



Ministerio de Agricultura,  
Ganadería y Pesca  
Presidencia de la Nación



Universidad Nacional  
de San Luis

**CHARLA:**

# Desarrollo AGTECH con Ciencia de Datos



**DÍA DE LA CHARLA:**

6 de Diciembre de 2019

**EVENTO:** Taller Argentino de  
Computación Científica (TAC<sup>2</sup>)

**LUGAR:** Universidad Nacional  
de San Luis, San Luis



**Yanina Bellini Saibene**  
**EEA Anguil**

# Mucho gusto...

Licenciada en Sistemas de Información

Magister en Explotación de Datos y Gestión del  
conocimiento



**Investigadora** en el Instituto Nacional de  
Tecnología Agropecuaria (21 años)

Organizadora Capítulo **R-Ladies** Santa Rosa (3  
años) y miembro de **R-Ladies Global** (1 año)

Socia Activa de **SADIO** (11 años)



# Actividad agropecuaria



# Evolución de las Tecnologías

## **Agricultura tradicional - Ag 1.0**

- Autosuficiencia y métodos naturales de fertilización, malezas y control de plagas.

## **Agricultura mecanizada - Ag 2.0**

- Mecanización y fertilización química y control de malezas / plagas.

## **Agricultura de precisión - Ag 3.0**

- Gestión de recursos en respuesta a las variabilidades inter e intra-campo en los cultivos.

## **Agricultura inteligente - Ag 4.0**

- Diversas definiciones, centradas en el bienestar humano, económico y ambiental.

**Producir más y mejor con los mismos o menos recursos cuidando el medio ambiente. Mirada a toda la cadena (producción, transformación, distribución/comercialización y consumo)**

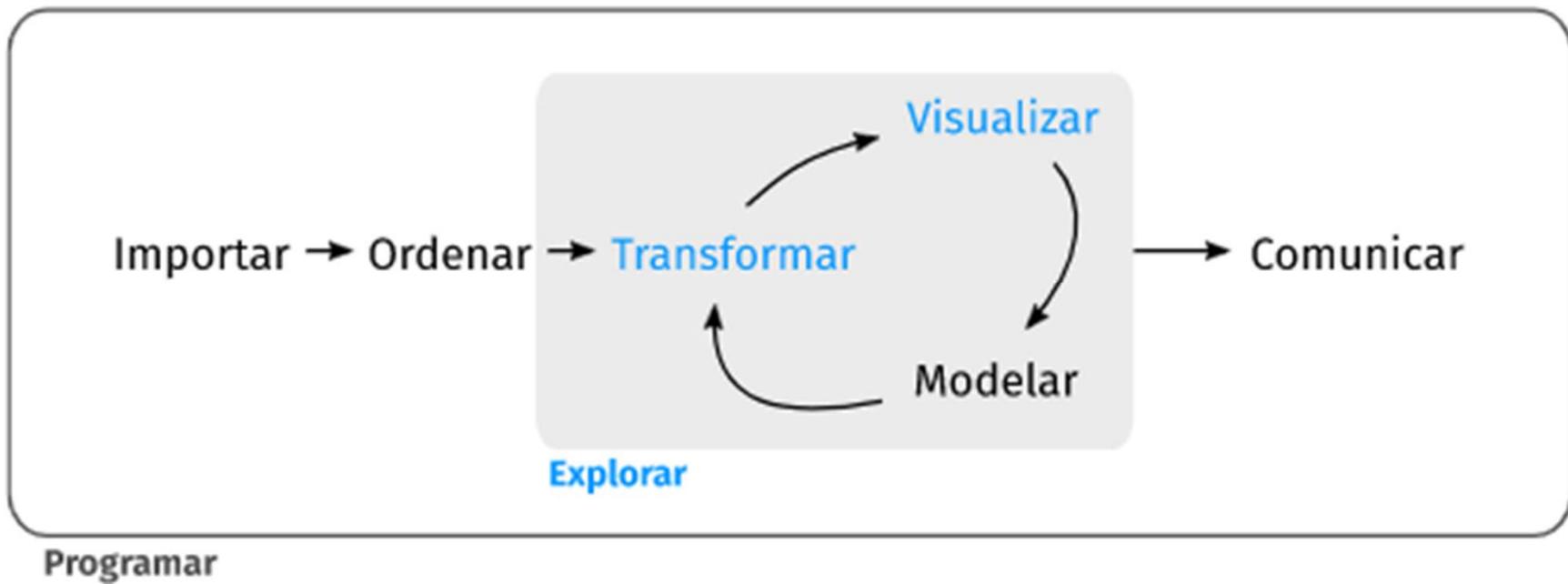
# AgTech AgroTICs

## Agricultural Technology

*Enmarca aplicaciones de **Ciencia de Datos** en el sector **Agropecuario** (y una serie de otras disciplinas relacionadas a lo digital-electrónico).*



# Ciencia de Datos



# Ciencia de Datos

## Dos “caminos”

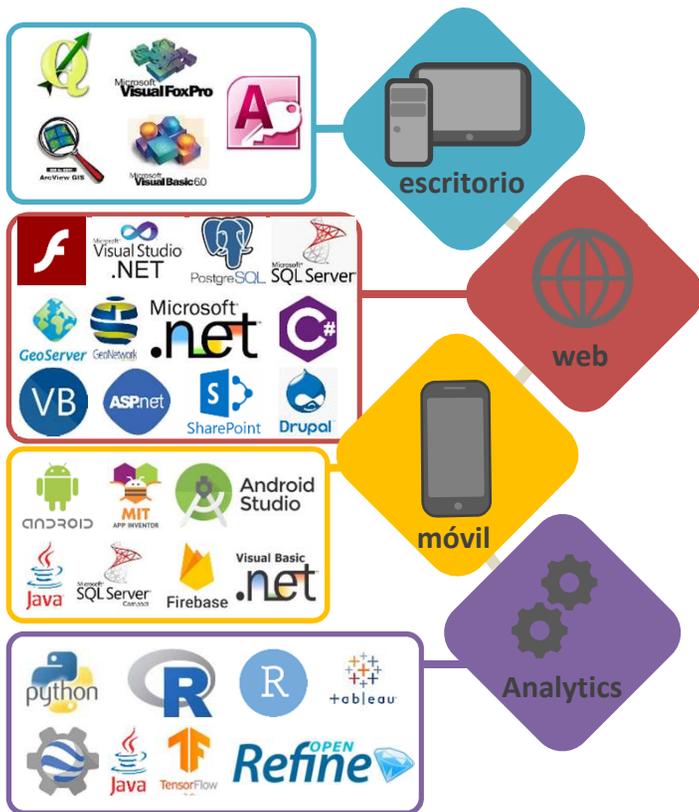
- Desarrollo de máquinas, componentes precisos, actuadores, sensores, etc.
- La explotación de datos disponibles y/o generados por estos equipos.



