



Ciencia de Datos aplicada al desarrollo de AgTech

R-Ladies La Paz – Junio 2020

Mucho gusto...

Licenciada en Sistemas de Información

Magister en Explotación de Datos y Gestión del
conocimiento

Rstudio Certified Instructor



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

Organizadora Capítulo **R-Ladies** Santa Rosa (3
años) y miembro de **R-Ladies Global** (2 años)

Fundadora de **MetaDocencia** (0.5 años)





Santa Rosa, La Pampa, Argentina

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)

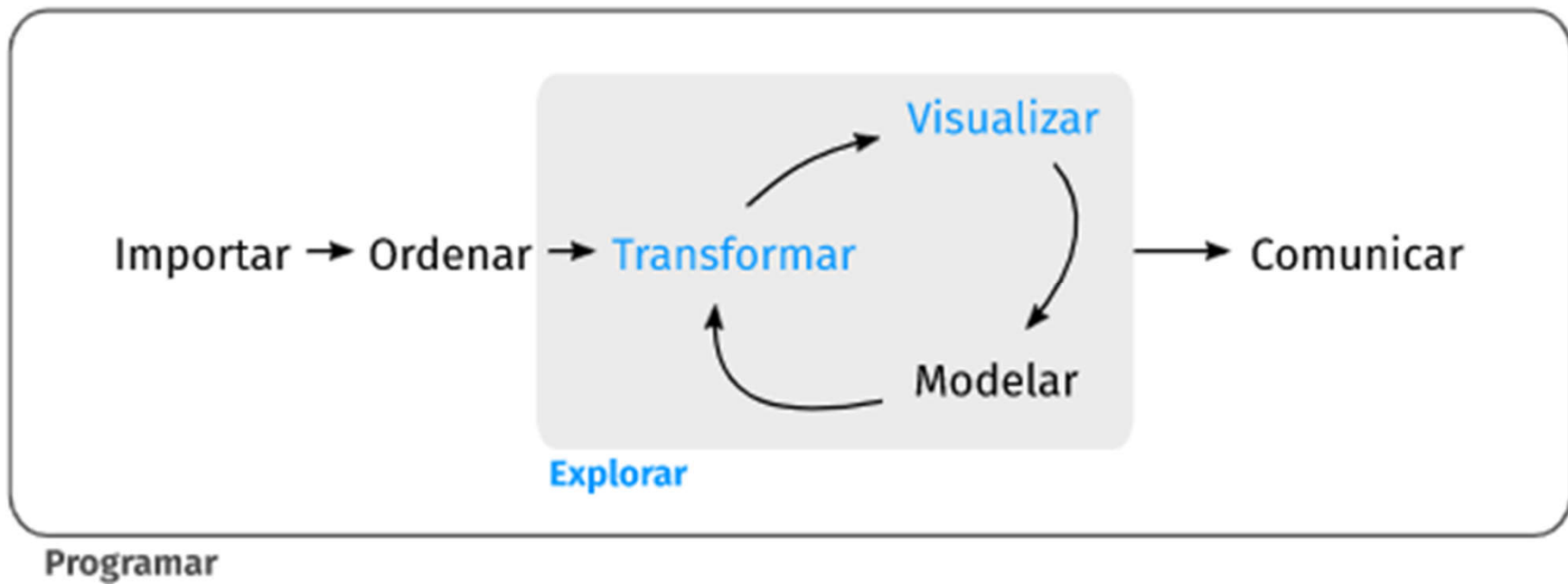
AgTech AgroTICs

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 y AgTech

Dos “caminos”

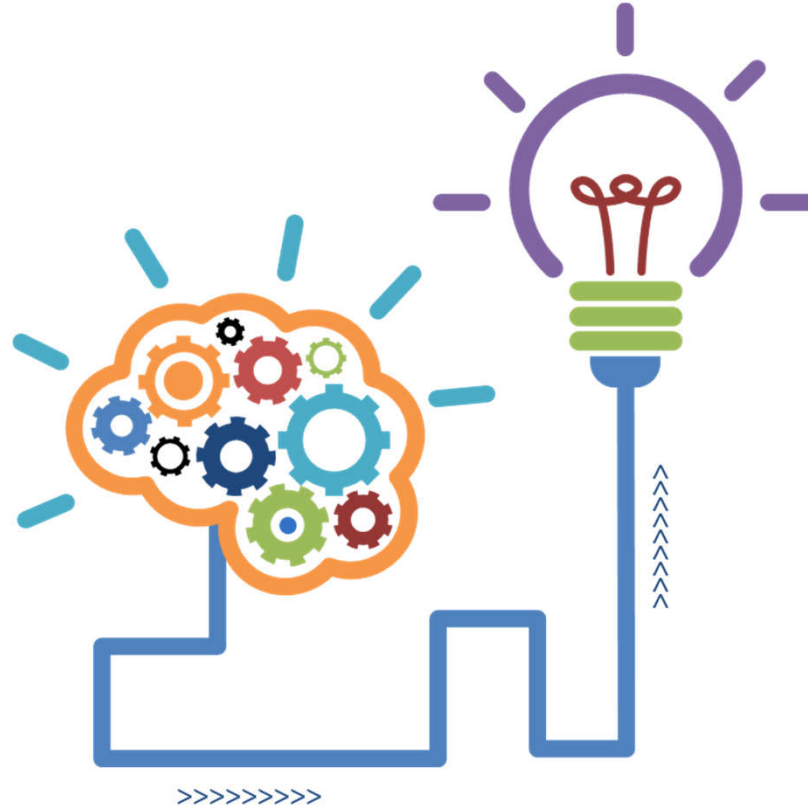
- Desarrollo de máquinas, componentes precisos, actuadores, sensores, etc.
- La explotación de datos disponibles y/o generados por estos equipos.

