



UNIVERSIDAD
DE LA RIOJA

Una breve introducción al aprendizaje automático

Grupo de Informática de la UR
Jónathan Heras



¿Quiénes somos?

Grupo de Informática de la Universidad de La Rioja:

- 22 investigadores del Dpto. Matemáticas y Computación de la UR
- Líneas de investigación:
 - Sistemas de cálculo simbólico
 - Algoritmos de Topología Algebraica
 - Verificación formal de algoritmos
 - Sistemas distribuidos
 - Procesamiento de datos
 - Provenance
 - Análisis de imagen biomédica
 - ...



Colaboración con universidades y empresas

- **Consortio Iberus Health-Tech: Big Data y Computación**
 - Universidad de Zaragoza
 - Universidad de Lleida
 - Universidad Pública de Navarra
 - Banco Santander
- **Colaboraciones estables (nacionales e internacionales)**
 - Universidad de Zaragoza, Universidad de Sevilla, y Universidad del País Vasco
 - Grenoble (Francia), Innsbruck (Austria), Rotterdam (Holanda), Southampton (GB)
- **Colaboraciones con empresas locales y nacionales:**
 - Proyectos ADER, OTRIs, contratos Artículo 83, etc



Índice de contenidos

1. Introducción
2. La importancia de los datos
3. Tipos de aprendizaje
4. Aprendizaje Automático en nuestro grupo



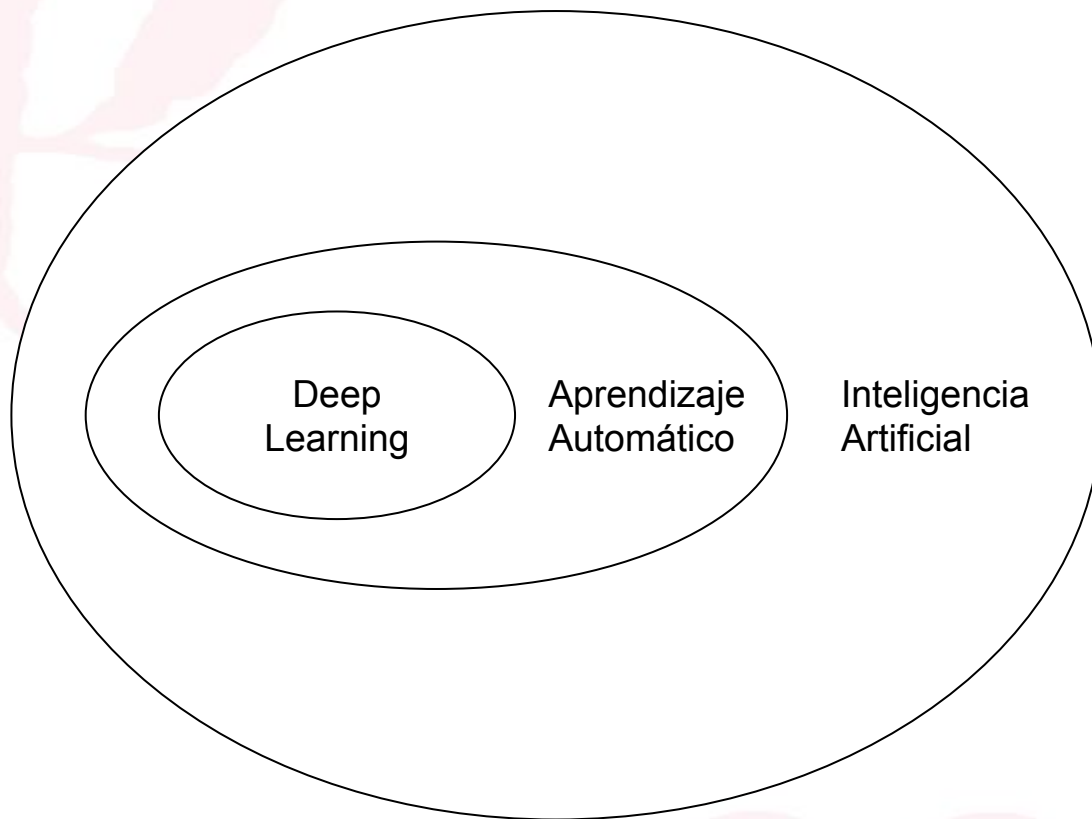
Índice de contenidos

- 1. Introducción**
2. La importancia de los datos
3. Tipos de aprendizaje
4. Aprendizaje Automático en nuestro grupo

Introducción



UNIVERSIDAD
DE LA RIOJA



¿Qué es la Inteligencia Artificial?

La ciencia de hacer máquinas inteligentes [J. McCarthy, 1957]

inteligencia¹

Del lat. *intelligentia*.

