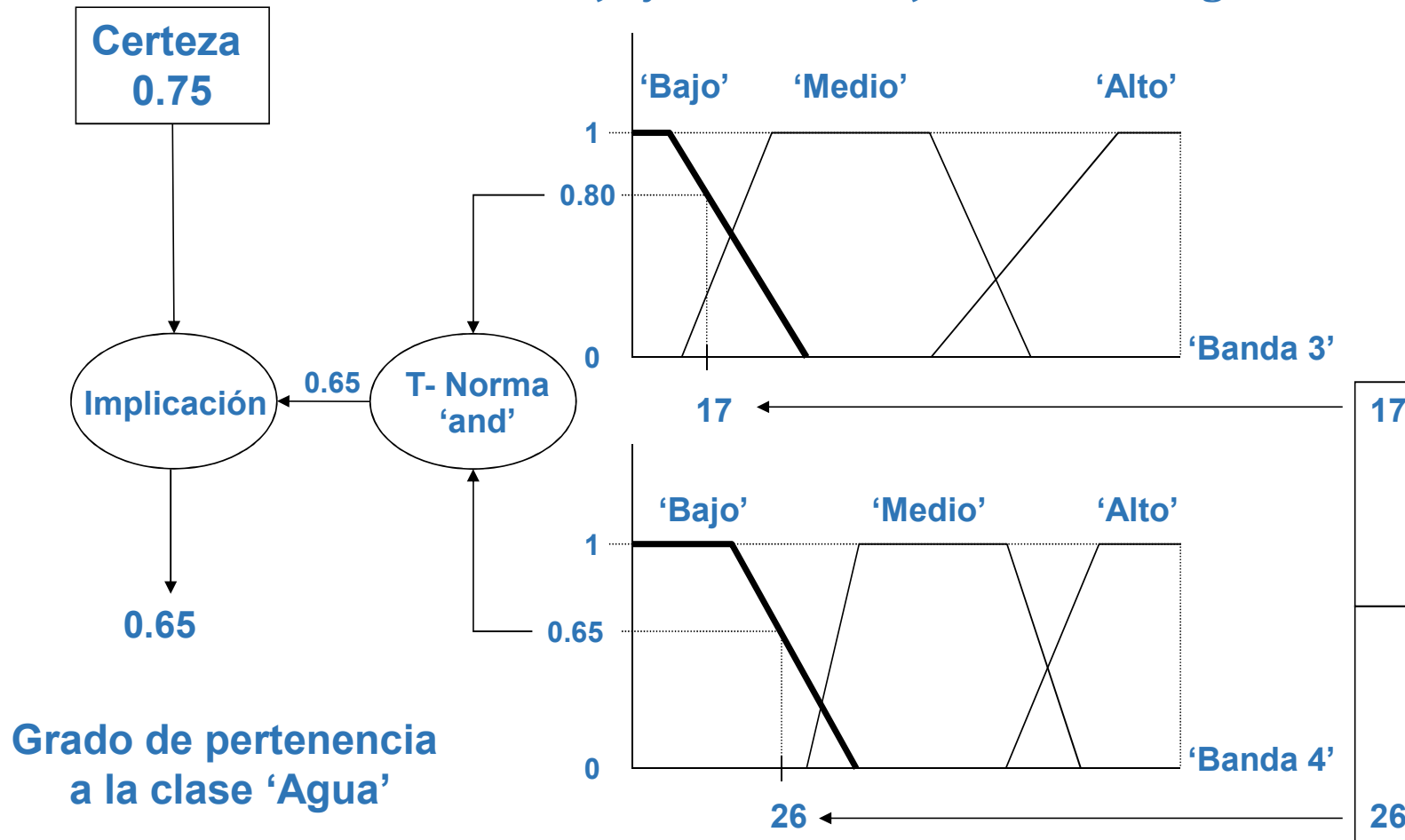
Clasificadores Hard y Soft (Fuzzy)

Las reglas de decisión son definidas como funciones miembro de cada clase

Las funciones miembro sitúan a cada píxel en un valor real entre 0 y 1 con un grado de pertenencia