Data Mining Machine Learning

**Aprendizaje
Supervisado**

**Aprendizaje
NO
Supervisado**

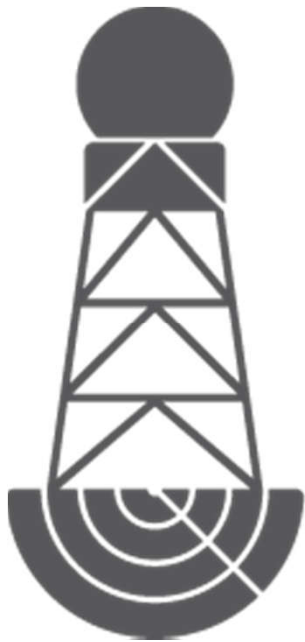
**Variables de
entrada**



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

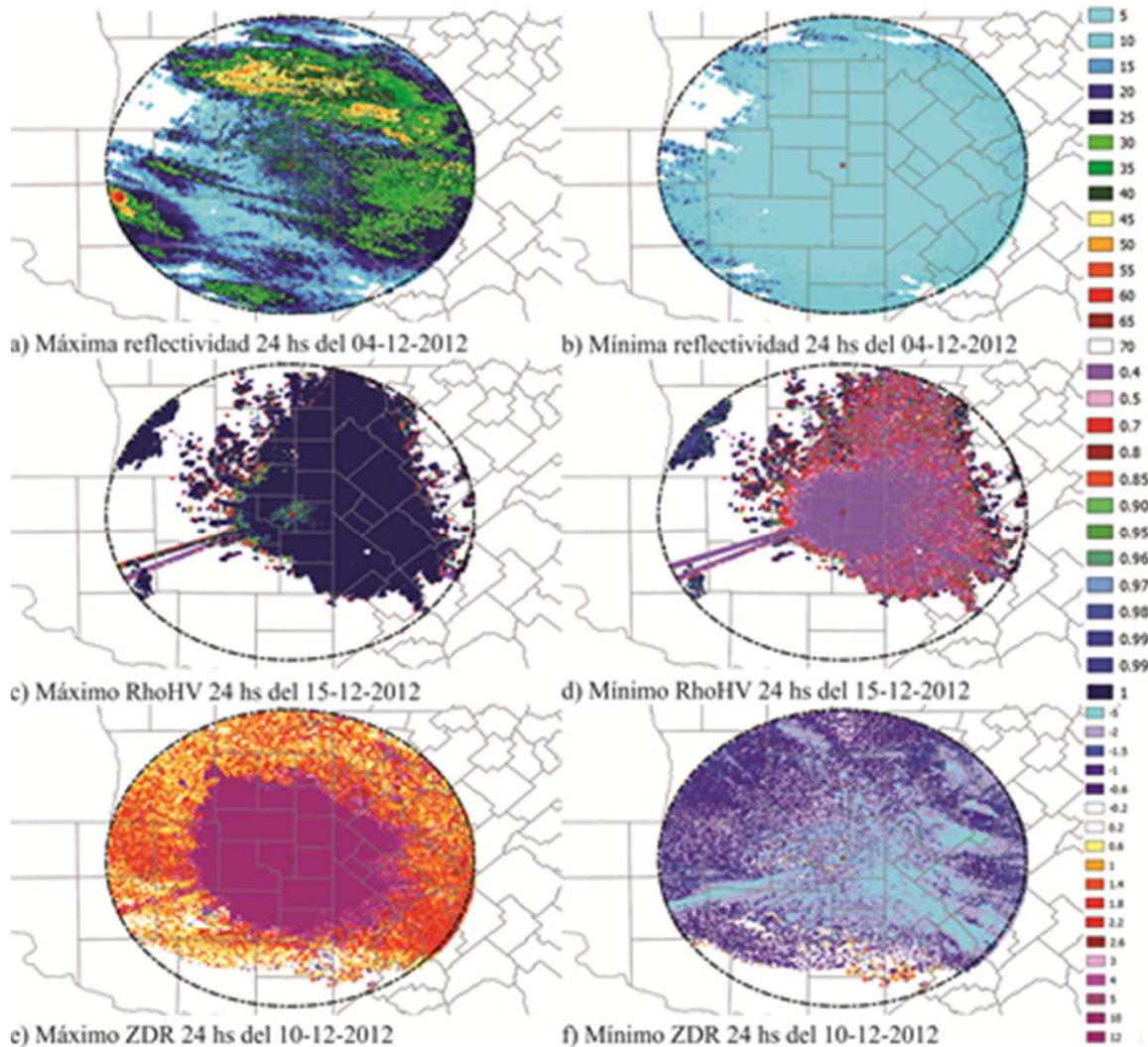


Aprendizaje supervisado

Ocurrencia de granizo: variables



Variables del RADAR



Variables del RADAR

5 variables

144 tomas al día

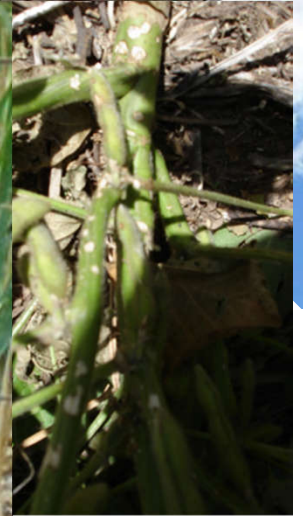
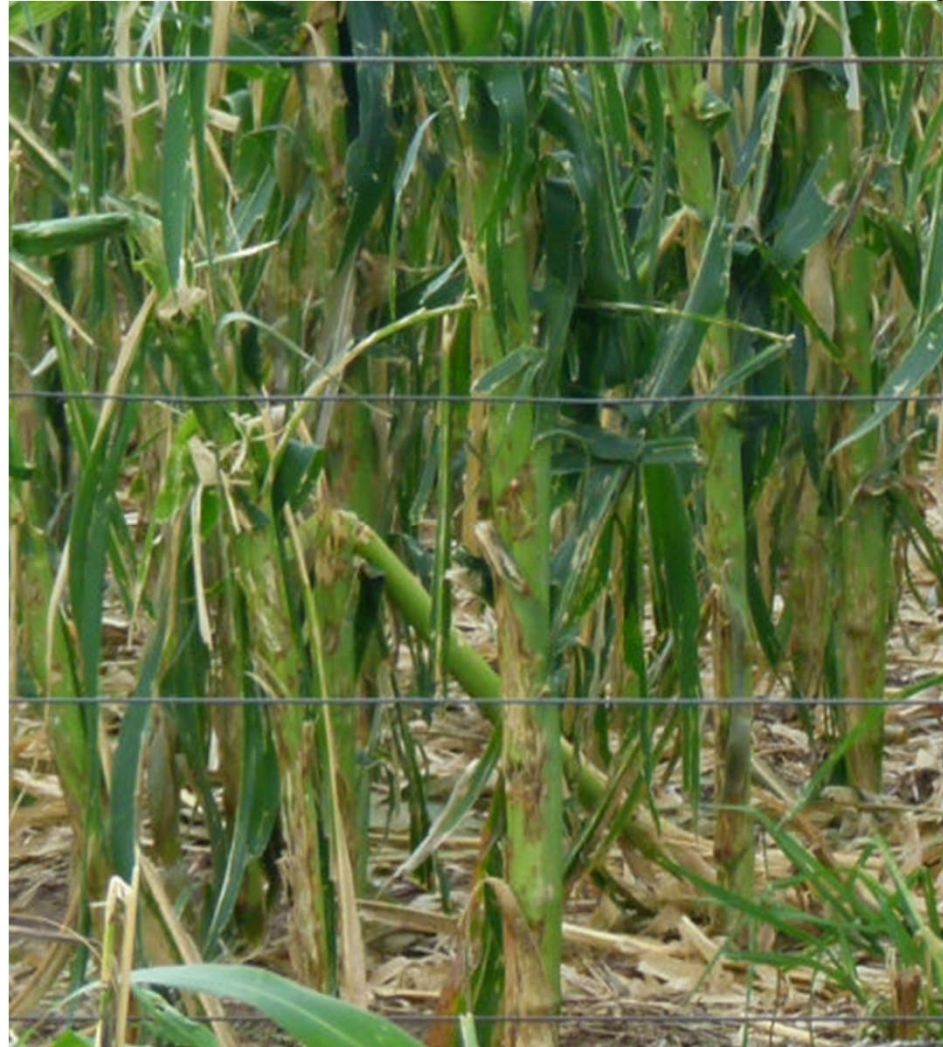
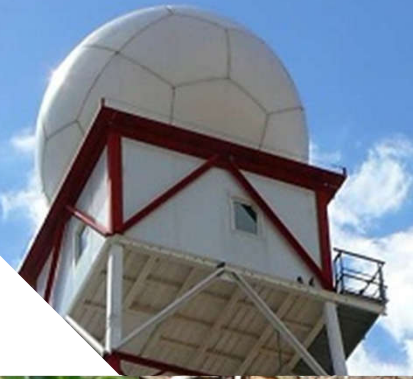
480x480 pixeles