1. f. Capacidad de entender o comprender.
2. f. Capacidad de resolver problemas.
3. f. Conocimiento, comprensión, acto de entender.
4. f. Sentido en que se puede tomar una proposición, un dicho o una expresión.
5. f. Habilidad, destreza y experiencia.
6. f. Trato y correspondencia secreta de dos o más personas o naciones entre sí.
7. f. Sustancia puramente espiritual.
8. f. **servicio de inteligencia.**

inteligencia artificial

1. f. *Inform.* Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico.



¿Qué es el aprendizaje automático?

Rama de la inteligencia artificial cuyo objetivo es desarrollar técnicas que permitan a los ordenadores aprender sin ser programados de manera explícita [A. Samuel, 1959]



Aproximación resolución problemas

Programación tradicional:

- Desarrollar programas que resuelvan una tarea

Limitaciones:

- Sólo resuelven problemas ya previstos
- Un sistema se considera inteligente si es capaz de observar su entorno y aprender de él
- La inteligencia reside en adaptarse, tener capacidad de integrar nuevo conocimiento, resolver nuevos problemas, y aprender de los errores



Aproximación resolución problemas

Aprendizaje automático:

- No modela el aprendizaje humano
- Es imposible prever todos los problemas desde el principio
- Buscamos dar a los programas la capacidad de adaptarse, es decir de **generalizar**, sin tener que ser reprogramados



¿Qué es el deep learning?

- Campo del aprendizaje automático
- Redes neuronales profundas
- Aprendizaje de múltiples niveles de representación incrementando su complejidad y abstracción
- Necesita grandes cantidades de datos y uso de GPUs



Índice de contenidos

1. Introducción
- 2. La importancia de los datos**
3. Tipos de aprendizaje
4. Aprendizaje Automático en nuestro grupo

Captura de datos

- Primer paso en cualquier aplicación de aprendizaje automático, puede ser complicado
- Capturar instancias que formen nuestro banco de datos (dataset) y anotarlas
- Avance de aprendizaje automático gracias a [grandes datasets](#) → permiten generalizar
- ¿Volumen de datos? → Cuantos más, mejor
- Distribución de las instancias debe ser cercana a la realidad



