



# Deep learning II & ética de la IA

Programa en Analítica

Curso *Capstone - Inteligencia Artificial & Deep Learning*  
Analítica Prescriptiva

Educación continua | Universidad de los Andes  
octubre 6 – noviembre 24  
2021

# Hoy

De qué vamos a hablar

## 1. Cómo encaja todo

*¿Para qué usamos ML y para qué optimización en la analítica prescriptiva?*

## 2. Un vistazo más a las redes neuronales:

*Diferentes formas para diferentes retos, no sólo en analítica prescriptiva.*

## 3. Ética de la IA:

*Lo que puede salir mal. Y algunas consideraciones clave.*

## 4. ¡Taller de redes neuronales en R!

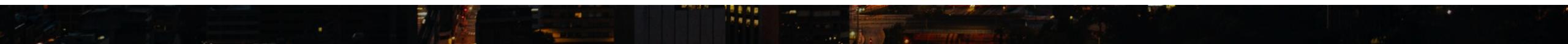
*Redes neuronales para clasificación*



Photo by [nappy](#) from [Pexels](#)

# Pero antes

- Mañana jueves podemos hablar si necesitan para facilitar la formalización de sus proyectos.
- Esta semana deberíamos tener en mente un problema de analítica prescriptiva formalizado.



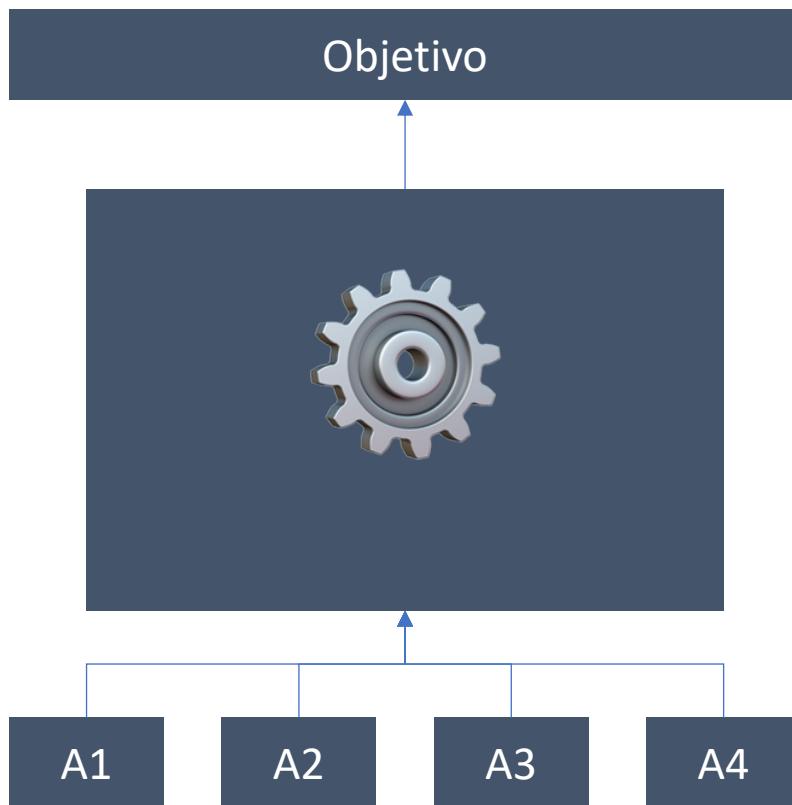
# ¿Cómo encaja todo?

Optimización y aprendizaje de máquinas...



Photo by [nappy](#) from [Pexels](#)

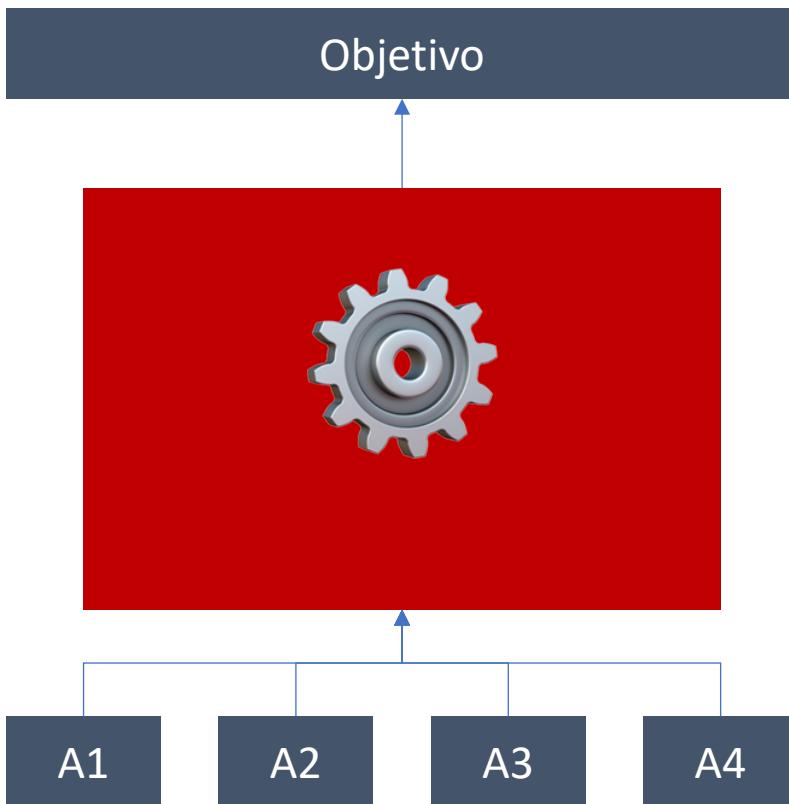
# Recuerden



Un sistema de analítica prescriptiva se construye con tres componentes básicos:

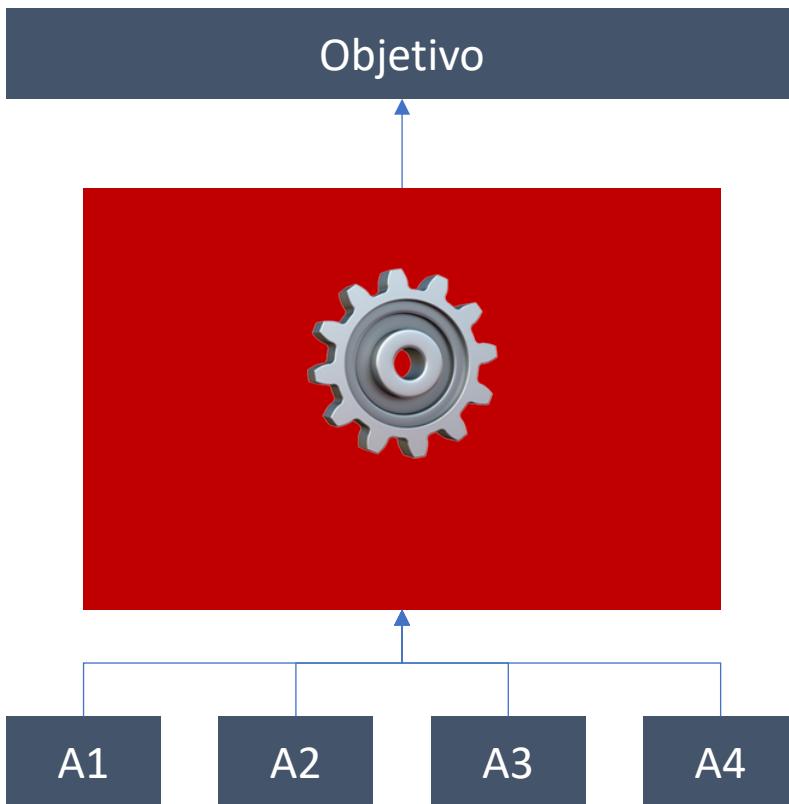
- **Acciones**
- **Objetivo**
- **Modelo**