165.888.000 datos
diarios

Aprendizaje supervisado

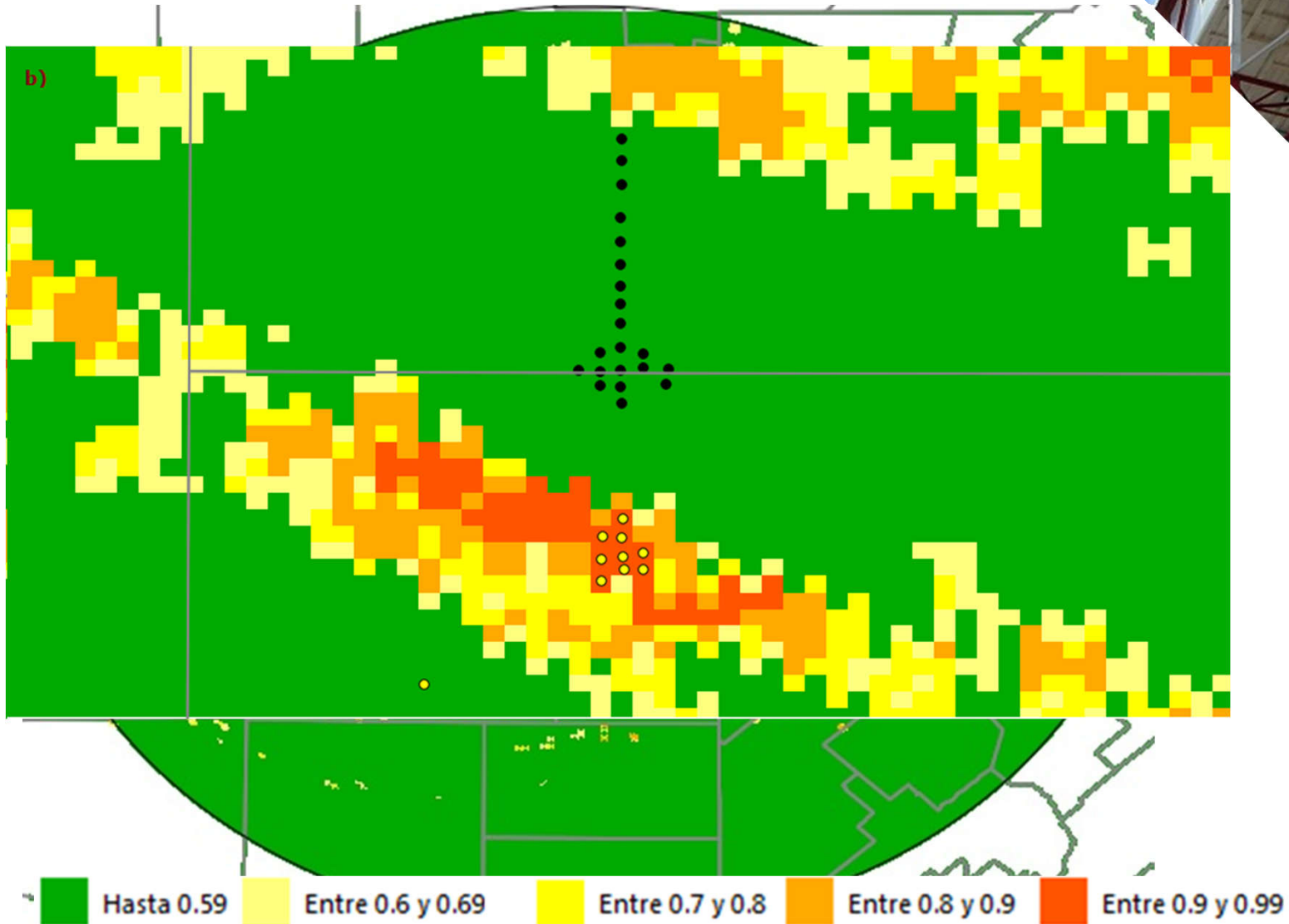
Ocurrencia de granizo: variables

Lotes con y sin daño por granizo



Aprendizaje supervisado

Modelo de granizo: resultados



Aprendizaje supervisado

Modelo de granizo y daño en cultivos



Herramientas



The Python ARM Radar Toolkit



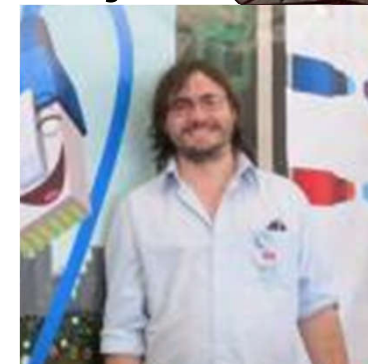
Equipo de trabajo



Juan Pablo
FERRANDEZ



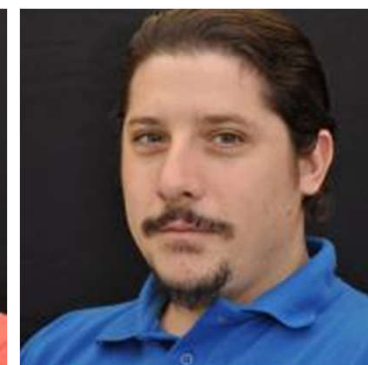
Romina
MEZHER



Santiago
BANCHERO



Laura
BELMONTE



Juan
CALDERA

<https://juanferr.shinyapps.io/raster/>
<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	
Total	

2) VERDEOS DE INVIERNO

Hectáreas	
Avena	
Cebada	
Centeno	
Otros	
Total	

3) CULTIVOS DIFERIDOS

Hectáreas	
Maíz	
Sorgo	
Otros	
Total	

4) RASTROJO

Hectáreas	
Rastrojo	
Barbecho	
Total	

5) PASTURAS PERENNES

Hectáreas	
Alfalfa pura	
Alf. Asociada	
Otros Leg. Puras	
Otros Leg. Cons.	
Pasto Llorón	
Otros Gram.Puras	
Total	

