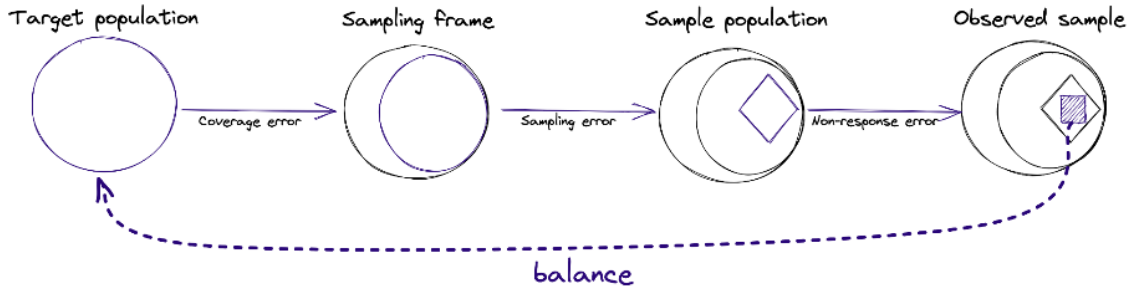
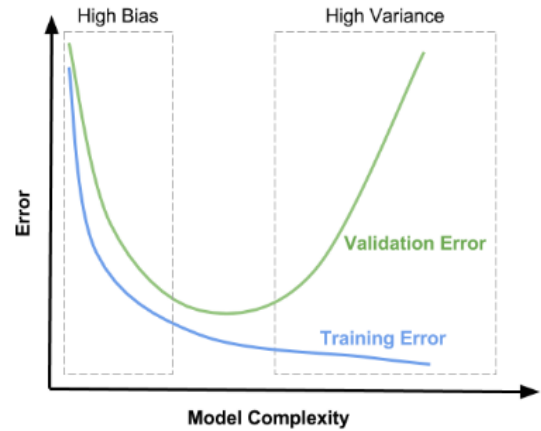


# Machine Learning- IA



## ¿Qué es el sesgo y cómo combatirlo?

# Índice

1.

¿Qué es el sesgo?

2.

Tendencias en Data Analytics para gestionar los sesgos

3.

Sesgos de selección ¿Cómo elegir una población representativa y trabajar con conjuntos de datos desbalanceados?

4.

Sesgo por falta de variables importantes. Caso de uso real

5.

I+D



# I 1. ¿Qué es el sesgo?

El problema del sesgo en IA empieza con la definición misma del término “sesgo”. Este término está sobrecargado y tiene un significado drásticamente distinto bajo diferentes contextos.

- **En estadística**, el sesgo es la diferencia entre el **valor esperado de un estimador (predicción) y su estimado (Valor real)**. El sesgo se refiere a **resultados que están sistemáticamente fuera de lugar**.

Si utilizamos el ejemplo de tiro con arco, e imaginamos que apuntamos incorrectamente. **El sesgo alto no significa que estemos disparando hacia todas partes** (eso en realidad significa que está generando una gran variación), pero que se están haciendo tiros perfectos, pero debajo de la diana todo el tiempo.



# I 1. ¿Qué es el sesgo?

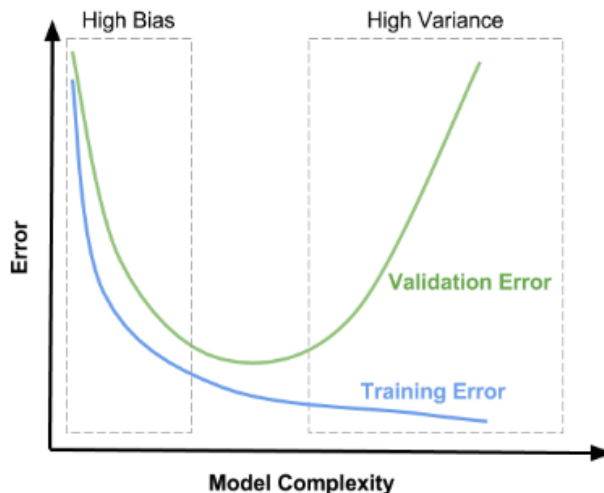
## Error de sesgo

**Los sesgos son los supuestos o simplificaciones que se hacen por un modelo para hacer que la función sea más fácil de aprender.**

- En general, los **algoritmos lineales tienen un alto sesgo** por lo que son rápidos de aprender y más fáciles de entender, pero en general son menos flexibles. A su vez, tienen un rendimiento predictivo más bajo en problemas complejos.

## Error de varianza

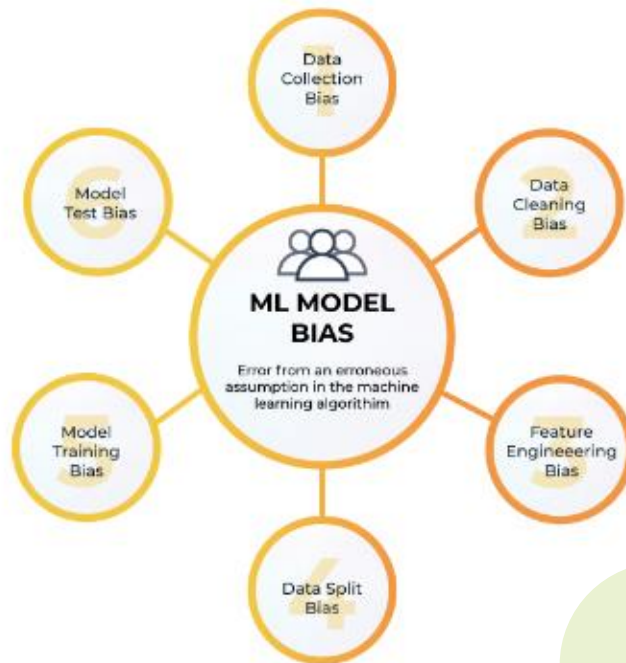
Los algoritmos de aprendizaje automático que tienen una **gran varianza** están fuertemente influenciados por las especificaciones de los conjuntos de datos de entrenamiento.



# I 1. ¿Qué es el sesgo?

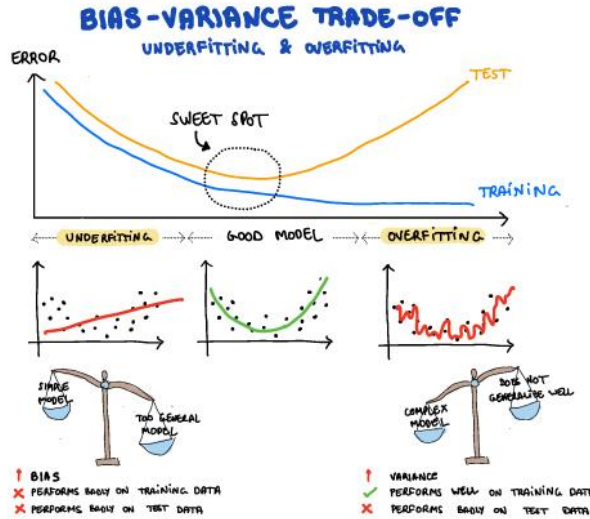
## Otros contextos:

- **En la recopilación de datos (y también en estadística)**, se refiere la manipulación de la recopilación de datos para que su muestra no sea representativa de su población de interés. **“Sesgo muestral”** es el nombre formal aquí. Este tipo de sesgo significa que no puede confiar en sus resultados estadísticos.
- **En lo que la mayoría de los expertos en IA piensan**, el sesgo se refiere al **sesgo algorítmico** que ocurre cuando un **sistema informático refleja los valores implícitos de los humanos que lo crearon**. (¿No es todo lo que los humanos crean un reflejo de valores implícitos?)



# 1. ¿Qué es el sesgo?

Dicotomía Sesgo Varianza está relacionada con cuan fiel sea la **muestra que tenemos de datos para entrenar el modelo** y la **complejidad del mismo**.



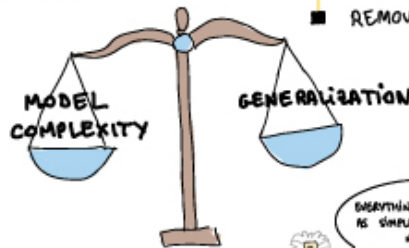
HOW TO TACKLE UNDERFITTING?

- ADD MORE FEATURES
- INCREASE MODEL COMPLEXITY
- INCREASE TRAINING TIME



HOW TO TACKLE OVERFITTING?

- EARLY STOPPING DURING TRAINING PHASE
- CROSS-VALIDATION
- ADD MORE DATA / DATA AUGMENTATION
- REGULARIZATION TECHNIQUES
- REMOVE FEATURES FROM DATA



EVERYTHING SHOULD BE MADE AS SIMPLE AS POSSIBLE, BUT NOT SIMPLER.

- ↓ BIAS    ↓ VARIANCE
- ✓ BALANCES UNDERFITTING & OVERFITTING
- ✓ GENERALIZES WELL

## 2. Tendencias para gestionar sesgos

---

Mencionamos a continuación técnicas disponibles a la hora de gestionar los sesgos en un proyecto real:

### 1. Preprocesamiento

Imputación de valores nulos

Ingeniería de variables

### 2. Corrección del algoritmo mediante modificación de función de coste

- Scale Positive Weights

### 3. Corrección del conjunto de entrenamiento (para conjuntos de datos desbalanceados, por ejemplo, tratar de igualar las clases para que no predominen unas sobre otras mediante Downsampling y Oversampling)

- SMOTE

- GAN (RN)

### 4. Métricas disponibles para la evaluación del sesgo

- Validación Cruzada

## 2. Tendencias para gestionar sesgos

---

### EJEMPLO CORRECCIÓN ALGORITMO

**SCALE POSITIVE WEIGHT** sobrerrepresentar la **clase minoritaria mediante corrección** por pesos class weight.

Existen diferentes opciones:

Generalmente **scale\_pos\_weight** es el ratio de elementos de la **clase negativa** dividido del número de elementos de la **clase positiva**.