# Si vamos a usar ML, lo usamos aquí:



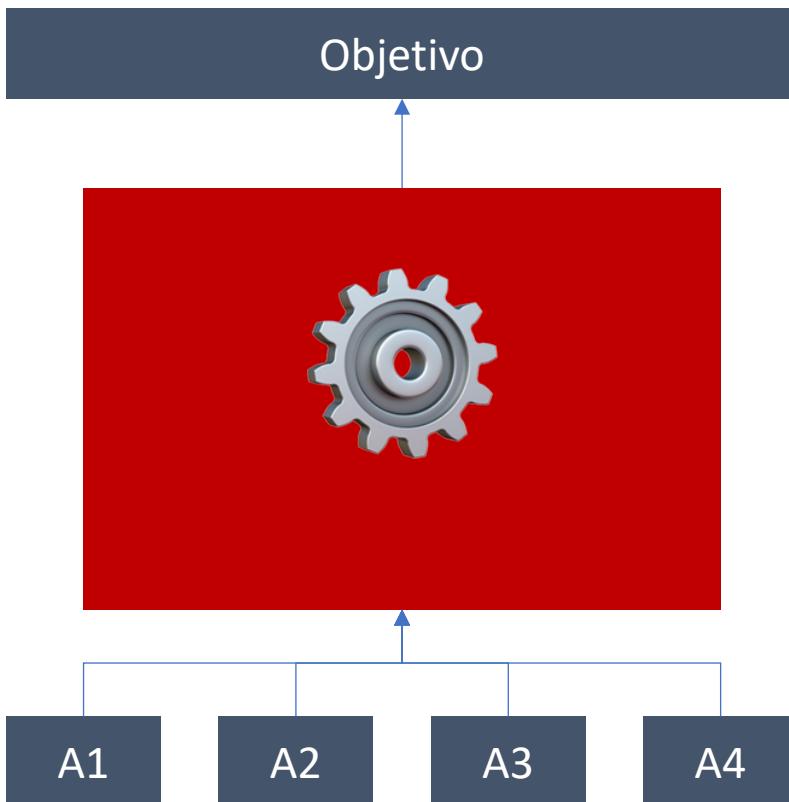
- Entrenamos **modelos que son capaces de predecir** cómo nuestras acciones van a afectar nuestro criterio de decisión.
- Con esas predicciones podemos optimizar:

# Si vamos a usar ML, lo usamos aquí:



- Entrenamos **modelos que son capaces de predecir** cómo nuestras acciones van a afectar nuestro criterio de decisión.
- Con esas predicciones podemos optimizar:
  - A1: qué nos predice la máquina?
  - A2: qué nos predice la máquina?
  - ...
  - A4: qué nos predice la máquina?

# Si vamos a usar ML, lo usamos aquí:

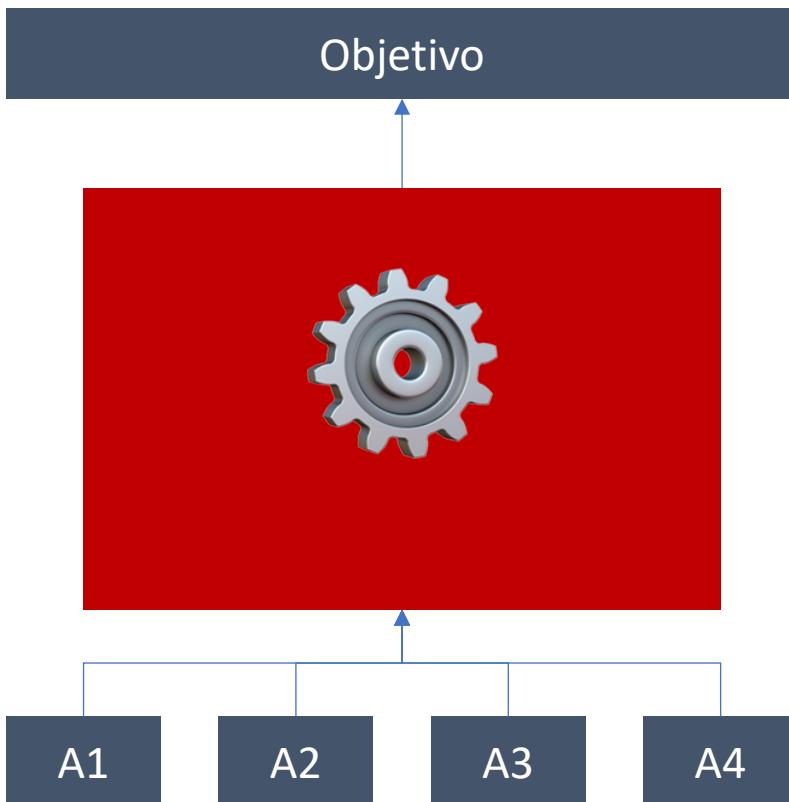


- Escogemos la que maximiza o minimiza nuestro criterio de decisión:

A3

(Esto es optimización discreta)

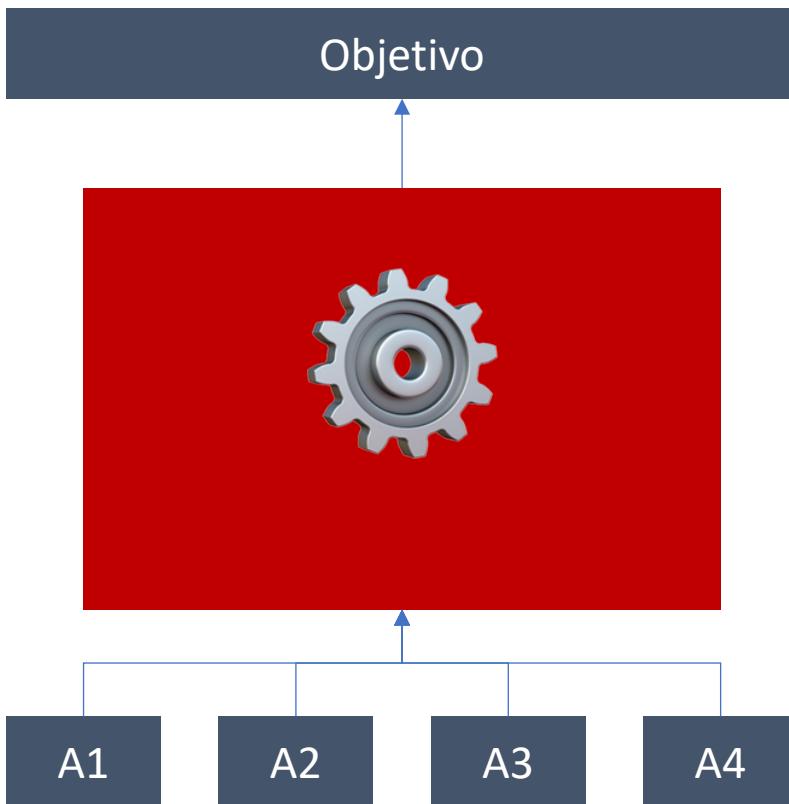
# Si tenemos muchas posibles decisiones...