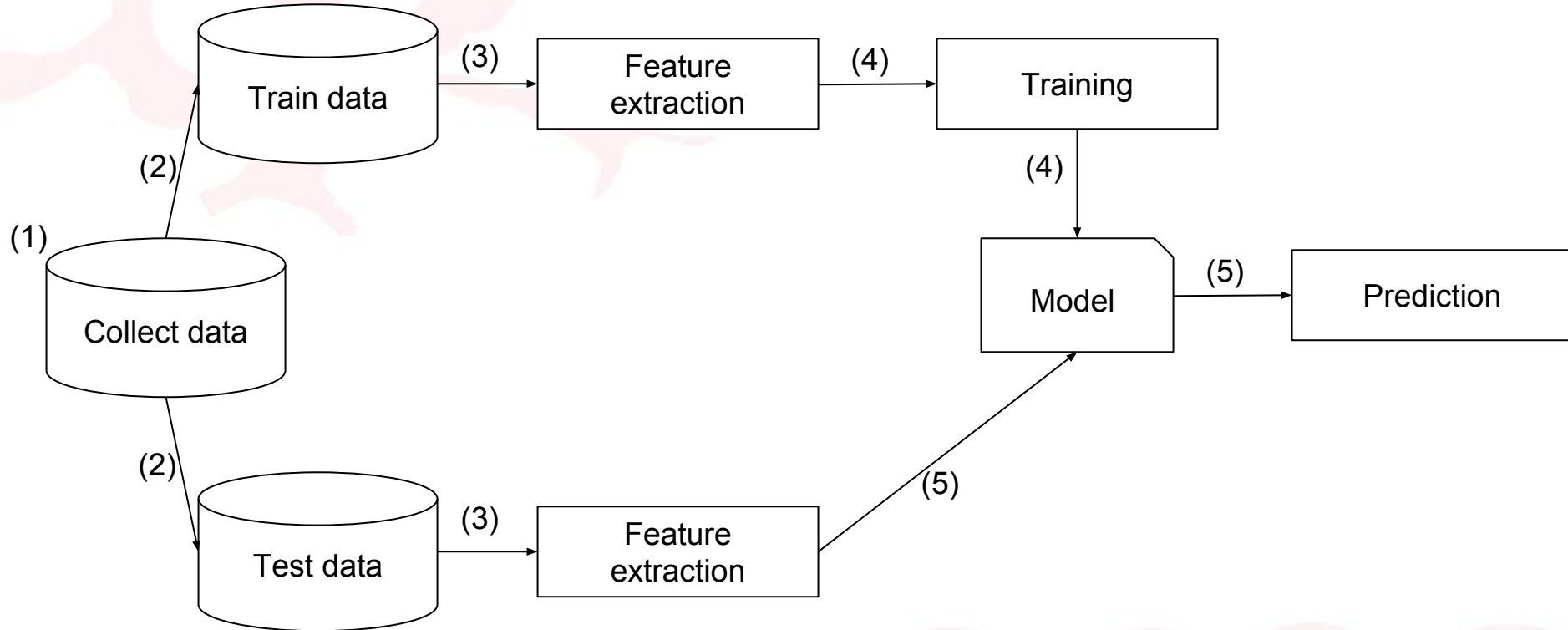
Índice de contenidos

1. Introducción
2. La importancia de los datos
- 3. Tipos de aprendizaje**
4. Aprendizaje Automático en nuestro grupo

Aprendizaje supervisado

- **Entrada:**
 - Conjunto de instancias con sus etiquetas asociadas
- **Objetivo:**
 - A partir de conjunto de entrenamiento etiquetado
 - Construir modelo capaz de predecir la etiqueta de nuevas instancias
- **Distintos tipos:**
 - Clasificación
 - Regresión
 - ...
- **Ejemplos:**
 - A partir de histórico de pasajeros, predecir número de pasajeros que volarán en navidades
 - Dada una imagen de un perro determinar la raza
 - A partir de información de una vivienda determinar su precio
 - Determinar si un mensaje es spam

Aprendizaje supervisado. Flujo de trabajo





Aprendizaje no supervisado

- Entrada: conjunto de instancias
- Objetivo: grupos de instancias similares entre ellas, no hay experto que diga los grupos a buscar
- Tipos:
 - Clustering
 - Clustering jerárquico
 - Reducción de la dimensionalidad
- Ejemplos:
 - Agrupar libros por título
 - Agrupar imágenes por distribución de colores
 - Agrupar palabras por contexto



Índice de contenidos

1. Introducción
2. La importancia de los datos
3. Tipos de aprendizaje
4. **Aprendizaje Automático en nuestro grupo**



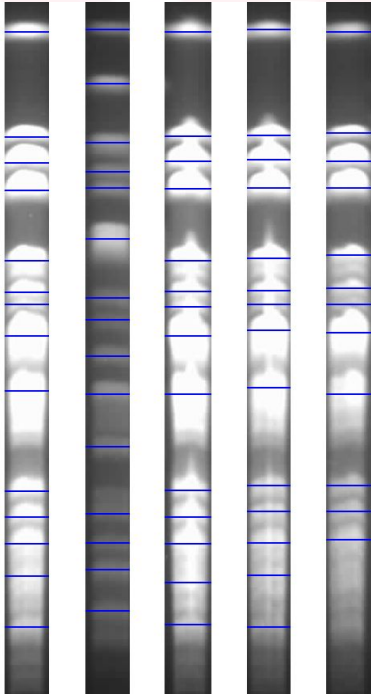
Nuestro grupo

Grupo de Informática de la Universidad de La Rioja:

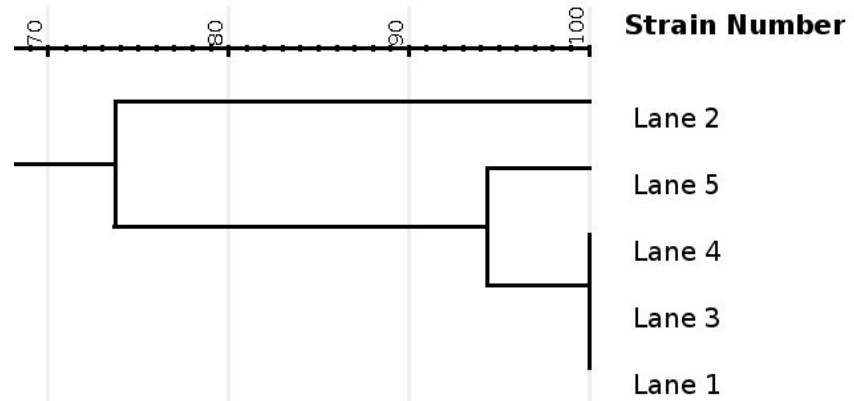
- 22 investigadores del Dpto. Matemáticas y Computación de la UR
- Líneas de investigación:
 - Sistemas de cálculo simbólico
 - Algoritmos de Topología Algebraica
 - Verificación formal de algoritmos
 - Sistemas distribuidos
 - Procesamiento de datos
 - Provenance
 - **Análisis de imagen biomédica**
 - ...