Para un ejemplo de 100 observaciones/muestras

- 90 muestras de la clase negativa
- 10 muestras de la clase positiva

Para este caso de forma general el valor de scale\_pos\_weight sería 9

**scale\_pos\_weight = count(negative examples)/count(Positive examples)**  
**= 90 / 10 = 9**

En casos en los que exista un desbalanceo extremo de datos utilizar la siguiente fórmula

**scale\_pos\_weight = sqrt( count(negative examples) / count(Positive examples) ) = sqrt(90 / 10 ) = sqrt(9) = 3**



## 2. Tendencias para gestionar sesgos

---

### SCALE POSWEIGHT SPW

```
1 | from xgboost import XGBClassifier
2 |
3 | smallest_class_count = y_train.sum()
4 | largest_class_count = len(y_train) - smallest_class_count
5 | spw = largest_class_count / smallest_class_count
6 |
7 | model = XGBClassifier(booster='gbtree',
8 |                       objective='binary:logistic',
9 |                       max_depth=12, learning_rate=0.1,
10 |                      n_estimators=10,
11 |                      scale_pos_weight=spw,
12 |                      random_state=101,
13 |                      n_jobs=-1)
14 |
15 | model.fit(X_train, y_train)
```

## 2. Tendencias para gestionar sesgos

Modificar el conjunto de entrenamiento **Downsampling** → Submuestreo de la clase mayoritaria y **Oversampling** → Sobre muestreo de la clase minoritaria

### Mencionamos 1 Técnica → SMOTE 100

Utilizar **SMOTE para generar nuevos casos minoritarios**. Agrega el mismo número de casos minoritarios que había en el conjunto de datos original.

Dado que **SMOTE no aumenta el número de casos mayoritarios**, la proporción de casos de cada clase ha cambiado.

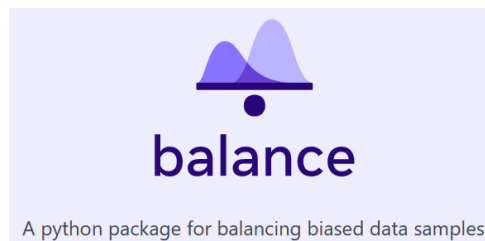
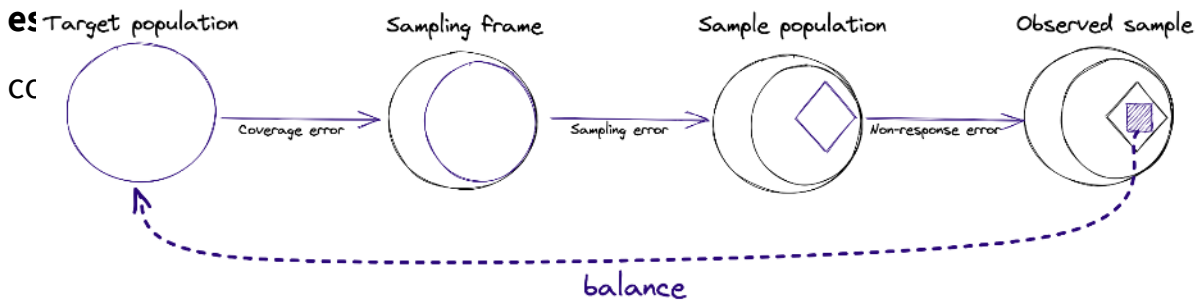
	Clase 0	Clase 1	total
Conjunto de datos original	570	178	748
(equivalente a porcentaje SMOTE = 0)	76%	24%	
Porcentaje SMOTE = 100	570	356	926
	62%	38%	
Porcentaje SMOTE = 200	570	534	1104
	52%	48%	
Porcentaje SMOTE = 300	570	712	1.282
	44%	56%	

## 2. Tendencias para gestionar sesgos

### Balance: Paquete de Python para el trabajo con conjuntos de datos sesgados

Su objetivo principal es **entrenar/ajustar y evaluar los pesos de una muestra**.

Para **cada elemento de la muestra (por ejemplo, una respuesta en una encuesta)**, la librería balance ajusta un peso que puede interpretarse (vagamente) **como el número de personas de la población objetivo que**

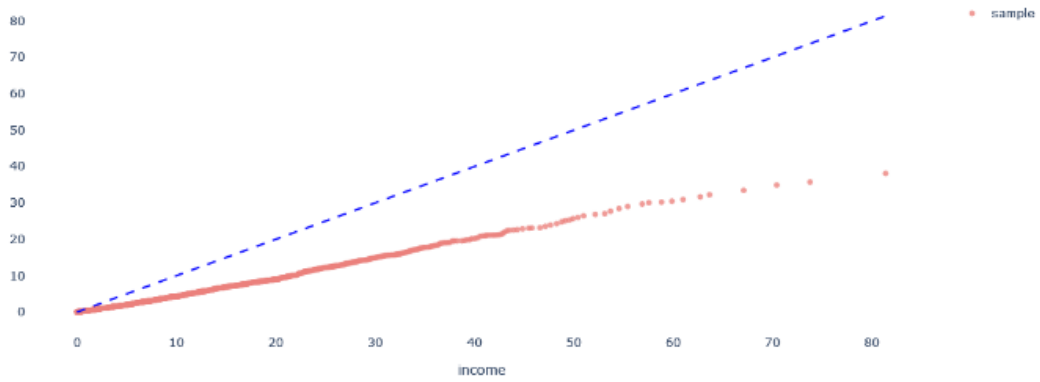


<https://github.com/facebookresearch/balance>

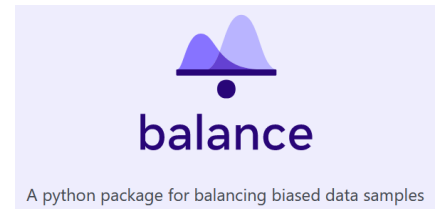
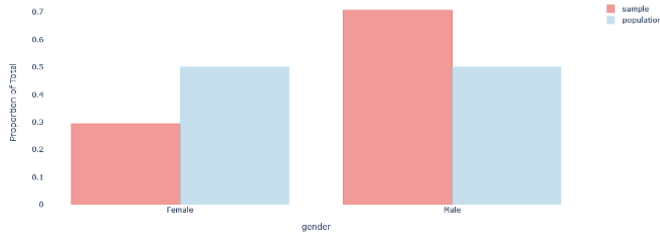
# 2. Tendencias para gestionar sesgos

## Muestra Población

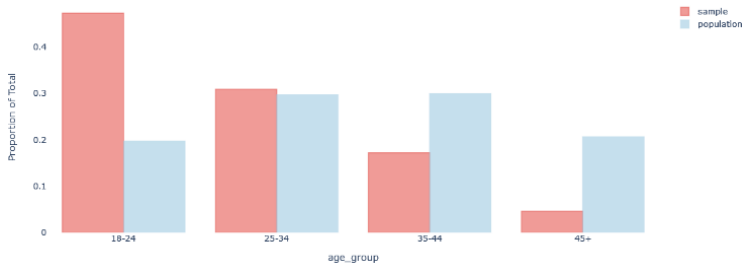
QQ Plots of Income



Sample Vs Target gender



Sample Vs Target age\_group



## 2. Tendencias para gestionar sesgos

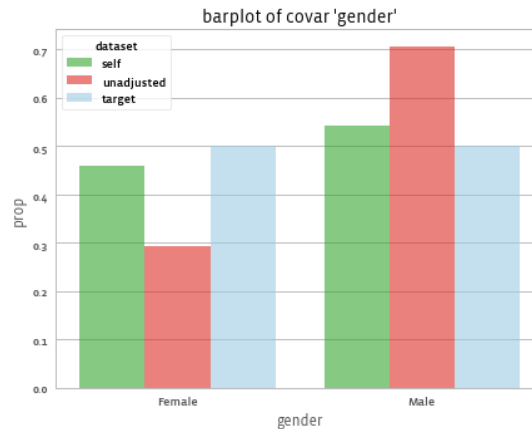
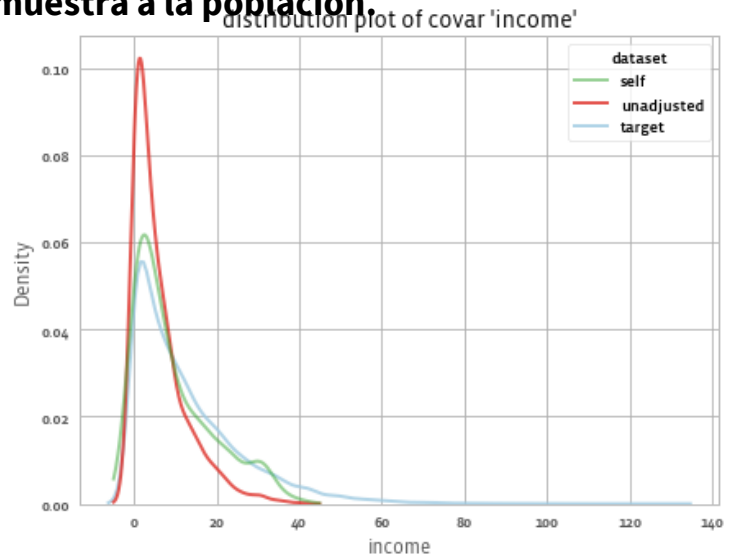
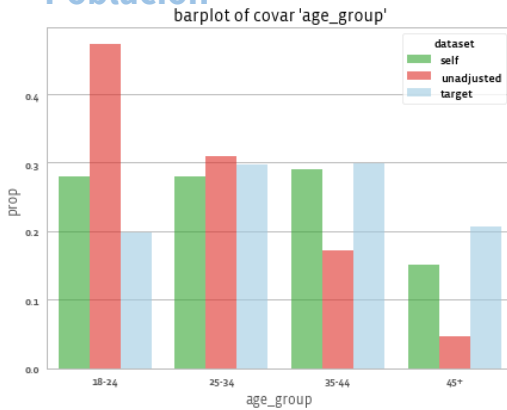
### Balance: Paquete de Python para el trabajo con conjuntos de datos sesgados

La librería nos ayuda a ajustar la muestra a la población.