- Podemos usar programación lineal
- Recordemos el problema del equipo de fútbol

ranking	“precio”	posición	país	nombre

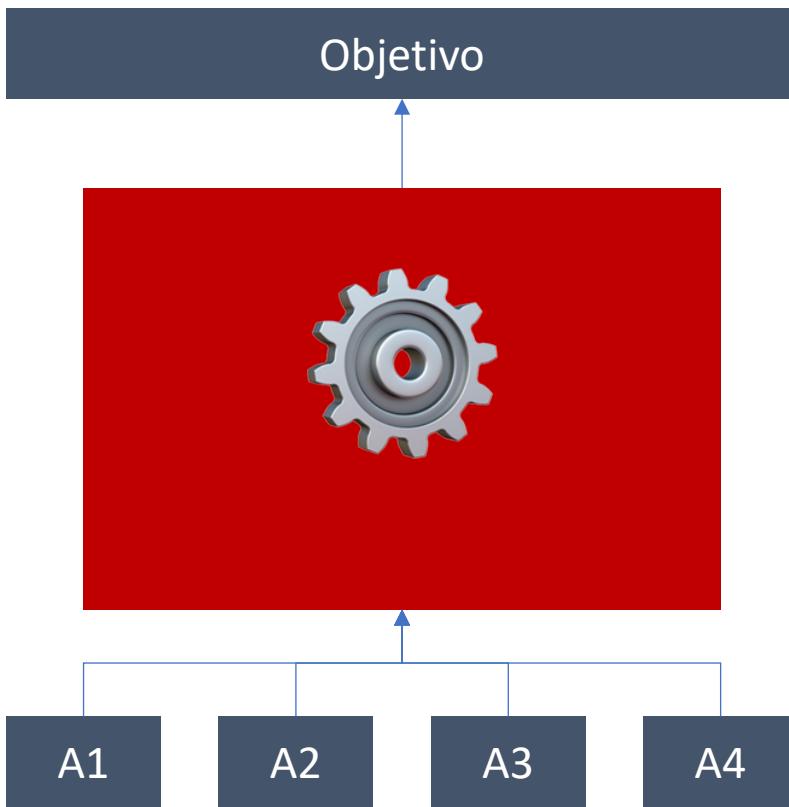
# Si tenemos muchas posibles decisiones...



- Podemos usar programación lineal
- Recordemos el problema del equipo de fútbol

ranking	“precio”	posición	país	nombre

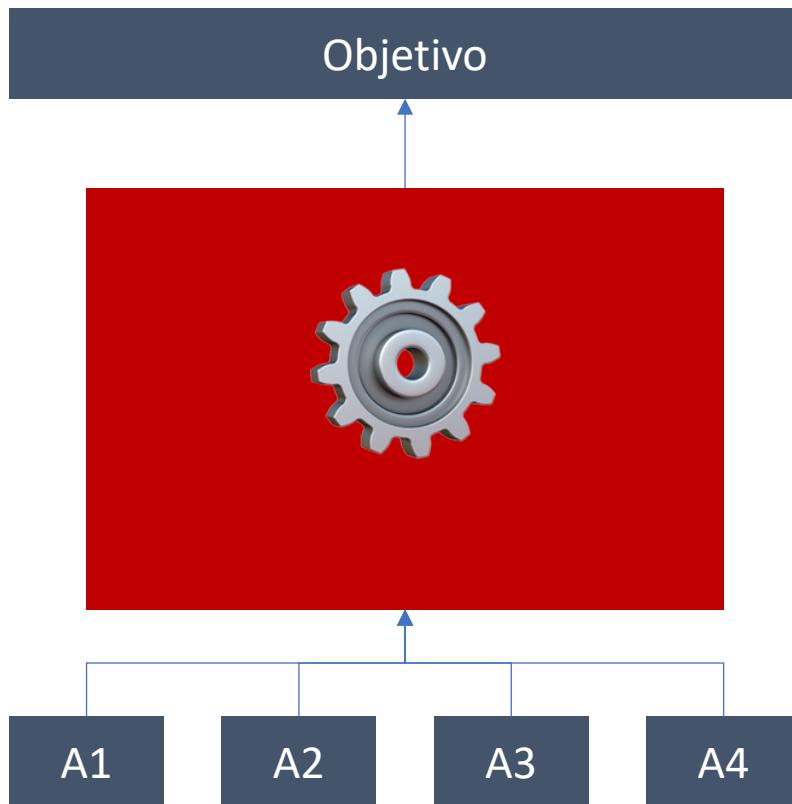
# Si tenemos muchas posibles decisiones...



- Podemos usar programación lineal
- Recordemos el problema del equipo de fútbol

ranking	“precio”	posición	país	nombre
			compra o no compra	
			1	
			0	
			1	
			1	

# Si tenemos muchas posibles decisiones...

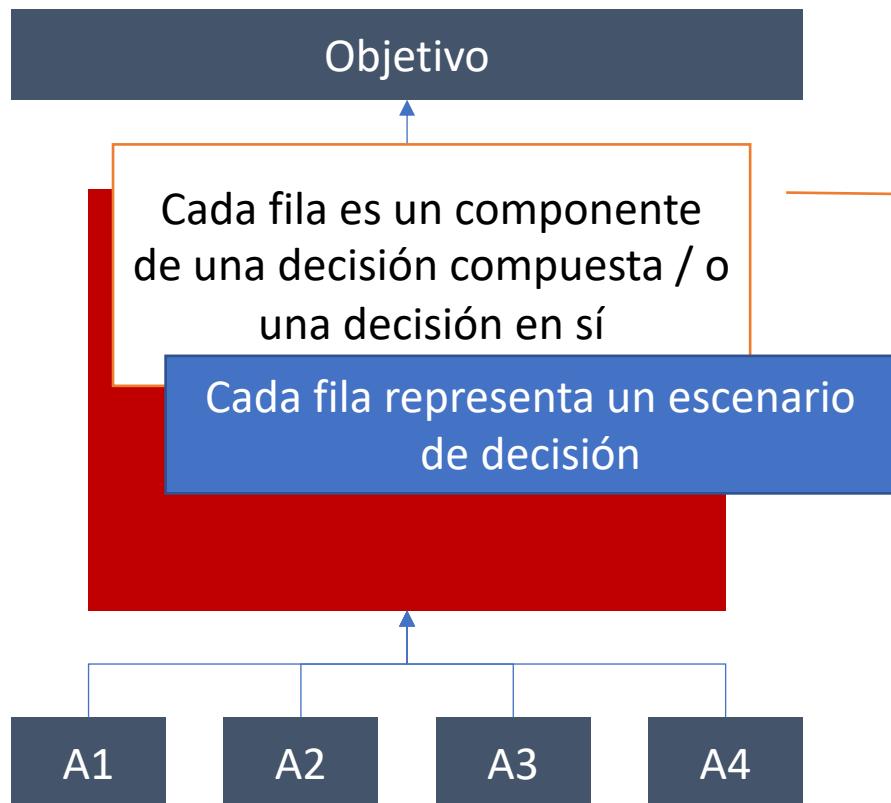


¿Cómo utilizamos ML cuando tenemos muchas opciones? Tengo que partir de una tabla en la que...

objetivo	res1	res2	...	decisión
				1
				0
				1
				1

*max      s.a.      s.a.      Otros features*

# Si tenemos muchas posibles decisiones...

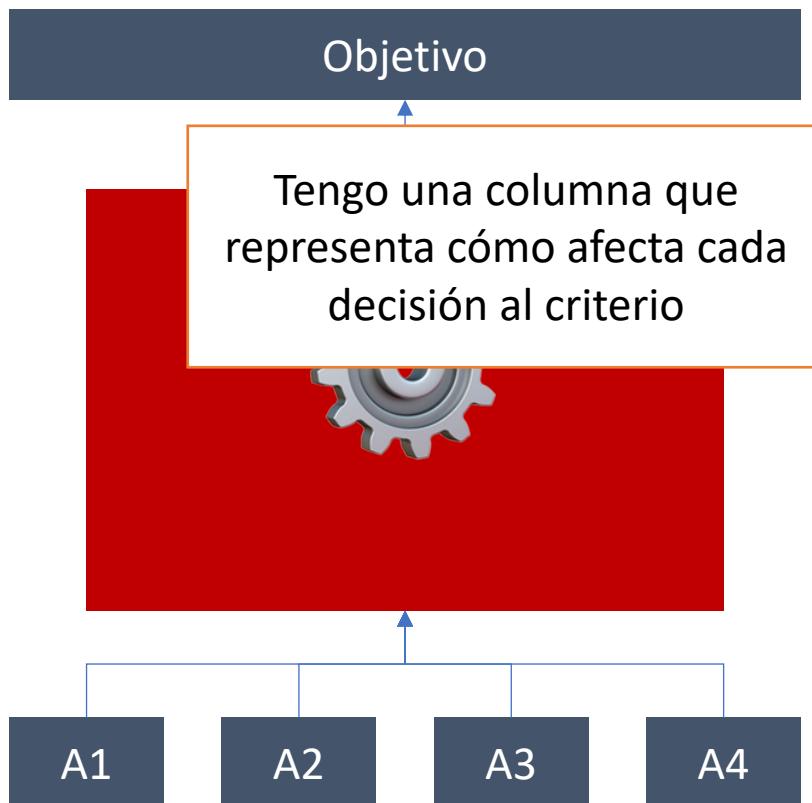


¿Cómo utilizamos ML cuando tenemos muchas opciones?