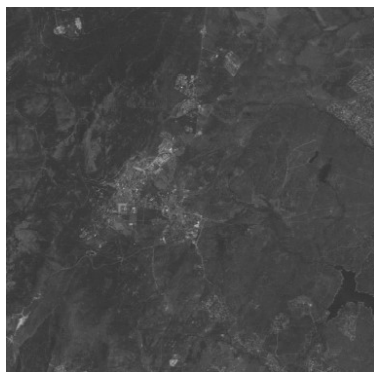
# LOGICA FUZZY PARA CLASIFICACION

SI Banda3 =Bajo y Banda4 = Bajo ENTONCES 'Agua'

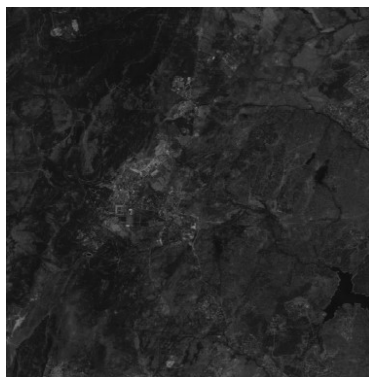


# BANDAS LANDSAT (Niveles de gris)

B1



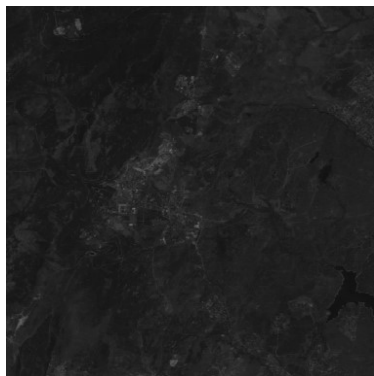
B3



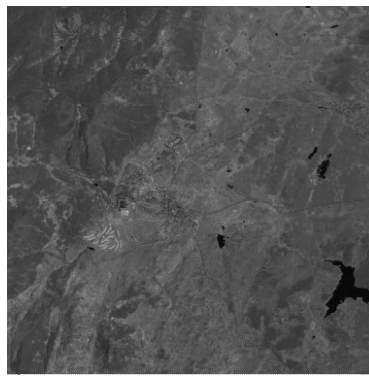
B5



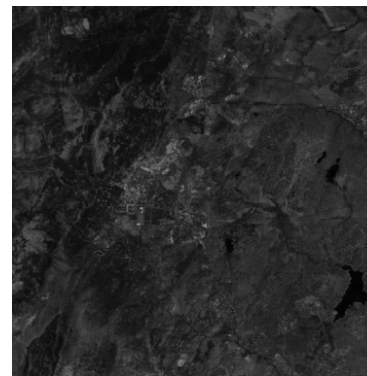
B2



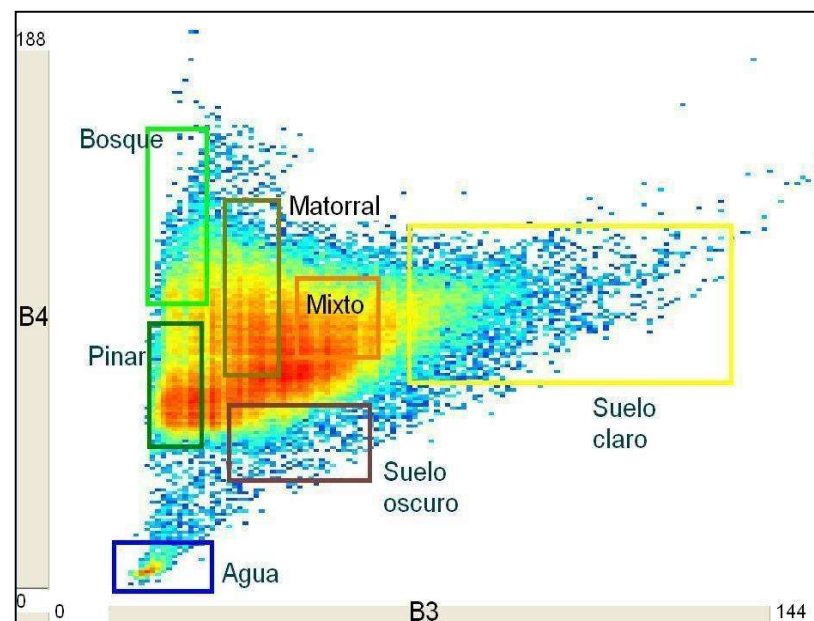
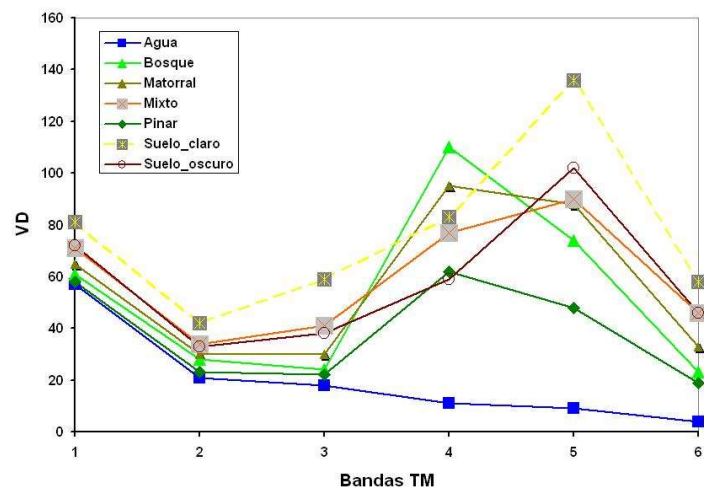
B4