- Muestra

- Muestra ajustada

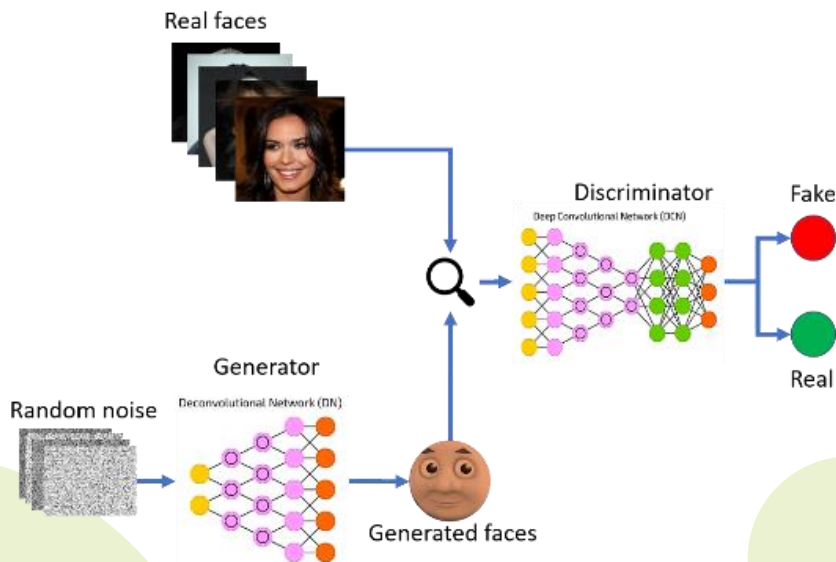
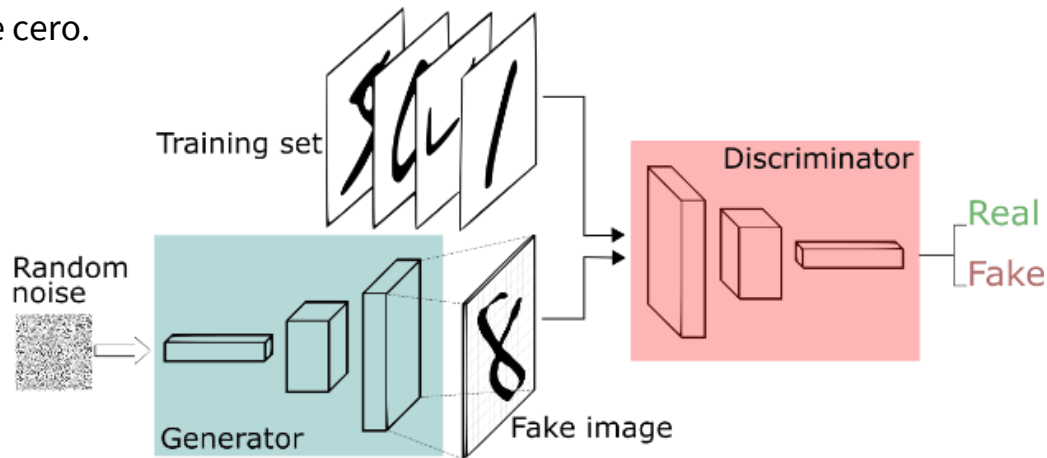
- Población



## 2. Tendencias para gestionar sesgos

### GAN- Generative Adversarial Networks - Redes Generativas Antagónicas

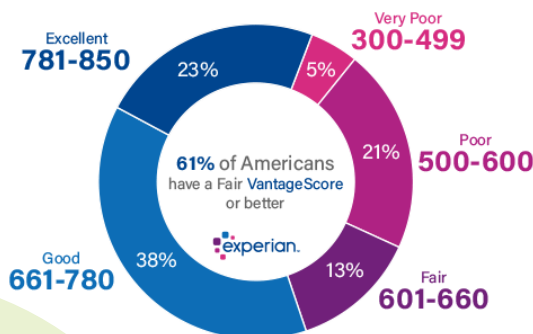
Actúan sobre la representación en generación de datos artificiales (con redes neuronales). GAN generación de caras falsas a partir de otros datos. Las GAN confieren a los algoritmos especie de sentido de imaginación para poder crear algo nuevo desde cero.



## 2. Tendencias para gestionar sesgos

### Clasificación del estado de un préstamo:

- **LoanID, CustomerID** : Id del cliente y del préstamo.
- **LoanStatus: Pagado (Fully Paid) o Impagado (Charged Off).**
- **CurrentLoanAmount:** Importe actual del préstamo
- **Term:** Corto o Largo Plazo
- **CreditScore:** Score de crédito (ejemplo Experian).
- **AnnualIncome:** Ingresos Anuales
- **Yearsincurrentjob:** Antigüedad en el trabajo.
- **HomeOwnership:** Hipoteca, Tiene Casa Propia, Alquiler.
- **Purpose:** Propósito del crédito.

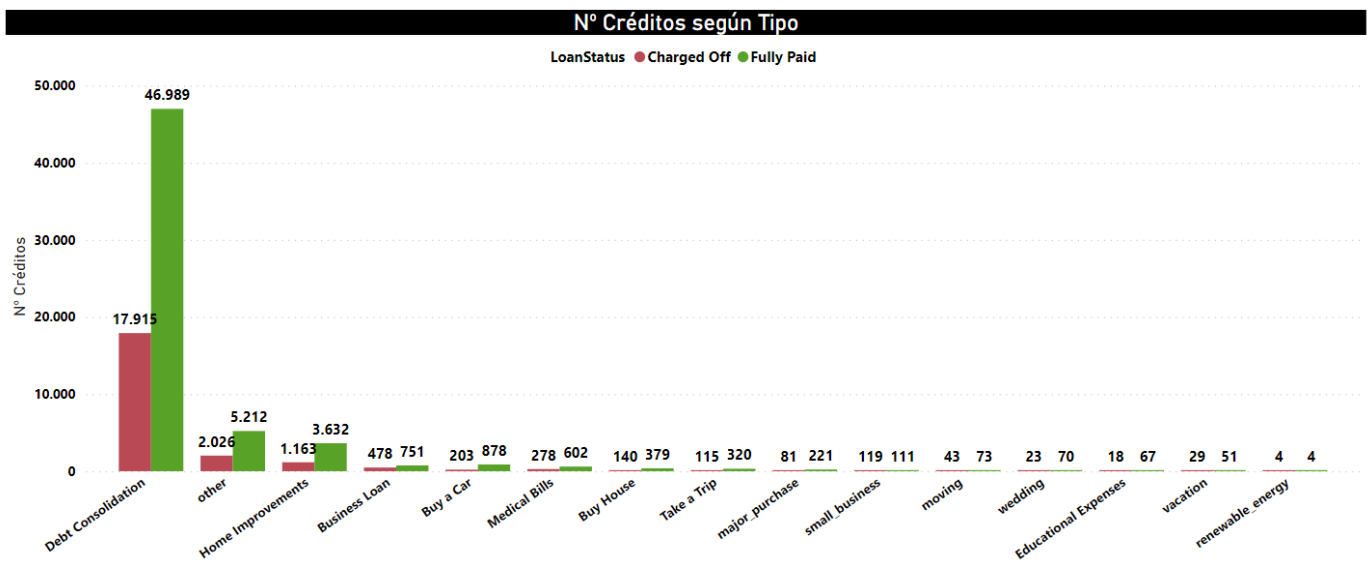


- **MonthlyDebt:** Cuota Mensual
- **YearsofCreditHistory**
- **Monthssincelastdelinquent:** Meses desde la última en mora.
- **NumberofOpenAccounts:** N° de cuentas abiertas.
- **NumberofCreditProblems:** N° de problemas de crédito previos.
- **CurrentCreditBalance:** Crédito actual.
- **MaximumOpenCredit:** Máximo Crédito.
- **Bankruptcies:** Bancarrotas.
- **TaxLiens:** Gravámenes fiscales.



## 2. Tendencias para gestionar sesgos

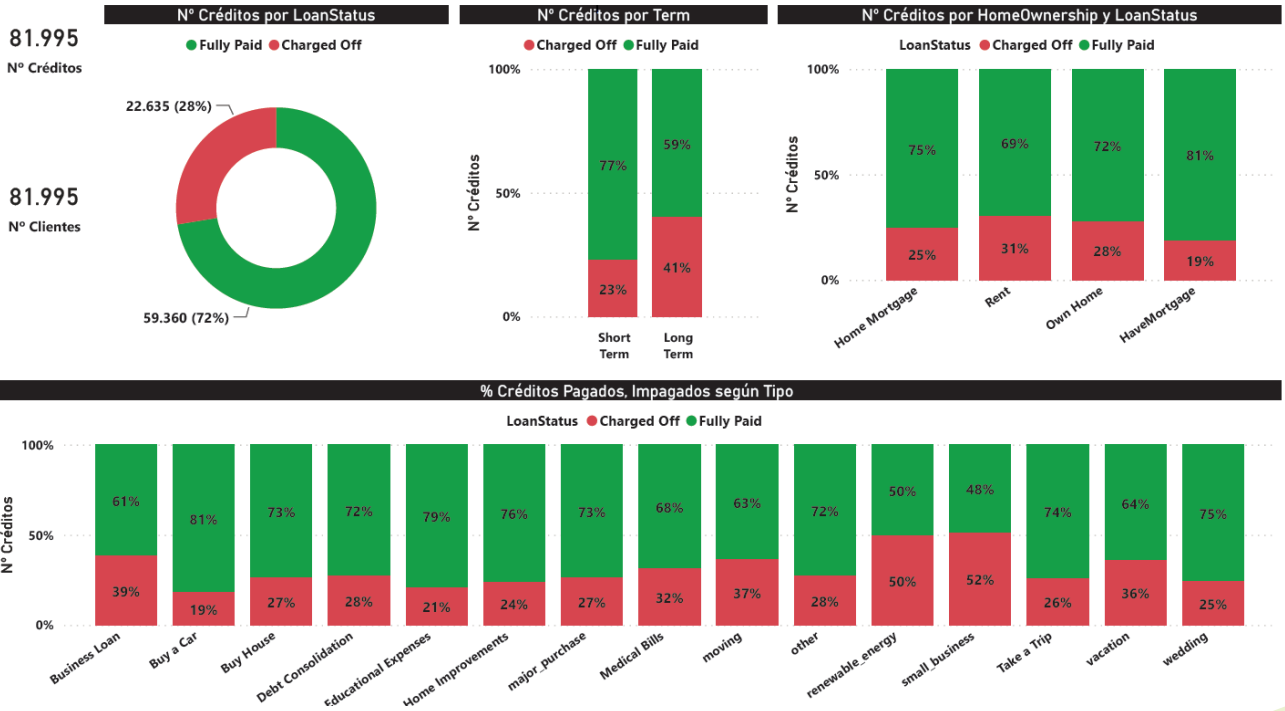
### Clasificación del estado de un préstamo – N° de préstamos por tipo y estado