The table represents the output of an ML model. It has columns for "objetivo", "res1", "res2", "...", and "decisión". There are four rows of data. Red annotations are present: "max" is written above the first column, and "s.a." is written above the second and third columns. "Otros features" is written above the last column. The "decisión" column contains values 1, 0, 1, and 1.

objetivo	res1	res2	...	decisión
				1
				0
				1
				1

# Si tenemos muchas posibles decisiones...

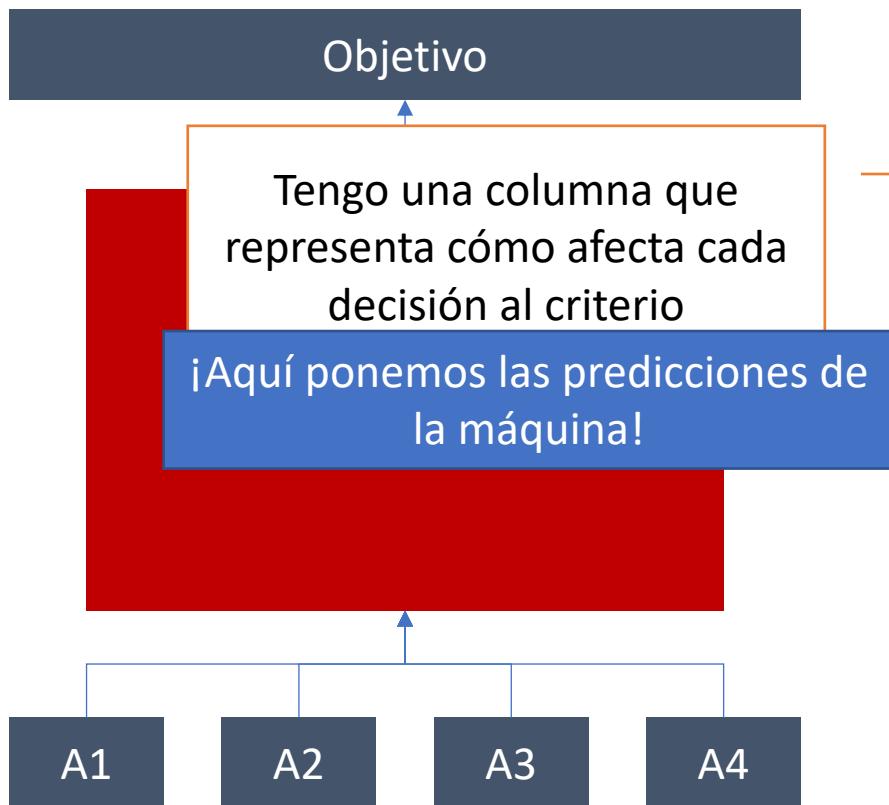


¿Cómo utilizamos ML cuando tenemos muchas opciones?

A table representing the output of a machine learning model. The columns are labeled "objetivo", "res1", "res2", "...", and "decisión". The "decisión" column contains values 1, 0, 1, 1. Above the table, arrows point from the text "max" and "s.a." to the first two rows of the "res1" column. An arrow also points from the text "s.a." to the third row of the "res1" column. A bracket labeled "Otros features" spans the "res1", "res2", and "..." columns.

objetivo	res1	res2	...	decisión
				1
				0
				1
				1

# Si tenemos muchas posibles decisiones...



¿Cómo utilizamos ML cuando tenemos muchas opciones?

max

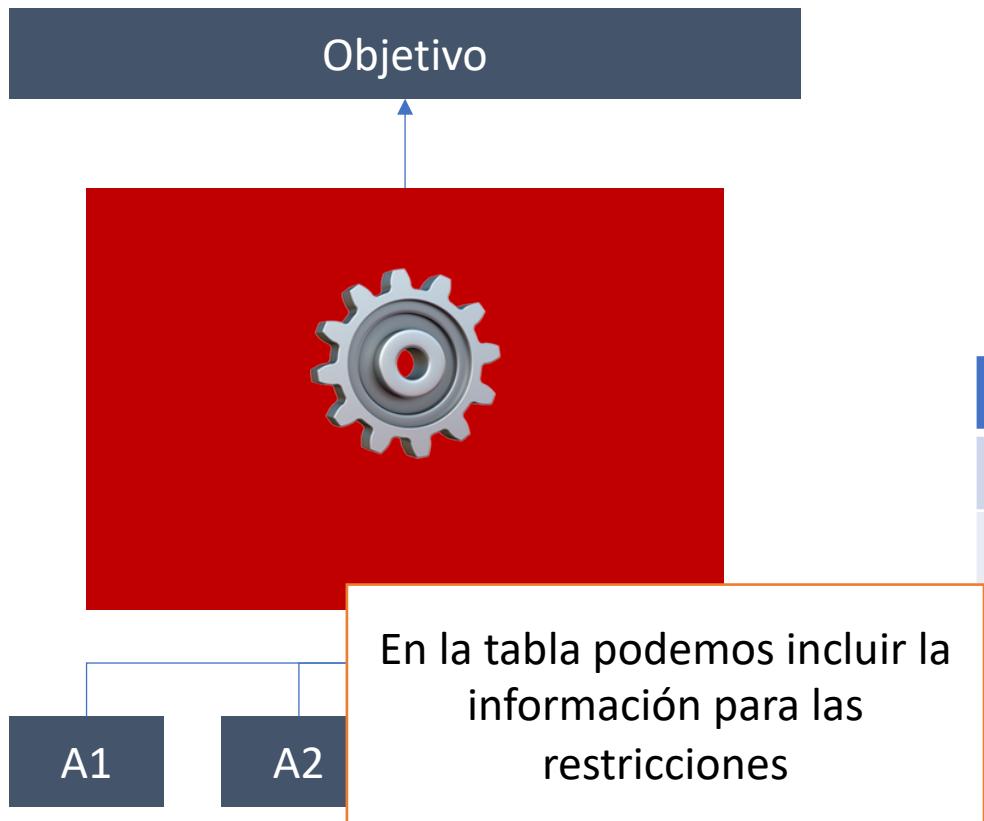
s.a.

s.a.