## Tecnología

## Audiencia

## Alcance



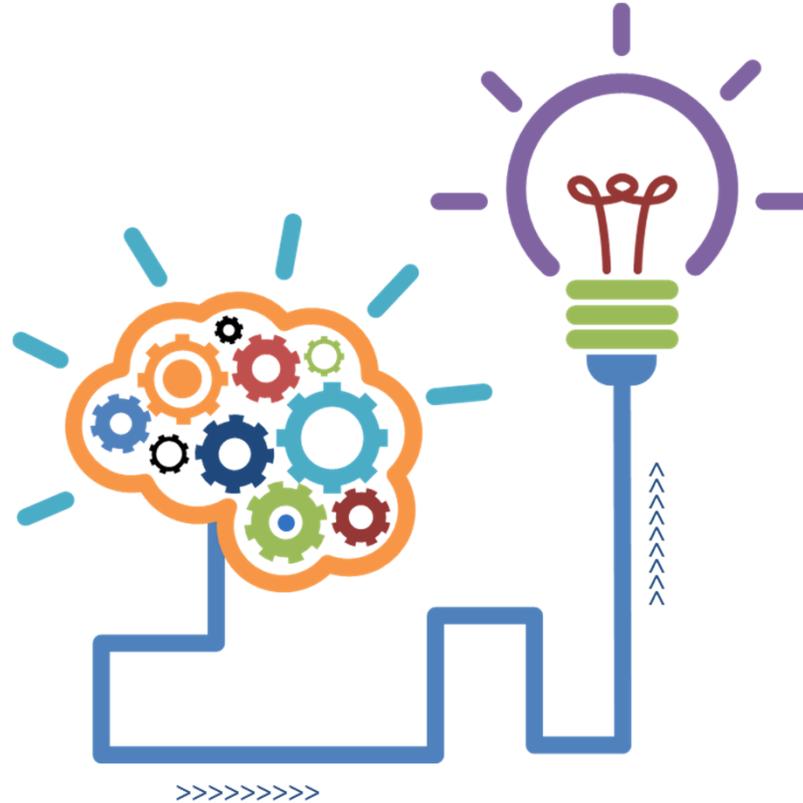
# Data Mining Machine Learning



**Aprendizaje  
Supervisado**

**Aprendizaje  
NO  
Supervisado**

**Variables de  
entrada**



**Modelo**

**Algoritmo**

# Aprendizaje Supervisado

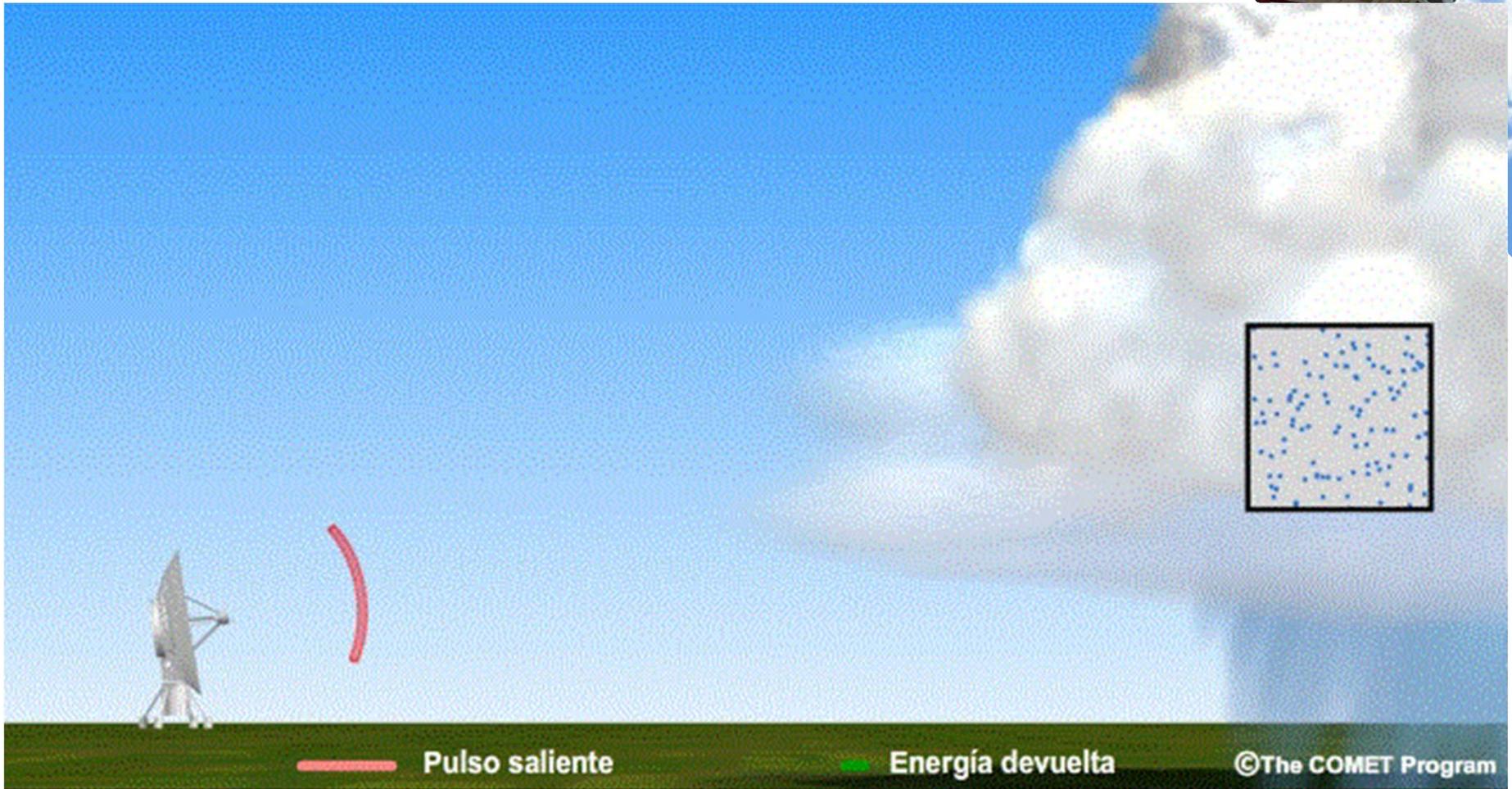


Estimación de ocurrencia  
de granizo en superficie,

mediante datos de radar meteorológico  
utilizando técnicas de **Data Mining**.