B7



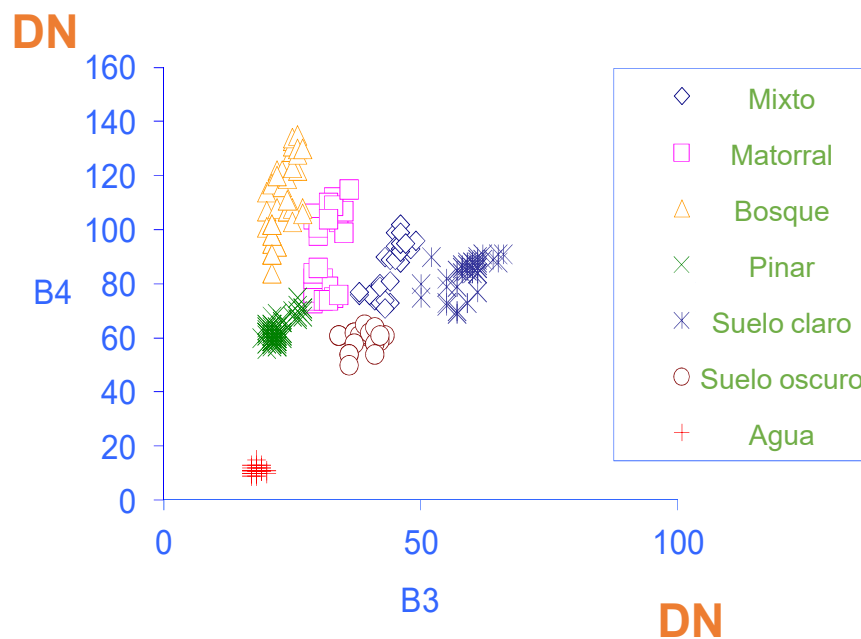
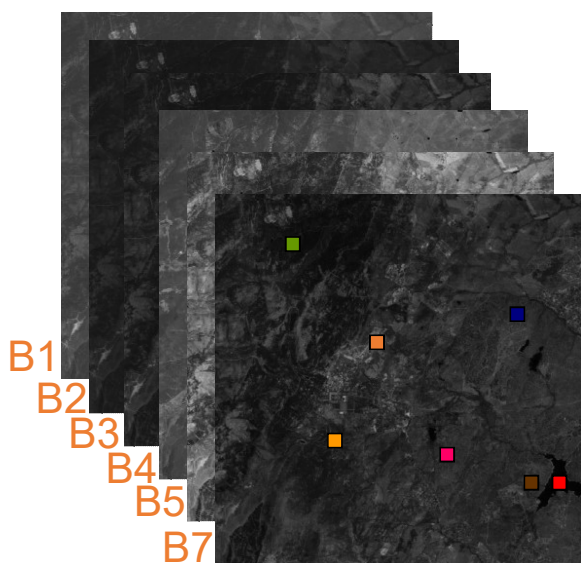
# Definición del espacio espectral de entrenamiento