Otros features

objetivo	res1	res2	...	decisión
				1
				0
				1
				1

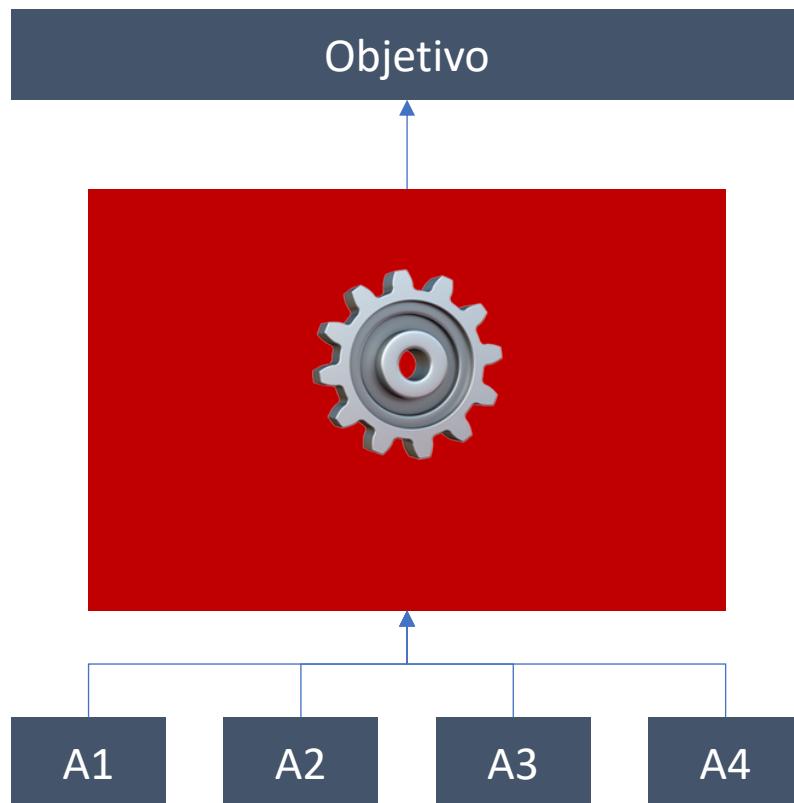
# Si tenemos muchas posibles decisiones...



¿Cómo utilizamos ML cuando tenemos muchas opciones?

objetivo	res1	res2	...	decisión
max	s.a.	s.a.	Otros features	1
				0
				1
				1

# Si tenemos muchas posibilidades...



Al final, el algoritmo de PL nos va a dar una determinada decisión para cada escenario (HABIENDO CONSIDERADO LAS RESTRICCIONES)

PL cuando tenemos muchas opciones?

max      s.a.      s.a.      Otros features

objetivo	res1	res2	...	decisión
				1
				0
				1
				1

# Recapitulando...

- Utilizamos ML para **predicir cómo nuestras decisiones afectan un criterio objetivo** que queremos maximizar o minimizar.
- Utilizamos optimización para escoger entre múltiples predicciones habiendo **considerado restricciones del problema**.



# Pregunta:

¿Por qué no utilizamos Machine Learning en el problema del equipo de fútbol?

# Pregunta:

¿Por qué no utilizamos Machine Learning en el problema del equipo de fútbol?

Porque no necesitábamos predecir cómo afectaba nuestra decisión a nuestro criterio de decisión:

Criterio: **maximizar** el ranking del equipo

Si yo contrato a alguien, su ranking....

# Pregunta:

¿Por qué no utilizamos Machine Learning en el problema del equipo de fútbol?

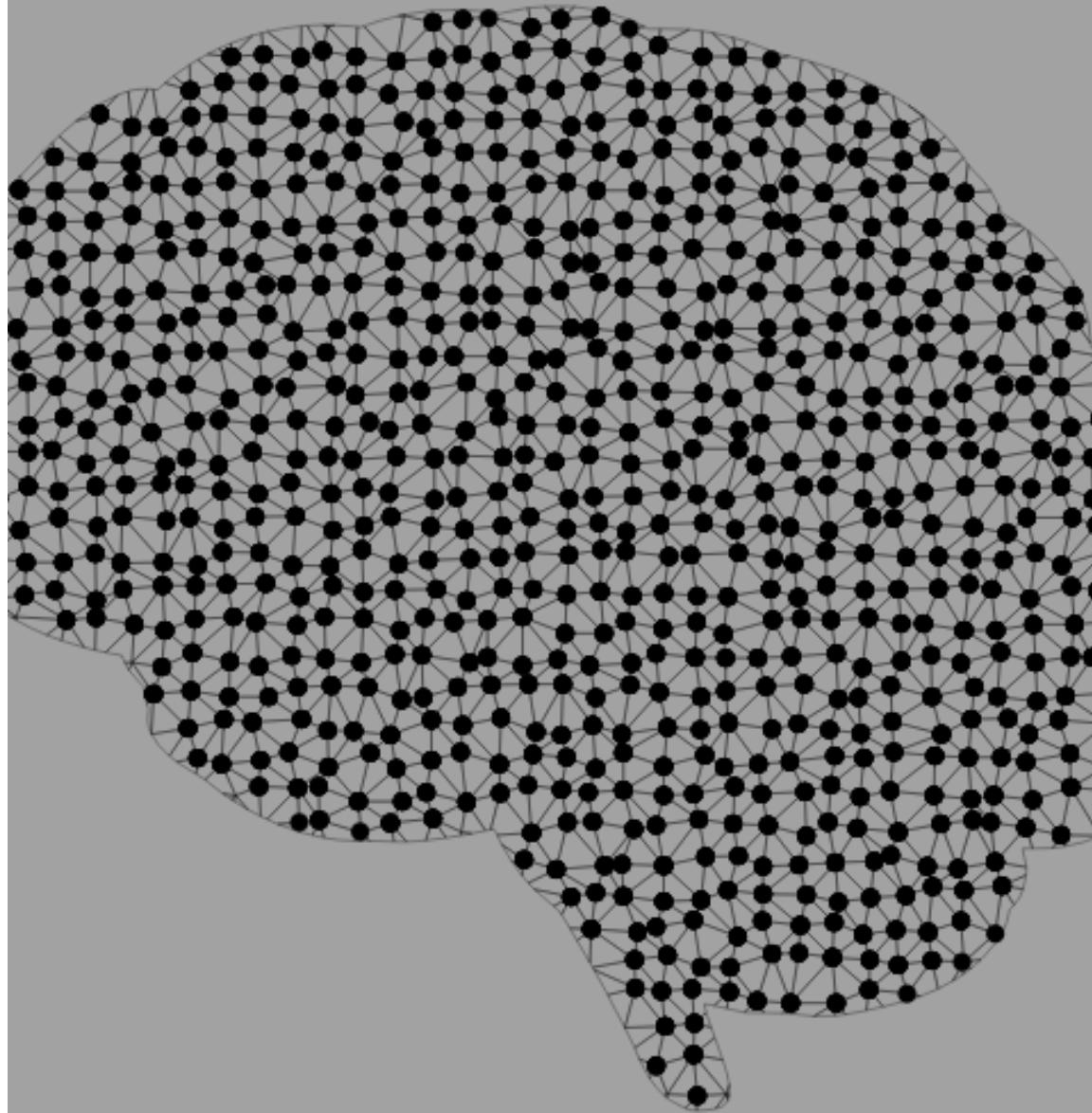
Porque no necesitábamos predecir cómo afectaba nuestra decisión a nuestro criterio de decisión:

Criterio: **maximizar** el ranking del equipo

Si yo contrato a alguien, su ranking.... suma al ranking del equipo.