# Aprendizaje supervisado

*Modelo de granizo: resultados*

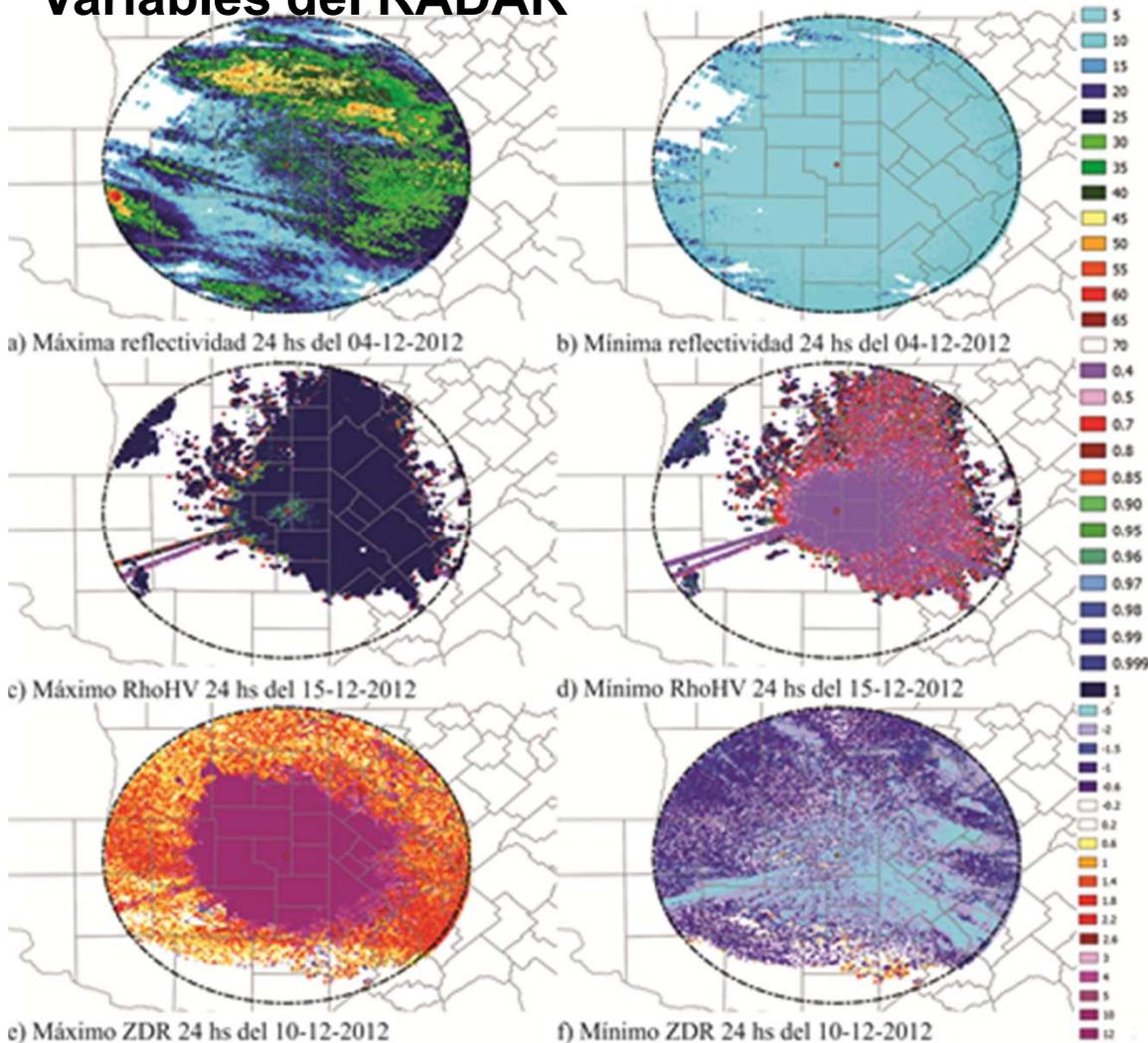


# Aprendizaje supervisado

## *Ocurrencia de granizo: variables*



### Variables del RADAR



### Variables del RADAR

5 variables

144 tomas al día

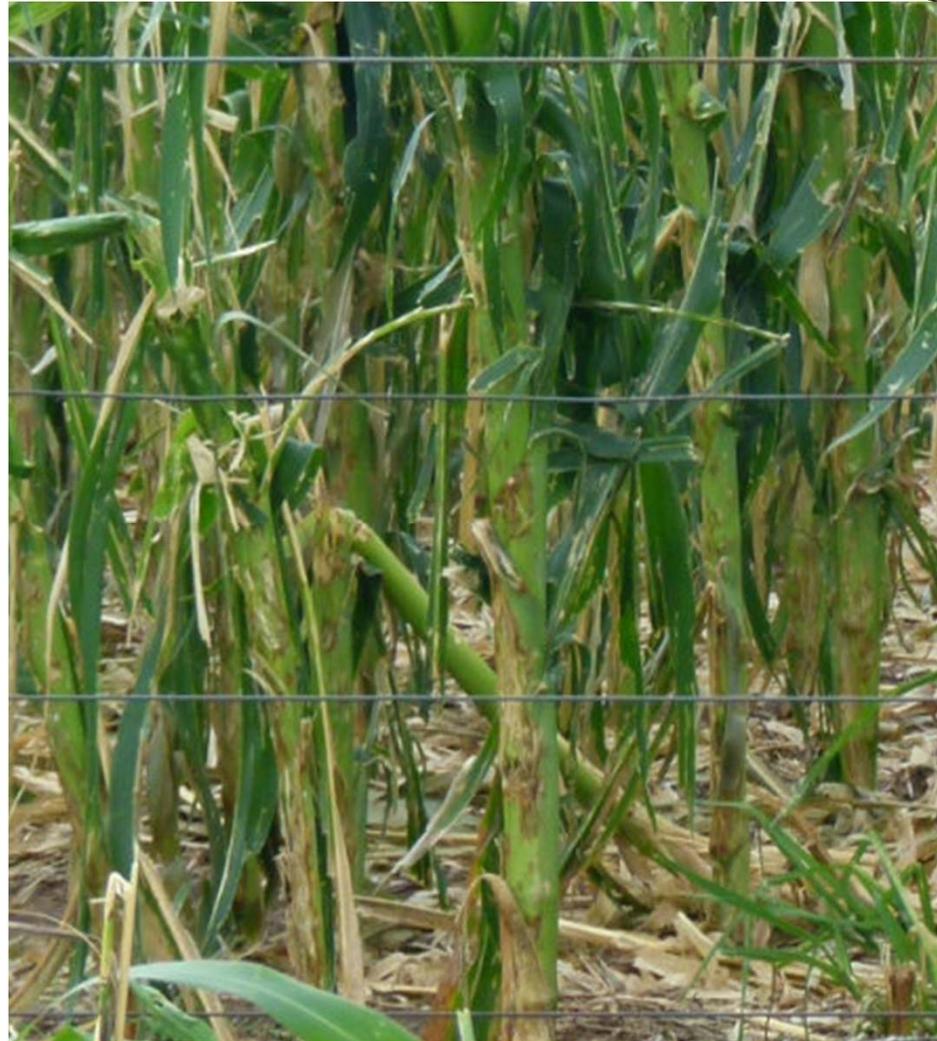
480x480 pixeles

165.888.000 datos  
diarios

# Aprendizaje supervisado

## *Ocurrencia de granizo: variables*

Lotes con y sin daño por granizo



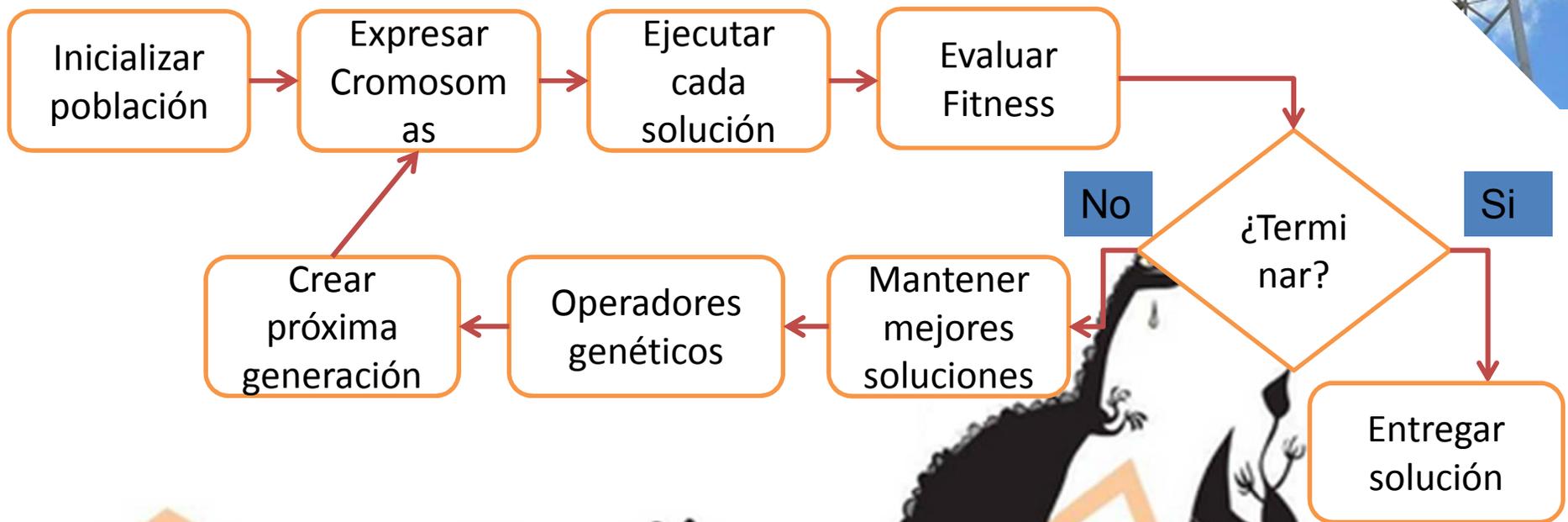
# Aprendizaje supervisado

## Ocurrencia de granizo

Metodología: Gene Expression Programing

Cromosomas

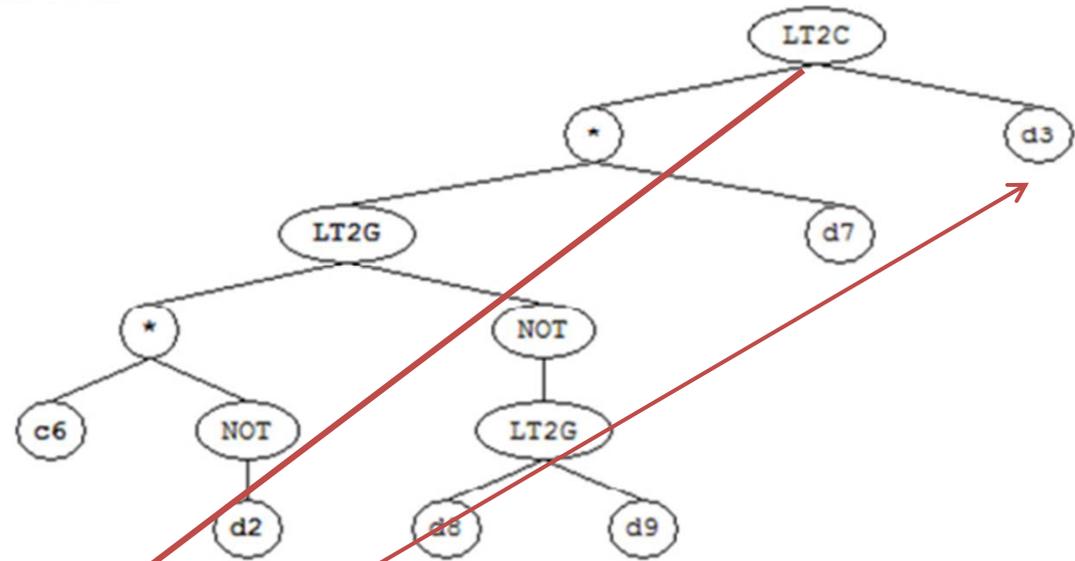
Árboles de Expresión



# Aprendizaje

## Modelo de granizo

Sub-ET 1



y = 0.0

```
y = (gepGOE2A((d[AvZDR]*gepGOE2C(G1(
y = y + (G2C3+(((gepGOE2G(gepLT2A(d
y = y + (((((gepAND1(G3C6,d[MxRho]))+((G3C0+G3C5)/2.0))/2.0)-gepGOE2G(gepAND1(d[MxZDR],d[MnZDR])).
y = y + (d[MnRho]-(gepGOE2E(G4C6,exp(d[MnRho])))*d[MnRho])))
```