6) MONTE NATURAL

Hectáreas	
Caldenol	
Renoval	
Jarillal	
Total	

7) CAMPO NATURAL SIN MONTE

Hectáreas	
Campos	
Total	

8) PLANTACION ARTIFICIAL

Hectáreas	
Plantaciones	
Total	

9) SUPERFICIE NO UTILIZABLE

Hectáreas	
Médanos	
Inundados	
Tasca en sup	
Otros	
Total	

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

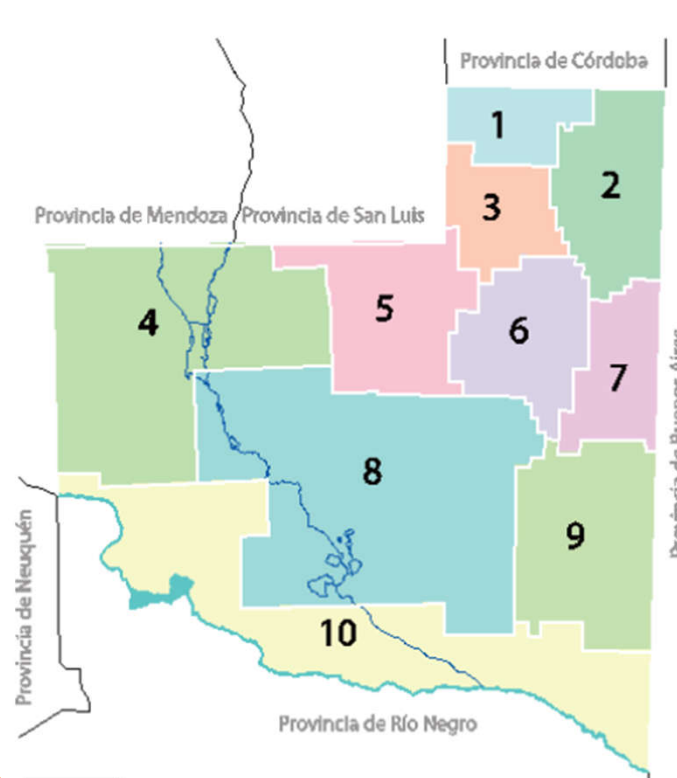
Hectáreas	
Superficie	
Total	

11) SUPERFICIE TOTAL

Hectáreas	
Superficie	
Total	

Sumatoria de los Totales de 1 a 11

Provincia de La Pampa



Aprendizaje NO supervisado

Sistemas productivos preponderantes



Algoritmo: 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 😊

Siluetas: 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

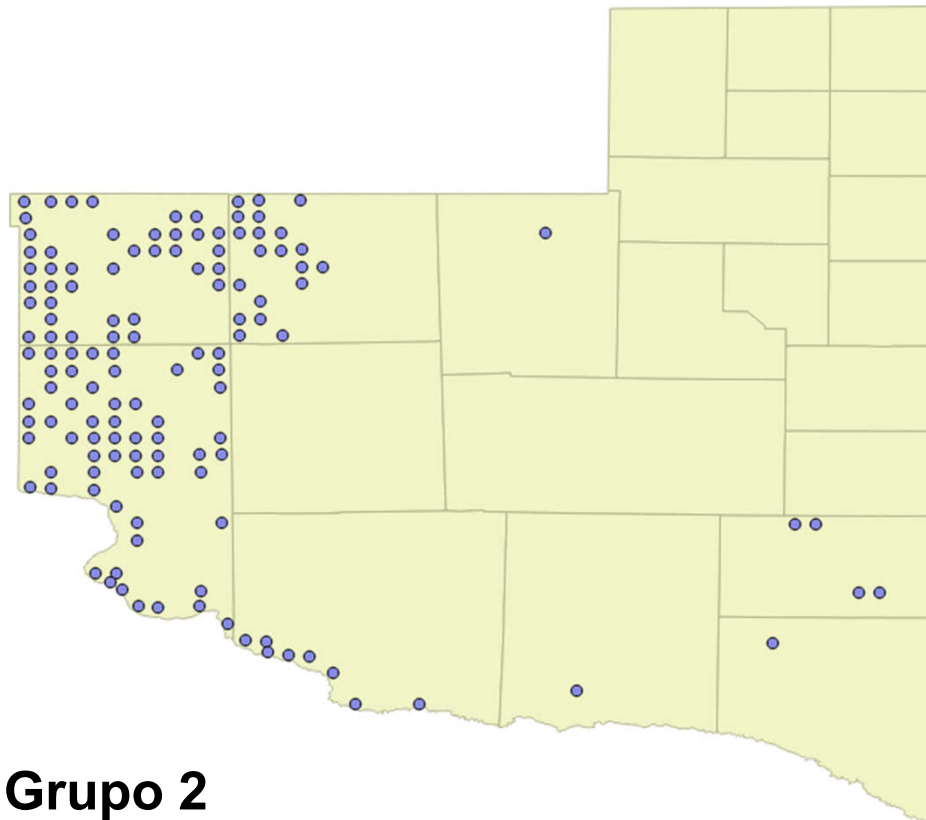
2. Buscamos una coherencia de:

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

Aprendizaje NO supervisado

Sistemas productivos preponderantes

Algunos resultados:



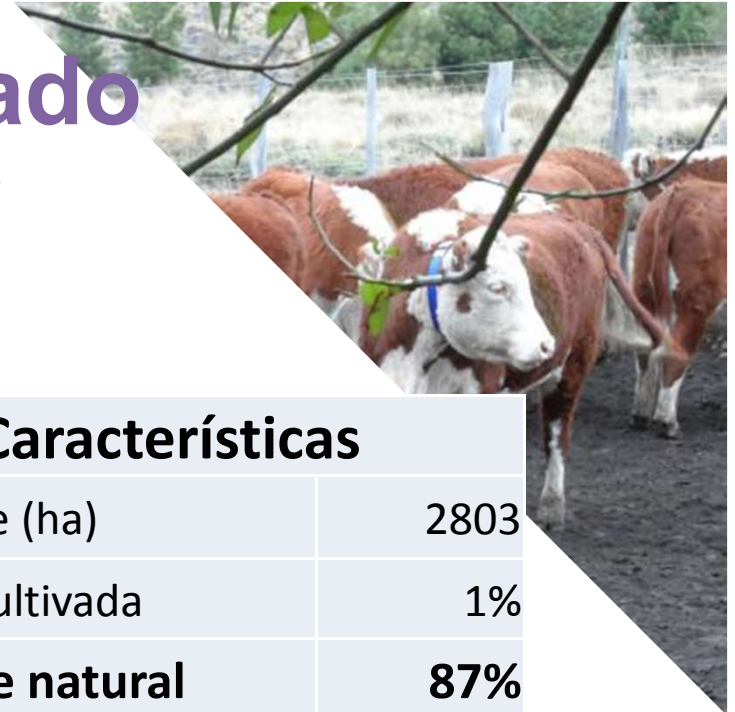
Grupo 2

N 213

Silueta 0.31

Jaccard 0.91

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

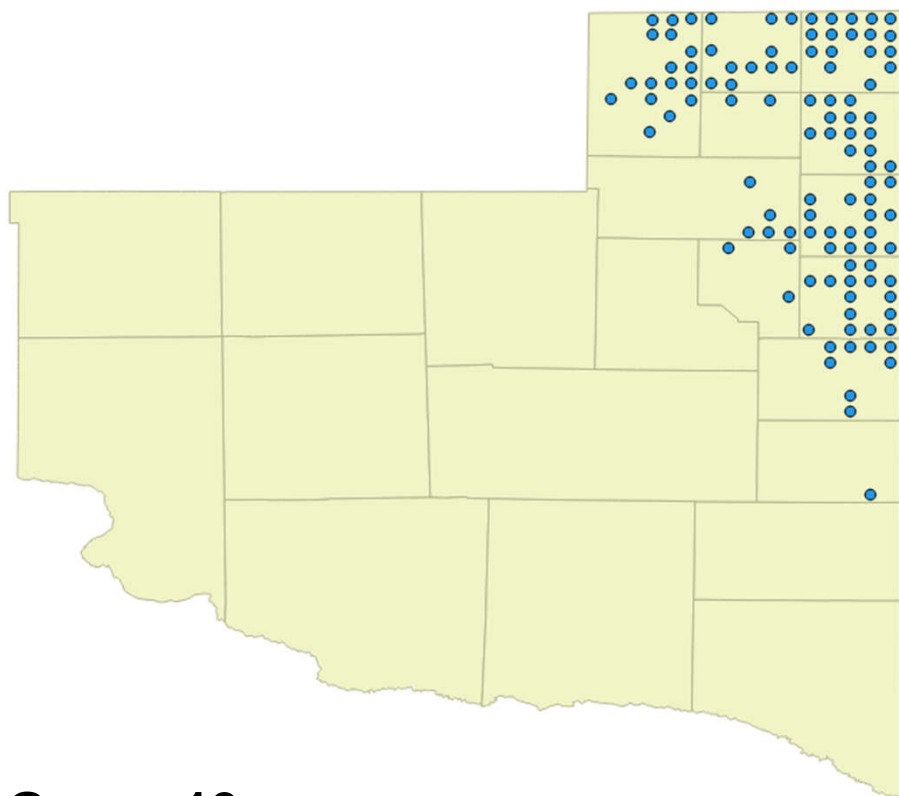


Sistema cría caprina y bovina

Aprendizaje NO supervisado

Sistemas productivos preponderantes

Algunos resultados:



Grupo 13

N 175

Silueta 0.52

Jaccard 0.92

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

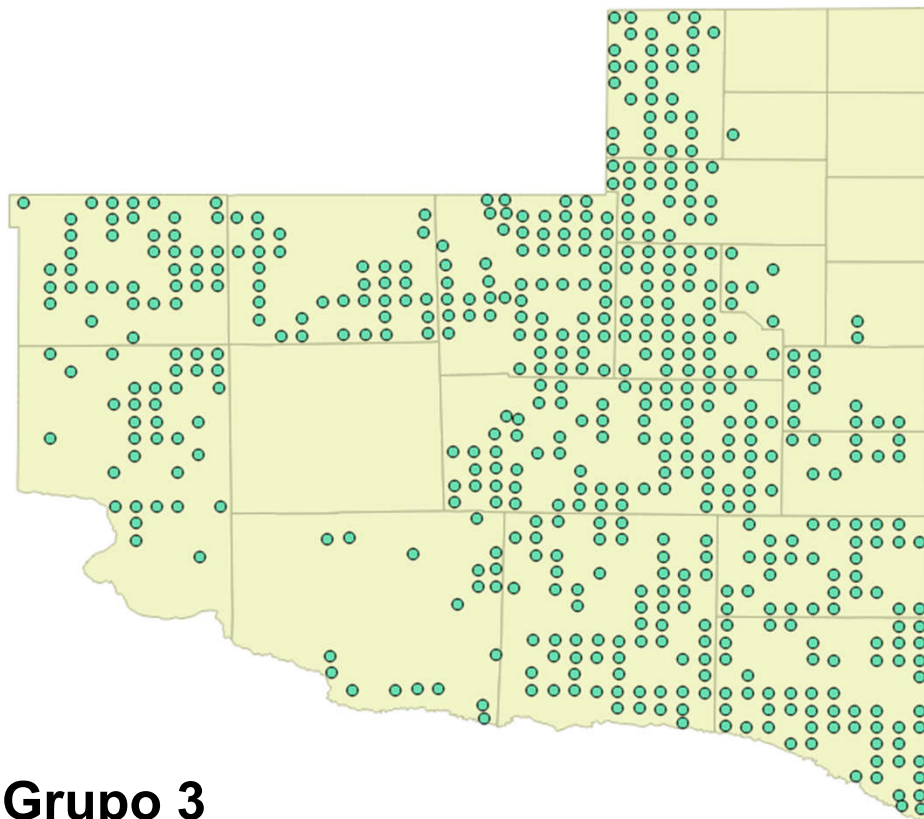


Sistema agricultura de verano

Aprendizaje NO supervisado

Sistemas productivos preponderantes

Algunos resultados:



Grupo 3

N 1065

Silueta 0.36

Jaccard 0.95

Sistema Cría Bovina

Características

Superficie (ha) 2000

% Area Cultivada 4%

% Monte natural 86%

Ganadería

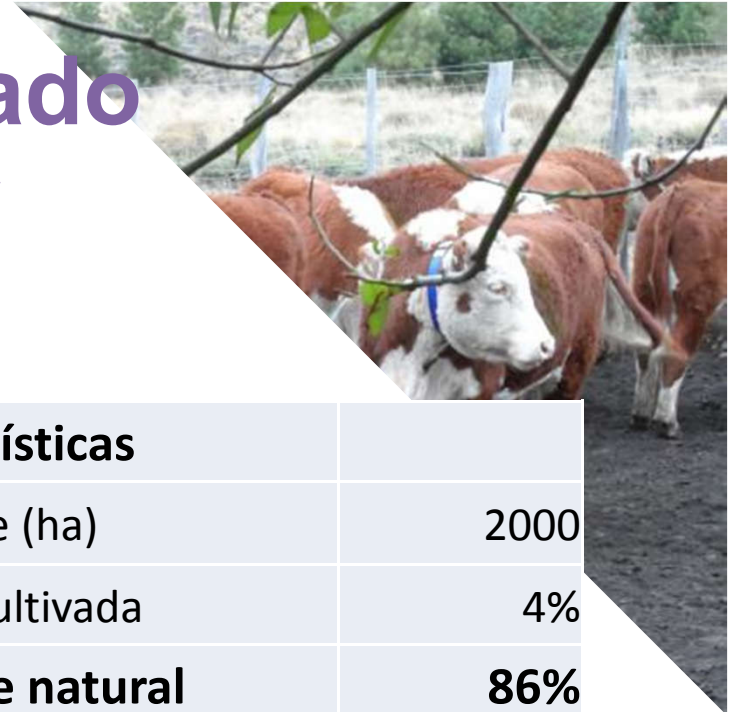
% Bovinos 93%

% Ovinos 3%

% Porcinos 0%

% Caprinos 1%

% Equinos 3%



Aprendizaje NO supervisado

Sistemas productivos preponderantes



Herramientas



fpc Flexible Procedures for Clustering

by ORPHANED

[View Source](#)

Equipo de trabajo



**Anabella
LOZZA**

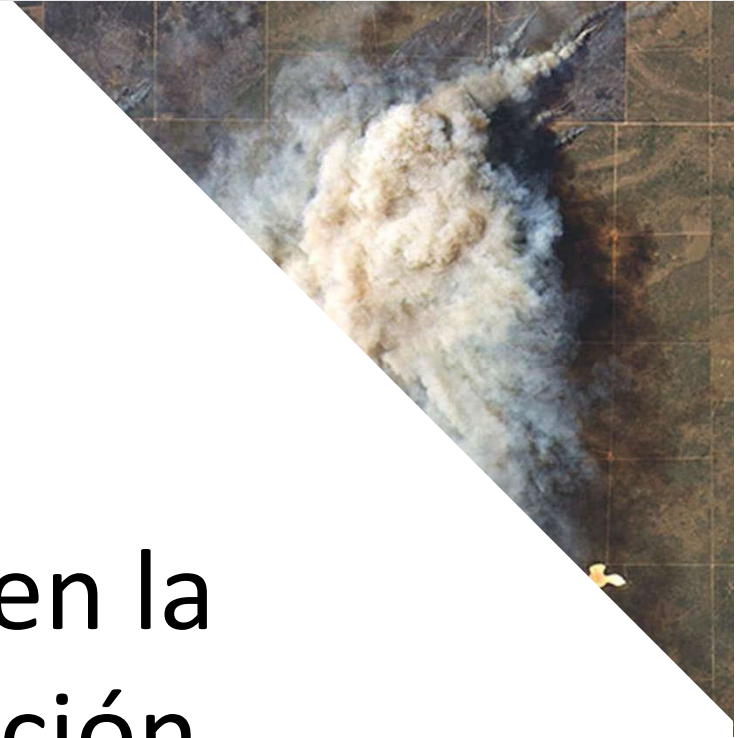


**Héctor
LORDA**



Aprendizaje Supervisado

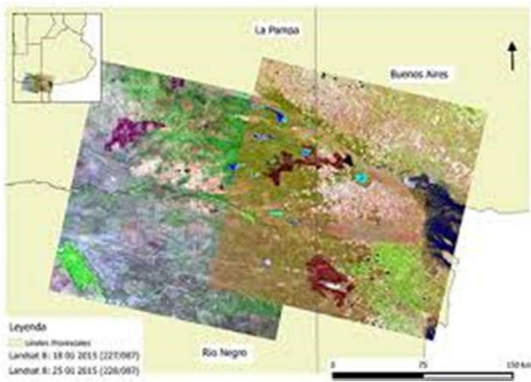
Geotecnologías en la
nube y su aplicación
en emergencias
agropecuarias