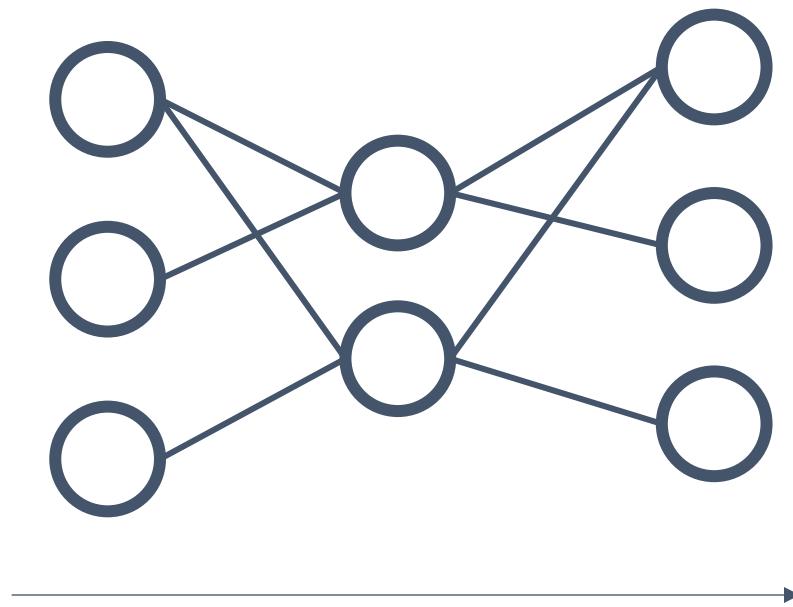
# Un vistazo más a las ANN

Aplicaciones más allá de la analítica  
prescriptiva



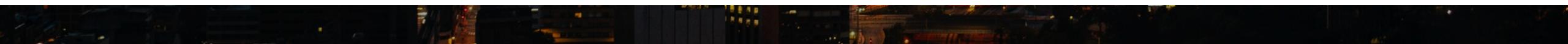
# Red Neuronales “Feed Forward”

- Hasta ahora hemos visto redes cuyos valores van de la capa inicial a las finales, siempre hacia adelante.



# El problema de clasificación de hoy

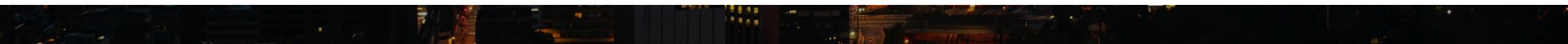
- Muy útiles para problemas de aprendizaje supervisado, como regresión, clasificación.
- Hoy vamos a trabajar con datos de una universidad que quiere clasificar a un aspirante en admitido/no admitido
- Utiliza tres datos para esto (GRE: exámenes, PGA: promedio general acumulado, Ranking en curso)
- Vamos a usar una red neuronal.



# ¿A dónde se va a mover el punto?



Y ahora.. ¿A dónde se va a mover el punto?

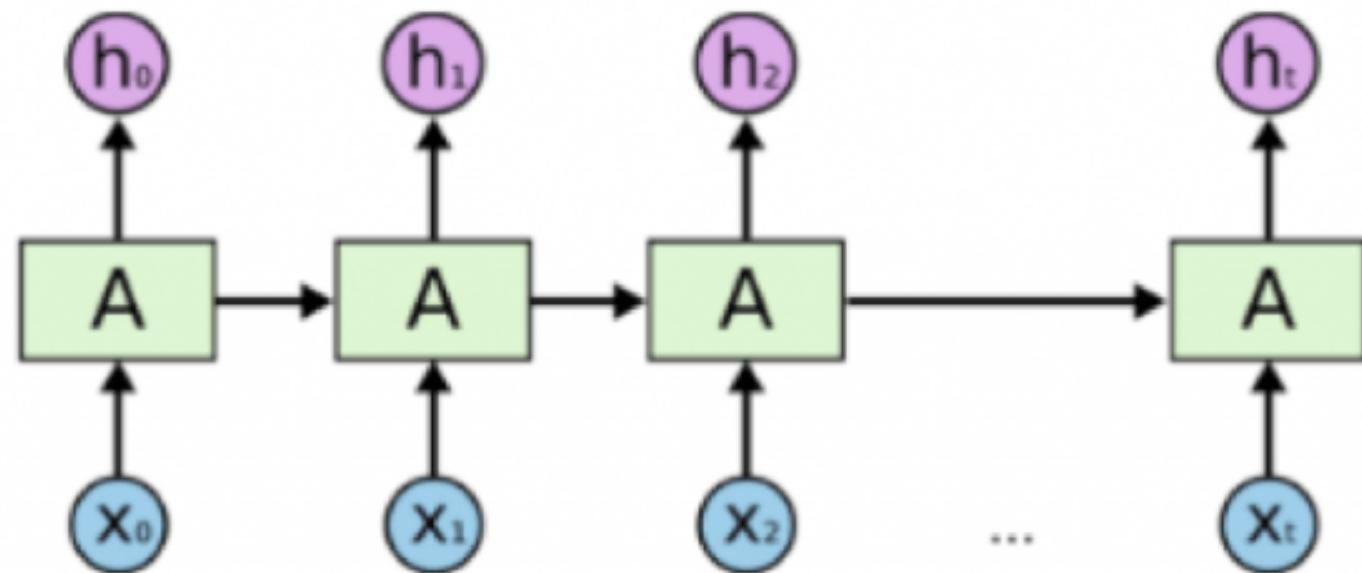


La información del pasado puede ayudar a predecir



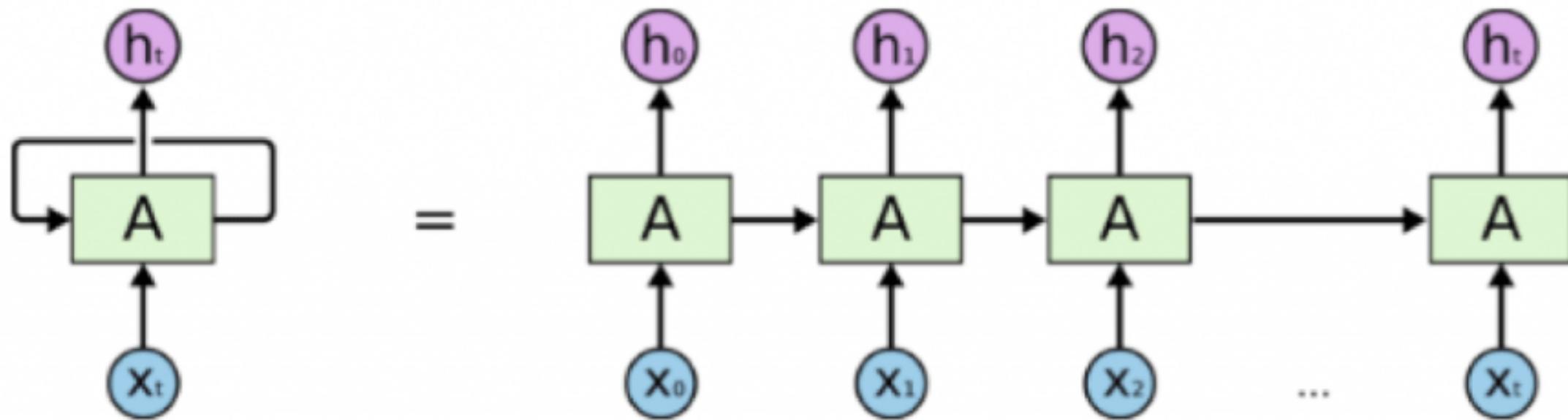
# Red neuronal recurrente

- Hay fenómenos que tienen ¡dinámica! (hay datos nuevos cada momento del tiempo).



# Red neuronal recurrente

- Hay fenómenos que tienen ¡dinámica! (hay datos nuevos cada momento del tiempo).