# 2. Tendencias para gestionar sesgos

## Clasificación del estado de un préstamo – Análisis Visual



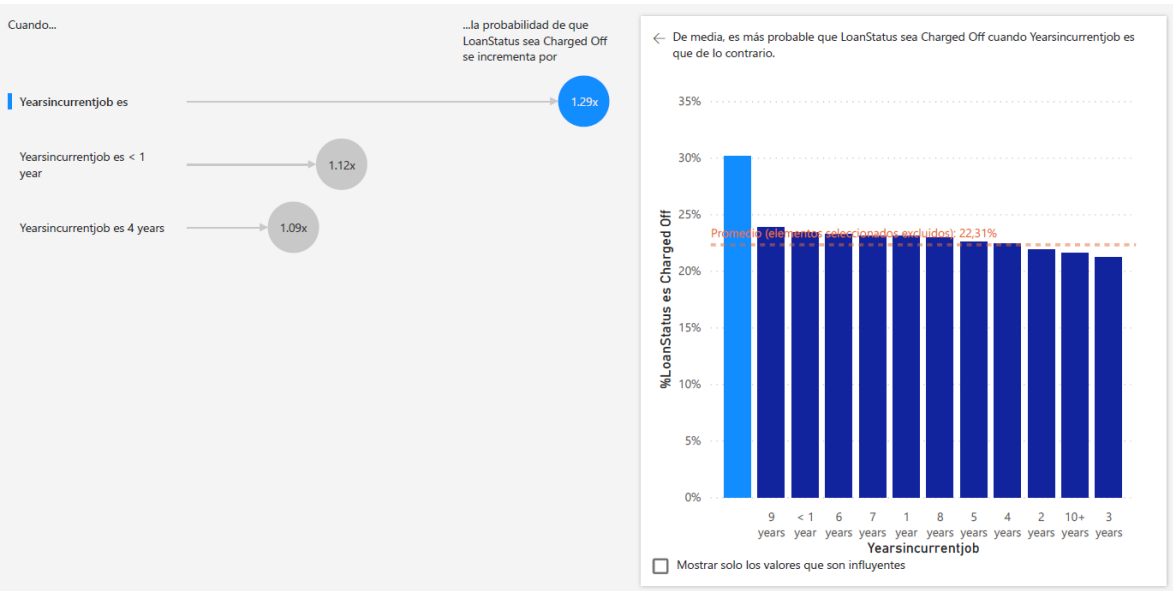
## 2. Tendencias para gestionar sesgos

### Clasificación del estado de un préstamo – Key Influencers en función de la antigüedad en el trabajo

Elementos influyentes clave Segmentos principales



Qué influye en LoanStatus para ser Charged Off ?

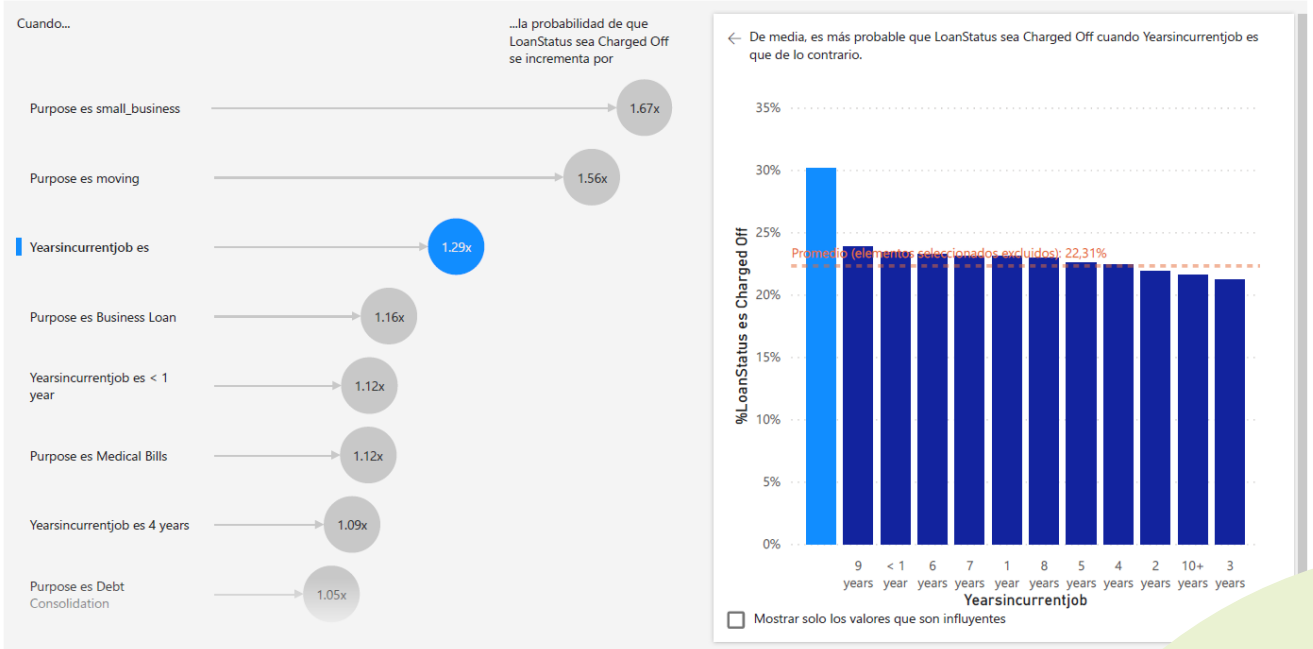


## 2. Tendencias para gestionar sesgos

### Clasificación del estado de un préstamo – Key Influencers en función de la antigüedad en el trabajo y propósito del crédito

Elementos influyentes clave Segmentos principales

Qué influye en LoanStatus para ser Charged Off ?



**Sesgo de Selección:** Error en la elección de los individuos o grupos que participan en un estudio. Idealmente los **sujetos de un estudio deberían ser muy similares entre sí y a la población general de la cual se seleccionan** (por ejemplo, todos los individuos con la misma enfermedad o afección). Si hay diferencias importantes, los resultados del estudio pueden no ser válidos.

Ejemplo: tratar de analizar la altura de la población de la Comunidad de Madrid utilizando como muestra la cantera de jugadores y jugadoras profesionales, juveniles y cadetes del **Club Baloncesto Estudiantes**.



**Comunidad  
de Madrid**



El **sesgo del superviviente** consiste en concentrarse en las personas o cosas que superaron un proceso de selección pasando por alto a aquellas que no lo hicieron, típicamente por su falta de visibilidad.

El sesgo de supervivencia **puede llevar a creencias demasiado optimistas porque se ignoran los fracasos**, como cuando se excluyen de los análisis del rendimiento financiero empresas que ya no existen.

Por ejemplo, **si 3 de 5 cinco estudiantes con las mejores notas universitarias fueron al mismo instituto** eso puede llevar a creer que ese instituto debe ofrecer una educación excelente. Esto podría ser cierto, pero la cuestión no puede ser respondida sin observar cuáles fueron las calificaciones de todos los demás estudiantes de ese instituto, no solo de los que "superaron" el proceso de selección de los 5 mejores.



En la 2ª Guerra Mundial la RAF Británica buscaba mejorar sus aviones analizando los impactos balísticos.

Imagina *que los impactos de bala en los aviones que consiguen «sobrevivir», los que no fueron derribados por los cazas enemigos, son como el avión de la imagen:*

¿Qué zonas del avión reforzarías para su blindaje?

### Una respuesta fácil sería: **TODO**

Lo ideal sería poder blindar todas las partes del avión, pero eso tenía dos problemas:

- 1) Costaba dinero
- 2) El blindaje debía ser **óptimo**: blindar lo justo ya que sino el avión pesaría demasiado

**Hay que decidir el menor blindaje posible para reforzar lo máximo.**

Por aquel entonces los ingleses recibían el apoyo de un grupo de estadísticos de Columbia, «**Statistical Research Group**» (SRG).

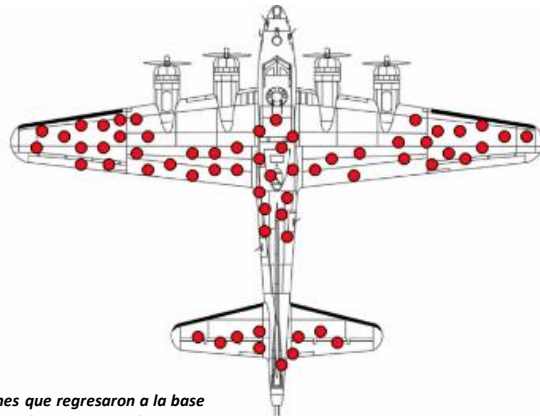
Los trabajos de este grupo propiciaron el nacimiento de una rama de la estadística, la **Investigación Operativa: maximizar/minimizar beneficios al menor coste posible.**

Los militares hicieron mapas de impactos en los aviones que volvían: muchos impactos en fuselaje, pocos en motores.

Esos dibujos se mostraron al SRG, preguntando a los sesudos matemáticos que parte de alas y fuselaje tenían que reforzarse.