Aprendizaje supervisado

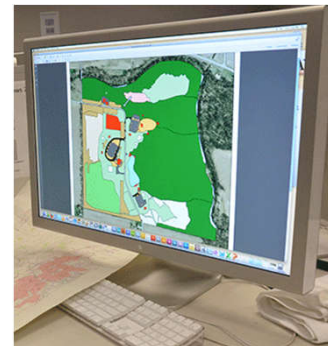
Teledetección

Un poco de historia: 1998

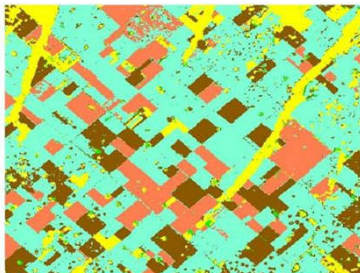


COMISIÓN NACIONAL DE
ACTIVIDADES ESPACIALES

\$\$



\$\$



Aprendizaje supervisado

Teledetección

Obteniendo información desde datos remotos: Anguil



1984

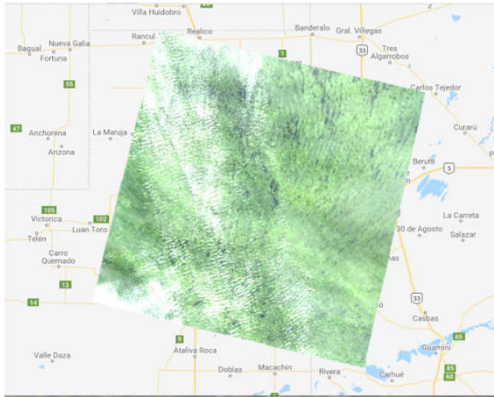
1997

2000

2002

Aprendizaje supervisado

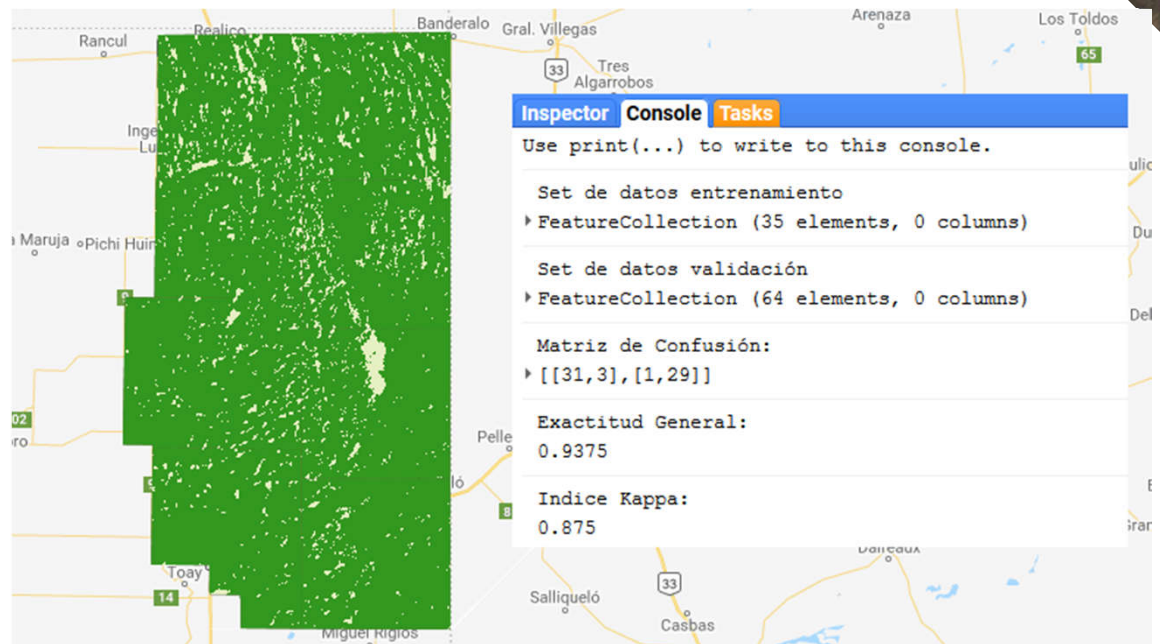
Teledetección



Obtener una imagen



Filtro una imagen



Aplico un algoritmo de clasificacion

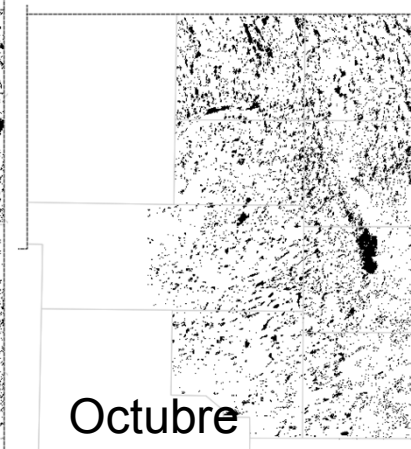
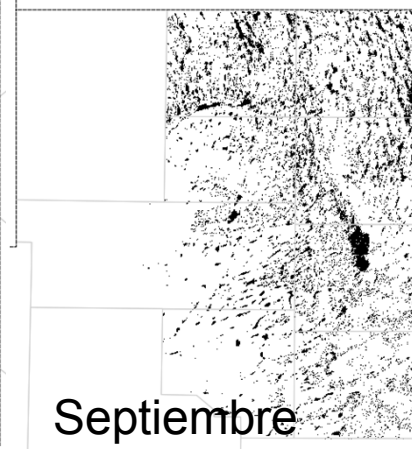
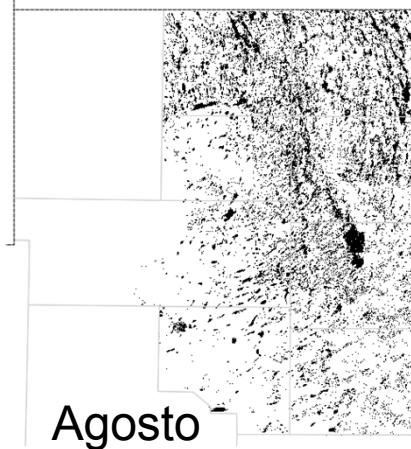
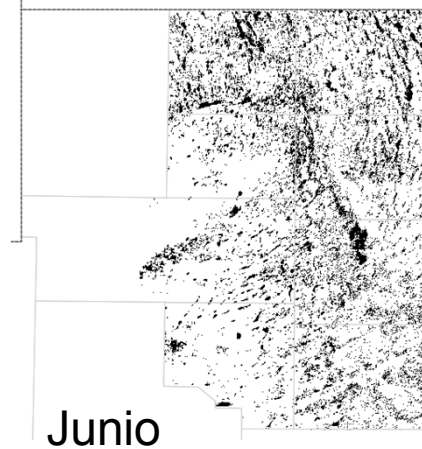
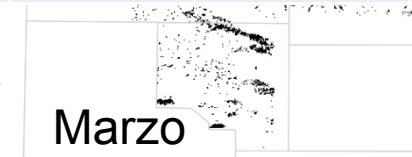
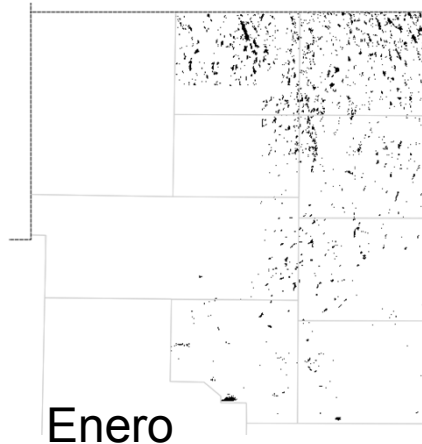
Aprendizaje supervisado