# Casos de aplicación

## Reconocimiento de lenguaje



Photo by [George Milton](#) from [Pexels](#)

Photo by [Roman Pohorecki](#) from [Pexels](#)

# Casos de aplicación

Reconocimiento de lenguaje



Photo by [George Milton](#) from [Pexels](#)

Predicción de texto

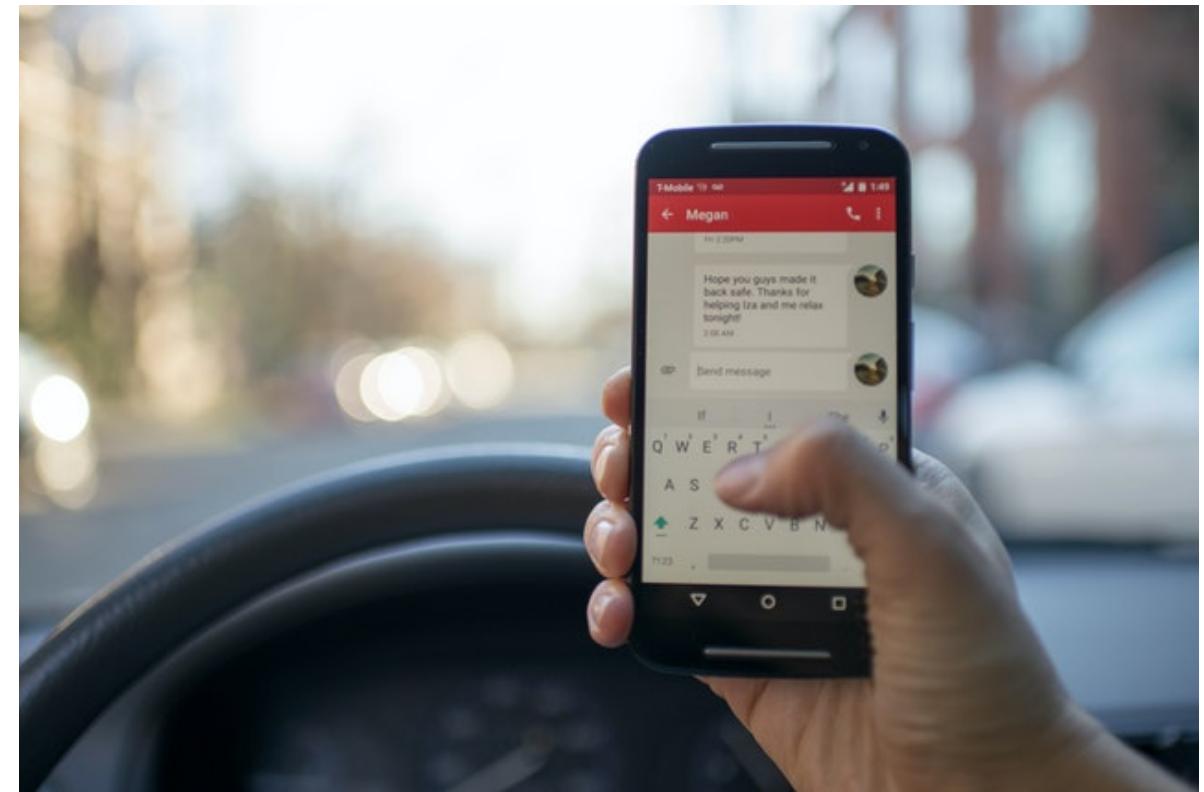


Photo by [Roman Pohorecki](#) from [Pexels](#)

# Redes generativas

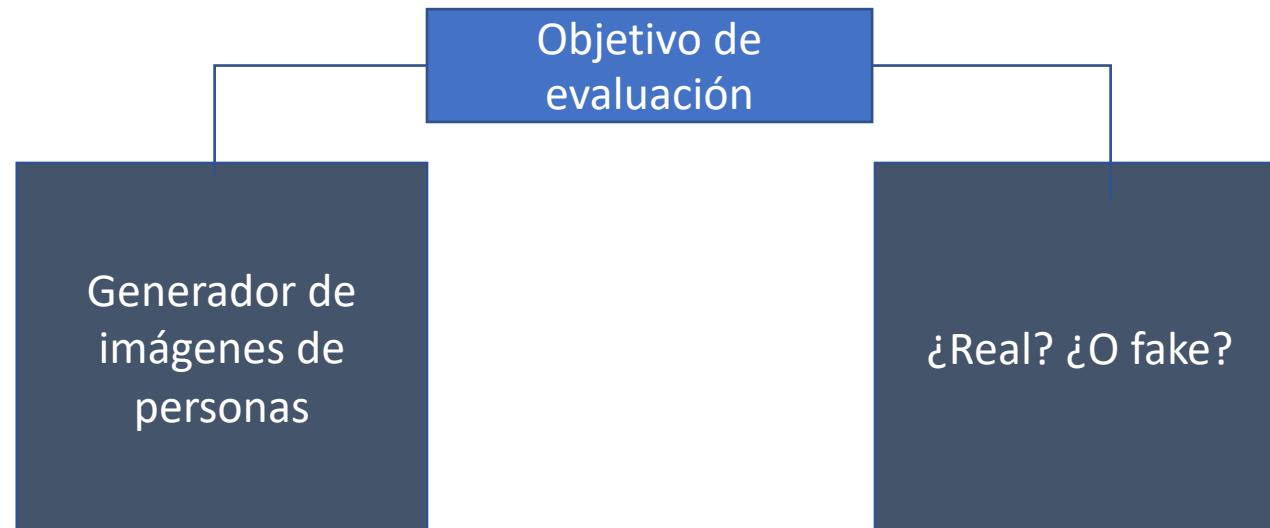
- En algunos casos queremos la predicción en sí como algo nuevo que quisiéramos que una máquina genere:
  1. Entrenamos un clasificador que reconoce un fenómeno (por ejemplo imágenes).
  2. Luego ese aprendizaje puede usarse para generar predicciones.

# iDeepfakes!



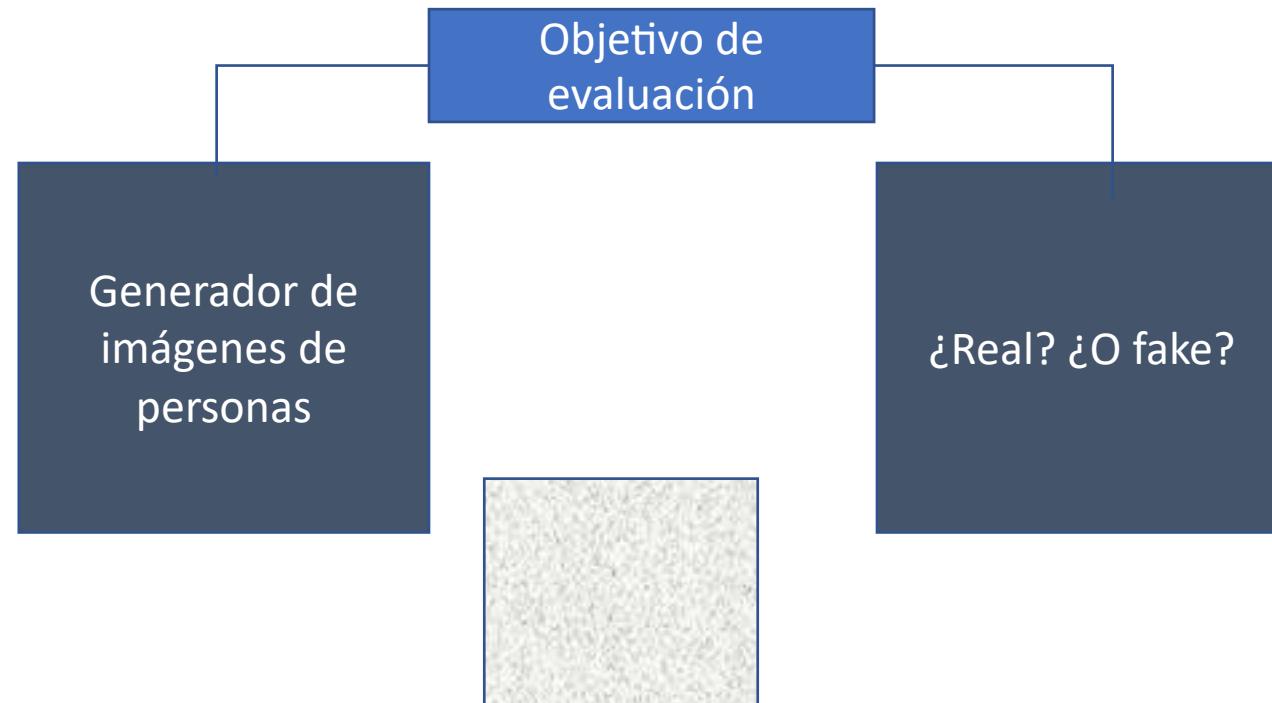
# Redes generativas adversarias

- Ponemos a competir a un clasificador y a una red generativa...