Quizás pienses que la respuesta es obvia: «**Reforzar las áreas de fuselaje y alas donde más impactos había en que volvían**»

**Pero... la respuesta no fue la esperada.**



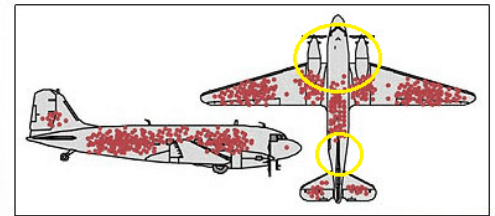
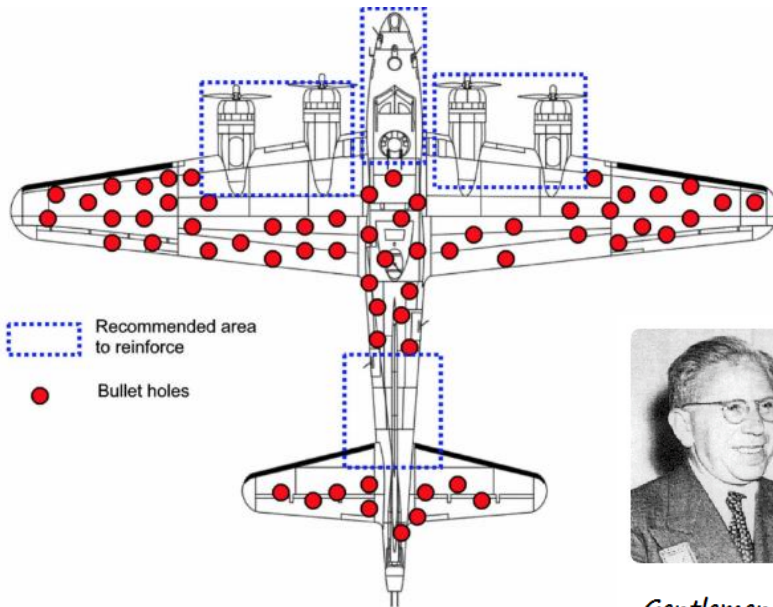
Aviones que regresaron a la base  
(Supervivientes)

La respuesta de un estadístico, Abraham Wald, les dejó blancos:

**¡Había que reforzar las zonas sin impactos!**

Abraham Wald, uno de los estadísticos del SRG, concluyó que lo óptimo era blindar **morro, motores y la parte media.**

Wald se preguntó: si estos son los **de los aviones que volvían... ¿Dónde están los agujeros de los aviones que faltan?**



*Gentlemen, you need to put more armour-plate where the holes aren't because that's where the holes were on the airplanes that didn't return - Abraham Wald 1942.*

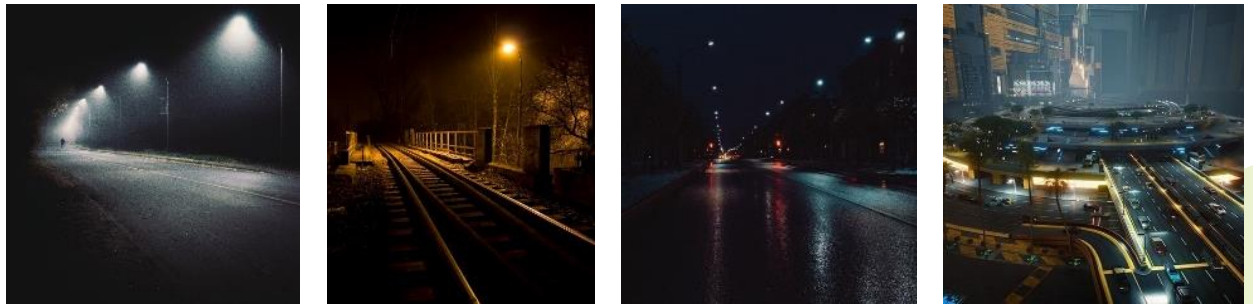
Si provisionamos los siguientes conjuntos etiquetados con Tanques y No Tanques **¿qué es lo que va a aprender una red neuronal?**

Etiqueta Si Tanque



*The Neural Net Tank Urban Legend*

Etiqueta No Tanque





Nuestro algoritmo diferenciaría el día y la noche, pero no detectaría correctamente los tanques, importante analizar resultados y fijarse en activaciones de la primera capa.



**Etiqueta Si Tanque**  
→ DIA

**Etiqueta No Tanque**  
→ NOCHE

Etiquetadas Jan Oblak



**Etiquetadas Macos Llorente**



00000000.jpg



00000001.jpg



00000002.jpg



00000003.jpg



00000004.jpg



00000010.jpg



00000011.jpg



00000012.jpg



00000013.jpg



00000014.jpg



00000021.jpg



00000022.jpg



00000023.jpg



00000024.jpg



00000025.jpg



00000031.jpg



00000032.jpg



00000033.jpg



00000034.jpg



00000035.jpeg



00000005.jpg



00000006.jpg



00000007.jpg



00000008.jpg



00000009.jpg



00000015.jpg



00000016.jpg



00000017.jpg



00000018.jpg



00000019.jpg



00000026.jpg



00000027.jpg



00000028.JPG



00000029.jpg



00000030.jpg



00000036.jpeg



00000037.jpg



00000038.jpg



00000039.jpg



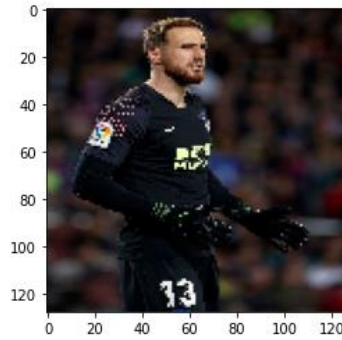
00000040.jpg

Escribir por el chat de Teams o quitar el silencio e indicar que partes de las imágenes van a ser utilizadas por la red neuronal para clasificar las imágenes de los jugadores del Atlético de Madrid.

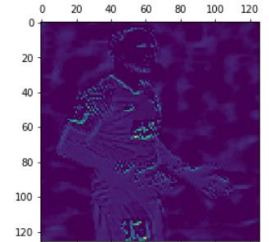
Os toca participar!!







3) 1º capa activaciones

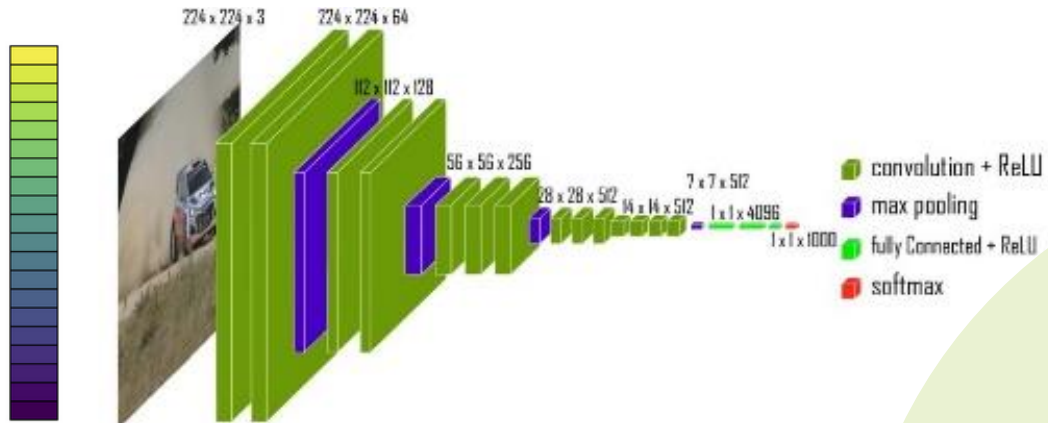


4) 2º capa activaciones

2) Imagen

redimensionada 128x128

1) Imagen original



## Early Stopping

La idea del **Early stopping** es sencilla: detener el entrenamiento en el momento que se observe un incremento en el valor del error de validación. Al hacer esto se logra entrenar el modelo para que este error sea mínimo y, por tanto, su precisión sea máxima, reduciendo así el sobreajuste (**overfitting**).

**En el ejemplo inferior en la época 7 → 0,6586 de validation accuracy y no mejora en las 3 siguientes (paciencia 3) por lo que el entrenamiento se para.**

```

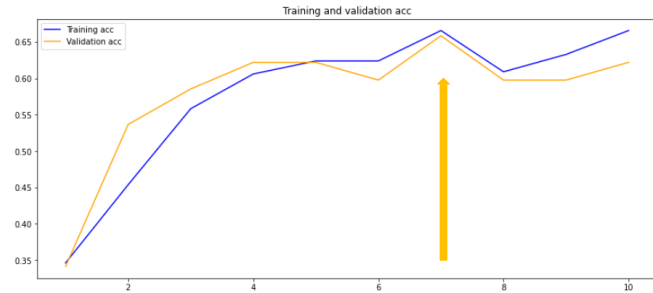
from tensorflow.keras.callbacks import EarlyStopping
train_steps = 15 #train_generator.n//train_generator.batch_size
validation_steps = 15 # val_generator.n//val_generator.batch_size

es = EarlyStopping(monitor='val_acc', mode='max', verbose=1, patience=3)
model = create_simple_model(IMG_SHAPE, NUM_CLASSES)

history = model.fit_generator(train_generator, epochs=20,
                             validation_data=val_generator, callbacks=[es], verbose=1)
    
```