SLOPE = 0.112107575724697  
INTERCEPT = -6.45746730193226

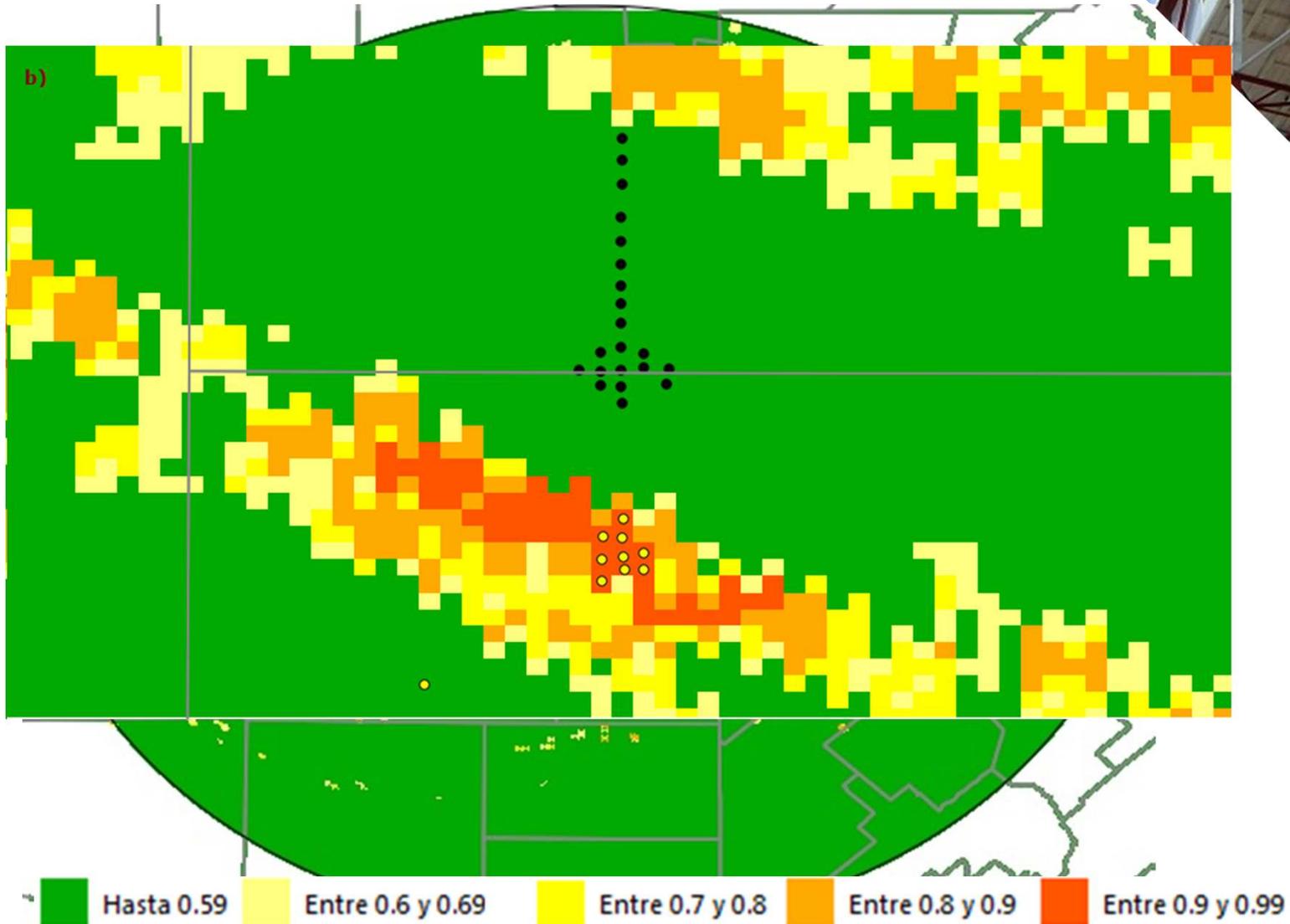
```
probabilityOne = 1.0 / (1.0 + exp(-(SLOPE * y + INTERCEPT)))
return probabilityOne
```

```
def gepLT2C(x, y):
    if (x < y):
        return (x+y)
    else:
        return (x-y)
```

**Figura 34.** Árbol de Expresión del modelo Pire (*Variables:* MxdBZ = d0, MndBZ = d1, TotdBZ = d2, AvdBZ = d3, MxRho = d4, TotRho = d6, AvRho = d7, MxZDR = d8, MnZDR = d9, TotZDR = d10, AvZDR = d11. *Constantes:* Sub-ET 1: C6 = 9.64666513565478, Sub-ET 2: C4 = -4.10822324594867, Sub-ET 4: C6 = 5.10116885891293).

# Aprendizaje supervisado

## *Modelo de granizo: resultados*



# Aprendizaje supervisado

## *Modelo de granizo y daño en cultivos*

### Herramientas



**Py-ART**

The Python ARM Radar Toolkit



### Equipo de trabajo



Romina  
MEZHER



Santiago  
BANCHERO



Laura  
BELMONTE



Juan  
CALDERA

<https://github.com/INTA-Radar>



# Aprendizaje NO Supervisado



**Clasificación de Sistemas  
Productivos Preponderantes  
utilizando técnicas de  
agrupamiento para la  
estimación de emisiones de  
Gases de Efecto Invernadero.**

# Aprendizaje NO supervisado

## Sistemas productivos preponderantes



### Insumos

#### Registro Provincial Agropecuario (REPAGRO)

**+1000** variables totales  
**7.766** casos totales para **2014**

B: USO DEL SUELO AL 30 DE JUNIO DE 2013

1) AGRICULTURA P/COSECHA

Hectáreas	
Trigo	
Avena	
Cebada	
Centeno	
Otros	
<b>Total</b>	

2) VERDEOS DE INVIERNO

Hectáreas	
Avena	
Cebada	
Centeno	
Otros	
<b>Total</b>	

3) CULTIVOS DIFERIDOS

Hectáreas	
Maíz	
Sorgo	
Otros	
<b>Total</b>	

4) RASTROJO

Hectáreas	
Rastrojo	
Barbecho	
<b>Total</b>	

5) PASTURAS PERENNES

Hectáreas	
Alfalfa pura	
Alf. Asociada	
Otros Leg. Puras	
Otros Leg. Cons.	
Pasto Llorón	
Otros Gram.Puras	
<b>Total</b>	