# Definición del espacio espectral de entrenamiento

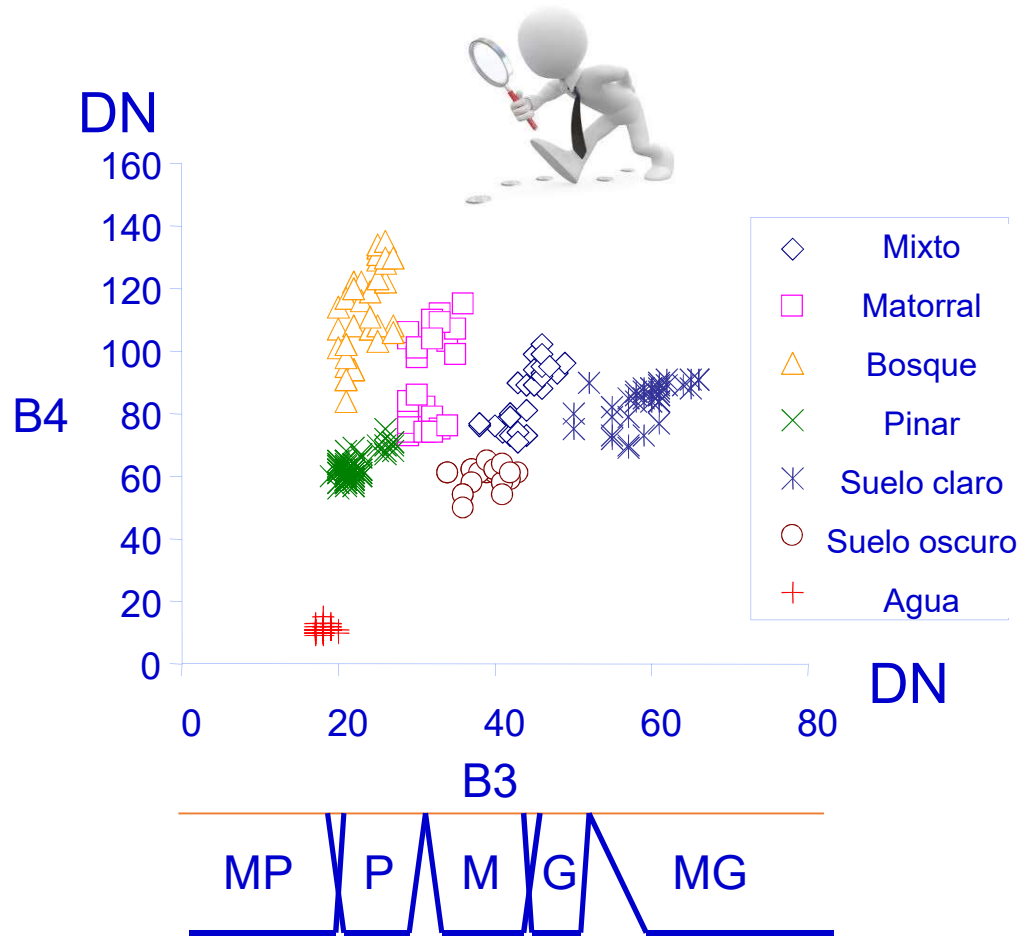
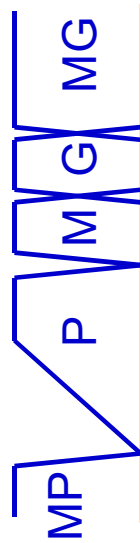


REDUCCION A UN ESPACIO DE DOS DIMENSIONES



# Definición de funciones miembro y términos lingüísticos

FUNCIONES MIEMBRO	
<b>B3</b>	
Muy Pequeño	0,0,19,20
Pequeño	18,20,28,30
Medio	30,32,42,44
Grande	42,43,49,50
Muy Grande	50,57,255,255
<b>B4</b>	
Muy Pequeño	0,0,16,20
Pequeño	20,56,76,80
Medio	80,84,100,104
Grande	96,100,116,120
Muy Grande	116,120,255,255