Epoch 1/20	11/11 [=====]	9s	798ms/step	loss: 3.9199	acc: 0.3463	val_loss: 1.2573	val_acc: 0.3415
Epoch 2/20	11/11 [=====]	8s	691ms/step	loss: 1.1143	acc: 0.4537	val_loss: 0.9701	val_acc: 0.5366
Epoch 3/20	11/11 [=====]	8s	714ms/step	loss: 0.9311	acc: 0.5582	val_loss: 0.9136	val_acc: 0.5854
Epoch 4/20	11/11 [=====]	8s	707ms/step	loss: 0.8492	acc: 0.6060	val_loss: 0.8447	val_acc: 0.6220
Epoch 5/20	11/11 [=====]	8s	688ms/step	loss: 0.7866	acc: 0.6239	val_loss: 0.7964	val_acc: 0.6220
Epoch 6/20	11/11 [=====]	8s	691ms/step	loss: 0.7629	acc: 0.6239	val_loss: 0.8343	val_acc: 0.5976
Epoch 7/20	11/11 [=====]	8s	726ms/step	loss: 0.7271	acc: 0.6657	val_loss: 0.8431	val_acc: 0.6585
Epoch 8/20	11/11 [=====]	8s	712ms/step	loss: 0.7716	acc: 0.6090	val_loss: 0.8347	val_acc: 0.5976
Epoch 9/20	11/11 [=====]	8s	724ms/step	loss: 0.7182	acc: 0.6328	val_loss: 0.8002	val_acc: 0.5976
Epoch 10/20	11/11 [=====]	8s	713ms/step	loss: 0.7290	acc: 0.6657	val_loss: 0.7600	val_acc: 0.6220

Epoch 00010: early stopping



## Existen múltiples artículos del uso de Inteligencia Artificial en el área de la salud.

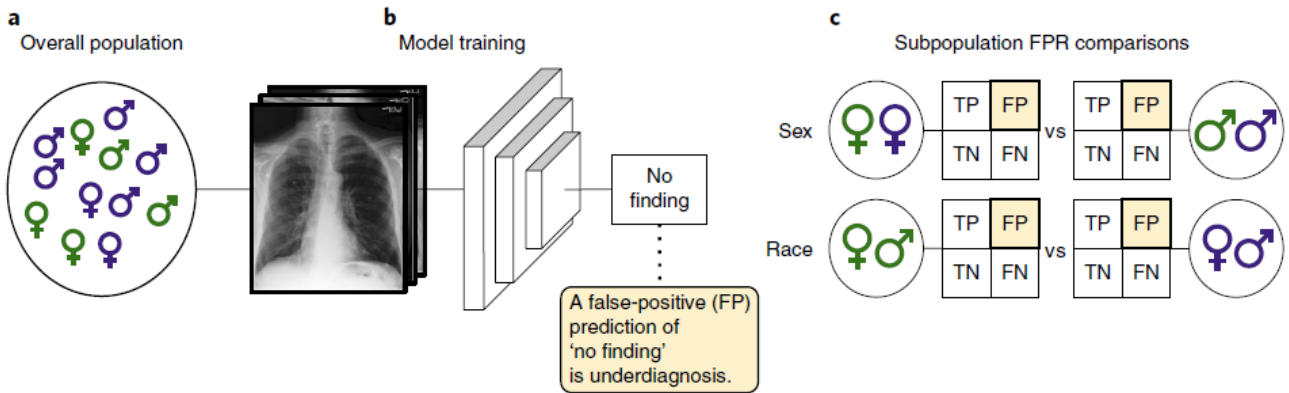
Por ejemplo, en un estudio se consigue adivinar la raza de una persona con tan sólo ver una radiografía, y los científicos no se explican cómo.

Los radiólogos están asombrados porque **la identificación racial no se puede realizar a través de este tipo de pruebas y, sin embargo, los algoritmos demostraron ser capaces de detectarla con precisión** durante un estudio que incluía a pacientes negros, blancos y asiáticos.

Como ya sucede en las tecnologías de reconocimiento facial, **los resultados han suscitado la preocupación de que la IA pueda amplificar la desigualdad en la atención sanitaria**, donde los estudios muestran que los pacientes negros y otros grupos raciales suelen recibir una peor atención.



Estudio inicial de inteligencia artificial de las imágenes de rayos X del pecho indicaba que tenían más probabilidades de **pasar por alto signos de enfermedad en las personas negras.**



*¿podría el software de computadora aplicar involuntariamente un sesgo racial al estudiar imágenes como estas?*

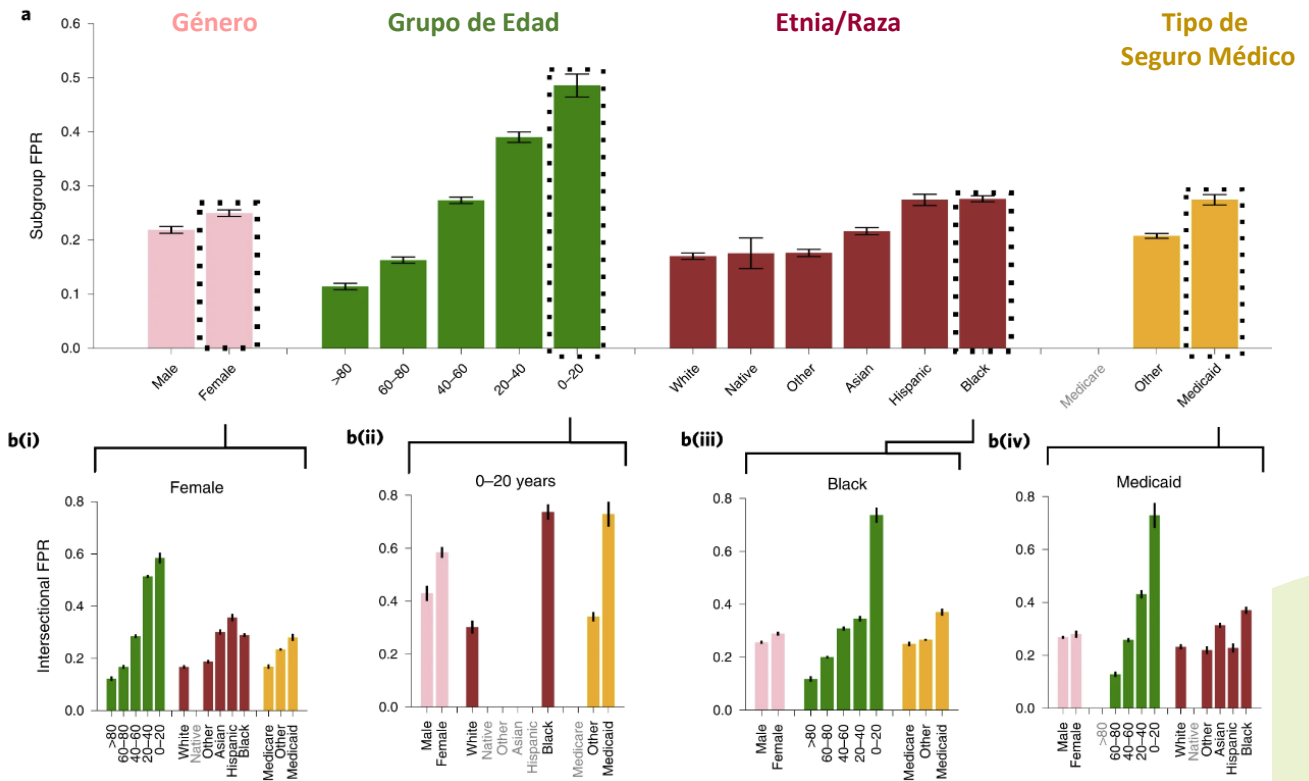
[Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations](#)



Sesgo de subdiagnóstico en función del:

- Género
- Grupo de Edad
- Etnia/Raza
- Tipo de Seguro Médico

From: [Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations](#)



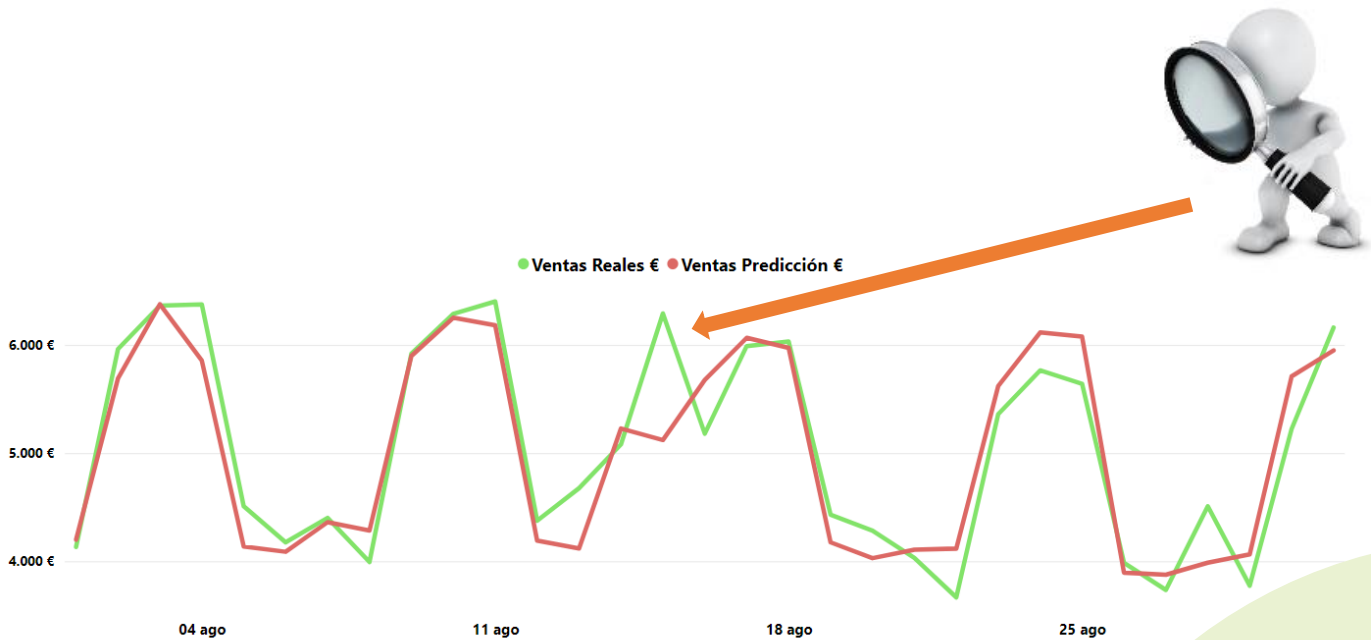
- 
- *Do no harm: a roadmap for responsible machine learning for health care. Nat. Med.*, (2019).
  - *Implementing machine learning in health care: addressing ethical challenges. N. Engl. J. Med.*, (2018).
  - *Treating health disparities with artificial intelligence. Nat. Med.*, (2020).
  - *Dissecting racial bias in an algorithm used to manage the health of populations. Science*, (2019).
  - *Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. Proc. Natl Ac. Sci. USA* (2020).
  - *CheXclusion: fairness gaps in deep chest X-ray classifiers. (World Scientific Publishing, 2021).*
  - *Hidden in plain sight: reconsidering the use of race correction in clinical algorithms. N. Engl. J. Med.*, (2020).
  - *Race and gender disparities are evident in COPD underdiagnoses across all severities of measured airflow obstruction. Chronic Obstr. Pulm. Dis.* (2018).
  - *Exploring gender disparities in time to diagnosis. In Machine Learning for Health (ML4H) at NeurIPS 2020* (2020).
  - *The quality of care delivered to patients within the same hospital varies by insurance type. Health Aff. (Milwood)* (2013).