Comparación de geles de ADN

Clustering jerárquico



A	B	C	D	E	F
	<u>Lane 1</u>	<u>Lane 2</u>	<u>Lane 3</u>	<u>Lane 4</u>	<u>Lane 5</u>
<u>Lane 1</u>	1.0	0.66	1.0	1.0	0.92
<u>Lane 2</u>	0.66	1.0	0.66	0.66	0.57
<u>Lane 3</u>	1.0	0.66	1.0	1.0	0.92
<u>Lane 4</u>	1.0	0.66	1.0	1.0	0.92
<u>Lane 5</u>	0.92	0.57	0.92	0.92	1.0



Lectura de códigos de antibióticos



Image analysis

Steps

- Step 1. Bacterial isolate data
- Step 2. Image preprocessing
- Step 3. Antimicrobial disk management
 - 3.1. Disk adjustment
 - 3.2. Fix disk size
 - 3.3. Read antimicrobial disk**
 - 3.4. Additional Information
- Step 4. Inhibition zone reading
- Step 5. Inhibition zone annotation

1: Ciprofloxacin 5

2: Tobramycin 10

3: Nalidixic acid 30

4: Trimethoprim-sulfamethoxazole 1.25-2:

5: Amikacin 30

6: Gentamicin 10

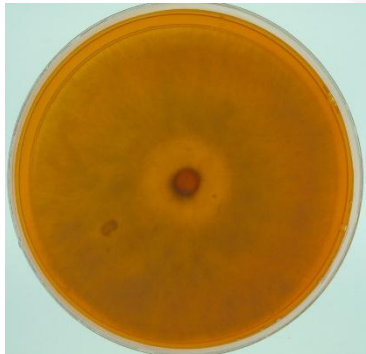
Show indices

Help
In this step, you must select the antimicrobial of each disk.

Decolorización de hongos

Problema biológico: uso de hongos para decolorar tintes

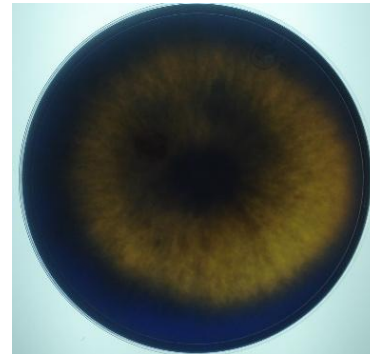
Objetivo: determinar de manera automática el nivel de decoloración de un tinte al cual se le ha inoculado un hongo



-



+



++



+++

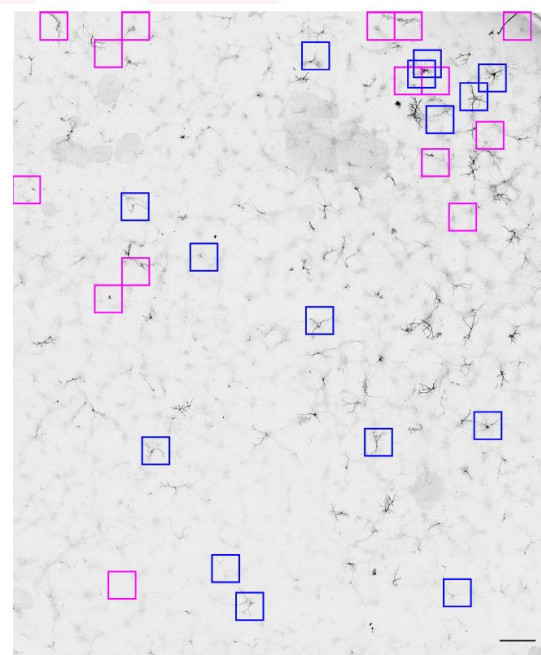
Colaboración con IDE Rioja



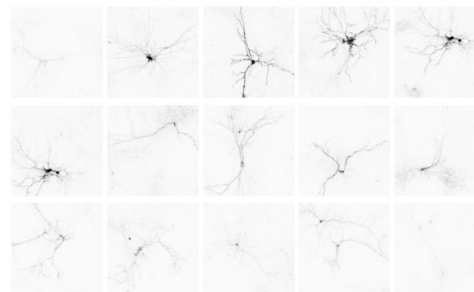
UNIVERSIDAD
DE LA RIOJA



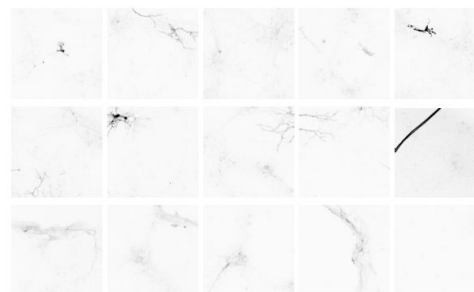
Detección de neuronas



(a) Example high-content image. Scale bar: 500 μ m.



(b) Example patches considered as positives (blue squares).



(c) Example patches considered as negatives (magenta squares).



UNIVERSIDAD
DE LA RIOJA

Una breve introducción al aprendizaje automático

Grupo de Informática de la UR
Jónathan Heras