6) MONTE NATURAL

Hectáreas	
Caldenol	
Renoval	
Jarillal	
<b>Total</b>	

7) CAMPO NATURAL SIN MONTE

Hectáreas	
Campos	
<b>Total</b>	

8) PLANTACION ARTIFICIAL

Hectáreas	
Plantaciones	
<b>Total</b>	

9) SUPERFICIE NO UTILIZABLE

Hectáreas	
Médanos	
Inundado	
Tasca en sup	
Otros	
<b>Total</b>	

10) SUPERFICIE CON VIVIENDAS, CAMINOS, PARQUES, GALPONES Y OTROS

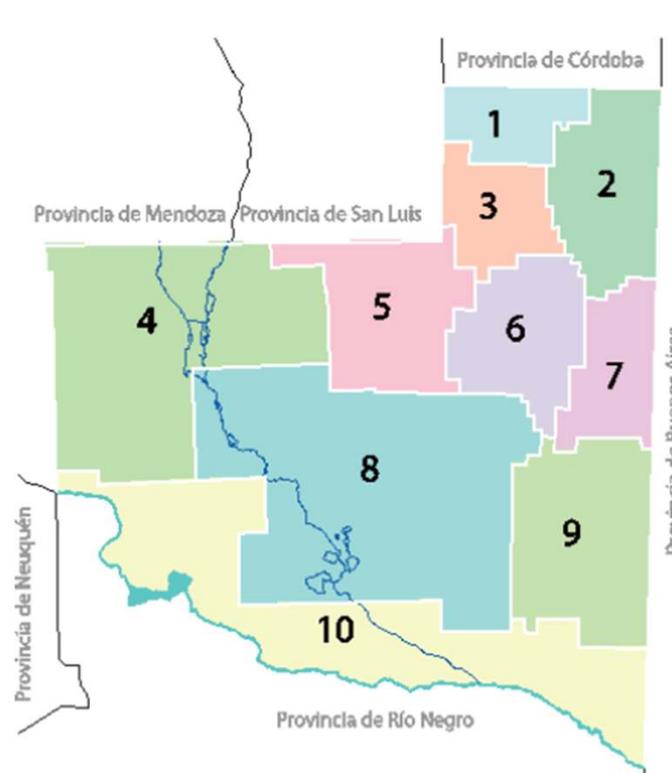
Hectáreas	
Superficie	
<b>Total</b>	

11) SUPERFICIE TOTAL

Hectáreas	
Superficie	
<b>Total</b>	

Sumatoria de los Totales de 1 a 11

#### Provincia de La Pampa



# Aprendizaje NO supervisado

*Sistemas productivos preponderantes*

## **Metodología: Clustering**

¿por qué  
atributo de  
similitud se  
podrían  
agrupar estos  
“casos”?

K-means

K-medoids



# Aprendizaje NO supervisado

*Sistemas productivos preponderantes*

*Resultados*

13 grupos

Indice Jaccard:  $> 0.80$  😊

Silueta: entre 0.1 😞 y 0.57 😊

Cantidad de casos: entre 64 y 1207 🤔



*La pregunta del millón:  
¿cómo sabemos si esos grupos  
representan grupos reales?*

# **Aprendizaje NO supervisado**

*Sistemas productivos preponderantes*



***1. Miramos los grupos con expertos del negocio***

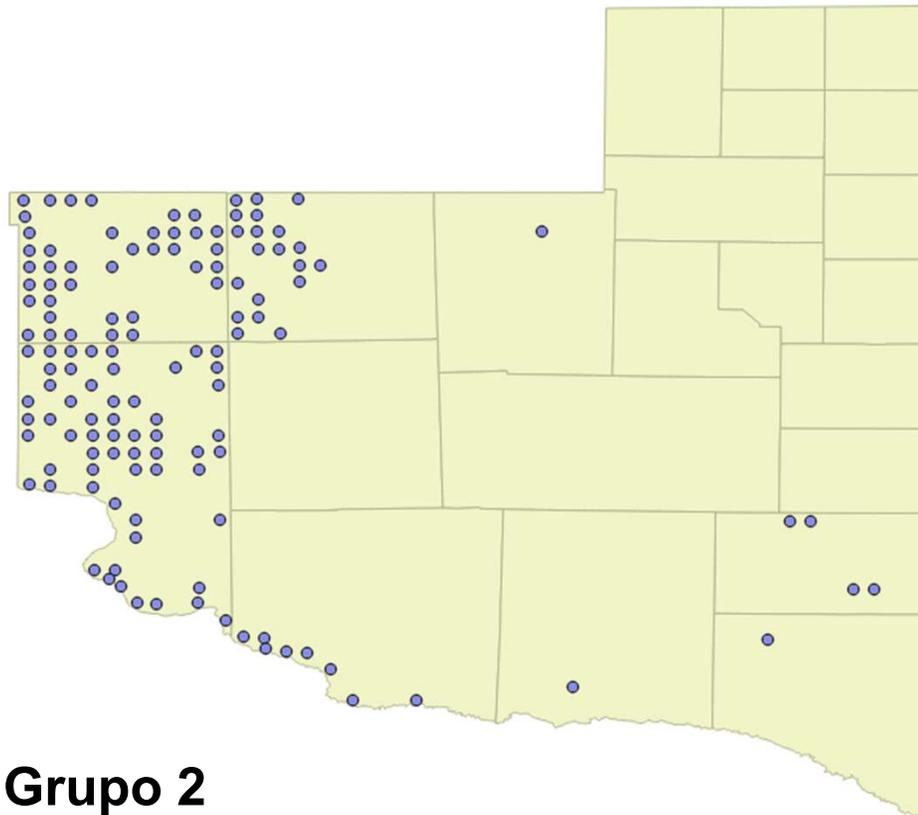
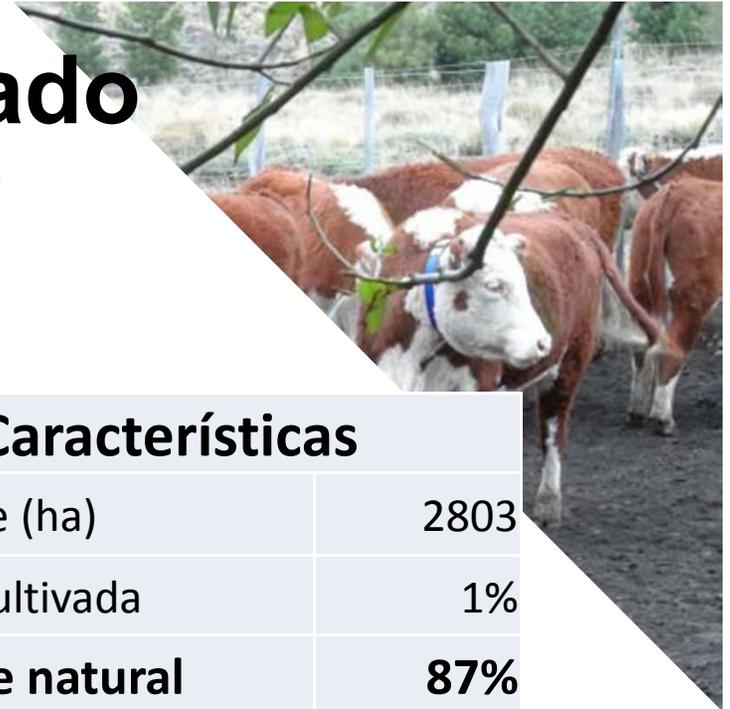
***2. Buscamos una coherencia de:***

- ***Tamaños***
- ***Actividades***
- ***Distribución espacial***

# Aprendizaje NO supervisado

*Sistemas productivos preponderantes*

*Algunos resultados:*



**Grupo 2**

**N 213**

**Silueta 0.31**

**Jaccard 0.91**

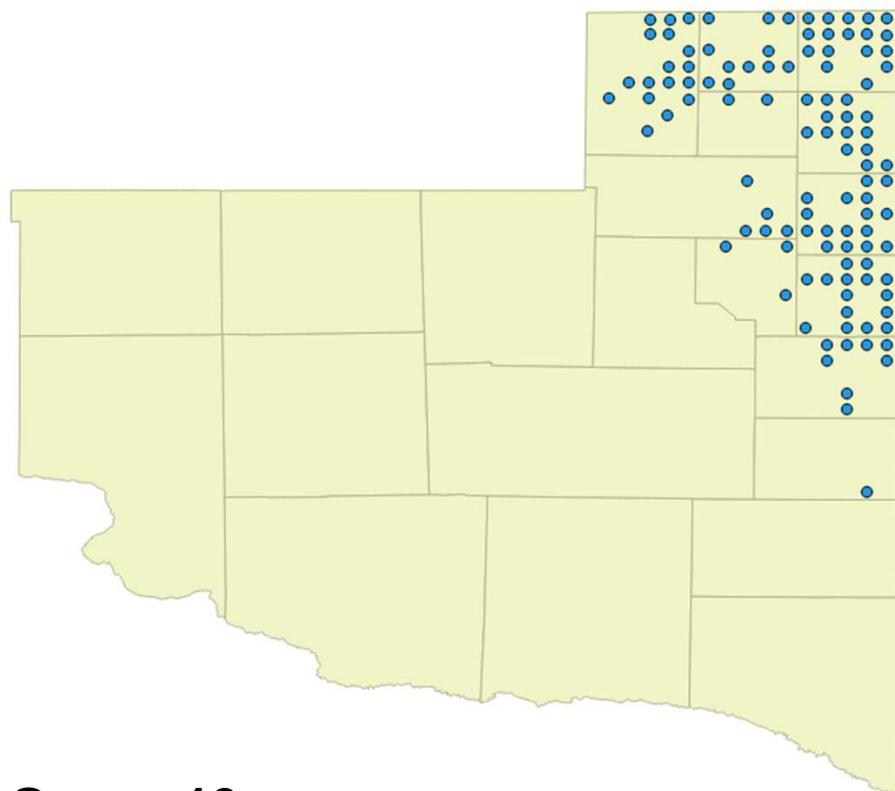
Características	
Superficie (ha)	2803
% Area Cultivada	1%
<b>% Monte natural</b>	<b>87%</b>
<b>Ganadería</b>	
% Bovinos	26%
% Ovinos	6%
% Porcinos	1%
<b>% Caprinos</b>	<b>56%</b>
% Equinos	11%