## Caso de uso real ejemplo distribución:

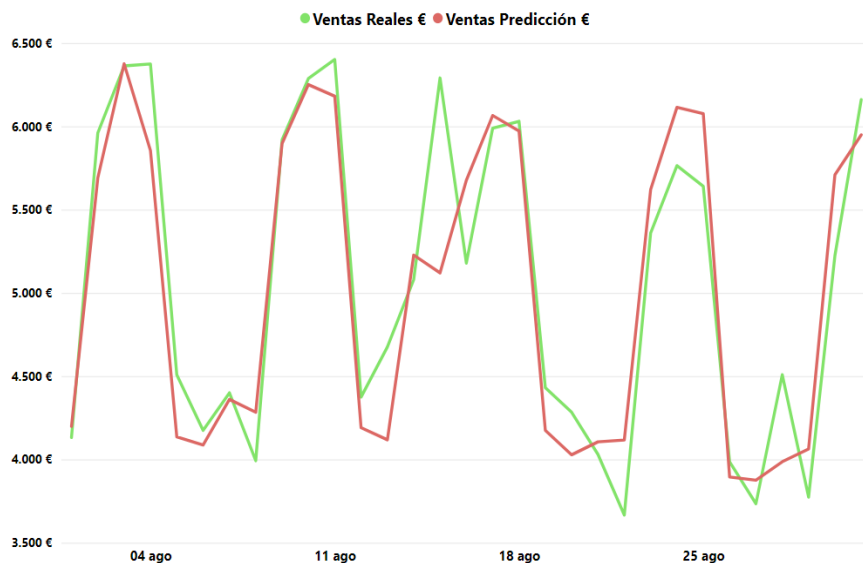
Ventas diarias en un determinado restaurante durante el mes de agosto.

Buenas predicciones en general **excepto día 15 de Agosto.**



Observamos lo anterior detallado día a día:

Fecha	Predicción	Reales	Error Absoluto	% Acierto
15 agosto	5.120 €	6.290 €	1.170 €	81,4 %
13 agosto	4.118 €	4.675 €	557 €	88,1 %
28 agosto	3.986 €	4.508 €	522 €	88,4 %
04 agosto	5.855 €	6.374 €	519 €	91,9 %
16 agosto	5.676 €	5.179 €	497 €	90,4 %
30 agosto	5.709 €	5.223 €	486 €	90,7 %
22 agosto	4.117 €	3.666 €	451 €	87,7 %
25 agosto	6.076 €	5.641 €	436 €	92,3 %
05 agosto	4.136 €	4.508 €	372 €	91,8 %
24 agosto	6.114 €	5.764 €	350 €	93,9 %
08 agosto	4.283 €	3.992 €	291 €	92,7 %
29 agosto	4.063 €	3.774 €	289 €	92,3 %
02 agosto	5.690 €	5.960 €	270 €	95,5 %
23 agosto	5.620 €	5.360 €	260 €	95,1 %
19 agosto	4.175 €	4.431 €	256 €	94,2 %
20 agosto	4.028 €	4.283 €	255 €	94,0 %
11 agosto	6.181 €	6.402 €	220 €	96,6 %
31 agosto	5.949 €	6.160 €	211 €	96,6 %
12 agosto	4.191 €	4.374 €	184 €	95,8 %
14 agosto	5.227 €	5.080 €	147 €	97,1 %
27 agosto	3.875 €	3.734 €	141 €	96,2 %
26 agosto	3.894 €	3.986 €	91 €	97,7 %
06 agosto	4.087 €	4.175 €	87 €	97,9 %
17 agosto	6.065 €	5.989 €	76 €	98,7 %
21 agosto	4.106 €	4.030 €	76 €	98,1 %
01 agosto	4.199 €	4.131 €	67 €	98,4 %
18 agosto	5.972 €	6.030 €	58 €	99,0 %
07 agosto	4.361 €	4.400 €	39 €	99,1 %
10 agosto	6.251 €	6.287 €	37 €	99,4 %
09 agosto	5.897 €	5.919 €	21 €	99,6 %
03 agosto	6.376 €	6.363 €	13 €	99,8 %
<b>Total</b>	<b>155.399 €</b>	<b>156.691 €</b>	<b>8.451 €</b>	<b>94,6 %</b>



**A raíz de las primeras ejecuciones del algoritmo se decidió incorporar diferentes variables que afectaban a las ventas:**

- Meteorología (Temperatura media, precipitaciones)
- Festivos Locales, Autonómicos y Provinciales
- Eventos cercanos al restaurante (Partidos de Fútbol, Liga ACB, Conciertos, Ferias)



- Festivos
- Nacionales
- Regionales
- Locales

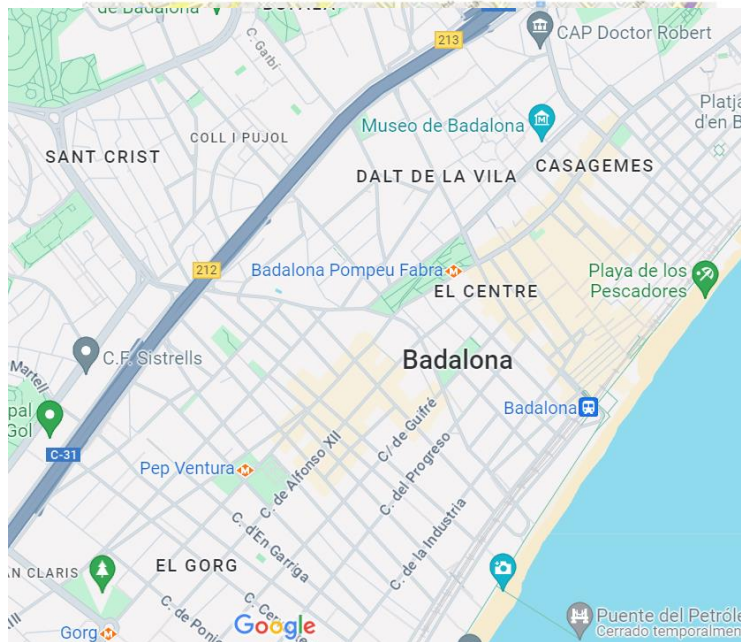


SUN	MON	TUE	WED	THU	FRI	SAT
☁	☁	☀	☀	☀	☀	☀
81%	31%		15.1/11.1	11.5/11.1		
77	78	82	85	86	89	91
60	58	60	62	65	69	



Restaurante afectado por musicales, circo y conciertos cercanos:

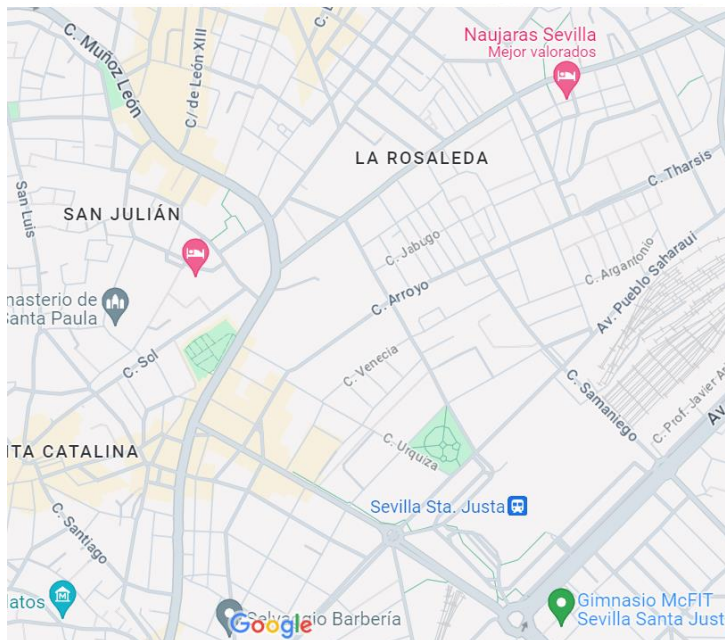
Restaurante	Categoría	Localización	Día	distancia	
ticketmaster®	Musical	Circo Raluy Tortosa	17/10/2019	403	
	Musical	Circo Raluy Tortosa	18/10/2019	403	
	Musical	Circo Raluy Tortosa	19/10/2019	403	
	Musical	Circo Raluy Tortosa	20/10/2019	403	
	Circo	Circo Raluy Tortosa	22/10/2019	403	
	Musical	Circo Raluy Tortosa	22/10/2019	403	
	Circo	Circo Raluy Tortosa	23/10/2019	403	
	Musical	Circo Raluy Tortosa	23/10/2019	403	
	Circo	Circo Raluy Tortosa	24/10/2019	403	
	Musical	Circo Raluy Tortosa	24/10/2019	403	
	Circo	Circo Raluy Tortosa	25/10/2019	403	
	Musical	Circo Raluy Tortosa	25/10/2019	403	
	Circo	Circo Raluy Tortosa	26/10/2019	403	
	Musical	Circo Raluy Tortosa	26/10/2019	403	
	Circo	Circo Raluy Tortosa	27/10/2019	403	
	Musical	Circo Raluy Tortosa	27/10/2019	403	
	StubHub International	Concierto	Harlem Jazz Club	23/10/2019	721
		Concierto	Harlem Jazz Club	07/11/2019	721
Concierto		Harlem Jazz Club	15/11/2019	721	
Concierto		Harlem Jazz Club	22/11/2019	721	
Concierto		Harlem Jazz Club	29/11/2019	721	
Curso		JARDIN DE HARA	26/09/2019	847	
Concierto		Sala Jamboree	16/10/2019	990	
Concierto		Sala Jamboree	17/10/2019	990	
Concierto		Sala Jamboree	18/10/2019	990	