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	



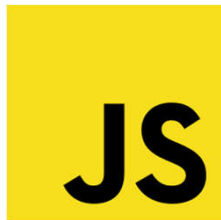
Aprendizaje supervisado

Teledetección

Herramientas



Google Earth Engine



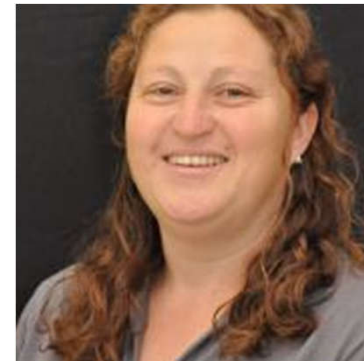
Equipo de trabajo



Julio
FERNANDEZ



Laura
BELMONTE

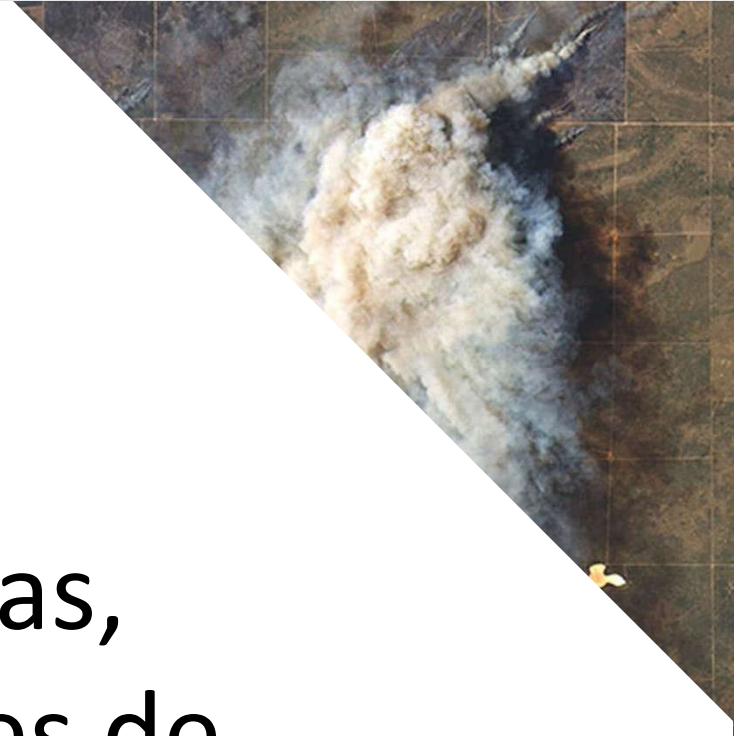


Mariela
FUENTES



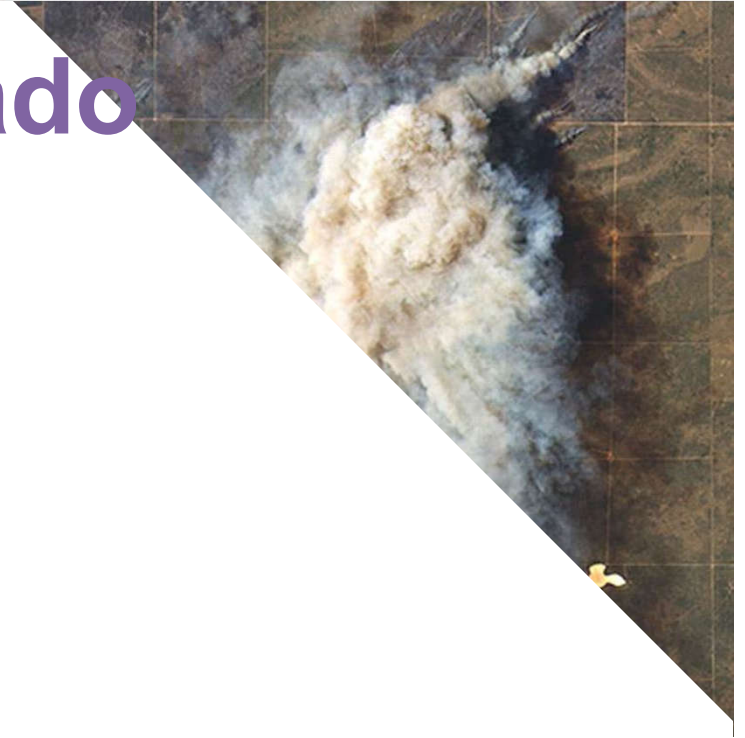
Aprendizaje Supervisado

Drones, cebadas,
girasoles e índices de
vegetación



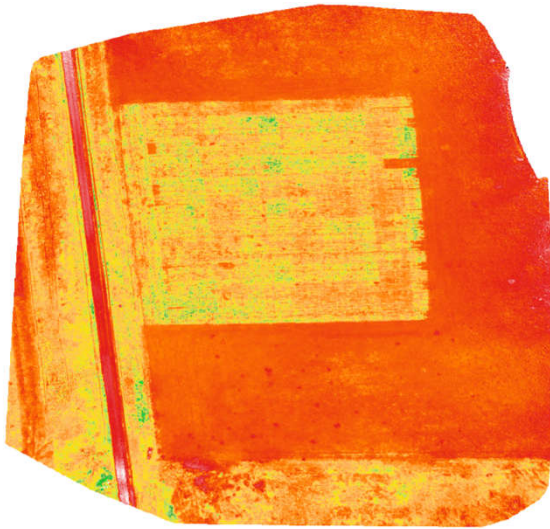
Aprendizaje supervisado

Teledetección

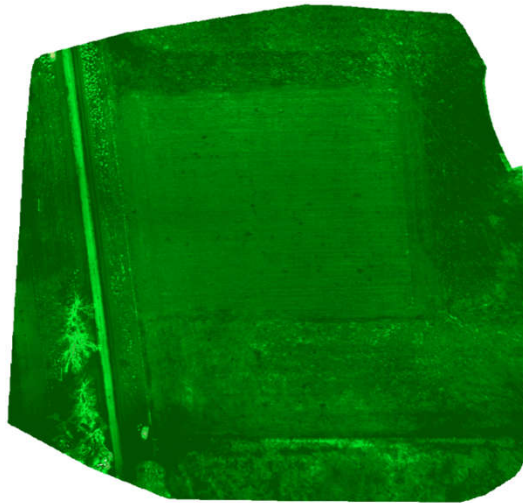
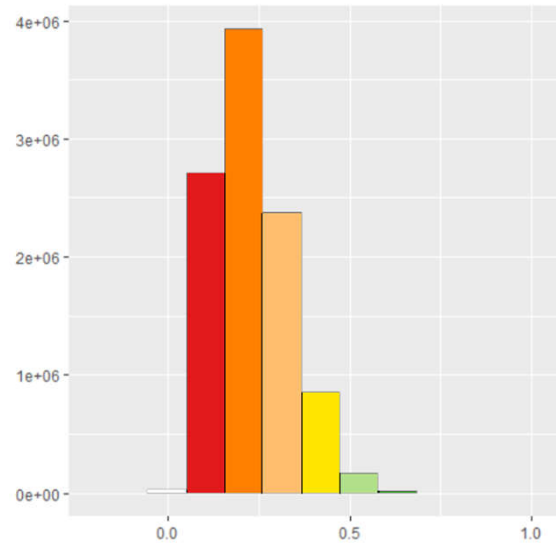


Aprendizaje supervisado

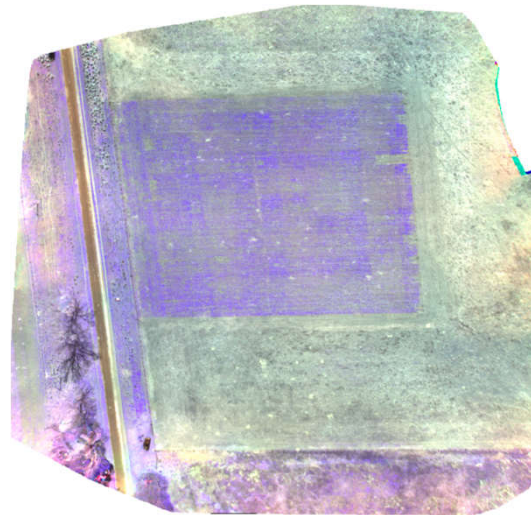
Lote con Cebada



NDVI



Indice de Brillo



Falso Color

Aprendizaje supervisado

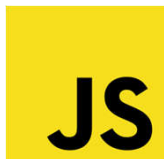
Lote con Cebada



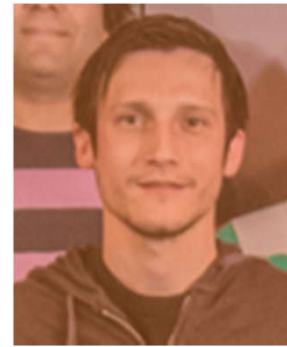
Aprendizaje supervisado

Teledetección