***Sistema cría caprina y bovina***

# Aprendizaje NO supervisado

*Sistemas productivos preponderantes*

*Algunos resultados:*



**Grupo 13**

**N 175**

**Silueta 0.52**

**Jaccard 0.92**

<b>Características</b>	
Superficie (ha)	403
<b>Agricultura</b>	
% Cosecha Fina	5%
<b>% Cosecha Gruesa</b>	<b>87%</b>
% Forrajeras Anuales	4%
% Forrajeras Perenne	1%
<b>Ganadería</b>	
% Bovinos	2%
% Otras producciones	4%

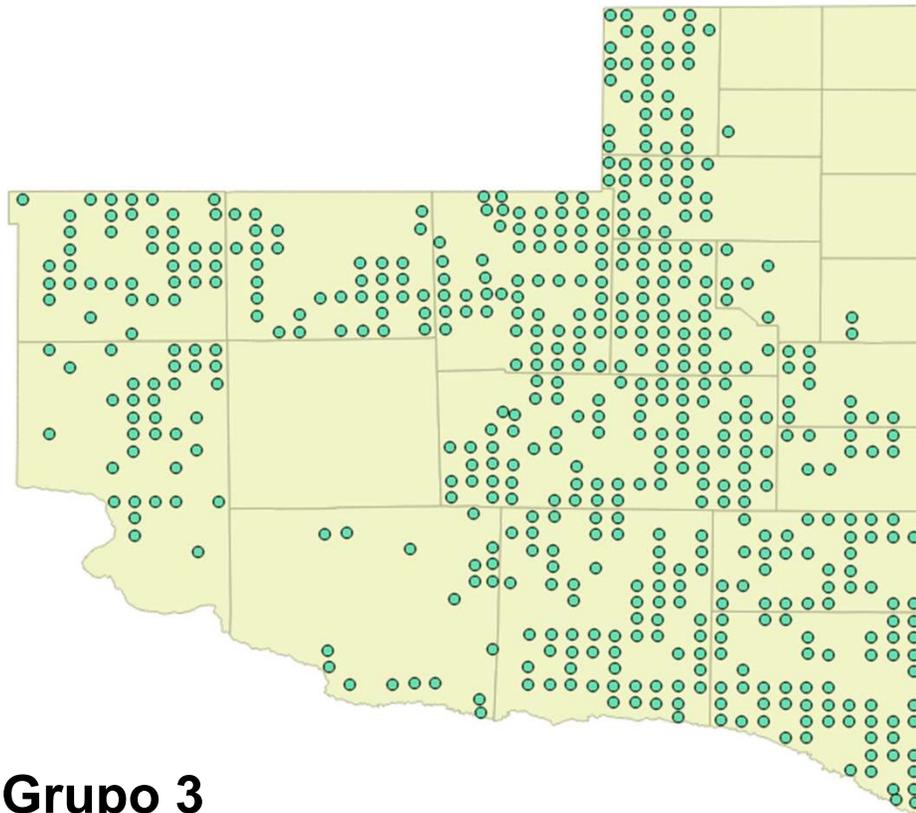
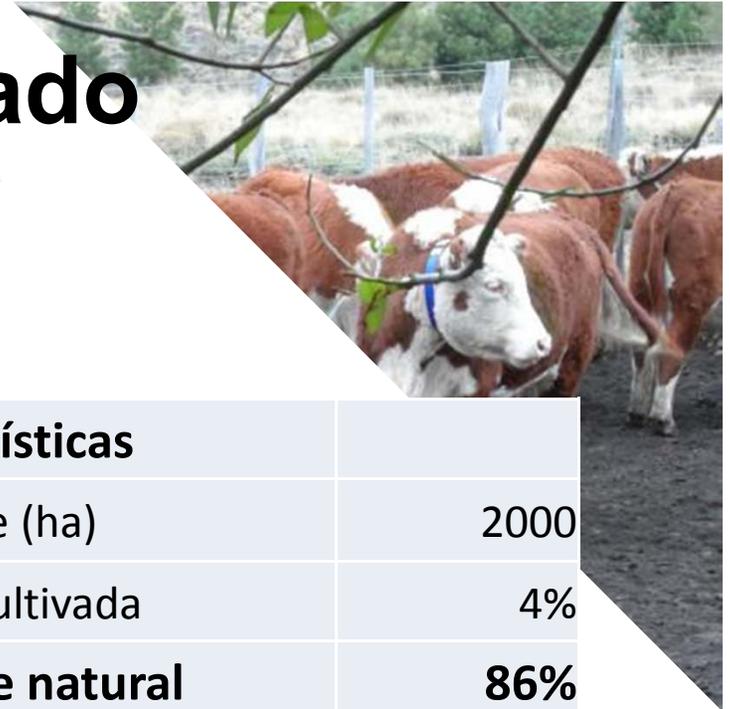


***Sistema agricultura de verano***

# Aprendizaje NO supervisado

*Sistemas productivos preponderantes*

*Algunos resultados:*



**Grupo 3**

**N 1065**

**Silueta 0.36**

**Jaccard 0.95**

<b>Características</b>	
Superficie (ha)	2000
% Area Cultivada	4%
<b>% Monte natural</b>	<b>86%</b>
<b>Ganadería</b>	
<b>% Bovinos</b>	<b>93%</b>
% Ovinos	3%
% Porcinos	0%
% Caprinos	1%
% Equinos	3%

***Sistema Cría Bovina***

# Aprendizaje NO supervisado

## *Sistemas productivos preponderantes*



### Herramientas



**fpc** Flexible Procedures for Clustering

by ORPHANED

[View Source](#)

### Equipo de trabajo



**Anabella  
LOZZA**

**Héctor  
LORDA**



# Aprendizaje Supervisado

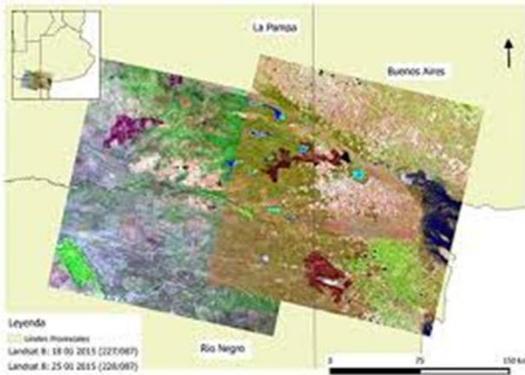
Geotecnologías en la  
nube y su aplicación  
en emergencias  
agropecuarias



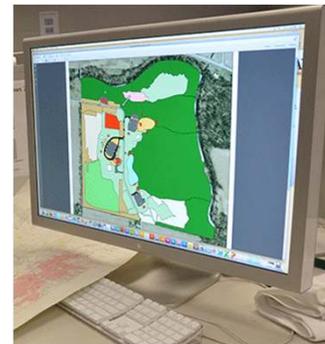
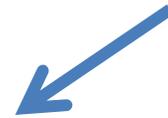
# Aprendizaje supervisado