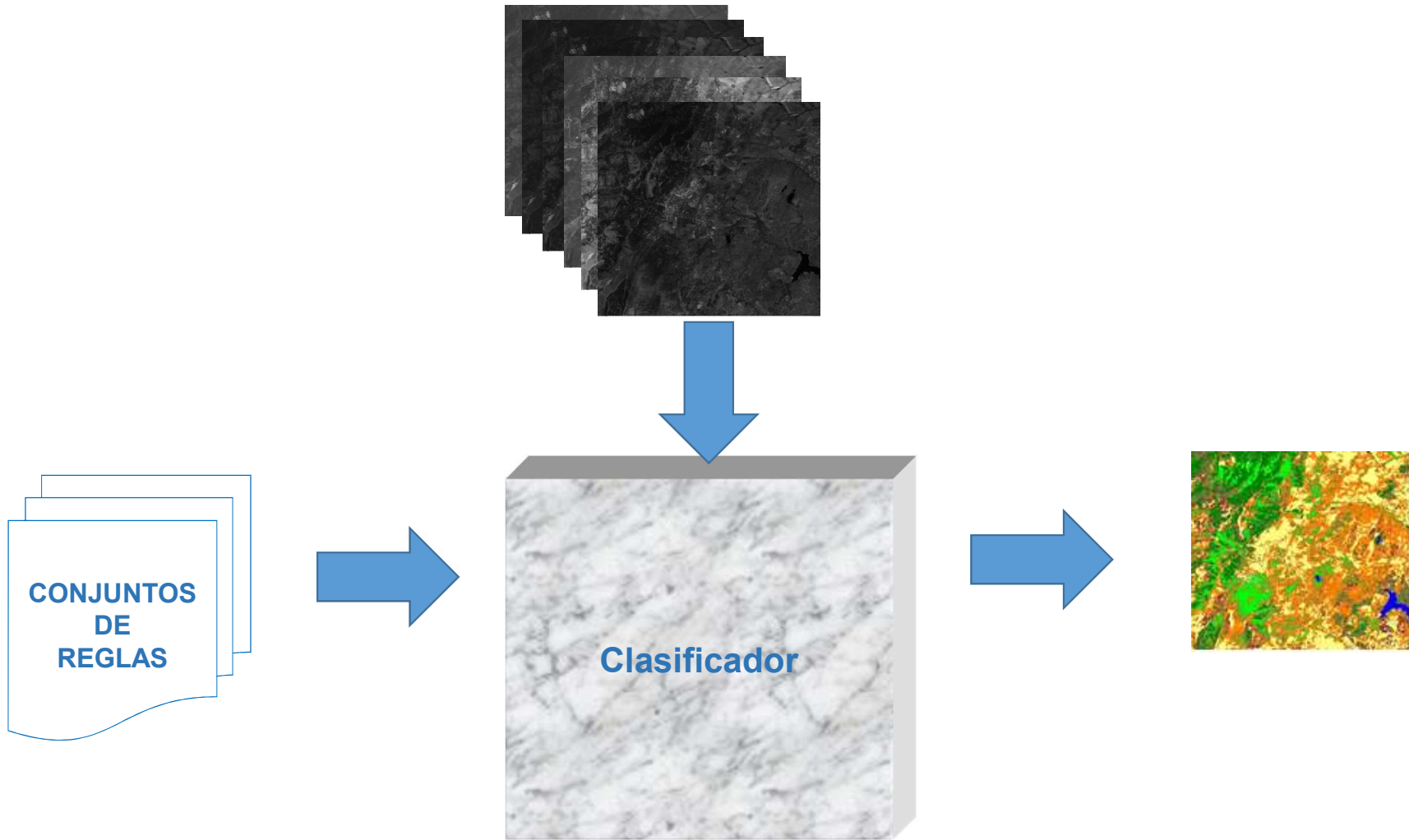
## Definición de reglas fuzzy



DEFINICION DE REGLAS
Regla 1: SI B3=pequeño , B4=pequeño ENTONCES pinar 1.0000
Regla 2: SI B3=muy grande , B4=pequeño ENTONCES suelo claro1.0000;
Regla 3: SI B3=muy grande , B4=medio ENTONCES suelo claro 1.0000;
Regla 4: SI B3=grande , B4=medio ENTONCES mixto 1.0000;
Regla 5: SI B3=medio , B4=medio ENTONCES matorral 1.0000;
Regla 6: SI B3=medio , B4=pequeño ENTONCES suelo oscuro 1.0000;
Regla 7: SI B3=grande , B4=pequeño ENTONCES suelo oscuro 1.0000;
Regla 8: SI B3=pequeño , B4=medio ENTONCES bosque 1.0000;
Regla 9: SI B3=muy pequeño , B4=muy pequeño ENTONCES agua 1.0000;
Regla 10: SI B3=pequeño , B4=muy pequeño ENTONCES agua 1.0000;
Regla 11: SI B3= muy pequeño , B4=pequeño ENTONCES pino 1.0000;
Regla 12: SI B3=medio , B4=pequeño ENTONCES matorral 1.0000;



# Clasificación



# Resultado de clasificación