# Redes generativas adversarias

- Ponemos a competir a un clasificador y a una red generativa...



# Redes generativas adversarias

- Ponemos a competir a un clasificador y a una red generativa...



# Ética de la IA

Lo que hacemos puede afectar a las personas



Photo by [Agung Pandit Wiguna](#) from [Pexels](#)

# Las máquinas fallan

- Las predicciones son adivinanzas informadas en datos.
- Pueden errar (y lo hacen frecuentemente)
- Aunque... a veces se equivocan menos que los humanos...
- No todas las máquinas son interpretables: ¿tomamos decisiones confiando en ellas ciegamente?



# Podemos afectar a otras personas con máquinas que fallan

- Existen tres grandes consideraciones éticas de la inteligencia artificial en este momento (sintetiza Sandel):
- Privacidad y vigilancia
- Sesgo y discriminación
- El rol del juicio humano en la sociedad



# Privacidad y vigilancia



Los datos alteran los procesos de poder

- Ahora podemos analizar una cantidad masiva de datos en tiempo real
- Riesgos asociados a la estabilidad de los sistemas políticos
- Modelo de China: disyuntiva vigilancia vs seguridad
- Además, nuestros datos dicen mucho de nosotros: caso *Cambridge Analytics*

# Sesgo y discriminación

- A primera vista las máquinas parecen “objetivas”
- Pero se ha identificado que pueden aprender y acentuar los sesgos preexistentes en la sociedad
- Si un sistema presenta discriminación contra un grupo, puede que esta discriminación se manifieste en los datos
- Las máquinas aprenden que estos son comportamientos “naturales del sistema”
- Ej. Sistemas de admisión a universidades, aprobaciones de crédito, etc.



No automatizar decisiones críticas:  
información complementaria



# ¿Nos reemplazarán las máquinas?



- En este momento sólo tenemos IA débil (especializada en una tarea)
- La automatización es real: tareas cada vez más creativas están siendo aprendidas por máquinas
- Pero no es nueva: revolución industrial
- Algunos argumentan que por cada trabajo desplazado por las máquinas se crearán otros

# Enfoque centrado en humanos para la IA

---

BRIEFING

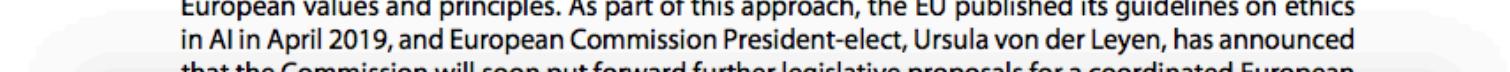
---



## EU guidelines on ethics in artificial intelligence: Context and implementation

### SUMMARY

The discussion around artificial intelligence (AI) technologies and their impact on society is increasingly focused on the question of whether AI should be regulated. Following the call from the European Parliament to update and complement the existing Union legal framework with guiding ethical principles, the EU has carved out a 'human-centric' approach to AI that is respectful of European values and principles. As part of this approach, the EU published its guidelines on ethics in AI in April 2019, and European Commission President-elect, Ursula von der Leyen, has announced that the Commission will soon put forward further legislative proposals for a coordinated European



# Enfoque centrado en humanos para la IA

- Siete principios que derivan en los lineamientos generales del Parlamento Europeo.

The key EU requirements for achieving trustworthy AI

- human agency and oversight
- robustness and safety
- privacy and data governance
- transparency
- diversity, non-discrimination and fairness
- societal and environmental well-being
- accountability

# ¡Manos al código!

Descarguemos las cosas

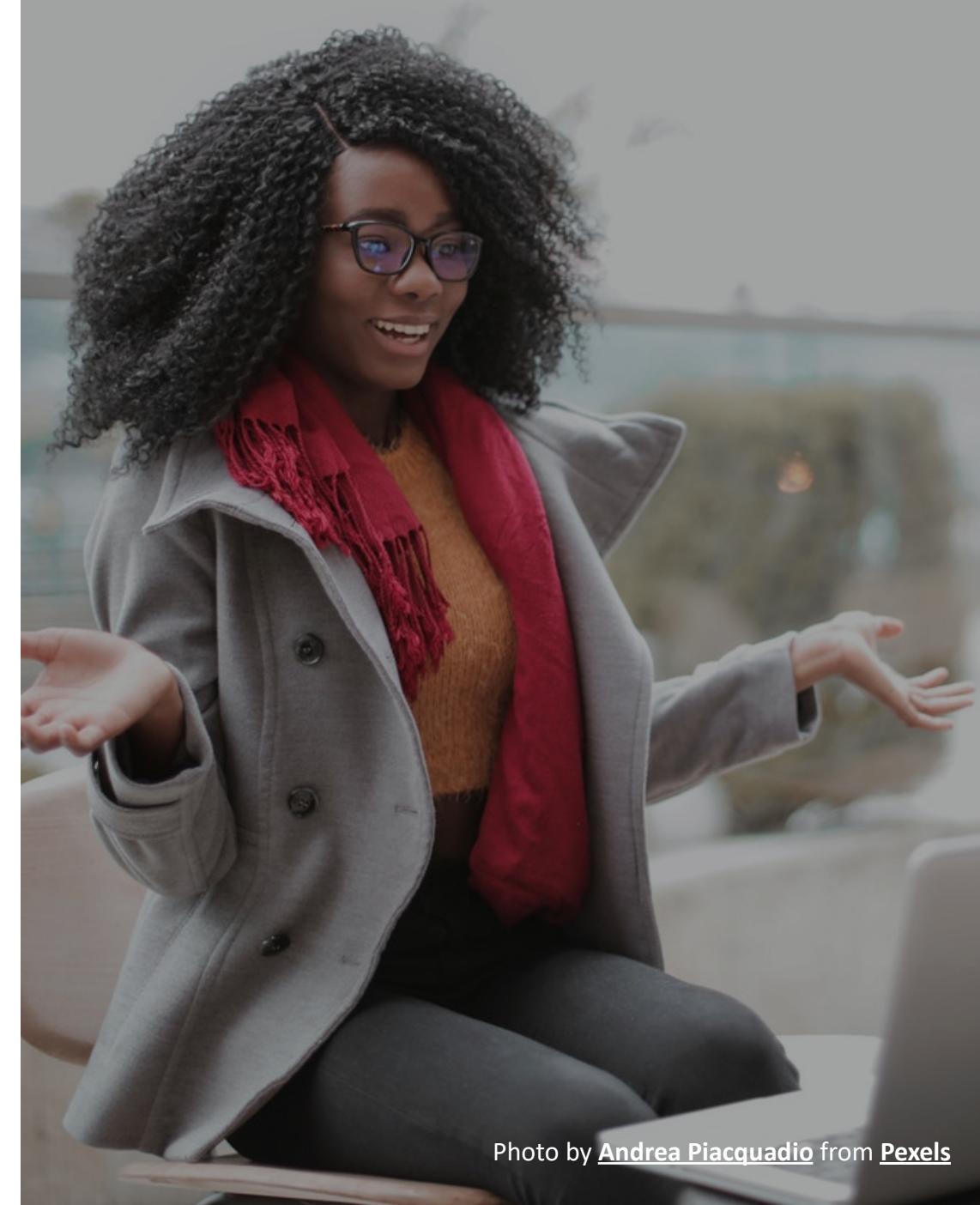


Photo by [Andrea Piacquadio](#) from [Pexels](#)