## Teledetección

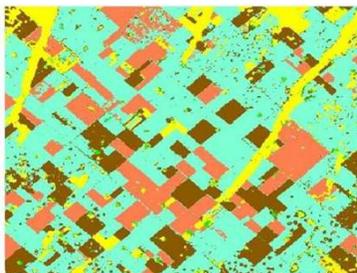
Un poco de historia: 1998



COMISIÓN NACIONAL DE  
ACTIVIDADES ESPACIALES



\$\$



# Aprendizaje supervisado

## *Teledetección*

Obteniendo información desde datos remotos: Anguil



1984

1997

2000

2002

# Aprendizaje supervisado

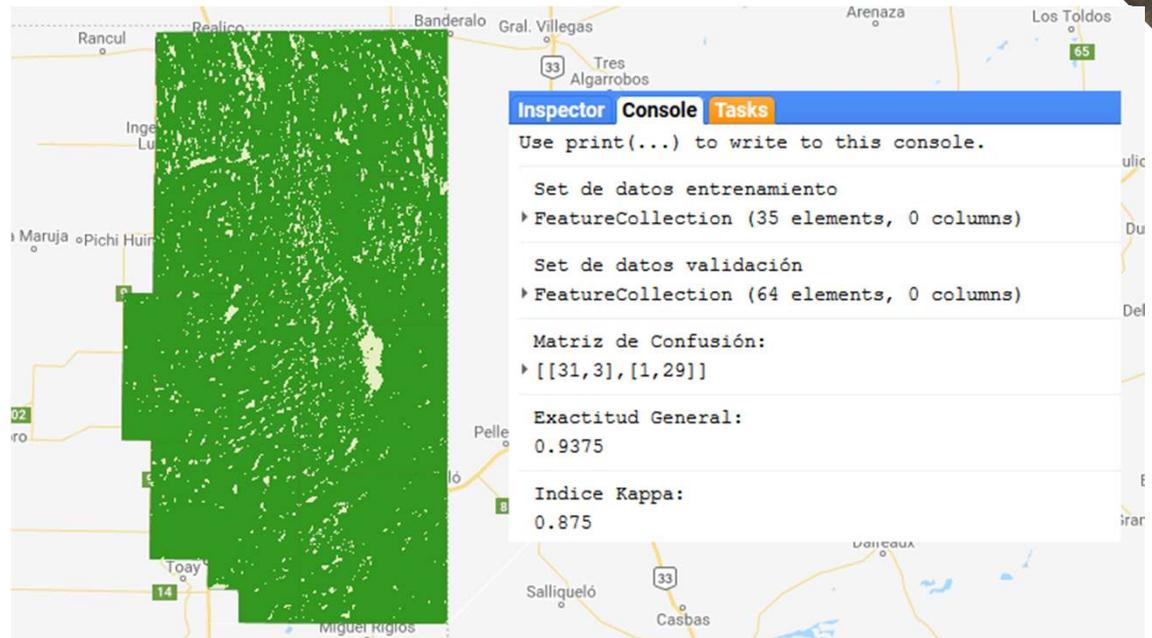
## Teledetección



Obtener una imagen



Filtro una imagen



Aplico un algoritmo de clasificacion

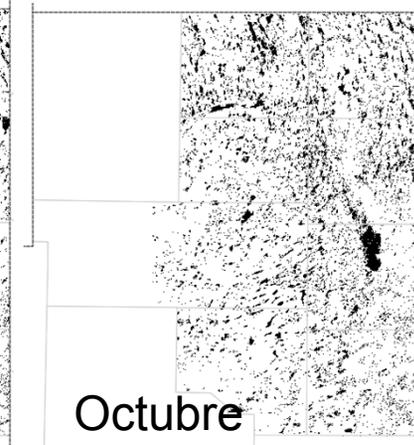
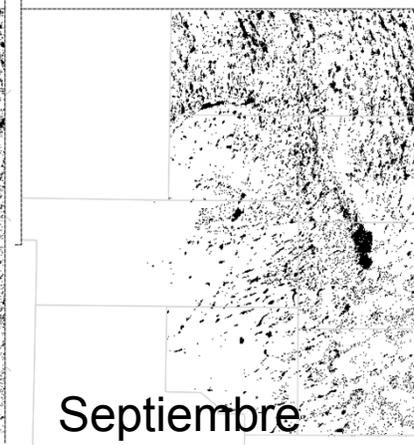
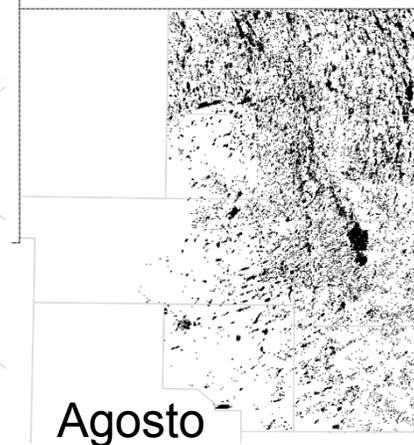
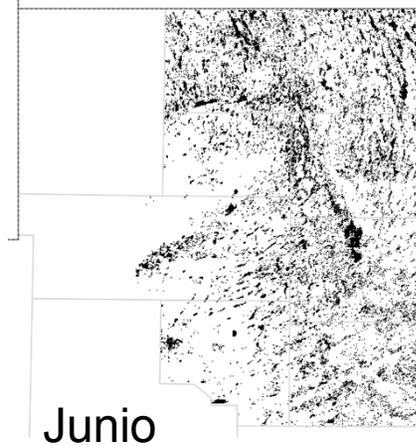
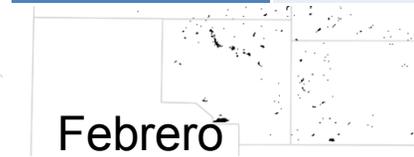
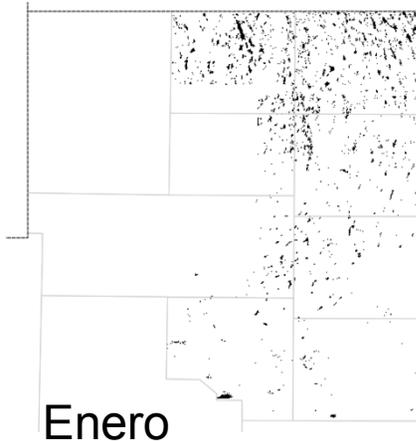
# Aprendizaje supervisado

## Teledetección



2017

Mes	Exactitud	Índice Kappa	Sup. afectada	Sup. analizada	%	Obs.
Enero	0,98	0,95	46.011	1.498.237	3	Nubes y humo
Agosto	0,99	0,97	300.337	2.219.654	14	



# Aprendizaje supervisado

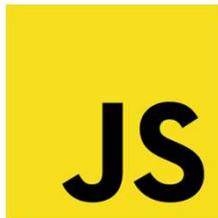
## Teledetección



### Herramientas



Google Earth Engine



### Equipo de trabajo



Julio  
FERNANDEZ



Laura  
BELMONTE



Mariela  
FUENTES



# Aprendizaje Supervisado

Drones, cebadas,  
girasoles e índices de  
vegetación



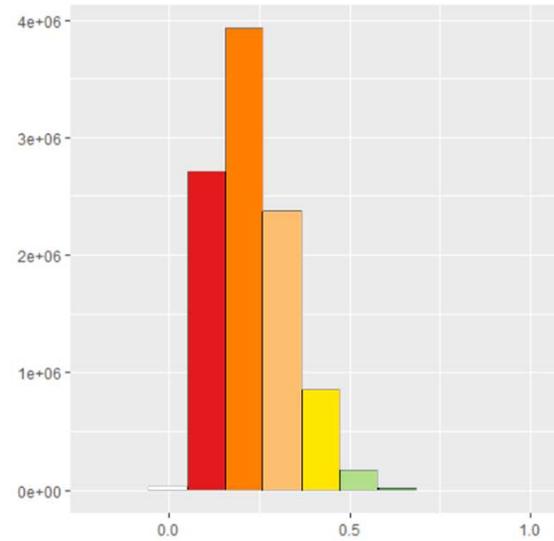
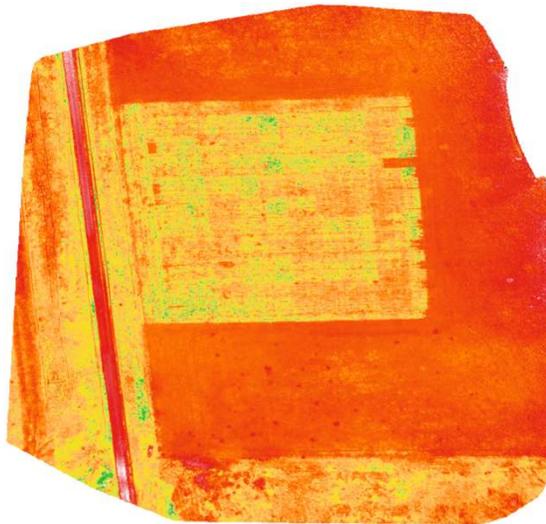
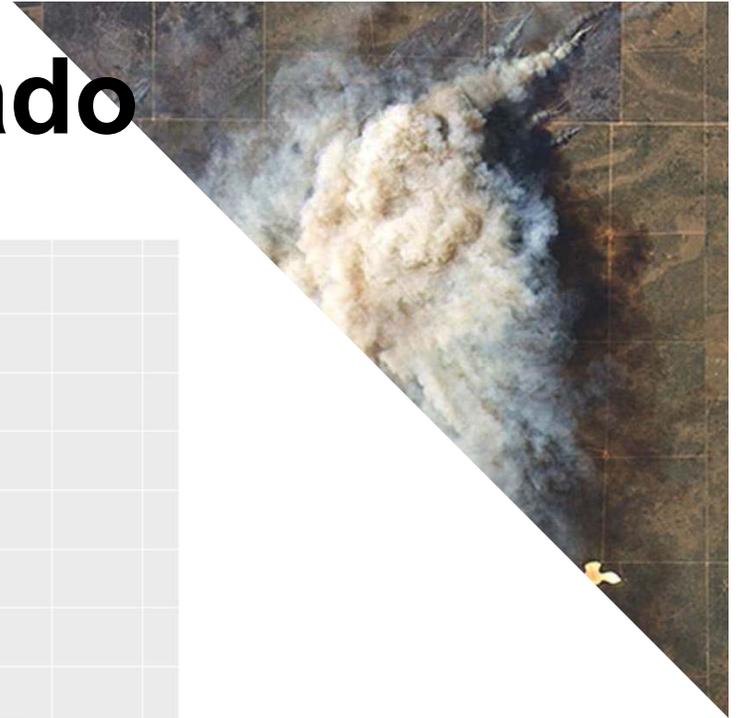
# Aprendizaje supervisado