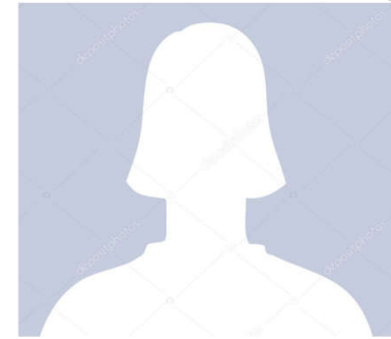
Herramientas



Equipo de trabajo



Matías
CAVILLA



Alexandra
DILLCHNEDER



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?



LatinR 2020, Primera semana de Octubre, On-Line, Gratuito para los asistentes

Comunidades

<http://49jaiio.sadio.org.ar/simposios/CAI>



**Ultimas dos semanas de Octubre
Gratuito para los asistentes**

¿Qué hacemos?



Mine CetinkayaRundel
@minebocek

R4DS translation team, what a fantastic effort!
@LatinR_Conf #rstats



es.r4ds.hadley.nz



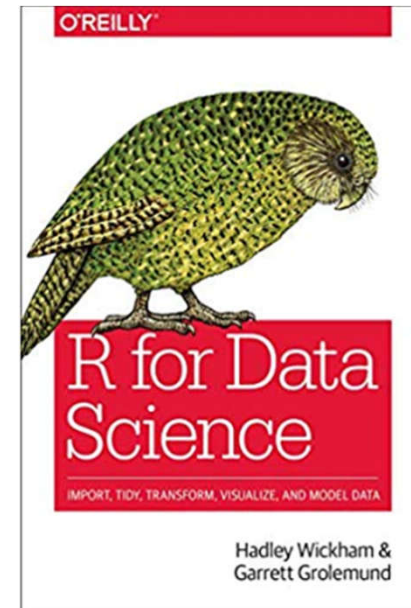
github.com/cienciadedatos...

Traducir Tweet



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

32 Retweets 153 Me gusta



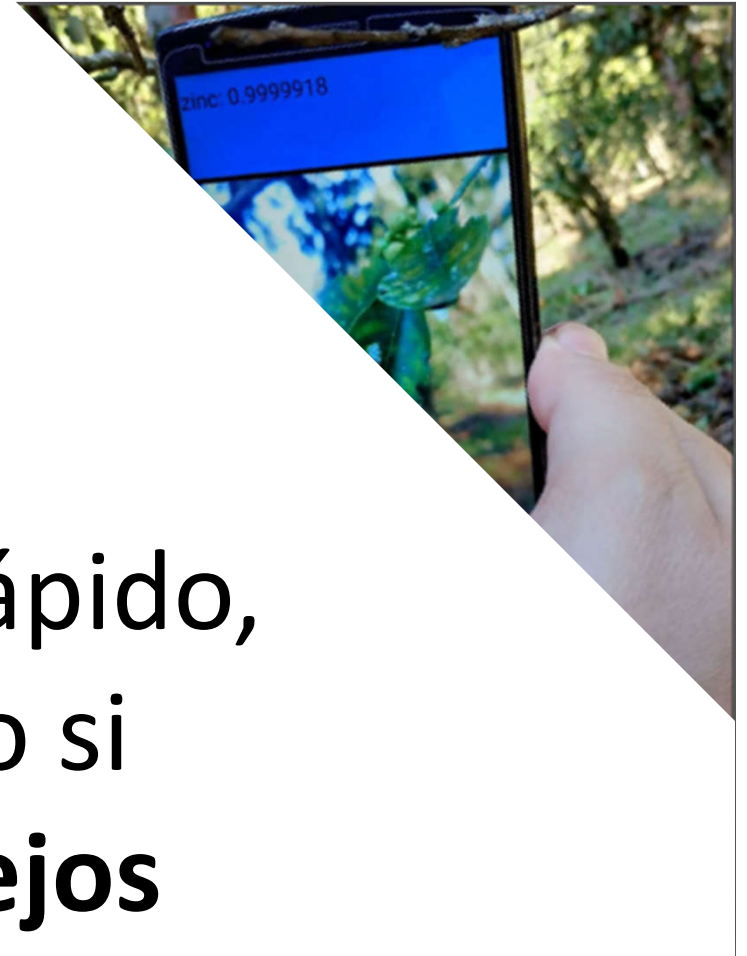
R para Ciencia de Datos

es.r4ds.hadley.nz



Comunidades

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



Proverbio africano

¿Qué preguntas tenés?

@yabellini
yabellini.netlify.com