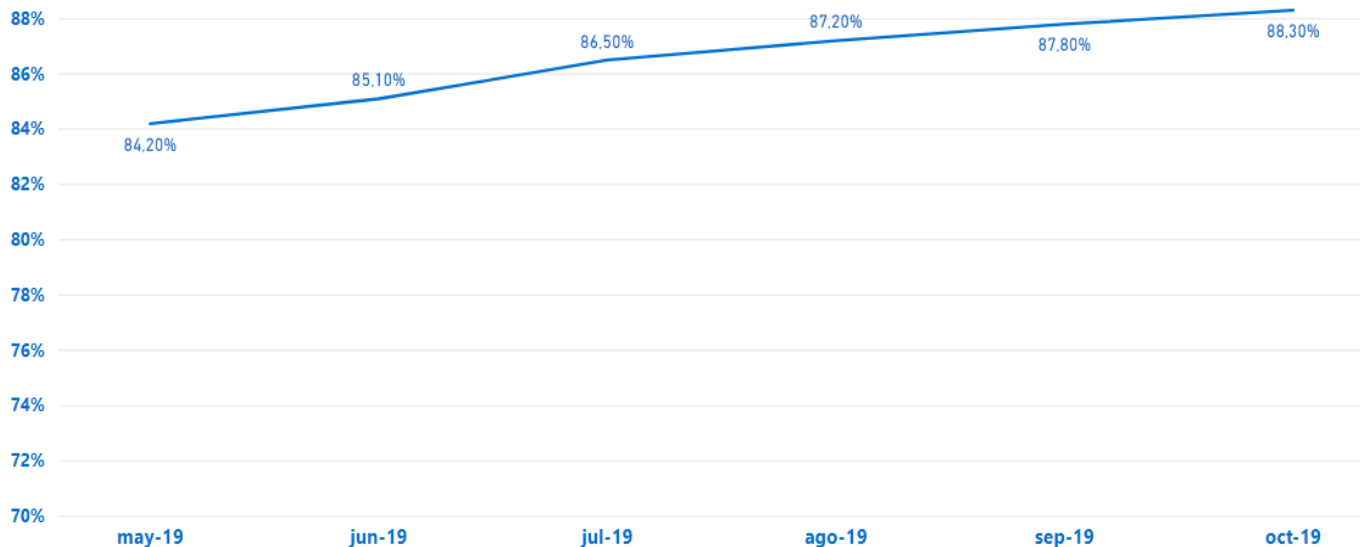


## Restaurante acertado por eventos deportivos:

Restaurante	Categoría	Localizacion	Día	distancia
21093	Europa League	Estadio Ramón Sánchez-Pizjuán	20/02/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	23/02/2019	54
21093	Europa League	Estadio Ramón Sánchez-Pizjuán	07/03/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	10/03/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	31/03/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	04/04/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	13/04/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	25/04/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	03/05/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	18/05/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	30/08/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	22/09/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	29/09/2019	54
21093	Europa League	Estadio Ramón Sánchez-Pizjuán	03/10/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	20/10/2019	54
21093	Europa League	Estadio Ramón Sánchez-Pizjuán	24/10/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	27/10/2019	54
21093	LaLiga Santander	Estadio Ramón Sánchez-Pizjuán	02/11/2019	54
21093	Europa League	Estadio Ramón Sánchez-Pizjuán	28/11/2019	54
21093	Curso	Casa de la Ciencia Sevilla	28/09/2019	983



A raíz de la incorporación de nuevas fuentes de información el % de acierto del algoritmo se incrementó.





## Chat GPT – Preguntas y Respuestas



Set up prompt text to introduce

```
In [4]: prompt=str(input())
```

```
what is the temperature in london
```

```
In [ ]:
```

```
In [ ]:
```

```
In [5]: response = openai.Completion.create(  
        model="text-davinci-003",  
        prompt=prompt,  
        temperature=0.7,  
        max_tokens=256,  
        top_p=1,  
        frequency_penalty=0,  
        presence_penalty=0  
    )
```

```
In [6]: response
```

```
Out[6]: <OpenAIObject text_completion id=cmpl-6Z0dyIPkf6LrrXW8u4EAwDJnE7MTB at 0x174141710e0> JSON: {  
  "choices": [  
    {  
      "finish_reason": "stop",  
      "index": 0,  
      "logprobs": null,  
      "text": "\n\nThe current temperature in London is 8 degrees Celsius (46 degrees Fahrenheit)."  
    }  
  ],  
  "created": 1673801974,  
  "id": "cmpl-6Z0dyIPkf6LrrXW8u4EAwDJnE7MTB",
```

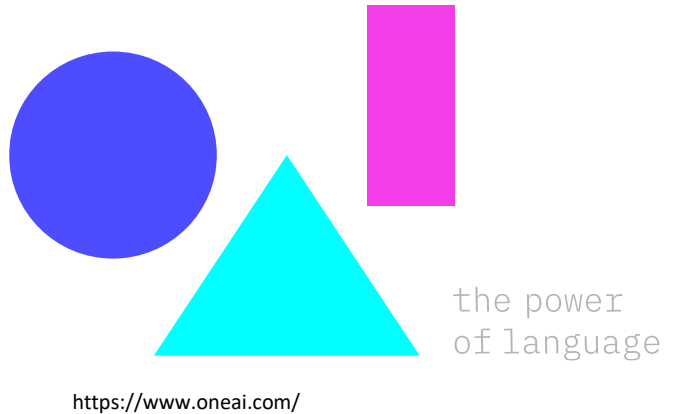
## One AI – Detectar Emociones, sentimientos, resumir, extraer palabras clave...

### Texto de Entrada (Web ING)

Fácil y rápida. O rápida y fácil. ¿Con tu huella o con tu mirada? Abre la app con un simple gesto, el que tú quieras. Llega antes a donde quieres. Si mantienes pulsado el icono de tu app, encontrarás accesos directos para llegar más rápido a lo que buscas.

### Resultado de Sentimientos

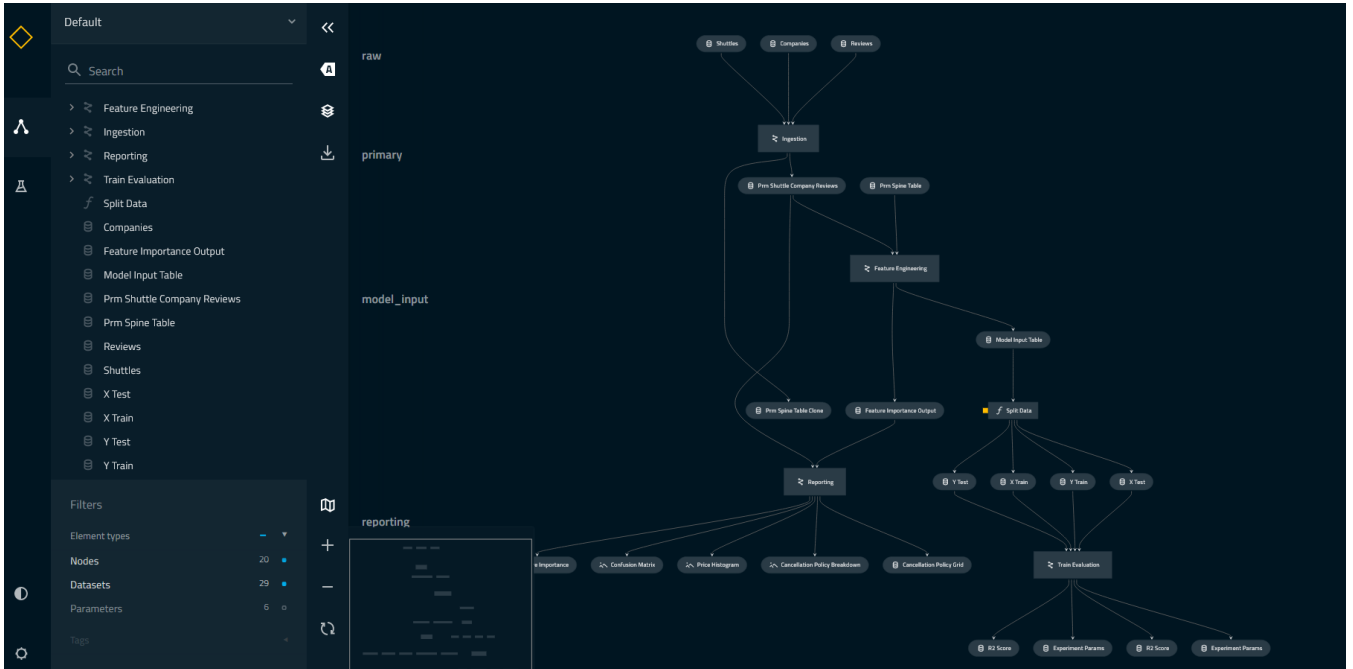
- **sentiments=**
- **type='sentiment'**
- **skill='sentiments'**
- **span\_text='Fácil y rápido'.**
- **value='POS'**



### Resultado de Emociones

Vacío

## Kedro librería experimentos visuales (pip install kedro)



[https://kedro.readthedocs.io/en/stable/get\\_started/new\\_project.html](https://kedro.readthedocs.io/en/stable/get_started/new_project.html)