## *Teledetección*

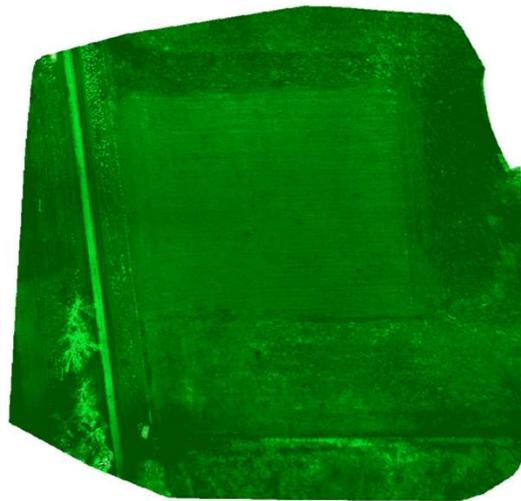


# Aprendizaje supervisado

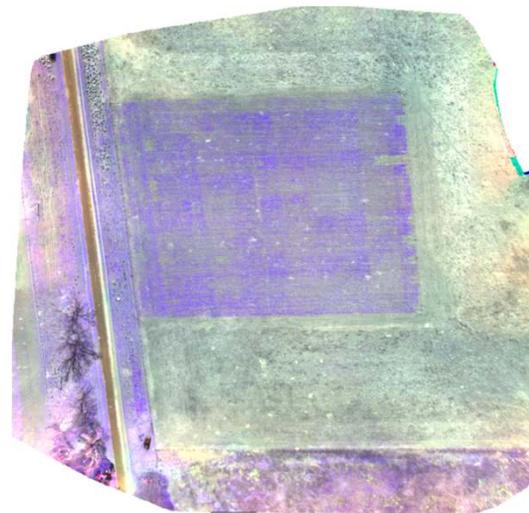
## *Lote con Cebada*



NDVI



Indice de Brillo



Falso Color

# Aprendizaje supervisado

## *Lote con Cebada*



# Aprendizaje supervisado

## *Lote con Cebada*



# Aprendizaje supervisado

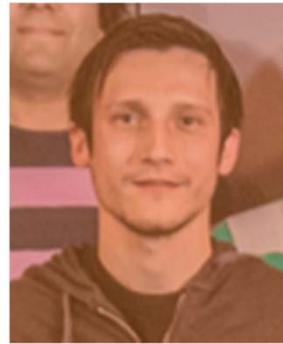
## Teledetección



### Herramientas



### Equipo de trabajo



Matías  
CAVILLA



Alexandra  
DILLCHNER



Priscilla  
MINOTTI



# Comunidades de Práctica

Grupo amplio de personas que comparten saberes y aprendizajes de un tema transversal a distintas profesiones



# ¿Qué comunidades hay?



THE  
CARPENTRIES



# ¿Qué hacemos?

 **Natalia da Silva** @pacocuak · 27 sept. ▼

Fin de #LatinR2019 @rivaquioga @\_lacion\_ @cantoflor\_87 @yabellini  
tremendas generadoras e impulsoras de cosas geniales en la comunidad R  
👏👏👏👏 @LatinR\_Conf @R4DS\_es @RLadiesValpo @RLadiesSantiago  
@RLadiesBA @RLadiesSR  
Nos vemos la próxima seguro en Montevideo #LatinR2020!



4 9 61

**LatinR 2020, Octubre  
Montevideo, Uruguay**



**latin-r.com**

# ¿Qué hacemos?



**Mine CetinkayaRundel**  
@minebocek

R4DS translation team, what a fantastic effort!  
[@LatinR\\_Conf](#) [#rstats](#)



[es.r4ds.hadley.nz](http://es.r4ds.hadley.nz)



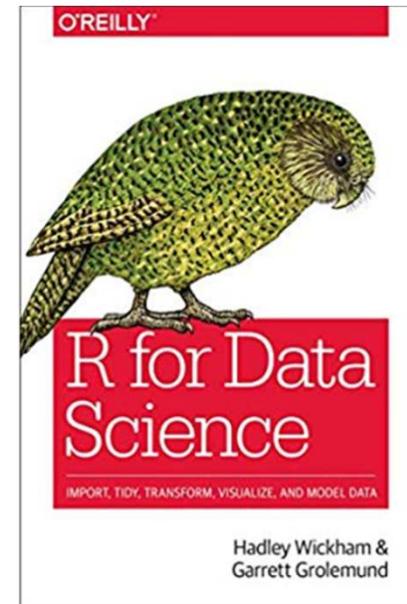
[github.com/cienciadedatos...](https://github.com/cienciadedatos...)

Traducir Tweet



5:03 p. m. · 27 sept. 2019 · Twitter for iPhone

32 Retweets 153 Me gusta



R para Ciencia de Datos

[es.r4ds.hadley.nz](http://es.r4ds.hadley.nz)



# ¿Qué hacemos?



## R-Ladies Argentina



R-Ladies Bariloche

R-Ladies Buenos Aires

R-Ladies Córdoba

R-Ladies La Plata

R-Ladies Mendoza

R-Ladies Resistencia

R-Ladies Santa Fe

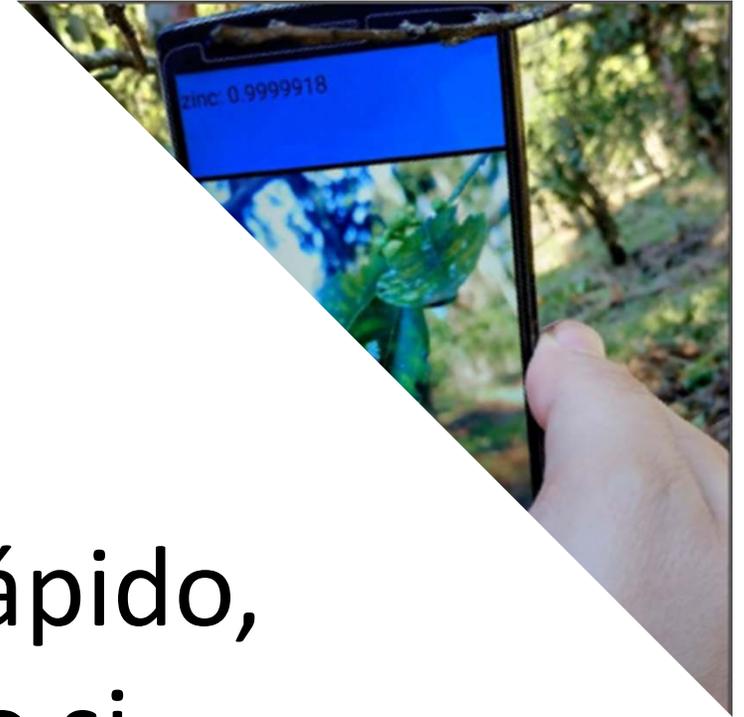
R-Ladies Santa Rosa

R-Ladies Ushuaia



# Comunidades

Si quieres llegar rápido,  
viajá solo, pero si  
**querés llegar lejos**  
**viajá acompañado**



Proverbio africano

# ¿Qué preguntas tenés?

***@yabellini***  
***yabellini.netlify.com***

